A Stochastic Control Perspective on Euclidean Quantum Field Theories

Sarah-Jean Meyer

Born 18th November 1997 in Saarlouis, Germany.

25th August 2022

Master's Thesis Mathematics

Advisor: Prof. Dr. Massimiliano Gubinelli

Second Advisor: Prof. Dr. Christian Brennecke

INSTITUT FÜR ANGEWANDTE MATHEMATIK

MATHEMATISCH-NATURWISSENSCHAFTLICHE FAKULTÄT DER

RHEINISCHEN FRIEDRICH-WILHELMS-UNIVERSITÄT BONN

Abstract

We construct the massive sine-Gordon measure on the infinite volume $\mathbb{R}^2$ for $\beta^2 < 4\pi$ using the variational method for Euclidean quantum field theories introduced by Barashkov and Gubinelli. Relying directly on the martingale structure of the renormalisation, the stochastic control problem is understood in terms of a forward-backward system of stochastic differential equations (FBSDE). The derived a priori estimates show that the FBSDE converges to a limit as the ultraviolet and the infrared cut-off are removed. For weak interactions, this limit is unique. As a result, we can construct the infinite volume sine-Gordon measure as the law of a coupled forward-backward system on $\mathbb{R}^2$. Similarly to previous results relying on the variational method, the estimates also enable a variational description for the Laplace transform in the infinite volume.

Acknowledgements

First and foremost, I would like to thank Massimiliano Gubinelli for his patience and encouragement and for introducing me to this exciting area of mathematics. This thesis would not have been possible without his exceptional support and I am very much looking forward to future projects.

I also wish to thank Christian Brennecke for agreeing to examine this thesis.

I thank Nikolay Barashkov for insightful comments on his preprint [2].

I am grateful to my family and friends for their support and countless discussions, mathematical and otherwise. Special thanks go to Yvonne Neuy and Cedric Kouayep Nguepnang for proofreading parts of this work.

The financial support provided by the Bonner Universitätsstiftung through the Deutschland Stipendium helped me free up more time for this thesis and is gratefully acknowledged.

Contents

1. Introduction

1.1 Euclidean Quantum Field Theories

Euclidean quantum field theories are special Borel measures on the space $\mathscr{S}'(\mathbb{R}^2)$ of Schwarz distributions. These measures can be understood as Gibbs measures on the continuum and they possess a great multitude of interesting characteristics. Like their discrete counterparts, they experience phase transitions and can arise as invariant measures of Hamiltonian dynamics and singular stochastic PDEs or scaling limits of discrete models. Originally, they were studied as Wick-rotations of the physically relevant Minkowski-space quantum field theories, which also gives them their name. The Wick-rotation comes down to evaluating quantum fields at an imaginary time to recover the Euclidean metric from the Minkowski metric on d-dimensional space-time. Euclidean quantum field theories are then the probability measures for which this transformation can be rigorously reversed. This idea is formalised by the Osterwalder-Schrader axioms and reconstruction theorem [53, 54] (or variants thereof). These axioms are conditions on the correlation functions of the probability measure to ensure that a quantum field theory (QFT; for short) can be recovered via analytic continuation and vice versa, see [31, Chapter 6] for a general introduction. For our purposes, we can content ourselves with the informal definition.

Definition. A Euclidean Quantum Field Theory is a probability measure ν on $\mathscr{S}'(\mathbb{R}^d)$ satisfying specific moment estimates in addition to Euclidean invariance and reflection positivity.

The combination of the three properties above puts serious constraints on the measure ν and the construction of "interesting" (i.e. interacting) quantum field theories poses a number of problems. Formally at least, a reflection positive measure can be constructed as a Gibbsian perturbation

$$\text{“ } \nu(d\varphi) = \frac{\exp(-S(\varphi))d\varphi}{\int \exp(-S(\varphi))d\varphi} \text{ ”}, \tag{1.1}$$

where $\varphi : \Lambda \to \mathbb{R}$ and S is an action functional, taking the form

$$S(\varphi) = \lambda \int_\Lambda V(\varphi) + m^2 \int_\Lambda \varphi^2 + \int_\Lambda |\nabla \varphi|^2,$$

for a suitable potential $V : \mathbb{R} \to \mathbb{R}$. As already indicated by the quotation marks, the notation in (1.1) is entirely suggestive and only used to convey the general idea: The problem already begins with the meaning of (1.1) in the absence of a Lebesgue measure $d\varphi$ in infinite dimensions. To still make sense of the measure ν, we observe that for the quadratic action, i.e. $\lambda = 0$, the measure $\mu(d\varphi) = \Xi^{-1} \exp(S_{\text{free}}(\varphi))d\varphi$ corresponds to a Gaussian measure, more precisely to the Gaussian free field. The free field is an example of a Euclidean QFT as was shown by Nelson [50], albeit a trivial one. However, the free field is well-understood and can also be realised on the infinite-dimensional space of Schwarz distributions $\mathscr{S}'(\mathbb{R}^d)$. As such, the free field presents itself as a starting point to construct more general QFTs and replaces the Lebesgue measure as a reference measure.

It is less clear how to proceed with the interaction $V(\varphi)$ when $\lambda \neq 0$. The first problem, an instance of the ultraviolet problem, arises from the irregularity at small scales. For $d \geq 2$, the construction of non-trivial Euclidean quantum fields is a difficult problem already on a finite volume due to the notorious and old problem of making sense of non-linear operations on distributions. The free field is only supported on a genuine distribution space, becoming more irregular as the dimension increases (with logarithmic divergencies in dimension $d = 2$, and polynomial divergencies for $d \geq 3$). For (1.1), this means that the value of the field φ at a point is not well-defined and we cannot make sense of $V(\varphi)$ in a pointwise manner. Initially, we might try to approximate the distribution φ with genuine functions φ_T by introducing a small-scale cut-off. Unfortunately, $V(\varphi_T)$ generally fails to converge to a non-trivial limit and the approximation picks up divergencies as the cut-off is removed. The standard procedure to save this seemingly hopeless situation is to combine the *regularisation* μ^T of the measure μ, with a *renormalisation* V_T of the potential V. This renormalisation should then lead to a theory which assigns finite values to the physically observable quantities. Whether this programme of ultraviolet-renormalisation can be carried out depends not only on the particular potential V but also on the dimension d. Additionally, applying a regularisation along all spacial directions breaks the reflection positivity of the measure and an additional argument is usually required to show that the limiting measure is reflection positive.

On the other end of the spectrum, the large-scale behaviour of the model introduces the largely unrelated infrared problem: To make the measure (1.1) invariant under translations, we would need to take $\Lambda = \mathbb{R}^d$. However, even if we can give a meaning to the random distribution $V_T(\varphi)$ for a typical sample $\varphi \sim \mu^T$, we cannot expect any decay in space and the integral $\int_{\mathbb{R}^2} V_T(\varphi)$ does not make sense. Therefore, also the density of the regularised measure ν^T is ill-defined at best in infinite volume and the analysis usually requires an additional cut-off $\xi : \mathbb{R}^2 \to [0,1]$ in space and yet another approximation $\nu^{T,\xi}$.

In summary, the first step towards a rigorous analysis of the measure ν is the construction of approximate problems on a finite volume whose solutions converge to a well-defined limit from which we hope to recover the measure ν. The objects of interest are thus the approximate measures

$$\nu^{T,\xi}(d\varphi) = \Xi_{T,\xi}^{-1} \exp\left(-\lambda \int \xi V_T(\varphi)\right) \mu^T(d\varphi),$$

and their limit as $T \to \infty$, $\xi \to 1$ in a suitable sense.

Despite the challenges involved in the construction of QFTs, much progress has been made and we cannot attempt to cover the literature here but will content ourselves with a few pointers. The mathematically rigorous construction of Euclidean QFTs of the form (1.1) dates back at least to the late '60s with the first constructions of the φ^4_2-theory on a finite volume by Nelson [49], which initially relied on the Markov property. Later on, the weaker condition of reflection positivity was introduced by Osterwalder Schrader [53, 54]. Over the next two decades, major progress was made ([7–9, 16, 17, 30]) and e.g.

resulted in a complete construction of the infinite volume φ^4_3 model with a full verification of the Osterwalder-Schrader axioms independently by Feldman and Osterwalder [29] and by Magnen and Sénéor [47] using cluster expansion techniques for small correlations λ. The constructions have since been revisited using various different methods. Thanks to the pioneering works of Da Prato and Debussche [61] in 2003 and more recent substantial extensions of their approach, building on ideas from the theory of rough paths, by Hairer [34] and by Gubinelli, Imkeller and Perkowski [33], the understanding of QFTs and their connections to singular stochastic PDEs increased rapidly in recent years. As a result, the field has seen great progress on problems previously well beyond reach and attracted considerable activity in the last decade. We only want to highlight the variational approach of Barahskov and Gubinelli [2–5] to study Gibbs measures as introduced (admittedly very vaguely) above, since this will be the point of view we take for this thesis. More generally, we refer again to [31] for a general introduction and a detailed treatment of the methods in the simpler setting of the quartic interaction in two dimensions and to [32] and the references therein for more recent developments.

1.2 The Variational Method for Euclidean QFTs

The starting point for this variational approach is a formula by Boué and Dupuis [13] for functionals of Brownian motion. By introducing an additional (fictitious) time parameter $t \in \mathbb{R}_+$ and a decomposition of the free field $W_T = \int_0^T Q_t dB_t$ in terms of a Brownian motion B_t, this formula provides a variational description for the logarithm of the Laplace transform of the approximate measure $\nu^{\xi,T}$ of the form

$$\mathcal{W}^{\xi,T}(g) := -\log \int e^{-g(\varphi)} \nu^{\xi,T}(d\varphi) = \mathcal{V}_T^{\xi,g} - \mathcal{V}_T^{\xi,0}, \tag{1.2}$$

where

$$\mathcal{V}_T^{\xi,g} := \inf_{u \in \mathcal{A}} \mathbf{E}\left[g(W_T + I_T(u)) + \lambda \int_\Lambda \xi V_T(W_T + I_T(u)) + \frac{1}{2}\int_0^T \|u_s\|^2_{L^2} ds \right]. \tag{1.3}$$

Here, ξ is a spacial cut-off which ensures convergence of the integral, T is the scale parameter and I_T is an operator behaving like $(-\Delta + m^2)^{-\frac{1}{2}}$, increasing the regularity by 1. The set of admissible controls $\mathcal{A}$ in this case is given by the space of all adapted, square-integrable stochastic processes. We can identify $\mathcal{V}_T^g$ as the value function of a stochastic control problem whose Hamilton-Jacobi-Bellmann equation is given by the Polchinski equation for the variance C_t of the Gaussian W_t,

$$\frac{d}{dt} v_t = \frac{1}{2}\Delta_{\dot{C}_t} v_t - \frac{1}{2}\|\nabla v_t\|^2_{\dot{C}_t}. \tag{1.4}$$

This equation can be understood as a continuous version of Wilson's renormalisation group ([63]) and is of particular interest because the (exact) renormalisation of the potential V should satisfy the equation above. First proposed by Polchinski [60], the systematic use of this continuous point of view had an important influence on the study of quantum field theories. We refer to [41] for more historical context and some relatively recent advances. In general, the direct analysis of (1.4) is difficult as the PDE is not only nonlinear but also infinite-dimensional. The stochastic control problem derived from the variational problem (1.3) offers an equivalent formulation and is well understood in the finite-dimensional setting. Unlike most of the current renormalisation group techniques, which are in some sense based on an expansion, this approach can capitalise on the special structure of the Polchinski equation, or more precisely, the martingale property of $V_T(W_T)$. As a result, this formulation is quite amenable to tools from stochastic analysis and enables a more direct study of the underlying QFT from a probabilistic point of view.

Originally, this program was applied to the φ^4_3-model ($V(\varphi) = \varphi^4$) in a finite volume [4], and has since proven fruitful in different contexts (see [5] and the references therein). More recently, the variational approach was also applied to infinite volume QFTs, which reveals another feature of the method. In the infinite volume, as we stressed earlier, the main difficulty comes from the integral $\int_{\mathbb{R}^2} V_T(\varphi)$ being ill-defined. However, the processes appearing in the stochastic control problem itself do not rely on the integral $\int_{\mathbb{R}^2} V_T(\varphi)$ and stay well-defined in infinite volume. This observation is leveraged in [2, 5], where uniform bounds on the optimal control in 2-dimensional Euclidean QFTs are obtained from stochastic calculus and Euler-Lagrange equations for the optimal control. We provide some more insight into the general strategy (adapted to our setting) in the next section.

More generally speaking, this Euler-Lagrange perspective on the variational method can also be understood as an instance of *stochastic quantisation* , whereby additional degrees of freedom are added to obtain an equation which characterises the measure (1.1): Once the existence of a minimiser u^T for $\mathcal{V}_T$ is shown, the variational description reduces to an equation for the measure ν^T. This point of view was first introduced using Langevin dynamics by Parisi and Wu [58], but has by now proven very fruitful with many different ways to stochastically quantise the target measure (see e.g. [21] for an early review). Compared to the more direct, Gibbsian interpretation (1.1), where the measure is understood via its density, this perspective is insensitive to the fact that the measures ν might not be absolutely continuous with respect to μ. The goal is then to use the quantisation equation describing the target measure ν as a tool to derive properties and uniform bounds to arrive at a description of the QFT.

1.3 Outline and General Strategy

In this thesis, the 2-dimensional massive sine-Gordon model with $\beta^2 < 4\pi$ serves as a test-bed to further explore the tradeoffs of the variational point of view. The sine-Gordon measure is the measure formally given by

$$\nu^{\xi}_{\mathrm{SG}}(d\varphi) = \frac{\exp(-V^{\xi}(\varphi))\mu(d\varphi)}{\int \exp(-V^{\xi}(\varphi))\mu(d\varphi)}, \quad \varphi \in \mathscr{S}'(\mathbb{R}^2),$$

where $\xi : \mathbb{R}^2 \to [0,1]$ is a smooth compactly supported spacial cut-off, $V^{\xi}(\varphi) = \lambda \int \xi \cos(\beta\varphi)$ is the interaction, μ is the Gaussian free field on $\mathbb{R}^2$ and $\mathscr{S}'(\mathbb{R}^2)$ is the space of Schwarz distributions. Note that, in contrast to the introduction, we have now included the spacial integral in the definition of V for notational convenience. The sine-Gordon model is a classic example of a non-Gaussian Euclidean QFT in two dimensions and some pointers and references will be provided at end of this section.

Our main goal is to show the following theorem using the stochastic control approach provided by the Boué-Dupuis formula.

Theorem 1.1. *There is a random variable $\mathcal{I}_\infty \in L^\infty(P, W^{1,\infty}(\mathbb{R}^2))$ such that*

$$\nu_{SG} = \mathrm{Law}(\mathcal{I}_\infty + W_\infty).$$

If $\lambda > 0$ is sufficiently small, $\mathcal{I}_\infty$ is unique and the measure ν_{SG} is non-Gaussian.

We now want to give a rough idea of how we are going to achieve this and mention some key estimates we obtain along the way. As already mentioned, the standard procedure to obtain the family of approximate measures comprises two components: A smooth approximation of the Gaussian free field μ^T and a renormalisation of the potential V_T. The renormalisation in the case of the sine-Gordon measure is given by the Wick-ordering (with respect to the covariance of μ^T) of the cosine and corresponds to a rescaling by a (divergent) factor α_T.

To transfer the problem to a stochastic control setting, we need a decomposition of the form

$$W_T = \int_0^T Q_t dB_t, \quad \mu^T := \mathrm{Law}(W_T), \tag{1.5}$$

where B is a cylindrical Brownian motion on $L^2(\mathbb{R}^2)$. While this decomposition is also linked to a scale decomposition, the more important point in the representation is the fact that we obtain an infinitely divisible decomposition of the Gaussian free field $\mu = \mathrm{Law}(W_\infty)$. This allows us to understand the problem via the Boué-Dupuis formula (1.2) and (1.3).

Once this decomposition and the variational description for $\nu^{T,\xi}$ are established, we can introduce the stochastic control problem

$$\begin{cases} X_t(u) = \varphi + \int_0^t Q_s u_s ds + \int_0^t Q_s dB_s, \\ Y_{t,T}(u) = (g + V_T^\xi)(X_T(u)) + \int_t^T \frac{1}{2}\|u_s\|_{L^2}^2 ds - \int_t^T Z_{s,T}(u) dBs, \end{cases}$$

where we want to minimise $Y_{0,T}(u)$. It is not hard to see that if an optimal control exists, it must satisfy

$$u_s^* = -Q_s \mathbf{E}[\nabla(g + V_T^\xi)(X_T(u^*))|\mathcal{F}_s].$$

On a finite volume, we can verify that there is a unique control satisfying this condition. Moreover, it follows from the variational description that the law of the optimally controlled process X (for a compactly supported cut-off ξ) is the approximate sine-Gordon measure $\nu_{\mathrm{SG}}^{\xi,T}$. Thus, showing Theorem 1.1 comes down to showing that the optimally controlled process has a limit as $T \to \infty$, $\xi \to 1$. Of course, all trivial bounds on the control and the optimally controlled process depend on V_T^ξ and degenerate in the limit as $T \to \infty$ and $\xi \to 1$. In other words, direct a priori estimates for the FBSDE will not be sufficient. To obtain better bounds, we essentially rely on two simple but important observations. First, the (optimal) dynamics of the system are captured entirely by

$$\begin{cases} X_{t,T} &= \varphi - \int_0^t Q_s^2 \nabla Y_{s,T} ds + \int_0^t Q_s dB_s, \\ \nabla Y_{t,T} &= \mathbf{E}[\nabla(g + V_T^\xi)(X_{T,T}) | \mathcal{F}_t], \end{cases}$$

which does not involve the potential $V_T^\xi(\varphi) = \int \xi \lambda \alpha_T \cos(\beta\varphi)$ but only its gradient $\nabla V_T^\xi(\varphi) = -\xi\beta\lambda\alpha_T \sin(\beta\varphi)$. The latter is well defined also for $\xi \equiv 1$ which allows us to derive bounds that are uniform in the volume. Secondly, the Wick-renormalisations $[\![\sin(W_T)]\!]$ and $[\![\cos(W_T)]\!]$ are martingales. Understanding that the process X is essentially a small perturbation of W motivates the change of variables

$$R_{t,T} = \nabla Y_{t,T} - \nabla V_t^\xi(X_{t,T}).$$

Using the martingale property of $\nabla V_t(W_t)$, we see that R is the unique solution to a BSDE of the form

$$R_{t,T} = \nabla g(X_{T,T}) + \int_t^T h^\xi(s, X_{s,T}, R_{s,T}) ds - \int_t^T \tilde{Z}_s dB_s.$$

Importantly, we will see that the data g, h in the BSDE do no longer dependent on T explicitly. This, combined with the estimate $\|Q_t f\|_{L^p} \lesssim \langle t \rangle^{-1} \|f\|_{L^p}$, allows us to derive bounds on the remainder R that do not depend on T and ξ. As a direct consequence, we also obtain uniform estimates on the drift $\mathcal{I}_t = \int_0^t Q_s^2 \nabla Y_s ds$ of X and thus convergence for $\nu_{\mathrm{SG}}^{T,\xi}$.

Theorem 1.2. *The family of random variables $\{\mathcal{I}^{T,\xi}\}_\xi \subset L^\infty(P, W^{1,\infty}(\mathbb{R}^2))$ such that*

$$\mathrm{Law}(W_T + \mathcal{I}_T^{T,\xi}) = \nu_{SG}^{T,\xi},$$

is uniformly bounded in $T \in [0, \infty]$ and $\xi \in C_c^\infty(\mathbb{R}^2; [0,1])$, or more precisely,

$$\sup_{T,\xi} \|\mathcal{I}^{T,\xi}\|_{W^{1,\infty}} < \infty$$

almost surely. Moreover, if $\lambda > 0$ is sufficiently small, the family $\{\mathcal{I}^\xi\}_{T,\xi}$ converges in $L^2(P, H^1(\langle x \rangle^{-\ell}))$ to a unique limit as $T \to \infty$ and $\xi \to 1$.

While the convergence to a unique limit requires $\lambda > 0$ to be small, the bound on the norm of $\mathcal{I}_\infty^\xi$ remains valid for any λ and still guarantees tightness and consequently convergence along suitable subsequences. Moreover, as $\mathcal{I}^{\xi,T}$ is bounded uniformly in $L^\infty(P, W^{1,\infty})$, this shows that any accumulation point of $\nu_{\mathrm{SG}}^{\xi,T}$ has Gaussian Tails.

On a finite volume and for $\beta^2 < 4\pi$, the sine-Gordon measure is absolutely continuous with respect to the free field, and the Gibbsian representation (1.1) makes sense. However, this property is lost for $\beta^2 \geq 4\pi$ or once the volume cut-off is removed. As a result, on the infinite volume, alternative descriptions gain significance. For functions g whose gradient is small outside a compact set, the bounds also allow us to derive a variational description for the Laplace transform on the infinite volume (c.f. Theorem 4.52), which we state below.

Theorem 1.3. *Assume that $\lambda > 0$ is sufficiently small and let $n \geq 2$. For a functional $g \in C_b^2(H^{-\delta}(\langle x\rangle^n)) \cap C_b^2(L^2(\langle x\rangle^n))$ with $\sup_{\varphi\in L^2(\mathbb{R}^2)} \|\nabla g(\varphi)\|_{L^2(\langle x\rangle^n)} < \infty$, we have the following variational description for the Laplace transform*

$$\mathcal{W}(g) = \lim_{\xi\to 1}\lim_{T\to\infty} \mathcal{W}^{\xi,T}(g) = \inf_{v\in\mathcal{A}(g)} \overline{J}^g(v),$$

with the cost functional

$$\overline{J}^g(v) = \mathbf{E}\left[g(X_\infty(\overline{u}+v)) + V_\infty(X_\infty(\overline{u}+v)) - V_\infty(X_\infty(\overline{u})) + \mathcal{E}(\overline{u},v)\right].$$

Here, $\mathcal{W}^{\xi,T}(g)$ is defined as in (1.2) and $X_\infty(u) := I_\infty(u) + W_\infty$ is the shifted free field. Moreover,

- *$\overline{u}$ is an adapted stochastic process which does not depend on g, v,*
- *I_∞ is a linear functional, which increases regularity by 1 and does not depend on g,*
- *$\mathcal{E}$ is a quadratic functional, also independent of g, and*
- *$\mathcal{A}(g)$ contains the adapted controls v such that $\mathbf{E}\int_0^\infty \|v_s\|^2_{L^2(\langle x\rangle^n)} ds \leq C_{\nabla g}$.*

Even though the approach sketched above very much follows the ideas in [2] and [5], we provide a different interpretation of the control problem in terms of BSDEs. This perspective further emphasises the stochastic analysis point of view and makes very explicit use of the martingale structure of the renormalisation. As a consequence, even with estimates that are quite elementary as far as the theory of BSDEs is concerned, we can obtain the results in [2]. While the estimates are then applied to derive e.g. the variational representation, the main contribution is the use of the a priori bounds on the BSDE for R to control the weak limit. The preprint [2] illustrates further applications of the estimates, including a full verification of the Osterwalder-Schrader axioms and a large deviation principle. As the approach we take here seems to simplify the estimates, it seems tractable to extend the results also to the second threshold $\beta^2 < 6\pi$ or beyond.

Finally, we want to make some remarks on the technical limitations in the approach taken here and provide some references for the sine-Gordon model. The estimates rely heavily on the bound $\beta^2 < 4\pi$ which implies that the renormalisation given by the Wick-orderd cosine converges in $H^{-1+\delta}(\langle x\rangle^{-\ell})$ for ℓ sufficiently large and in turn also that $V_\infty(W_\infty + I_\infty(u))$ is a well-defined distribution via the dual pairing of the Sobolev spaces H^1 and H^{-1}. Additionally, we use the explicit bounds on the renormalisation constant which also relies on the restriction on the parameter β^2 to $(0, 4\pi)$. The above problems are an expression of the fact that the parameter β^2 plays an essential role in the small-scale behaviour (i.e. the limit of $T \to \infty$): Knowing that the correlations of the free field are logarithmic, we expect that "$\cos\beta W = \Re\exp(i\beta W)$" has polynomial correlations, whose degree depends on β. In this way, the parameter β influences the severity of the divergencies and plays a similar role as the dimension for the polynomial φ^4_d-model. As β^2

passes the thresholds $8\pi(1-\frac{1}{2n})$, $n=1,2,\dots$, the partition function picks up additional divergencies which require additional renormalisation until the theory is no longer renormalisable for $\beta \geq 8\pi$. For a detailed discussion of the influence of β^2, we refer to [7, 43, 51]. At least on a finite volume (i.e. the Torus $\mathbb{T}^2$), the convergence of the approximate measures $\nu^T \to \nu$ for the optimal regime $\beta^2 < 8\pi$, was achieved by various different approaches (e.g. [7, 17, 24, 25]). Additional references and context on the sine-Gordon model can e.g. be found in [6] or [43].

It would be interesting to also extend the approach taken here also to the next threshold $\beta^2 < 6\pi$ or even the optimal regime $\beta^2 < 8\pi$. Finally, to pass to the infinite volume $\mathbb{R}^2$, we rely on the positive mass $m > 0$ to obtain the required decay in space. In the future, it might also be interesting to try and refine this approach to cover the formally massless sine-Gordon model (that is the limit $m \to 0$).

Structure

In Chapter 2 and Chapter 3, we give a brief introduction to stochastic control and how stochastic control problems can be studied via forward-backwards stochastic differential equations (FBSDE, for short). Many proofs are omitted but some instructive proofs are provided in detail even if the standard results contain only minor modifications to the Hilbert space setting. While we do not claim originality, we present some additional examples and context to prepare for the applications in the next chapter.

In Chapter 4, we return to the sine-Gordon EQFT. The first goal is to obtain the decomposition of the free field and thus the approximate measures $\nu_{\mathrm{SG}}^{T,\xi}$ in Section 4.1. We then draw the connection to the stochastic control problem and BSDEs in Section 4.2. By relaxing the variational problem, we show the existence of a minimiser and derive a coupled system for the optimally controlled process X_t and the value function $\mathcal{V}_T$ in Section 4.3. Using the martingale property of the Wick-ordered cosine, we derive uniform a priori estimates in Section 4.4. These estimates are applied in Section 4.5 to show convergence in T and ξ. Along the lines of [2], we apply our estimates to derive a variational description for the Laplace transform on the infinite volume and show that the limiting measure is non-trivial in Section 4.6 and Section 4.7. We also include some initial estimates beyond the $\beta^2 < 4\pi$ threshold.

The Appendix A and Appendix B contain some definitions and results on Besov spaces and some elementary properties of Gaussian measures.

Notation

The Lebesgue measure of a set $A \subset \mathbb{R}^2$ will be denoted by $|A|$. Since we will be concerned with the infinite volume model, we also have to introduce some weighted spaces. Let $\langle x \rangle := (1+|x|^2)^{\frac{1}{2}}$ and $\rho_\ell(x) := \langle x \rangle^\ell$. For $w \in L^1_{\mathrm{loc}}$, we define the weighted space $L^p(w(x))$ with the norm

$$\|u\|^2_{L^p(w(x))} := \int_{\mathbb{R}^2} u^p(x) w^p(x) dx.$$

The weighted Sobolev spaces $W^{s,p}(w(x))$ and $H^s(w(x))$ for $s \in \mathbb{R}$ are defined analogously. We will also use the Besov spaces $B^s_{p,r}$, for which we provide the required background and notation in the Appendix.

Integration over the variable $x \in \mathbb{R}^d$ is usually left implicit, i.e. for a function $f \in L^1(\mathbb{R}^d)$ we write $\int_{\mathbb{R}^d} f := \int_{\mathbb{R}^d} f(x)dx$ as long as no ambiguities arise. Given a measure μ and $f \in L^1(\mu)$, we may also use the common shorthand $\mu(f) := \int f \, d\mu$.

For real and separable Hilbert spaces J, H, K, we denote by $C^k(K; H)$ the space of all continuous maps $F : H \to K$ which are k-times continuously Fréchet differentiable. We write $\nabla F : K \to L(K, H)$ and $\operatorname{Hess} F : K \to L(K, L(K, H))$ for the first and second-order Fréchet derivatives respectively to reserve the symbol D for the Malliavin derivative. For $H = \mathbb{R}$, we identify $\nabla F(k)$ with $k \in K$ with the unique element in K provided by the Riesz representation theorem. Likewise, the second derivative $\operatorname{Hess} F(k)$ is understood as a self-adjoint linear operator $K \to K$. The space of Hilbert-Schmidt operators from $K \to H$, equipped with the Hilbert-Schmidt norm $\|z\|^2_{\mathcal{L}_2} = \operatorname{Tr}_H(zz^*)$ will be denoted by $\mathcal{L}_2(K, H)$. For $f : J \times H \to K, (x, y) \mapsto f(x, y)$, we may also use the notation $\partial_x f$ (respectively ∂_y) to denote the Fréchet derivative of f with respect to the first (respectively second) component. Given elements x_1, x_2 of a generic Hilbert space H, we denote their inner product by $\langle x_1, x_2 \rangle$, and the induced norm by $|x_1|^2 := \langle x_1, x_1 \rangle$, where we leave the Hilbert space implicit to avoid additional clutter as long as no ambiguities arise.

The process B will always denote a cylindrical Brownian motion on a separable Hilbert space and unless specifically indicated otherwise, all considerations are with respect to the augmentation of the filtration $\{\mathcal{F}_t\}_{t \geq 0}$ generated by B.

2. Background on BSDEs

Backward stochastic differential equations (BSDE, for short) are well-suited to study stochastic control problems and have been studied extensively for finite-dimensional control problems. First introduced in 1973 by Bismut [10, 11] in the context of stochastic optimal control, well-posedness for a broader class of SDEs with a terminal condition was only shown by Pardoux and Peng in 1990 [56] which initiated the systematic study of these equations. Motivated by the fruitful applications and connections to mathematical finance and stochastic control theory, the theory progressed at a rapid pace in the 1990s and has since seen widespread applications.

In this chapter, we give a (very incomplete) overview of the theory and recall some basic properties and definitions for BSDEs, their connection to Hamilton-Jacobi-Bellman equations, and stochastic control. The material in this chapter is very standard in the literature and we refer to the survey articles [27, 55] and the book [65] for the theory of BSDEs. For a more general perspective on stochastic control, we refer to [64] for the classical theory and to [28] for an adaptation to the infinite-dimensional case. As a general reference for stochastic calculus on Hilbert spaces, we refer to [20, 45].

2.1 Motivation

SDEs with a terminal condition

To motivate the need for a dedicated theory for SDEs with a terminal condition, we can look at a simple example. Given a Brownian motion B on a probability space $(\Omega, \mathcal{F}, P)$ with its natural filtration, consider the differential equation on $[0, T]$ with a terminal condition

$$\begin{cases} dY_t = 0, & 0 \leq t \leq T, \\ Y_T = \xi. \end{cases}$$

If ξ is deterministic, the unique solution is of course given by the constant function $Y_t \equiv \xi$. For a general random variable $\xi \in \mathcal{F}_T$, this is still the only possible candidate for a solution. However, we usually require a solution to an SDE to be adapted. If we want to keep adaptedness for the SDE with a terminal condition, then a sensible adaptation of the solution is given by the square-integrable martingale $Y_t = \mathbf{E}[\xi|\mathcal{F}_t]$. By the martingale representation theorem, there exists a unique square-integrable, adapted process Z such that

$$Y_t = Y_0 + \int_0^t Z_s dB_s.$$

This leads to the following reformulation of the SDE with a terminal condition

$$\begin{cases} dY_t = Z_t dB_t, \\ Y_T = \xi, \end{cases} \qquad \text{or equivalently} \quad Y_t = \xi - \int_t^T Z_s dB_s.$$

Notice that a solution now consists of a *pair* (Y,Z) of adapted processes. It is thanks to the additional martingale part Z that we can find adapted solutions to stochastic terminal value problems. Allowing for more generality in the differential equation, we arrive at the following problem formulation for terminal value problems: Given a *generator* f, a terminal time T and a terminal condition ξ, we look for a pair (Y,X) of adapted processes such that

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dB_s, \quad 0 \le t \le T.$$

Non-uniqueness without square-integrability

We may be concerned to lose uniqueness because we are now looking for two processes (Y,Z) as a solution to the differential equation. Indeed, without requiring the processes to be square-integrable, we cannot expect uniqueness. This can already be observed in this simplest example for $(f,\xi) = (0,0)$. Then, the unique square-integrable solution is given by $(Y,Z) = (0,0)$. Clearly, this is also the solution provided by the martingale representation theorem as we have seen in the example above. However, as $k(t,x) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{2t}}$ solves the heat equation, Itô's formula shows that

$$k(T-t, B_t) = -\int_t^T \partial_x k(T-s, B_s) dB_s.$$

In the class of square-integrable processes, the martingale representation theorem guarantees uniqueness in this simple case. It is a well-known fact that for Lipschitz BSDEs square-integrability still provides uniqueness (c.f. Theorem 2.2).

2.2 Adapted Solutions to SDEs with a Terminal Condition

Let K, H be two real and separable Hilbert spaces. For a cylindrical Brownian motion $\{B_t;\, t \ge 0\}$ on K, we denote by $\{\mathcal{F}_t\}$ the augmentation of the filtration generated by B. For Hilbert spaces $\mathcal{H}, \mathcal{H}'$, we recall the notation $\mathcal{L}_2(\mathcal{H}, \mathcal{H}')$ for the space of Hilbert-Schmidt operators from $\mathcal{H} \to \mathcal{H}'$, equipped with the Hilbert-Schmidt norm $\|z\|^2 := \|z\|^2_{\mathcal{L}_2(\mathcal{H},\mathcal{H}')} := \mathrm{Tr}_{\mathcal{H}'}(zz^*)$. Given $\theta \in \mathbb{R}$, $p > 0$, a terminal time $T \in (0, \infty]$ and an adapted process $\varphi : \Omega \times [0,T] \to \mathcal{H}$, let

- $L^p(\mathcal{F}_t; \mathcal{H}) = L^p(\Omega, \mathcal{F}_t, P; \mathcal{H})$ be the space of all $\mathcal{F}_t$-measurable random variables $\Omega \to \mathcal{H}$ with finite p-th moments,
- $\|\varphi\|^2_\theta := \mathbf{E} \int_0^T e^{\theta t} |\varphi_t|^2_{\mathcal{H}} dt$, and $\mathbb{H}^2_{T,\theta}(\mathcal{H})$ the space of all adapted processes φ such that $\|\varphi\|^2_\theta < \infty$,
- $\|\varphi\|^2_{H^1_{T,\theta}} := \mathbf{E} \sqrt{\int_0^T e^{\theta t} |\varphi_t|^2_{\mathcal{H}} dt}$, and $\mathbb{H}^1_{T,\theta}(\mathcal{H})$ the space of all adapted processes φ such that $\|\varphi\|_{\mathbb{H}^1_{T,\theta}(\mathcal{H})} < \infty$,

- $\|\varphi\|_{H^\infty_{T,\theta}} := \mathbf{E}\left[\sup_{t\in[0,T]}|\varphi_t|^2_{\mathcal{H}}\right]$ and $\mathbb{H}^\infty_{T,\theta}(\mathcal{H})$ the space of all adapted processes φ such that $\|\varphi\|_{\mathbb{H}^\infty_{T,\theta}(\mathcal{H})} < \infty$.

We may also write $\mathbb{H}^2_{\infty,\theta}(\mathcal{H}) := \mathbb{H}^2_\theta(\mathcal{H})$ and omit the subscript θ for $\theta = 0$. With this notation, we can associate to any process $Z \in \mathbb{H}^2_T(\mathcal{L}_2(K,H))$ a H-valued stochastic integral

$$\int_0^t Z_s dB_s; \quad 0 \le t \le T,$$

and for any orthonormal basis e_i of K, the sequence $\{\sum_{i=1}^n \int_0^t Z_s e_i \langle e_i, dB_s\rangle\}_{n\in\mathbb{N}}$ approximates the integral in $L^2(P;H)$. The particular choice of filtration means that we will also have access to the martingale representation theorem Theorem 2.4. Given a *generator* $f : \Omega \times \mathbb{R}^+ \times H \times \mathcal{L}_2(K,H) \to H$, a terminal time $T > 0$ and a terminal condition $\xi \in \mathcal{F}_T$, we are interested in studying adapted solutions to

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dB_s. \tag{2.1}$$

Definition 2.1. A *solution* to (2.1) is a pair (Y,Z) of adapted processes such that

- $t \mapsto Y_t$ is continuous,
- $Z \in \mathcal{L}_2(K,H)$ with $\int_0^T \|Z_s\|^2 ds < \infty$ almost surely, and,
- the equation (2.1) holds almost surely.

As discussed earlier, we can in general not expect uniqueness without asking for additional integrability. Thus, we will always assume that the terminal condition is square-integrable, that is $\xi \in L^2(\mathcal{F}_T;H)$. For our purposes, it is sufficient to consider generators f such that $f(\cdot,0,0) \in \mathbb{H}^2_T(H)$ and which are uniformly Lipschitz in y,z; more precisely, there is a constant $L > 0$ such that $dP \otimes dt$-almost surely,

$$|f(\omega,t,y_1,z_1) - f(\omega,t,y_2,z_2)| \le L(|y_1-y_2| + \|z_1-z_2\|).$$

We will also call such a pair (f,ξ) *standard parameters*. There are now many results with weaker assumptions on the generator to account for typical applications (e.g. only assuming continuity and linear growth [44], allowing for quadratic growth in Z initiated in [40], assuming stochastic Lipschitz conditions [26]). While BSDEs arising in the context of stochastic control usually have quadratic growth in Z, we are only concerned with the simplest quadratic special case and thus only present the classical setting of Lipschitz generators.

We recall the result and its proof in the Hilbert space setting here.

Theorem 2.2 (Pardoux-Peng 1990). *Given standard parameters (f,ξ), there is a unique solution $(Y,Z) \in \mathbb{H}^2_T(H) \times \mathbb{H}^2_T(\mathcal{L}_2(K,H))$.*

A simple and direct proof can be given using a priori estimates, which are also useful beyond this setting, and a fixed point argument. We first state some preliminary lemmata and standard tools from stochastic calculus which also allow us to fix some notation.

Lemma 2.3 (Itô's formula). *Let $\{X_t; 0 \le t \le T\}$ be a H-valued semimartingale satisfying*

$$X_t = X_0 + \int_0^t F_s ds + \int_0^t Z_s dB_s,$$

for some $F \in \mathbb{H}^2_T(H)$ and $Z \in \mathbb{H}^2_T(\mathcal{L}_2(K,H))$. Then, for any $\Psi \in C^2(H;J)$ and $t \in [0,T]$,

$$\Psi(X_t) = \Psi(X_0) + \int_0^t \nabla\Psi(X_s) F_s ds + \int_0^t \nabla\Psi(X_s) Z_s dB_s + \frac{1}{2}\int_0^t \tilde{\mathrm{tr}}(\mathrm{Hess}\,\Psi(X_s)) Z_s ds.$$

Here, we use the notation

$$\tilde{\mathrm{tr}}(\Theta)Z := \sum_{i=1}^{\infty} \Theta(Zk_i, Zk_i),$$

where $\{k_i\}_i$ an orthonormal basis for B. In particular, for $K = H$ and $J = \mathbb{R}$, we have

$$\Psi(X_t) = \Psi(X_0) + \int_0^t \langle \nabla\Psi(X_s), F_s \rangle ds + \int_0^t \langle \nabla\Psi(X_s), Z_s dB_s \rangle + \frac{1}{2}\int_0^t \mathrm{Tr}(Z_s Z_s^* \,\mathrm{Hess}\,\Psi(X_s)) ds.$$

Proof. See [19, Theorem 3.8 and Corollary 3.34]. □

Theorem 2.4 (Martingale Representation Theorem). *For $\xi \in L^2(\mathcal{F}_T; H)$, there is a unique $Z \in \mathbb{H}^2_T(\mathcal{L}_2(K,H))$ such that*

$$\xi = \mathbf{E}[\xi] + \int_0^T Z_s dB_s.$$

Proof. Verbatim from finite-dimensional version, c.f. [39, Theorem 3.4.2] □

Theorem 2.5 (Burkholder-Davis-Gundy). *Let M be a H-valued càdlàg local martingale with $M_0 = 0$. For any $p \in [1, \infty)$, there are universal constants $\tilde{c}_p, c_p$ such that*

$$\tilde{c}_p \,\mathbf{E}\langle M \rangle_t^{\frac{p}{2}} \le \mathbf{E}[\sup_t |M|^p] \le c_p \,\mathbf{E}\langle M \rangle_t^{\frac{p}{2}},$$

where $\langle M \rangle_t$ denotes the quadratic variation of M.

Proof. See e.g. [48] for a proof with constants independent of the dimension. □

Lemma 2.6. *For $\varphi \in \mathbb{H}^1_T(\mathcal{L}_2(K,H))$, the process $M = \{\int_0^t \varphi_s dB_s,\ t \in [0,T]\}$ is a martingale.*

Proof. Since $P(\int_0^T \|\varphi_t\|^2 dt < \infty) = 1$, the process M is a local martingale. Moreover, by the Burkholder-Davis-Gundy inequality,

$$\mathbf{E}[\sup_{t\le T} |M_t|] \le c\,\mathbf{E}\left[\left(\int_0^T \|\varphi_s\|^2 ds\right)^{\frac{1}{2}}\right] < \infty.$$

Then, we may conclude by dominated convergence applied to a localising sequence $\{M_{\tau_n \wedge t}, n \in \mathbb{N}\}$. □

Lemma 2.7. *Suppose* (Y, Z) *is a solution to* (2.1) *with standard parameters* (f, ξ)*. Then* $\sup_{t\le T} \|Y_t\| \in L^2(\mathcal{F}_T; H)$ *and the process* $\{\int_0^t \langle Y_s, Z_s dB_s\rangle,\ 0 \le t \le T\}$ *is a martingale.*

Proof. By Jensen's inequality and the Lipschitz continuity of f,

$$\mathbf{E}\left[\left(\int_0^T |f(s, Y_s, Z_s)|\, ds\right)^2\right] \le TL^2(\|Y\|_0^2 + \|Z\|_0^2) + 2T\|f(\cdot, 0, 0)\|_0^2 < \infty.$$

For the stochastic integral, by Itô's isometry and the Burkholder-Davis-Gundy inequality

$$\begin{aligned} &\mathbf{E}\left[\sup_{t\le T}\left|\int_t^T Z_s dB_s\right|^2\right] \\ \le& 2\,\mathbf{E}\left[\left|\int_0^T Z_s dB_s\right|^2\right] + 2\,\mathbf{E}\left[\sup_{t\le T}\left|\int_0^t Z_s dB_s\right|^2\right] \\ \le& C\,\mathbf{E}\int_0^T \|Z_s\|^2 ds < \infty. \end{aligned}$$

Since

$$\sup_{t\le T} |Y_t| \le |\xi| + \int_0^T |f(s, Y_s, Z_s)|\, ds + \sup_{t\le T}\left|\int_t^T Z_s dB_s\right|,$$

and $\xi \in L^2(\mathcal{F}_T; H)$ by assumption, we conclude $\mathbf{E}[\sup_{t\le T} |Y_t|^2] < \infty$.

Finally, the stochastic integral is a martingale as

$$\mathbf{E}\left[\left(\int_0^T |Y_s|^2 \|Z_s\|^2 ds\right)^{\frac{1}{2}}\right] \le \mathbf{E}\left[\sup_{t\le T} |Y_t|^2\right]^{\frac{1}{2}} \mathbf{E}\left[\int_0^T \|Z_s\|^2 ds\right]^{\frac{1}{2}} < \infty,$$

which implies the claim by Lemma 2.6. □

Remark 2.8. In the same way, we can show finite p-th moments for any $p \ge 2$ assuming the data (f, ξ) has finite p-th moments.

We are now ready to derive the a priori bounds and Theorem 2.2.

Proposition 2.9. *For $i \in \{0,1\}$, let (Y^i, Z^i) be a square-integrable solution to (2.1) with standard parameters (f^i, ξ^i) respectively and suppose that L is a Lipschitz constant for f^1. We introduce the notation for the spread between the two solutions*

$$(\delta Y_t, \delta Z_t) := (Y_t^1 - Y_t^2, Z_t^1 - Z_t^2) \;\; \text{and} \;\; \delta_2 f_t := f^1(t, Y_t^2, Z_t^2) - f^2(t, Y_t^2, Z_t^2).$$

Then,

(i) for any (λ, μ, θ) such that $\mu > 0$, $\lambda^2 > K$ and $\theta \geq K(2+\lambda^2)+\mu^2$, it follows that

$$\begin{aligned} \|\delta Y\|_\theta^2 &\leq T\left[e^{\theta T}\,\mathbf{E}[|\delta Y|^2] + \frac{1}{\mu^2}\|\delta_2 f\|_\theta^2\right], \\ \|\delta Z\|_\theta^2 &\leq \frac{\lambda^2}{\lambda^2 - L}\left[e^{\theta T}\,\mathbf{E}[|\delta Y_T|^2] + \frac{1}{\mu^2}\|\delta_2 f\|_\theta^2\right], \end{aligned}$$

(ii) there is a constant $C > 0$ such that

$$\mathbf{E}\left[\sup_{t\leq T}|\delta Y_t|^2 + \int_0^T \|\delta Z_t^2\| ds\right] \leq C\,\mathbf{E}\left[|\delta Y_T|^2 + \int_0^T |\delta_2 f_t|^2\,dt\right].$$

Proof. For for any $t \leq s \leq T$, the difference $(\delta Y, \delta Z)$ satisfies the SDE

$$\delta Y_s = \delta Y_t - \int_t^s (f^1(r, Y_r^1, Z_r^1) - f^2(r, Y_r^2, Z_r^2))dr + \int_t^s \delta Z_r dB_r.$$

Note that in the above we use the *forward* version of the equation so that we may apply Itô's formula in the usual way for this first proof without worrying about sign changes. In this way, we obtain after integrating from t to T,

$$\begin{aligned} e^{\theta T}|\delta Y_T|^2 - e^{\theta t}|\delta Y_t|^2 = &\int_t^T e^{\theta s}\left(\theta\,|\delta Y_s|^2 + \|Z_s\|^2\right)ds \\ &-2\int_t^T e^{\theta s}\langle \delta Y_s, f^1(s, Y_s^1, Z_s^1) - f^2(s, Y_s^2, Z_s^2)\rangle ds \\ &+2\int_t^T e^{\theta s}\langle \delta Y_s, Z_s dB_s\rangle. \end{aligned} \tag{2.2}$$

Thanks to the Lipschitz continuity we have for any $\lambda, \mu > 0$,

$$\begin{aligned} &2\langle \delta Y_s, f^1(s, Y_s^1, Z_s^1) - f^2(s, Y_s^2, Z_s^2)\rangle \\ \leq &2L\,|\delta Y_s|^2 + 2L\,|\delta Y_s|\,\|Z_s\| + 2\,|\delta Y_s|\,|\delta_2 f_s| \\ \leq &2L\,|\delta Y_s|^2 + L\left(\lambda^2\,|\delta Y_s|^2 + \lambda^{-2}\|Z_s\|^2\right) + \mu^2\,|\delta Y_s|^2 + \mu^{-2}\,|\delta_2 f_s|. \end{aligned}$$

Gathering terms, this implies

$$
\begin{aligned}
e^{\theta T}|\delta Y_T|^2 \leq & e^{\theta t}|\delta Y_t|^2 + (L(2+\lambda^2)+\mu^2-\theta)\int_t^T e^{\theta s}|\delta Y_s|^2\,ds \\
&+ \frac{L-\lambda^2}{\lambda^2}\int_t^T e^{\theta s}\|\delta Z_s\|^2 ds + \mu^{-1}\int_t^T e^{\theta s}|\delta_2 f_s|^2\,ds \\
&+ 2\int_t^T e^{\theta s}\langle \delta Y_s, \delta Z_s dB_s\rangle.
\end{aligned} \tag{2.3}
$$

(i) By Lemma 2.7, the stochastic integral in (2.2) is a martingale. For $t=0$, rearranging (2.3) and taking expectation yields the estiamte on $\|\delta Z\|_\theta$, provided $\lambda^2 > L$. To obtain the estimate on $\|\delta Y\|_\theta^2$, we choose $\theta > L(2+\lambda^2)+\mu^2$ to obtain

$$
\mathbf{E}\left[e^{\theta t}|\delta Y_t|^2\right] \leq \mathbf{E}\left[e^{\theta T}|\delta Y_T|^2\right] + \mu^{-1}\|\delta_2 f\|_\theta^2.
$$

Then, the claim follows after integrating from $t=0$ to $t=T$.

(ii) By the Burkholder-Davis-Gundy inequality, in the same way as in the proof of Lemma 2.7,

$$
\mathbf{E}\left[\sup_{t\leq T}\int_t^T \langle \delta Y_s, \delta Z_s dB_s\rangle\right] \leq \frac{1}{4}\mathbf{E}\left[\sup_{t\leq T}|\delta Y_t|^2\right] + 4c\|Z\|_0^2.
$$

Thus, taking the supremum and expectation in (2.3), we obtain for $\lambda^2 = 2L$, $\mu = 1$ and $\theta = 0$,

$$
\begin{aligned}
\mathbf{E}\left[\sup_{t\leq T}|\delta Y_T|^2\right] \leq & \mathbf{E}|\delta Y_T|^2 + \|\delta_2 f\|_0^2 + C\left(\|\delta Y\|_0^2 + \|\delta Z\|_0^2\right) \\
&+ \frac{1}{2}\mathbf{E}\left[\sup_{t\leq T}|\delta Y_T|^2\right]
\end{aligned}
$$

By the equivalence of the norms $\|\cdot\|_\theta$ on the compact time interval, we know from (i), that for some $C>0$,

$$
\|\delta Y\|_0^2 + \|\delta Z\|_0^2 \leq C\,\mathbf{E}|\delta Y_T|^2 + \|\delta_2 f\|_0^2.
$$

Since all quantities are finite due to Lemma 2.7, the claim follows.

□

In practice, for non-standard parameters, one may often try to follow the general idea of this proof for a given BSDE to obtain estimates on the solution. In the Lipschitz case, choosing the parameter θ appropriately, we can employ a fixed point argument to directly conclude global existence and uniqueness for the BSDE and prove Theorem 2.2.

Proof of Theorem 2.2. We will construct a contraction

$$\Phi : \mathbb{H}^2_T(H) \times \mathbb{H}^2_T(\mathcal{L}_2(K,H)) \to \mathbb{H}^2_T(H) \times \mathbb{H}^2_T(\mathcal{L}_2(K,H)),$$

such that a pair (Y,Z) solves (2.1) if and only if it is a fixed point of Φ. By the equivalence of the norms $\|\cdot\|_\theta$ on $[0,T]$, this proves Theorem 2.2.

For a pair $(y,z) \in \mathbb{H}^2_T(H) \times \mathbb{H}^2_T(\mathcal{L}_2(K,H))$, define the martingale $\overline{M}_t = \mathbf{E}[\xi + \int_0^T f(s,y_s,z_s)ds|\mathcal{F}_t]$. In the same way as before, we see that $\overline{M}_t \in L^2(\mathcal{F}_t,H)$. Hence, by the martingale representation theorem there is a continuous version M_t of the martingale $\overline{M}_t$ and a unique $Z \in \mathbb{H}^2_T(\mathcal{L}_2(H))$ such that

$$M_t = M_0 + \int_0^T Z_s dB_s.$$

Let $Y_t = M_t - \int_0^t f(s,y_s,z_s)ds$. Then, $Y \in \mathbb{H}^2_T(H)$ and

$$Y_t = M_0 - \int_0^t f(s,y_s,z_s)ds + \int_0^t Z_s dB_s.$$

After rearranging, this yields (2.1). Thus, the map Φ is well-defined and a pair $(Y,Z) \in \mathbb{H}^2_T(H) \times \mathbb{H}^2_T(\mathcal{L}_2(K,H))$ is a solution to (2.1) if and only if it is a fixed point of Φ.

To see that Φ is a contraction in $\|\cdot\|_\theta$ for θ large enough, consider $(y^i,z^i) \in \mathbb{H}^2_T(H) \times \mathbb{H}^2_T(\mathcal{L}_2(K,H))$ and define $(Y^i,Z^i) = \Phi(y^i,z^i)$ for $i = 1,2$. Using the notation from Proposition 2.9, the functions f^i do not depend on Y^i, Z^i and the terminal values ξ^1, ξ^2 coincide. Applying the aforementioned theorem, we have with $L = 0$ and $\mu^2 = \theta$ sufficiently large,

$$\|\delta Y\|^2_\theta + \|\delta Z\|^2_\theta \le \frac{2L^2(1+T)}{\theta}(\|\delta y\|^2_\theta + \|\delta z\|^2_\theta) \le \|\delta y\|^2_\theta + \|\delta z\|^2_\theta,$$

where L is a Lipschitz constant for f. ☐

Remark 2.10. As is to be expected from the contraction argument, the proof gives a construction of the solution by a Picard iteration, inductively defining

$$\begin{cases} (Y^0,Z^0) = (0,0) \\ Y^{(n+1)}_t = \xi + \int_t^T f(s,Y^{(n)}_s,Z^{(n)}_s)ds - \int_t^T Z^{(n+1)}_s dB_s \quad n \ge 1. \end{cases}$$

Because the convergence is even geometric, the sequence converges not only in the semimartingale norm but also almost surely by the Borel-Cantelli lemma.

Remark 2.11. Instead of introducing the discounted norms $\|\cdot\|_\theta$, we could have also followed a local approach to construct a solution to the BSDE. Note how *(ii)* in Proposition 2.9 implies that there is a $\delta > 0$, depending only on the generator f, such that the map Φ we constructed above is a contraction on $\mathbb{H}^2_T(H) \times \mathbb{H}^2_T(\mathcal{L}_2(K,H))$ whenever $T < \delta$. To extend the result to arbitrary time intervals $[0,T]$, consider a partition

$0 = t_N < t_{N-1} < \ldots < t_0 = T$ with $t_n - t_{n+1} < \delta$. Set $Y_T = Y_{t_0} = \xi$ and inductively define (Y, Z) as the unique solution to the BSDE on $[t_n, t_{n-1}]$

$$Y_t = Y_{t_{n-1}} + \int_t^{t_{n-1}} f(s, Y_s, Z_s)ds + \int_t^{t_{n-1}} Z_s dB_s. \tag{2.4}$$

By construction, the pair (Y, Z) is a solution to the BSDE with data (f, ξ). Moreover, any other solution $(\tilde{Y}, \tilde{Z})$ solves (2.4) the time intervals $[t_n, t_{n-1}]$ and by the uniqueness on the time intervals $[t_n, t_{n-1}]$, we see that the solution must be unique. We should emphasise that the independence of the stepsize δ from the terminal condition is indispensable for this approach: Without it, the step size could deteriorate (i.e. $\sum_{n=0}^{\infty} t_n - t_{n+1} < \infty$) and the solution cannot be extended to the entire time interval. For coupled forward-backward systems, the stepsize will usually not be independent of the terminal condition which may prevent the argument to cover the entire time interval $[0, T]$ (see also Remark 4.34).

2.3 Linear BSDEs and a Comparison Theorem

For linear BSDEs, a variation of constants type argument provides a more explicit formula for the continuous component of the BSDE.

Proposition 2.12. *Let (β, γ) be a bounded $\mathbb{R} \times K$-valued process. For $\varphi \in \mathbb{H}^2_T(\mathbb{R})$ and $\xi \in L^2(\mathcal{F}_T; \mathbb{R})$, the linear BSDE*

$$Y_t = \xi + \int_t^T (\varphi_s + \beta_s Y_s + \langle \gamma_s, Z_s \rangle)ds - \int_t^T \langle Z_s, dB_s \rangle,$$

has a unique solution (Y, Z) and Y_t admits the representation

$$Y_t = \mathbf{E}\left[\Gamma_T^t \xi + \int_t^T \Gamma_s^t \varphi_s ds \middle| \mathcal{F}_t \right], \tag{2.5}$$

where

$$\Gamma_s^t = 1 + \int_t^s \beta_u \Gamma_u^t du + \int_t^s \Gamma_u^t \langle \gamma_u, dB_u \rangle.$$

In particular, if ξ and φ are nonnegative, so is Y. If in addition $Y_0 = 0$, then for any t, also $Y_t = 0$ almost surely, $\xi = 0$ almost surely and $\varphi_t = 0$ $dP \otimes dt$-almost surely.

Proof. Since the coefficients are bounded, the linear generator of the backward equation is Lipschitz. As $\varphi \in \mathbb{H}^2_T(H)$ and $\xi \in L^2(F_T; H)$, there is a unique solution (Y, Z) by Theorem 2.2.

Applying Itô's formula to $\Gamma_s^t Y_s$,

$$\begin{aligned} d(\Gamma_s^t Y_s) &= \Gamma_s^t(-\varphi_s - \beta_s Y_s - \langle \gamma_s, Z_s \rangle)ds + \Gamma_s^t Y_s(\beta_s ds + \langle \gamma_s, dB_s \rangle) + \Gamma_s^t \langle \gamma_s, Z_s \rangle ds \\ &= -\varphi_s \Gamma_s^t ds + \Gamma_s^t \langle Z_s + Y_s \gamma_s, dB_s \rangle. \end{aligned}$$

Integrating from 0 to T,

$$\Gamma^t_T Y_T - \Gamma^t_t Y_t = -\int_t^T \varphi_s \Gamma^t_s \, ds + \int_t^T \Gamma^t_s \langle Z_s + Y_s \gamma_s, dB_s \rangle.$$

By standard moment estimates for SDEs (see also Proposition 2.22) and the boundedness of the coefficients, we have $\mathbf{E}[\sup_s |\Gamma^t_s|^2] < \infty$. Since the solution (Y, Z) is square-integrable, Hölder's inequality implies $\Gamma^t_s(Z_s + Y_s\gamma_s) \in \mathbb{H}^1_T(K)$ and by Lemma 2.6, the stochastic integral is a martingale. Using $\Gamma^t_t Y_t = Y_t$ and $Y_T = \xi$, we obtain the claim after conditioning on $\mathcal{F}_t$.

For the second part, let us note that

$$\Gamma^t_s = \exp\left\{ \int_t^s (\beta_u + \frac{1}{2}|\gamma_u|^2) du + \int_t^s \langle \gamma_u, dB_u \rangle \right\},$$

is almost surely positive. If $Y_0 = 0$, then the representation (2.5) implies $\xi = 0$ almost surely and $\varphi_t = 0$ $dP \otimes dt$- almost surely and consequently also $Y_t = 0$ almost surely. □

A simple, yet powerful consequence of this representation is the following comparison theorem. For a real-valued BSDE, it amounts to a monotonicity of the solution Y with respect to the data (f, ξ). The theorem is a key component for the weak formulation of stochastic control problems, and monotone approximation is an essential tool in the theory for real-valued BSDEs (see e.g. [40, 44] for two prominent examples).

Theorem 2.13. *For $i = 1, 2$, let (f^i, ξ^i), be real-valued standard parameters and denote the solution to the associated BSDE by (Y^i, Z^i). If*

- *$\xi^1 \geq \xi^2$ almost surely,*
- *$f^1(t, Y^2_t, Z^2_t) - f^2(t, Y^2_t, Z^2_t) \geq 0$ holds $dP \otimes dt$-almost surely,*

then for any time t, we have $Y^1_t \geq Y^2_t$. Moreover, the comparison is strict: Whenever $P(\xi^1 > \xi^2) > 0$, or $f^1(t, Y^2_t, Z^2_t) > f^2(t, Y^2_t, Z^2_t)$ on a set of positive $dP \otimes dt$- measure, it follows that $Y^1_0 > Y^2_0$.

Proof. Essentially, we only need to apply the representation for linear BSDEs to the difference $Y^1 - Y^2$. Let $\{k_n\}$ be an orthonormal basis for K and define

$$\overline{Z}^n := \sum_{i=1}^n \langle Z^2, k_i \rangle + \sum_{i=n+1}^\infty \langle Z^1, k_i \rangle.$$

Since

$$\begin{aligned}
f^1(t, &Y^1_t, Z^1_t) - f^2(t, Y^2_t, Z^2_t) \\
&= f^1(t, Y^1_t, Z^1_t) - f^1(t, Y^2, Z^1_t) \\
&\quad + \sum_{i=0}^\infty f^1(t, Y^2, \overline{Z}^i) - f^1(t, Y^2_t, \overline{Z}^{(i+1)}) + f^1(t, Y^2_t, Z^2_t) - f^2(t, Y^2_t, Z^2_t),
\end{aligned}$$

we see that $(\delta Y, \delta Z) := (Y^1 - Y^2, Z^1 - Z^2)$ satisfies the linear BSDE

$$\delta Y_t = \delta\xi + \int_t^T \varphi_s + \beta_s \delta Y_s + \langle \gamma, Z\rangle ds - \int_t^T \langle \delta Z_s, dB_s\rangle,$$

where

$$\begin{aligned}
\delta\xi &= \xi^1 - \xi^2,\\
\varphi_t &= f^1(t, Y_t^2, Z_t^2) - f^2(t, Y_t^2, Z_t^2),\\
\beta_t &= \begin{cases} (f^1(t, Y_t^1, Z_t^1) - f^2(t, Y_t^2, Z_t^1))(\delta Y_t)^{-1}, & \text{if } \delta Y_t \neq 0,\\ 0 & \text{otherwise,}\end{cases}\\
\langle \gamma_t, k_i\rangle &= \begin{cases} (f^1(t, Y_t^2, \overline{Z}^i) - f^1(t, Y_t^2, \overline{Z}^{i+1}))(\langle \delta Z_t, k_i\rangle)^{-1} & \text{if } (\langle \delta Z_t, k_i\rangle) \neq 0,\\ 0 & \text{otherwise.}\end{cases}
\end{aligned}$$

Now, f is Lipschitz and thus the coefficients of the linear BSDE are bounded. Furthermore, by the integrability of f, also $\varphi \in \mathbb{H}_T^2(\mathbb{R})$. By assumption $\delta\xi \geq 0$ and $\varphi \geq 0$ $dP \otimes dt$-almost surely, and hence Proposition 2.12 concludes. □

Remark 2.14. For forward SDEs, similar comparison theorems are only available under much stronger assumptions and require for example that the diffusion coefficients coincide (see e.g. the classical result by Ikeda and Watanabe [36] or Chapter 5 of [39]). BSDEs of course owe this property to their close connection to control problems.

If the terminal condition is not just $\mathcal{F}_T$-measurable, the solution can reach the terminal condition also before the terminal time T.

Proposition 2.15. *Let (Y, Z) be the solution to (2.1) with standard parameters (f, ξ) and let τ be a stopping time with $P(\tau < T) = 1$. If ξ is $\mathcal{F}_\tau$-measurable and $\mathbb{1}_{t \geq \tau} f(t, y, z) = 0$, then also $(Y, Z) = (\xi, 0)$ on $\{t \geq \tau\}$. Thus, the stopped solution $(Y^\tau, Z^\tau) = (Y_{\tau\wedge t}, \mathbb{1}_{t\leq\tau} Z_t)$ is the unique solution to the BSDE with the random terminal time τ*

$$Y_t^\tau = \xi + \int_{t\wedge\tau}^{\tau} f(s, Y_s^\tau, Z_s^\tau) ds - \int_{t\wedge\tau}^{\tau} Z_s^\tau dB_s. \tag{2.6}$$

Proof. By the assumption that f vanishes on $\{t \geq \tau\}$,

$$Y_\tau = \xi - \int_\tau^T Z_s dB_s.$$

The stochastic integral is a martingale and conditioning on $\mathcal{F}_\tau$ yields the desired claim for Y as $Y_{t\vee\tau} = \mathbf{E}[\xi | \mathcal{F}_{t\vee\tau}] = \xi$. To see that Z vanishes on $\{t \geq \tau\}$, apply Itô's formula on $[\tau, T]$ to obtain

$$|Y_\tau|^2 = |\xi|^2 - 2\int_\tau^T \langle Y_s, Z_s dB_s\rangle + \int_\tau^T \|Z_s\|^2 ds.$$

The stochastic integral is again a martingale and upon rearranging and conditioning on $\mathcal{F}_\tau$, we see $\mathbf{E}[\int_\tau^T \|Z_s\|^2 ds|\mathcal{F}_\tau] = 0$. Hence, also $\mathbf{E}\int_\tau^T \|Z_s\|^2 ds = 0$ and $Z = 0$ on $\{t \le \tau\}$. Then, inserting (Y^τ, Z^τ) back into the BSDE, we see that (2.6) is satified on $[0, \tau]$. Finally, uniqueness follows since any solution to (2.6) also solves (2.1). □

Remark 2.16. Conversely, this also amounts to a 'flow'-property for BSDEs: If (Y, Z) is the unique solution to (2.1) with standard data (f, ξ), then for any stopping time $\tau \le T$, the pair (Y, Z) is a solution to the BSDE on $[0, \tau)$ with data (f, Y_τ). The Lipschitz assumption is only to ensure the existence of a unique solution. If this is established under weaker assumptions, then this flow property also follows immediately.

With this observation, we can extend the solution (Y, Z) to (2.1) on $[0, T]$ to the entire time interval $[0, \infty)$ by extending the generator and the solution

$$f(t, x, y) = 0, \text{ and } (Y_t, Z_t) = (\xi, 0) \text{ for } t \ge T,$$

without introducing any ambiguities. From now on, we will implicitly understand that all solutions are continued in this manner.

Remark 2.17. We will interpret Proposition 2.15 as the definition for BSDEs with a (bounded) random terminal time. It is also possible to weaken the assumption that $\tau \le T$ almost surely to consider more general BSDEs with unbounded random terminal times, see e.g. [22].

2.3.1 Continuity and Dependence upon Parameters

Let $(f^a, \xi^a)_{a\in\mathbb{R}}$ be a family of standard parameters and denote the solutions to the associated BSDEs by (Y^a, Z^a). Assuming that the data (f^a, ξ^a) is continuous or differentiable in a with respect to the semi-martingale norms, the a priori estimates Proposition 2.9 suggest that also (Y^a, Z^a) should depend continuously on or even be differentiable in $a \in \mathbb{R}$. We assume that

(H1) The Lipschitz constant of f^a can be chosen independent of a, that is there is a $L > 0$ such that $dt \otimes dP$-almost surely

$$|f^a(t, y_1, z_1) - f^a(t, y_2, z_2)| \le L(|y_1 - y_2| + \|z_1 - z_2\|) \quad \forall a \in \mathbb{R},$$

(H2) the map $\mathbb{R} \to \mathbb{H}_T^2(H) \times L^2(\mathcal{F}_T), a \mapsto (f^a, \xi^a)$ is continuous in the sense that $\xi^{a_1} \to \xi^{a_2}$ in $L^2(\mathcal{F}_T)$ as $a_1 \to a_2$ and

$$\lim_{a_1 \to a_2} \|f^{a_1}(t, Y_t^{a_2}, Z_t^{a_2}) - f^{a_2}(t, Y_t^{a_2}, Z_t^{a_2})\|_0 \to 0,$$

(H3) f is two times continuously differentiable with bounded derivatives in y and z,

(H4) the maps in (H2) are also differentiable in a with derivatives which we denote by $\partial_a f^a$ and $\partial_a \xi^a$.

Proposition 2.18. *(i) Under* **(H1)** *and* **(H2)**, *the solution map* $a \mapsto (Y^a, Z^a)$ *is continuous as a map* $\mathbb{R} \to \mathbb{H}_T^2(H) \times \mathbb{H}_T^2(\mathcal{L}_2(K,H))$.

(ii) If the parameters also satisfy **(H3)** *and* **(H4)**, *then* $a \mapsto (Y^a, Z^a)$ *as in (i) is also differentiable in* a *and the derivatives* $(\nabla_a Y^a, \nabla_a Z^a)$ *satisfy the linear BSDE*

$$\nabla_a Y_t^a = \nabla_a \xi^a - \int_t^T \nabla_a Z_t^a dB_t$$
$$+ \int_t^T \left(\partial_y f(a,t,Y_t^a,Z_t^a) \nabla_a Y_t^a + \partial_z f(a,t,Y_t^a,Z_t^a) \nabla_a Z_t^a + \partial_a f(a,t,Y_t^a,Z_t^a) \right) dt.$$

Proof. The first claim is a direct consequence of the stability Proposition 2.9 (ii). For the second claim, we refer to [26, Proposition 2.4], using the same method as in the proof of Theorem 2.13. □

Remark 2.19. If the parameters satisfy the conditions also for the norm on $\mathbb{H}_T^\infty$ in y, then the claims of course also hold in this norm. The condition **(H3)** is more restrictive than necessary, but satisfied for our particular case. Moreover, the same result stays true if we replace $a \in \mathbb{R}$ with a more general complete metric space.

2.4 Forward-Backward SDEs

In this section, we consider the important special case where the randomness in the generator is induced by an Itô diffusion process. These equations arise naturally in stochastic control problems. In the finite-dimensional case, their connection to the Hamilton-Jacobi-Bellman equation of the stochastic control problem has been studied extensively.

Let again J, H, K be real and separable Hilbert spaces. For a forward SDE with deterministic coefficients $b : [0,T] \times J \times H \times K \to J$ and $\sigma : [0,T] \times J \times H \times K \to \mathcal{L}_2(K,J)$, on $[t,T]$ and a forward SDE,

$$X_s^{t,x} = x + \int_t^s \mathbb{1}_{\{s \le T\}} b(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) dr + \int_t^s \mathbb{1}_{\{s \le T\}} \sigma(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) dB_r \quad (2.7)$$

we consider the BSDE

$$Y_s^{t,x} = g(X_T^{t,x}) + \int_s^T \mathbb{1}_{\{r \ge t\}} f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) dr + \int_s^T Z_r^{t,x} dW_r, \quad (2.8)$$

where $g : J \to H$ and $f : [0,T] \times J \times H \times K \to H$ are continuous. The system (2.7, 2.8) is also called a forward-backward SDE (FBSDE, for short). If the coefficients of the forward equation b, σ do not depend on the backward equation we call the system *decoupled*. Otherwise, we say that the system is *coupled*. We will generally omit the superscript t, x if $t = 0$ and no ambiguities arise. For simplicity, we will assume that for some $L > 0$, the coefficients b, σ, f and g are uniformly Lipschitz in x, y, z and that the norms of $b(\cdot, 0), \sigma(\cdot, 0), f(\cdot, 0), g(0)$ are bounded uniformly by L. This will guarantee a uniform Lipschitz condition for all coefficients as well as linear growth. As a shorthand, we introduce the notation $\Theta^{t,x} := (X^{t,x}, Y^{t,x}, Z^{t,x})$.

We stress that, in contrast to the generic BSDE we considered earlier, the generator f and the terminal condition $g(X_T^{t,x})$ depend on Ω only as a function of X. If X is a Markov process, (e.g. if the coefficients b and σ depend only on (r,x)), these equations are called 'Markovian'.

Some remarks on coupled FBSDEs.

Coupled FBSDEs arise naturally from stochastic control problems, as a necessary condition for the optimal control (e.g. via a stochastic maximum principle) and are thus particularly important in applications. In general, results for FBSDEs often make use of connections with stochastic control and are often relatively one-dimensional [46]. If the FBSDE is coupled, sufficient conditions for the existence of a solution are no longer as straightforward and are usually quite ad-hoc. Indeed, since we prescribe the initial as well as the terminal condition of the system, the expectation of the FBSDE must satisfy a boundary value problem. While the difficulty of existence is thus not entirely due to the stochastic nature of the equations, the randomness does still introduce additional problems. The following two simple examples illustrate this point.

Example 2.20. We consider the one dimensional ($K = H = J = \mathbb{R}$) FBSDE

$$\begin{cases} X_t = \int_0^t Y_s ds, \\ Y_t = X_T + \int_t^T X_t dt - \int_t^T Z_s dB_s. \end{cases}$$

Upon taking expectations, we see that $(x,y) := (\mathbf{E}[X], \mathbf{E}[Y])$ satisfies the ODE

$$\begin{bmatrix} \dot{x}(t) \\ \dot{y}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}.$$

As the solutions to this equation are of the form $\lambda_1 \cos(t) + \lambda_2 \sin(t)$, choosing $T = \pi + \frac{\pi}{4}$ implies that there are no solutions for $x(0) \neq 0$, and infintely many if $x(0) = 0$.

Even where a solution to the BVP for the expectation exists, it is not guaranteed that (unique) adapted solutions to the FBSDE exist.

Example 2.21. For $\xi \in L^2(\mathcal{F}_T)$, consider the one-dimensional FBSDE,

$$\begin{cases} X_t = \int_0^t Z_s dB_s, \\ Y_t = X_T + \xi - \int_t^T Z_s dB_s. \end{cases}$$

If (X,Y,Z) is a solution, it statisfies $Y_0 = X_T + \xi - \int_0^T Z_s dB_s = \xi$, and unless ξ is deterministic, this implies that Y cannot be adapted. On the other hand, if ξ is deterministic, we are free to choose any $Z \in \mathbb{H}_T^2(\mathbb{R})$ and letting $(X_t, Y_t) = (\int_0^t Z_s dB_s, X_t + \xi)$ the triple (X,Y,Z) defines a solution. Hence, even in this simple case, there is either no solution or infinitely many solutions to the FBSDE.

2.4.1 Decoupled FBSDEs

We now assume that the forward equation (2.7) does not depend on the backward equation (2.8), that is

$$b(s,x,y,z) = \tilde{b}(s,x) \quad \sigma(s,x,y,z) = \tilde{\sigma}(s,x).$$

In this case, thanks to standard moment estimates for the solutions to SDEs, well-posedness is immediate from Theorem 2.2. We recall them here without proof. The Burkholder-Davis-Gundy inequalities extend the same estimates also to $\mathbf{E}|X_t|^p$ for $p \geq 2$ provided the coefficients are sufficiently integrable themselves.

Proposition 2.22. *For $x, x' \in J$, $0 \leq t' < t \leq T$, the solution to the forward equation (2.7) statisfies the moment estimates*

$$\mathbf{E}\left[\sup_{t\leq s\leq T} \left|X_s^{t,x}\right|^2\right] \leq C(1+|x|^2),$$

$$\mathbf{E}\left[\sup_{t\leq s\leq T} \left|X_s^{t',x'} - X_s^{t,x}\right|^2\right] \leq C(\left|x-x'\right|^2 + (1+|x|^2)\left|t-t'\right|).$$

Corollary 2.23. *For any $(x,t) \in [0,T] \times J$, the FBSDE (2.7, 2.8) has a unique square-integrable solution.*

Proof. Since all coefficients are Lipschitz and grow at most linearly, the existence of a solution to (2.7) follows from a standard fixed-point iteration. Thanks to the moment estimates Proposition 2.22 on the solution $X^{t,x}$ and the linear growth assumptions on f and g,

$$\mathbf{E}\left[\left|g(X_T^{t,x})\right|^2 + \int_0^T \left|f(s, X_s^{t,x}, 0, 0)\right|^2 ds\right] < \infty.$$

Adaptedness follows from the adaptedness of X. In other words, the parameters for (2.8) are standard parameters and Theorem 2.2 concludes. □

Under the assumptions on the coefficients f, g in the backward equation, the estimates in Proposition 2.22 also transfer to the solution of the backward equation. We omit the straightforward proof.

Lemma 2.24. *For $x, x' \in J$, and $0 < t' < t \leq T$, the following moment bounds apply for the solution the the backward equation*

$$\mathbf{E}\left[\sup_{t\leq s\leq T} \left|Y_s^{t,x}\right|^2 + \int_t^T \|Z_s^{t,x}\|^2 ds\right] \leq C(1+|x|^2),$$

$$\mathbf{E}\left[\sup_{t\leq s\leq T} \left|Y_s^{t,x} - Y_s^{t',x'}\right|^2 + \int_t^T \|Z_s^{t,x} - Z_s^{t',x'}\|^2 ds\right] \leq C(1+|x|^2)(\left|x-x'\right|^2 - \left|t-t'\right|).$$

2.4.2 The Markov property

In the Markovian setting, the measurability of the coefficients transfers to the solution of the backward component.

Proposition 2.25. *There is a continuous deterministic function $u : [0,T] \times J \to H$ such that the solution to (2.7-2.8) satisfies for any $s \geq t$ up to indistingushability, $Y_s^{t,x} = u(s, X_s^{t,x})$.*

Proof. We only sketch the argument and omit (the mostly technical) details. We fix the initial time t and we consider the FBSDE (2.7-2.8) with a *random* initial condition x, independent of B. Denote by $\{\tilde{\mathcal{F}}_s\}_s$ the augmentation of the filtration generated by x and $B_{s\wedge t} - B_t$. Thanks to the Markov property of the Brownian motion, the restarted Brownian motion $B_{s\wedge t} - B_t$ is again a Brownian motion in the filtration $\{\mathcal{F}_{s\geq 0}\}$ independent of $\mathcal{F}_t$. The standard results for SDEs imply that there is a solution to the forward equation

$$X_s^t = x + \int_0^s b(r+t, X_s^t) + \int_0^s \sigma(r+t, X_s^t) d(B_{s\wedge t} - B_t).$$

But as $\{\tilde{\mathcal{F}}_s\}_{s\geq 0} \subset \{\mathcal{F}_s\}_{s\geq 0} \vee \sigma(x)$, the uniqueness for (2.1) implies $X_{s-t}^t = X_s^{t,x}$. Consequently, $X^{t,x}$ is $\{\tilde{\mathcal{F}}_s\}_{s\geq 0}$-adapted.

In the same way, we can solve the backward equation with respect to the filtration $\{\tilde{\mathcal{F}}_s\}_{s\geq 0}$ which by uniqueness again implies that the solution $(Y^{t,x}, Z^{t,x})$ is $\{\tilde{\mathcal{F}}_s\}_{s\geq 0}$-adapted. In particular, $Y_t^{t,x}$ is almost surely $\sigma(x)$-measurable. Thus, the Doob-Dynkin Lemma implies that for each time t there is a Borel-measurable function $x \mapsto u(t,x)$ such that $Y^{t,x} = u(t,x)$.

The uniqueness for the equation (2.7) and a factorisation lemma implies that $X^{t,x}$ sastisfies the Markov- or flow-property

$$X^{t,x} = X^{t', X_{t'}^{t,x}},$$

for all $t < t'$. Since the solution $Y^{t,x}$ is unique, we conclude that almost surely

$$u(s, X_s^{t,x}) = Y^{s, X_s^{t,x}} = Y_s^{t,x}.$$

Lemma 2.24 and $u(t,x) = Y_t^{t,x}$ now also imply that u is jointly continuous in (t,x). So far we have seen that the processes $Y_s^{t,x}$ and $u(s, X_s^{t,x})$ coincide up to modification. But both processes are continuous and we can conclude they must already be indistinguishable. □

Remark 2.26. The argument only relies on the facts that the time shifted coefficients $b(r+t, \Theta_{r+t}), \sigma(r+t, \Theta_{r+t}), f(r+t, \Theta_{r+t})$ are adapted to $\{\tilde{\mathcal{F}}_s\}_{s\geq 0}$ and that there is a unique solution to the FBSDE, but not explicitly on the assumption that b, σ are independent of y, z. If existence and uniqueness are established under weaker conditions, then Proposition 2.25 follows verbatim.

Remark 2.27. If x is deterministic, the above implies that the same is true for $Y_t^{t,x}$. Hence, we can extend the solution on $[t,T]$ to $[0,T]$ by defining

$$(X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x}) := (x, Y_t^{t,x}, 0), \quad s \leq t,$$

without losing adaptedness. From now on, we will always assume this extension implicitly. With this convention, the moment estimates in Lemma 2.24 also extend to $[0,T]$.

Remark 2.28. In the same way, we can argue that there is a function v such that $Z_s^{t,x} = v(s, X_s^{t,x})$. This connection is more naturally understood in relation to PDEs and stochastic control. Note, however, that the moment estimates for the martingale part Z in Lemma 2.24 are not sufficient to conclude continuity for v.

2.4.3 Relation to semilinear parabolic PDEs

Denote by

$$\mathscr{L} f = \langle b, \nabla f \rangle + \frac{1}{2} \mathrm{Tr}(\sigma\sigma^* \,\mathrm{Hess}\, f),$$

the infinitesimal generator of the diffusion X and suppose that u is a classical solution to

$$\begin{cases} \partial_t u + \mathscr{L} u + f(t,x,u,\nabla u\sigma) = 0, \\ u(T,x) = g(x), \end{cases} \tag{2.9}$$

with polynomial growth in space. We want to stress that the nonlinearity is not in the most general form: f can depend only on $\nabla u\sigma$ and not ∇u directly. Then it follows from Itô's formula that

$$(Y_s^{t,x}, Z_s^{t,x}) := (u(s, X^{t,x}), (\sigma^* \nabla u)(s, X^{t,x})),$$

is the unique square-integrable solution to the BSDE (2.8). In particular, the following generalisation of the Feynman-Kac formula holds

$$u(t,x) = Y_t^{t,x} = \mathbf{E}\left[g(X_T^{t,x}) + \int_t^T f(s, \Theta_s^{t,x}) ds \right].$$

We want to point out that the decoupled FBSDE (2.7, 2.8) has a solution under fairly mild regularity assumptions on the coefficients. The same is not true for the PDE (2.9), at least if we are looking for classical solutions. Relaxing the notion of solutions to (2.9), we may hope to recover also the reverse connection, that is identify $Y_t^{t,x}$ as a solution to (2.9) in a suitable sense. In finite dimension, the relation of u and the PDE (2.9) is well-studied and we mostly refer to [55] and the references therein for the classical finite-dimensional theory. Under fairly general assumptions, $u(t,x) := Y_t^{t,x}$ is the unique viscosity solution to (2.9), which was already explored in the pioneering work of Pardoux and Peng [57]. The major drawback is that viscosity solutions are not differentiable, and thus the process Z (corresponding to the control) cannot be identified as $(\nabla u\sigma)(t,X)$. To get better control on the martingale part, we therefore often rely on stronger regularity conditions on the coefficients which may also imply that $Y_t^{t,x}$ solves (2.9) in a stronger sense. We provide some results in this direction in the next section. Generalising the relation with viscosity solutions to infinite dimensions introduces complications in the choice of the test functions and is not straightforward. We refer to [15] and the references therein for adaptations to the infinite-dimensional setting.

2.4.4 Regularity

As a result of the stability of SDEs and BSDEs, if the coefficients are in addition also continuously Fréchet-differentiable and sufficiently integrable, the solution $\Theta^{t,x} = (X^{t,x}, Y^{t,x}, Z^{t,x})$ is not just continuous in t, x but also differentiable. For simplicity, we assume that the coefficients and their derivatives are uniformly bounded by some constant L. This condition is far from necessary and can for example be relaxed without additional technical complications to requiring finite 4^{th}-moments, see e.g. [27].

Proposition 2.29. *Assume that the coefficients b, σ and f are twice continuously differentiable in x, y, z with bounded derivatives. Then, for each $t \geq 0$,*

$$J \to \mathbb{H}_T^\infty(J) \times \mathbb{H}_T^\infty(H) \times \mathbb{H}_T^2(\mathcal{L}_2(K, H));\ x \mapsto (X^{t,x}, Y^{t,x}, Z^{t,x}),$$

is differentiable in x and the derivative $(\nabla X, \nabla Y, \nabla Z)$satisfies the FBSDE

$$\begin{cases} \nabla X_s^{t,x} = 1 + \int_t^s \partial_x b^{t,x}(r, X_r) \nabla X_r^{t,x} dr + \int_t^s \partial_x \sigma(r, X_r^{t,x}) \nabla X_r^{t,x} dB_r \\ \nabla Y_s^{t,x} = \partial_x g(X_T^{t,x}) \nabla X_T^{t,x} + \int_s^T F(r, \nabla X_r^{t,x}, \nabla Y_r^{t,x}, \nabla Z_r^{t,x}) dr - \int_s^T \nabla Z_r^{t,x} dB_r, \end{cases}$$

where $F(s, x_s, y_s, z_s) = \partial_x f(s, \Theta_s^{t,x}) x_s + \partial_y f(s, \Theta_s^{t,x}) y_s + \partial_z f(s, \Theta_s^{t,x}) z_s$.

Proof. We only have to check that the assumptions of Proposition 2.18 (ii) are satisfied. Under the assumptions on the coefficients b and σ, the map $x \mapsto X^{t,x}$ is differentiable. This can be seen from a priori estimates for SDEs (see e.g. [65, Theorem 3.2.2]) in the same way as in Proposition 2.18 and $\nabla X^{t,x}$ satisfies the forward equation as required. From this, we also see that ∇X has moments of any order by Proposition 2.22. Therefore, using the boundedness of $\partial_x f$, the map in the hypothesis **(H4)** is differentiable with

$$\frac{d}{dx} f(s, X_s^{t,x}, y_s, z_s) = \partial_x f(s, X_s^{t,x}, y_s, z_s) \nabla X_s^{t,x}.$$

In the same way, we see that $g(X_T^{t,x})$ is differentiable in x with

$$\frac{d}{dx} g(X_T^{t,x}) = \partial_x g(X_T^{t,x}) \nabla X_T^{t,x}.$$

The condition **(H3)** is satisfied by assumption and we can conclude by Proposition 2.18. □

Remark 2.30. We know from Proposition 2.25 that $Y_s^{t,x} = u(s, X_s^{t,x})$. Additionally, we have just seen that under suitable assumptions on the coefficients, the map $x \mapsto Y_s^{t,x}$ is differentiable (in the sense of Proposition 2.18). Under these assumptions Proposition 5.3 from [27] implies that there is a version of the Malliavin derivative such that $Z_s = D_s Y_s^{t,x}$. Thus, by the chain rule ([52, Proposition 2.4] or [42]),

$$Z_s = D_s Y_s^{t,x} = \nabla u(s, X^{t,x}) D_s X_s^{t,x} = (\nabla u \sigma)(s, X_s^{t,x}).$$

This relation can also be established under weaker assumptions (See e.g. [57] in the Markovian case, and [27] for a generalisation).

2.5 Generalised BSDEs

Motivated by control problems, we consider a generalisation of the BSDEs on a probability space with a filtration that is not necessarily generated by a Brownian motion. To obtain the existence of a minimiser, it may be useful to relax the variational problem to obtain the required compactness and continuity. In this setting, the underlying filtration is no longer generated by the Brownian motion and we are therefore interested in studying BSDEs with respect to a general right-continuous, complete filtration $\mathcal{F}_t$. Of course, this implies that we do not have access to the Brownian martingale representation theorem. Fortunately, we still have the following more general version of the martingale representation theorem.

Lemma 2.31. *If $L^2(\mathcal{F}_T)$ is a separable Hilbert space, there is a at most countable sequence of square-integrable and pairwise orthogonal $\{F_t\}$-martingales $\{M^n\}$ such that for any $\xi \in L^2(\mathcal{F}_T)$ there are are predictable processes $\{Z^n\}$ satisfying*

$$\mathbf{E}\sum_{n=0}^{\infty}\int_0^T \|Z_s^n\|^2 d\langle M^i\rangle_u < \infty,$$

and

$$\mathbf{E}[\xi|\mathcal{F}_t] = \mathbf{E}[\xi] + \sum_{n=0}^{\infty}\int_0^t Z_u^i dM_u^i.$$

Proof. See e.g. [23] or [18, Theorem 2.1]. □

This suggests that we should consider the following formulation for the generalised BSDE

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dB_s - \int_t^T dM_s. \tag{2.10}$$

Here, Y is an H-valued adapted càdlàg process, Z is predictable and $\mathcal{L}_2(K,H)$-valued. Finally, M is an H-valued local càdlàg-martingale starting in 0 orthogonal to the Brownian motion, that is $\langle M, B\rangle_s = 0$. Under the same assumptions on the parameters as in the setting with a Brownian filtration, existence and uniqueness follow from a fixed point argument. In contrast to the estimates for (2.1), we now have to rely on martingale inequalities instead of Itô's formula. Consequently, we cannot introduce the discounted norms $\|\cdot\|_\theta$ and we first obtain existence only for small terminal times T. This solution can then be extended to the entire time interval $[0, T]$ for any finite $T > 0$. The point of this section is thus the following theorem.

Theorem 2.32. *For standard parameters (f,ξ), there is a unique pair $(Y,Z) \in \mathbb{H}_T^\infty(H) \times \mathbb{H}_T^2(\mathcal{L}_2(K,H))$ and a unique orthogonal martingale M such that (2.10) holds for (Y,Z,M).*

Lemma 2.33. *For $i = 1, 2$, let (f^i, ξ^i) be standard parameters, and suppose that (Y^i, Z^i, M^i) are solutions to (2.10) (in the sense of Theorem 2.32). Denote the spread between the two solutions by $(\delta Y, \delta Z, \delta M)$ and define $\delta_2 f = f^1(t, Y_t^2, Z_t^2) - f^2(t, Y_t^2, Z_t^2)$. If $T > 0$ is* *sufficiently small, there is a positive constant $C > 0$ such that*

$$\mathbf{E}\left[\sup_t |\delta Y_t|^2 + \int_0^T \|\delta Z_s\|^2 ds + \langle \delta M \rangle_T\right] \leq C\,\mathbf{E}[|\delta Y_T|^2] + T\|\delta_2 f\|_0^2.$$

Proof. We note that the norm on (Y,Z,M) is exactly the classical semimartingale-norm and that we can therefore directly apply the semimartingale inequalities (see e.g. [59, Lemma 2.2]), to obtain

$$\mathbf{E}\left[\sup_t |\delta Y_t|^2\right] \leq C\,\mathbf{E}\left[|\delta Y_T|^2 + \int_0^T |\delta f_s|^2\, ds\right],$$

where as before $\delta f := f^1(t, Y_t^1, Z_t^1) - f^2(t, Y_t^2, Z_t^2)$. Estimating the difference in f as in the proof of Proposition 2.9, this implies the claim for Y. The full statement follows in exactly the same way as before with the Burkholder-Davis-Gundy inequality applied to the martingale $\int_0^\cdot Z_s dB_s + \int_0^\cdot dM_s$. ☐

Proof of Theorem 2.32. Fix $(y,z) \in \mathbb{H}_T^\infty(H) \times \mathbb{H}_T^2(\mathcal{L}_2(K,H))$. Similarly to the proof of Theorem 2.2, we define Y_t as the càdlàg-version of the semimartingale $\mathbf{E}[\xi + \int_t^T f(s, y_s, z_s)|\mathcal{F}_s]$ and (Z,M) by the orthogonal decomposition from Lemma 2.31 for the martingale $\mathbf{E}[\xi + \int_0^T f(s, y_s, z_s)ds|\mathcal{F}_t]$ with respect to the Brownian motion B, so that

$$Y_0 + \int_0^t Z_s dB_s + M_s = \mathbf{E}\left[\xi + \int_0^T f(s, y_s, z_s)\middle|\mathcal{F}_t\right].$$

Rearranging the above we see that then (Y,Z,M) satisfies the BSDE (2.10) with generator $f(\cdot, y_\cdot, z_\cdot)$.

Letting $(f^2, \xi^2) = (0,0)$, the estimate in Lemma 2.33 shows that then $(Y,Z) \in \mathbb{H}_T^\infty(H) \times \mathbb{H}_T^2(\mathcal{L}_2(K,H))$. In other words, Φ maps the space $\mathbb{H}_T^\infty(H) \times \mathbb{H}_T^2(\mathcal{L}_2(K,H))$ into itself.

For (y^i, z^i) as above and $i = 1,2$, let (Y^i, Z^i, M^i) be the associated solutions. Applying Lemma 2.33 to this situation and keeping in mind that the terminal conditions coincide, we have thanks to the Lipschitz continuity of f,

$$\mathbf{E}\left[\sup_t |\delta Y_t|^2 + \int_0^T \|\delta Z_s\|^2 ds\right] \leq CT\,\mathbf{E}\left[\int_0^T |\delta y_t|^2\, dt + \int_0^T \|\delta z_s\|^2 ds\right].$$

For $T < 1$ this means

$$\mathbf{E}\left[\sup_t |\delta Y_t|^2 + \int_0^T \|\delta Z_s\|^2 ds\right] \leq CT\,\mathbf{E}\left[\sup_t |\delta y_t|^2 + \int_0^T \|\delta z_s\|^2 ds\right],$$

and Φ is a contraction for $T < \frac{1}{C} =: \delta$. Then, there is a unique fixed point (Y,Z) of Φ and defining M as before via the orthogonal decomposition given by Lemma 2.31, we see that (Y,Z,M) is the unique (square-integrable) solution on $[0,T]$.

To extend the solution to arbitrary time intervals, we follow the idea sketched in Remark 2.11: Choose a partition $0 = t_N < t_{N-1} < \cdots < t_0 = T$ of $[0, T]$ with mesh size $t_n - t_{n+1} < \delta$ and define (Y, Z, M) on $[t_n, t_{n-1}]$ as the unique solution to

$$Y_t = Y_{t_{n-1}} + \int_t^{t_{n-1}} f(s, Y_s, Z_s) ds - \int_t^{t_{n-1}} Z_s dB_s - \int_t^{t_{n-1}} dM_s,$$

and the claim follows. □

Remark 2.34. The results Proposition 2.18 and Proposition 2.29 transfer verbatim also to the generalised setting thanks to the a priori bounds in Lemma 2.33.

3. Stochastic Optimal Control

We want to briefly cover some relevant aspects of stochastic control theory, focusing mainly on the direct connection to BSDEs, while also providing a verification theorem without proof. The BSDE interpretation roughly follows [26] and for a more general perspective, we again refer to [64] and [28].

As before, we assume that the Hilbert spaces H, J, K are real and separable and that the *control space* Π is a Banach space. For a fixed terminal time $T \in [0, \infty]$, define the set of admissible controls

$$\mathcal{A} := \mathcal{A}^{t,T} := \{u : [t, T] \times \Omega \to \Pi \text{ is adapted}\}.$$

In general, we can of course impose additional constraints on the set of admissible controls such as square-integrability. We usually leave the dependence of $\mathcal{A}$ on t, T implicit to simplify the notation. Consider a family of stochastic differential equations parameterised by $u \in \mathcal{A}^{t,T}$,

$$X_s^{t,x}(u) = x + \int_t^s b(r, X_r^{t,x}, u_r)dr + \int_t^s \sigma(r, X_r^{t,x})dB_r, \tag{3.1}$$

where we assume that b, σ satisfy conditions to ensure well-posedness for every admissible control $u \in \mathcal{A}$. If $X^{t,x}(u)$ is the solution to (3.1) for a given control $u \in \mathcal{A}$, we call $(u, X(u))$ an *admissible pair*. Notice that the SDE is not in the most general form since the diffusion does not depend on the control $u \in \mathcal{A}$.

We seek to minimise the (undiscounted) cost functional

$$J^{t,x}(u) = \mathbf{E}\left[g(X_T^{t,x}(u)) + \int_t^T \ell(s, X_s^{t,x}(u), u_s)ds\right],$$

where g and ℓ are appropriate real-valued functionals called the *terminal cost* and the *running cost* of the stochastic control problem respectively. The *value function* of the stochastic control problem is defined as

$$\mathcal{V}^{t,x} := \inf_{u \in \mathcal{A}} J^{t,x}(u).$$

Writing the problem in terms of BSDEs, this means we consider the controlled forward-backward system

$$\begin{cases} X_s^{t,x}(u) = x + \int_t^s b(r, X_r^{t,x}(u), u_r)dr + \int_t^s \sigma(r, X_r^{t,x}(u))dB_r, \\ Y_s^{t,x}(u) = g(X_T^{t,x}(u)) + \int_s^T \ell(r, X_r^{t,x}(u), u_r)dr - \int_s^T Z_r^{t,x}(u)dB_r, \end{cases} \tag{3.2}$$

with $\mathcal{V}^{t,x} = \inf_u Y_t^{t,x}(u)$.

Remark 3.1. Typical cost functions ℓ arising in applications are often quadratic in the control u (the simplest case being $\ell(s, x, a) = \tilde{\ell}(a) = \|a\|^2$). This usually implies that the optimal BSDE associated with the control problem is also quadratic in Z, and consequently does not satisfy the conditions of Theorem 2.2. Fortunately, the backward component Y is only one-dimensional, which means we have access to the comparison principle. This fact was first leveraged in [40], to derive well-posedness for an important class of BSDEs with quadratic growth in Z. If Y is not scalar, some results exist but their scopes are rather specialized and the conditions under which they apply generally remain contrived.

3.1 Verification and the HJB equation

From the relation with PDEs, we can often identify a possible candidate for the optimal (feedback) control. Verification theorems can be used to verify that this candidate is indeed optimal. As this will not be our primary tool, we only give a verification theorem that is applicable in the situation we will be interested in later on in Chapter 4.

The *Hamilton-Jacobi-Bellman* equation (HJB, for short) associated to this stochastic control problem (3.2), is

$$\begin{cases} \partial_t v(t,x) + \inf_{u\in\Pi} H(t,x,\nabla v, \operatorname{Hess} v, u) = 0 \\ v(T,x) = g(X_T), \end{cases} \tag{3.3}$$

where

$$H(t,x,p,Z,u) := \frac{1}{2}\operatorname{Tr}((\sigma\sigma^*)(t,x)Z) + \langle p, b(t,x,u)\rangle + \ell(t,x,u),$$

is the (current-value) Hamiltonian. If we can solve (3.3), we have access to the following verification theorem.

Theorem 3.2 (Verification). *Suppose that v is a classical solution to (3.3) and let (u^*,X^*) be an admissible pair. If for almost every $s \in [t,T]$*

$$u_s^* \in \operatorname{argmin}_{a\in\Pi} H(s, X_s^*, \nabla v(s, X_s^*, \operatorname{Hess} v(s,X^*)), a), \qquad P\text{-almost surely,} \tag{3.4}$$

then $v(t,x) = \mathcal{V}^{t,x}$ and (u^*,X^*) is optimal for (3.2).

Proof. See e.g. [28, Theorem 2.36]. □

Remark 3.3. If $\mathcal{V}$ satisfies the Hamilton-Jacobi-Bellman equation, then the condition (3.4) is also necessary (see e.g. [28, Corollary 2.37]). The assumption that v be a classical solution to (3.3) can be relaxed significantly.

3.2 A Weak Formulation

If we cannot or do not want to rely on the verification, it is often helpful to relax the variational problem to guarantee the existence of minimisers. In many applications, it can be shown that the value functions coincide. Once the existence of a minimiser is established for the relaxed problem, in special cases, the weak optimal control also gives rise to an optimal strong solution. In this section, we give the general construction of a weak formulation for stochastic control problems, which is in full analogy to weak solutions in the theory of SDEs. The argument will rely heavily on the fact that there is no diffusion control, i.e. that $\sigma(t,x,a) = \sigma(t,x)$ for any $a \in \Pi$.

Suppose there is a θ satisfying $b(t,x,a) = \sigma(t,x)\theta(t,x,a)$, for all $a \in \Pi$. and consider the (uncontrolled) process

$$X_t = \varphi + \int_0^t \sigma(s,X_s)dB_s,$$

We assume that $\theta_t^u := \theta(t,X_t,u_t)$ can be chosen such that for any $a \in \mathcal{A}$, the associated Doléans-Dade exponential

$$M_t^u = \exp\left(\int_0^t \langle\theta_s^u, dB_s\rangle - \frac{1}{2}\int_0^t \left|\theta_s^u\right|^2 ds\right),$$

is a martingale, which enables the Girsanov theorem (see e.g. [20, Theorem 10.14] for our setting). For an admissible control $a \in \mathcal{A}$, we can then define the equivalent martingale measures

$$dP^u = M_T^u dP.$$

Under P^u, the process $B_t^u = B_t - \int_0^t \theta_s^u ds$, is a Brownian motion by Girsanov's theorem and thanks to the equivalence of the measures P^u and P, we have P^u-almost surely

$$X_t = \varphi + \int_0^t b(s,X_s,u_s)ds + \int_0^t \sigma(s,X_s)dB_s^u.$$

The weak control problem then becomes

$$\mathcal{V}^w = \inf_{a\in\mathcal{A}} J^w(u) \text{ where } J^w(u) = \mathbf{E}^{P^u}\left[g(X_T) + \int_0^T f(t,X_t,u_t)dt\right],$$

and if $\mathcal{A}$ contains only processes adapted to the filtration generated by X, we have $\mathcal{V}^w \leq \mathcal{V}$. While the control problems look similar, they are conceptually quite different: In the strong formulation, the underlying Brownian motion is fixed and we control the paths of X. In contrast, the weak formulation fixes the paths X and we control the distribution P^u. We would like to apply the martingale representation theorem (with respect to the P^u-Brownian motion B^u), to obtain the backward equation

$$Y_t^u = g(X_T) + \int_t^T f(s,X_s,u_s)ds - \int_t^T Z_s^u dB_s^u, \quad P^u\text{-almost surely,}$$

so that $J(u) = Y_0^u$. However, $X_T \hat{\in} \mathcal{F}_T^B$ and in general the inclusion $\mathcal{F}_T^{B^u} \subset \mathcal{F}_T^B$ is strict. This implies that we cannot directly apply the martingale representation theorem but require an additional approximation step, which we note in the following lemma.

Lemma 3.4. *If M^u is a true martingale, then for any $\xi \in L^2(\mathcal{F}_T, P^u)$ there is a unique $\mathcal{F}_t^B$-adapted process Z such that*

$$\xi = \mathbf{E}^{P^u}[\xi] + \int_0^T Z_t dB_t^u, \quad \mathbf{E}^{P^u}\int_0^T \|Z_t\|^2 dt < \infty.$$

Proof. Since we were not able to locate a reference with a complete proof, we provide it here for the sake of completeness. We instead show that the BSDE

$$Y_t = \xi - \int_t^T Z_t dB^u_t,$$

has a $\{\mathcal{F}_t\}$-adapted solution, which implies the claim directly. A formal computation using Itô's formula and the definition of the shifted Brownian motion B^u yields using $dM^u_t = M^u_T \theta^u_t dB_t$,

$$\begin{aligned} M^u_T Y^u_T &= M^u_T \xi - \int_t^T \theta^u_s M^{\theta^u}_s Z_s ds - \int_t^T M^{\theta^u}_s \theta^u_s Y_s dB_s - \int_t^T M^{\theta^u}_s Z_s dB^u_s \\ &= M^u_T \xi - \int_t^T M^{\theta^u}_s (\theta^u_s Y_s + Z_s) dB_s. \end{aligned}$$

Thus, we want to show that the BSDE for $(\overline{Y}_t, \overline{Z}_t) = (M^{\theta^u}_t Y_t, M^{\theta^u}_t(\theta^u_s Y_s + Z_s))$ with the $\mathcal{F}_T$-measurable terminal condition $\overline{\xi} = M^u_T \xi$ has a solution and then recover the solution (Y, Z) to the original equation. Note that $\overline{\xi}$ need not be in $L^2(\mathcal{F}_T, P)$.

Bounded terminal condition. Let $\xi \in L^2(\mathcal{F}_T; P^u)$ be bounded and introduce the sequence of stopping times

$$\tau_n := \inf\{t \geq 0 : M^{\theta^u}_t \geq n\} \wedge T.$$

Since $M^{\theta^u}_{\tau_n} \mathbf{E}[\xi | \mathcal{F}_{\tau_n}] \in L^2(\mathcal{F}_{\tau_n}, P)$, there is a unique pair $(\overline{Y}^n, \overline{Z}^n)$ such that

$$\overline{Y}^n_t = M^{\theta^u}_{\tau_n} \mathbf{E}[\xi | \mathcal{F}_{\tau_n}] - \int_t^{\tau_n} \overline{Z}_t dB_t,$$

where we understand the BSDE with a (bounded) random terminal time (as defined in Proposition 2.15). Passing to the measure P^u and the P^u-Brownian motion B^u, we define

$$Y^n_t := (M^{\theta^u}_{t\wedge\tau_n})^{-1} \overline{Y}^n_{t\wedge\tau_n}, \quad Z^n_t := (M^u_t)^{-1}(\overline{Z}^n_t - \theta^u_t \overline{Y}_t)\, \mathbb{1}_{[0,\tau_n]}(t).$$

The same computation as before then shows that

$$Y^n_t = \mathbf{E}[\xi | \mathcal{F}_{\tau_n}] - \int_t^T Z^n_t dB^u_t.$$

By assumption, $\tau_n \to \infty$ and thus $\mathbf{E}[\xi | \mathcal{F}_{\tau_n}] \to \xi$ P-almost surely and by equivalence also P^u-almost surely. Thanks to the boundedness, the convergence also holds in $L^2(P^u)$ by dominated convergence and we can use the stability Proposition 2.9 to conclude

$$\mathbf{E}^{P^u}\left[\sup_t \left|Y^n_t - Y^m_t\right|^2 + \int_0^T \|Z^n_t - Z^m_t\|^2 dt\right] \to 0.$$

This means the limit of the Cauchy sequence is the desired pair (Y,Z).

General terminal condition. The general case follows by approximation with $\xi_n = (-n) \vee \xi \wedge n$. As $|\xi_n| \leq \xi \in L^2(\mathcal{F}_T, P^u)$ and $\xi_n \to \xi$ P^u-almost surely, dominated convergence implies $E^{P^u}[|\xi_n - \xi|^2] \to 0$ as $n \to \infty$. If (Y^n, Z^n) are the solutions to the BSDEs with the bounded terminal conditions ξ_n constructed in the first step, the same argument shows that (Y^n, Z^n) is Cauchy and with the limit satisfying the desired BSDE and measurability. □

Returning to the original P-Brownian motion B, again using the equivalence of P^u and P,

$$Y_t^u = g(X_T) + \int_t^T f(s,X_s,u_s) + Z_s^u \theta(s,X_s,u_s)ds - \int_t^T Z_s^u dB_s, \quad P\text{-almost surely.}$$

Define the Hamiltonians

$$H^*(s,x,z) := \inf_{a\in\Pi} H(s,x,z,a) := \inf_{a\in\Pi} f(s,x,a) + z\theta(s,x,a),$$

and the optimally controlled weak formulation for the BSDE

$$Y_t^* = g(X_T) + \int_t^T H^*(s,X_s,Z_s^*)ds - \int_t^T Z_s^* dB_s.$$

Proposition 3.5. *Let M^{θ^u} be a true martingale for every $u \in \mathcal{A}$. If the BSDEs with parameters (g,H) and (g,H^*) have unique solutions, then*

$$Y_0^* = \mathcal{V}^w.$$

Proof. By the comparison theorem (c.f. Theorem 2.13) we have $Y^* \leq Y^u$ for any $u \in \mathcal{A}$. Note that the comparison theorem may be used in this context since the terminal values $g(X_T)$ are fixed and independent of the control a. This is not the case for the strong formulation, where both the generator and the terminal condition need to be optimised simultaneously.

For the remaining inequality, let $\varepsilon > 0$. By the definition of the Hamiltonians, there is a Borel measurable function I^ε, such that

$$H(t,x,z,I^\varepsilon(t,x,z)) \leq H^*(t,x,z) + \varepsilon.$$

The a priori estimates (Proposition 2.9) and straightforward computation show that for $u^\varepsilon = I^\varepsilon(t,x,z)$, the difference satisfies $Y^{u^\varepsilon} - Y^* \leq T\varepsilon$. As this holds for any $\varepsilon > 0$ we conclude $Y_0^* = \mathcal{V}^w$. □

Remark 3.6. Suppose there is a control $u^* = I^*(t,X_t,Z_t)$ such that

$$H^*(t,x,z) = H(t,x,z,u^*).$$

Then by uniqueness $Y^* = Y^{u^*}$. In this case, given a solution and an optimal control u^* in the weak formulation, we can return to the original problem and ask whether the control is also optimal in the strong formulation. Roughly speaking, the difference between the strong and weak formulation comes down to the difference between strong and weak solutions to SDEs: For an optimal feedback control in the weak formulation $u^*(X_\cdot)$, the SDE

$$X_t = \varphi + \int_0^t b(s, X_s, u_s^*(X_s))ds + \int_0^t \sigma(s, X_s)dB_s,$$

in general does not admit a strong solution: The control u^* need not even be continuous in (s, x). The weak formulation is therefore much more likely to have a minimiser than the strong formulation even if the value functions may often coincide.

Remark 3.7. In general, $\mathrm{Law}_{P^{u^w}}(u^w, B^{u^w}) \neq \mathrm{Law}_P(u^w, B)$ and thus $J(u^w) \neq J^w(u^w)$. But if we suppose that $u^w \in \mathcal{A}$ is $\mathcal{F}^{B^u}$-adapted, then there is a function $v(B^u) = u^w$ and consequently $J(v(B^u)) = J^w(u)$. In the same way, if the a control u^s for the strong formulation happens to be $\mathcal{F}^{X^u}$-adapted, there is a v such that $J^w(v(B)) = J(u^s)$. Combined, if u^s and u^w are optimal for the strong and weak formulation respectively and they satisfy the measurability properties above,

$$\mathcal{V} = \mathcal{V}^w.$$

3.3 Relating the Weak and Strong Formulation

We want to understand the relation between the weak and strong value function more directly without relying on the Hamilton-Jacobi-Bellman equation. As indicated in Remark 3.7, this comes down to finding representations of the controls as functions of the Brownian motion on the path space.

For this section only, we introduce some additional conventions. We assume for simplicity that the Brownian motion B is the canonical process $B_t(\omega) = \omega_t$ on the canonical probability space $(C([0,T];K), \{\mathcal{F}_t\}, P)$. Let us also introduce the translations

$$\tau_u(\omega)_t = \omega_t - \int_0^t u_s(\omega)ds,$$

by a stochastic process u on this probability space. Using these definitions, we see that $B^u = \tau_u(B)$ and $P^u \circ \tau_u^{-1} = P$.

Definition 3.8. We say a stochastic process $u \in \mathcal{A}$ is *simple* if there is an increasing sequence $\{0 = t_0 < t_1 < \ldots < t_N = T\}$ and bounded random variables $\{\xi_j\}_j$ such that

$$u_t(\omega) = \xi_0(\omega)\,\mathbb{1}_0(t) + \sum_{j=1}^{N} \xi_j(\omega)\,\mathbb{1}_{(t_j,t_{j+1}]}(t).$$

We denote the class of simple processes in $\mathcal{A}$ by $\mathcal{A}_S$.

With this notation, we introduce the value functions

$$\tilde{\mathcal{V}} := \inf_{u \in \mathcal{A}_S} J(u); \qquad \tilde{\mathcal{V}}^w := \inf_{u \in \mathcal{A}_S} J^w(u).$$

As $\mathcal{A}_S \subset \mathcal{A}$, we also have $\tilde{\mathcal{V}} \geq \mathcal{V}$ and $\tilde{\mathcal{V}}^w \geq \mathcal{V}^w$. We will assume that the SDE

$$X_t^u = \varphi + \int_0^t b(s, X_s^u, u) + \int_0^t \sigma(s, X_s) dW_s,$$

has a unique strong solution for $u \in \mathcal{A}_s$. For now, let us take note that for simple controls, the value functions of the strong and weak formulation coincide.

Lemma 3.9. $\tilde{\mathcal{V}} = \tilde{\mathcal{V}}^w$.

Proof. We only show $\tilde{\mathcal{V}}^w \leq \tilde{\mathcal{V}}$. The other direction follows in the same way. Suppose that $u \in \mathcal{A}_S$ can be written as

$$u_t(\omega) = \xi_0(\omega) \mathbb{1}_{\{0\}}(t) + \sum_{j=1}^N \xi_j(\omega) \mathbb{1}_{(t_j, t_{j+1}]}(t),$$

where we used the same notation as in Definition 3.8. We will show that there is a control $\tilde{u} \in \mathcal{A}$ such that

$$\mathrm{Law}_P(u, B) = \mathrm{Law}_{P^{\tilde{u}}}(\tilde{u}, B^{\tilde{u}}). \tag{3.5}$$

By the uniqueness of the solution to the SDE (2.7), this implies $J^w(\tilde{u}) = J(u)$ and consequently $\mathcal{V}^w \leq \mathcal{V}$. To show (3.5), let $\tilde{\xi}_0 = \xi_0$ and define recursively

$$\tilde{\xi}_{j+1}(\omega) = \xi_{j+1}(\omega) - \sum_{i=1}^j \tilde{\xi}_i(\omega)(t_i - t_{i-1})).$$

Then, $\tilde{\xi}_j \hat{\in} \mathcal{F}_{t_j}$ and we can introduce the simple process

$$\tilde{u}_t(\omega) = \tilde{\xi}_0(\omega) \mathbb{1}_{\{0\}}(t) + \sum_{i=1}^j \tilde{\xi}_i(\omega) \mathbb{1}_{(t_{i-1}, t_i]}(t) = u(\tau_{\tilde{u}}(\omega))_t$$

Thus, denoting again $B_t^{\tilde{u}} = B_t - \int_0^t \tilde{u}_s ds$,

$$\begin{aligned}
&P^{\tilde{u}}(B^{\tilde{u}} \in U, \tilde{u} \in V) \\
&= P^{\tilde{u}}(\{\omega |\, \tau_{\tilde{u}}(\omega) \in U;\ u(\tau_{\tilde{u}}(\omega)) \in V\}) \\
&= P^{\tilde{u}}(\{\omega |\, \omega \in \tau_{\tilde{u}}^{-1}(U);\ u(\omega) \in \tau_{\tilde{u}}^{-1}(V)\}) \\
&= P(\{B \in U; u \in V\}),
\end{aligned}$$

for Borel measurable sets $U \subset C([0,T];K)$ and $V \subset L^2([0,T];K)$. But this proves (3.5) and hence the claim. $\square$

Finally, we would like to remove the restriction to simple functions and show that the value functions of the weak and strong formulation coincide. This of course requires appropriate continuity assumptions on the cost functional J. In the strong formulation, this continuity is usually a simple consequence of the continuity of f and g and the density of the simple processes. Indeed, in our special case this is the content of Lemma 4.21. The situation in the weak formulation is more subtle and requires additional technical assumptions to make the approximation rigorous. As we will only require the lower bound later, this is not our concern. We state an approximation result for future reference.

Lemma 3.10. *If X is a bounded and adapted process, then there is a sequence $\{X^n\}_{n\in\mathbb{N}}$ of simple processes such that*

$$\lim_{n\to\infty}\int_0^T \left|X_t^n - X_t\right|^2 dt = 0.$$

Proof. See e.g. [39, Lemma 3.2.4]. □

4. A Stochastic Control Problem for the Sine-Gordon Model

We are now ready to return to the main object of interest for this thesis, which is concerned with the construction of the sine-Gordon model on $\mathbb{R}^2$. As a first step, we have to construct a family of approximate measures

$$\nu_{\mathrm{SG}}^{T,\xi} := \frac{\exp(-V_T^{\xi}(\varphi))\mu^T(d\varphi)}{\int \exp(-V_T^{\xi}(\varphi))\mu^T(d\varphi)},$$

where $T \in [0, \infty)$ is a small-scale cut-off, $\xi : \mathbb{R}^2 \to [0, 1]$ is a smooth spacial cut-off and $V_T^{\xi} = \int \xi \alpha_T \cos(\beta\varphi)$ is the renormalisation of V corresponding to the regularisation μ^T of the free field.

4.1 Decomposing the Free Field

The objective in this section is the construction of a convenient decomposition of the free field and, in turn, a renormalisation of the potential V which ensures a non-trivial limit. We also collect some technical estimates on the covariance of the regularised free field. For some additional background and references on Gaussian measures, we refer to the appendix.

4.1.1 The heat kernel decomposition

Given two independent Gaussian random variables with covariance $\mathcal{C}_1, \mathcal{C}_2$, the stability of the Gaussian distribution means that their sum is a Gaussian measure on $\mathscr{S}'(\mathbb{R}^2)$ with covariance $\mathcal{C} = \mathcal{C}_1 + \mathcal{C}_2$. In other words, a (discrete) decomposition of a Gaussian can be obtained by simply decomposing its variance $\mathcal{C}$ as a sum of covariances. In complete analogy, a (continuous) decomposition of a covariance operator $\mathcal{C} = \int_0^{\infty} \dot{\mathcal{C}}_t dt$ in terms of positive semi-definite operators $\mathcal{C}_t$, yields a decomposition of the Gaussian random variable with covariance $\mathcal{C}$ as the Wiener integral $\int_0^{\infty} \sqrt{\dot{\mathcal{C}}_t} dB_t$ thanks to Itô's formula and Lemma B.6.

Thus, the decomposition (1.5) comes down to suitably decomposing the covariance $(-\Delta + m^2)^{-1}$ of the free field. We will use a heat-kernel decomposition of the form

$$(-\Delta + m^2)^{-1} = \int_0^{\infty} Q_t^2 dt \quad \text{with} \quad Q_t := \left(\frac{1}{t^2} e^{-\frac{-\Delta + m^2}{t}}\right)^{\frac{1}{2}},$$

mainly because this decomposition plays nicely with exponentially weighted L^p-spaces (see Lemma 4.6). By the considerations above, we would like to define the process

$$W_T := \int_0^T Q_s dB_s, \tag{4.1}$$

as an approximation to the free field.

Proposition 4.1. *For any $T < \infty$, the random variables W_T are Gaussian with covariance $\int_0^T Q_s^2 ds$ and $W_T \in L^2(\rho_\ell)$ almost surely. Moreover, the sequence $\{W_T;\, T \in [0,\infty)\}$ converges in $L^2(P, H^{-\delta}(\rho_\ell))$ to $W_\infty \sim \mu$. Here, μ denotes the Gaussian free field, that is the centred Gaussian measure on $\mathscr{S}'(\mathbb{R}^2)$ with covariance $(-\Delta + m^2)^{-1}$.*

Proof. As the exponential of self-adjoint operators, the operators Q_t defined in this way are again self-adjoint and positive. Furthermore, since the kernel k_t is continuous we can compute the trace on L^2 by integrating over the diagonal

$$\begin{aligned}
\|Q_t\|^2_{\mathcal{L}_2(L^2(\rho_\ell))} &= \mathrm{Tr}_{L^2(\rho_\ell)}\left(\int_0^t Q_s^2 ds\right) \\
&= \int_0^t ds \int_{\mathbb{R}^2} dx k_s(x,x)\langle x\rangle^{-2\ell} \\
&= \int_0^t k_s(0,0) ds \int_{\mathbb{R}^2} \langle x\rangle^{-2\ell} dx \\
&\leq C(\log(t \vee 1) + 1) < \infty,
\end{aligned}$$

provided $\ell \geq \frac{d}{2}$. Therefore, the stochastic integral (4.1) is well-defined and the process $\{W_T\}_{T\geq 0}$ induces Gaussian measure on $L^2(\rho_\ell)$ with covariance

$$\mathcal{C}_t = \int_0^t Q_s^2 dt = (-\Delta + m^2)^{-1} e^{-\frac{-\Delta+m^2}{t}}.$$

It remains to check the convergence in $H^{-\delta}(\rho_\ell)$. To this end, note that by Itô's formula,

$$\mathbf{E}\|W_T - W_\infty\|^2_{H^{-\delta}(\rho_\ell)} = \int_T^\infty \|Q_s\|^2_{\mathcal{L}_2(H^{-\delta}(\rho_\ell))} ds.$$

Hence, the claim will follow once we compute the trace of Q_s^2 on $H^{-\delta}(\rho_\ell)$. Using the relation between the inner products on H^s and L^2

$$\langle f, g\rangle_{H^s(\rho_\ell)} = \langle(-\Delta+1)^{\frac{s}{2}} f, (-\Delta+1)^{\frac{s}{2}} g\rangle_{L^2(\rho_\ell)},$$

we compute for $f \in H^s(\rho_\ell)$

$$\begin{aligned}
&(-\Delta+1)^{\frac{s}{2}} Q_s^2 (-\Delta+1)^{\frac{s}{2}} f(x) \\
=&\mathscr{F}^{-1}((|\xi|^2+1)^s \frac{1}{t^2} e^{\frac{(-|\xi|^2-m^2)}{t}} \mathscr{F}(f))(x) \\
=& \int_{\mathbb{R}^2} dy f(y) \left(\int_{\mathbb{R}^2} \frac{d\xi}{(2\pi)^d} (|\xi|^2+1)^s \frac{1}{t^2} e^{\frac{(-|\xi|^2-m^2)}{t}} e^{i\langle y-x,\xi\rangle}\right),
\end{aligned}$$

where $\mathscr{F}(f) = \hat{f}$ denotes the Fourier transform of f. In other words, the integral kernel of Q_s^2 on $H^s(\rho_\ell)$ is given by

$$\tilde{K}_t(x,y) = \int_{\mathbb{R}^d} \frac{1}{(2\pi)^d} (|\xi|^2+1)^s \frac{1}{t^2} e^{-\frac{-|\xi|^2-m^2}{t}} e^{i\langle y-x,\xi\rangle} d\xi.$$

Since $\tilde{K}_t$ is continuous in x, y, the trace is again computed as the integral over the diagonal and

$$\begin{aligned}
&\int_T^\infty \|Q_s\|^2_{\mathcal{L}_2(H^s(\rho_\ell))} ds \\
&= \int_T^\infty dt \int_{\mathbb{R}^2} dx \langle x\rangle^{-2n} \tilde{K}_t(0,0) \\
&= \int_{\mathbb{R}^2} d\xi \frac{1}{(2\pi)^2} (|\xi|^2+1)^s \int_T^\infty dt \frac{1}{t^2} e^{-\frac{|\xi|^2+m^2}{t}} \int_{\mathbb{R}^2} dx \langle x\rangle^{-2\ell} \\
&= \int_{\mathbb{R}^2} d\xi \frac{1}{(2\pi)^2} \frac{(|\xi|^2+1)^s}{|\xi|+m^2} (1-e^{-\frac{|\xi|^2+m^2}{T}}) \int_{\mathbb{R}^2} dx \langle x\rangle^{-2\ell}.
\end{aligned}$$

Now,

$$\int_{\mathbb{R}^2} \frac{1}{(2\pi)^2} \frac{(|\xi|^2+1)^s}{|\xi|+m^2} d\xi < \infty,$$

if and only if $s < 1 - \frac{d}{2} = 0$. Finally, $(1 - e^{-\frac{|\xi|^2+m^2}{T}})$ is bounded and converges to 0 as $T \to \infty$. Dominated convergence then concludes the argument. □

Remark 4.2. The same computation also shows that the Gaussian free field can be realised as the canonical random variable on the Sobolev space $H^s(\langle x\rangle^{-\ell})$ with $s < \frac{2-d}{2}$ and $\ell > \frac{d}{2}$.

By the considerations above, the random variables $\{W_T, T \in [0,\infty)\}$ are genuine functions and we can define the measures $\mu^T := \mathrm{Law}(W_T)$ as a smooth approximation to the free field.

4.1.2 Estimates on the Kernel

In this section, we collect some useful technical estimates and properties of the heat kernel decomposition.

To derive explicit estimates for $p \neq 2$, the following representation of Q_s will be helpful.

Lemma 4.3. *The integral kernel k^{Q_t} of Q_t and $k^{Q_t^2}$ of Q_t^2 on $L^2(\mathbb{R}^2)$ are given by*

$$k_t^{(\frac{1}{2})}(x,y) := k^{Q_t}(x,y) = \frac{1}{2\pi} e^{-\frac{m^2}{2t}} e^{-2t|x-y|^2},$$

and

$$k_t(x,y) := k^{Q_t^2}(x,y) = \frac{1}{4\pi t} e^{-\frac{m^2}{t}} e^{-\frac{t}{4}|x-y|^2}.$$

In particular, k_t is constant on the diagonal $x = y$ and, for a constant C_m depending on the mass, $k_t(0,0) \leq C_m \langle t \rangle^{-1}$ and

$$\int_0^t k_s(0,0)ds \leq \frac{1}{4\pi} \log(t \vee 1) + C.$$

Proof. This is an immediate consequence of the fact that $e^{t\Delta}$ corresponds to the heat kernel on $L^2(\mathbb{R}^d)$, defined by

$$\kappa_t(x,y) = \frac{1}{(4\pi t)^{d/2}} e^{-|x-y|^2/4t},$$

and the definition of Q_t. □

Lemma 4.4. *For any ℓ,*

$$\|Q_t f\|_{L^2(\rho_\ell)} \leq C\langle t \rangle^{-1} \|f\|_{L^2(\rho_\ell)}, \text{ and } \|Q_t f\|_{L^\infty} \leq \langle t \rangle^{-1} \|f\|_{L^\infty}.$$

In particular, Q_t is a bounded operator $L^2(\rho_\ell) \to L^2(\rho_\ell)$.

Proof. As

$$\|k_t^{(\frac{1}{2})}\|_{L^1} = \int_{\mathbb{R}^2} \frac{1}{2\pi} e^{-\frac{m^2}{t}} e^{-2t|x|^2} dx = \frac{C}{t} e^{-\frac{m^2}{t}} \leq C\langle t \rangle^{-1},$$

the L^∞ bound is a direct consequence of Young's inequality

$$\|Q_t f\|_{L^\infty} = \|k_t^{(\frac{1}{2})} * f\|_{L^\infty} = \|k_t^{(\frac{1}{2})}\|_{L^1} \|f\|_\infty.$$

To show the L^2-bound, we compute for any $f \in \mathscr{S}(\mathbb{R}^2)$ (so that also $Q_t f \in \mathscr{S}(\mathbb{R}^2)$) from Parseval's identity and the fact that $\rho_\ell \in \mathscr{S}'(\mathbb{R}^2)$,

$$\begin{aligned}
\|Q_t f\|_{L^2(\rho_\ell)} &= (2\pi)^2 \|\mathscr{F}(\rho_\ell Q_t f)\|_{L^2} \\
&= (2\pi)^2 \|\mathscr{F}(\rho_\ell) * \frac{1}{t} \exp(-\frac{m^2 + |\xi|^2}{2t}) \mathscr{F}(f)(\xi)\|_{L^2} \\
&= \frac{(2\pi)^2}{t} e^{-\frac{m}{2t}} \|\int_{\mathbb{R}^2} \mathscr{F}(\rho_\ell)(k-\xi) e^{-\frac{|\xi|^2}{2t}} \mathscr{F}(f)(\xi) d\xi\|_{L^2_k} \\
&\leq C\langle t \rangle^{-1} \|\mathscr{F}(\rho_\ell) * \mathscr{F}(f)\|_{L^2} \\
&= C\langle t \rangle^{-1} \|f\|_{L^2(\rho_{2\ell})}.
\end{aligned}$$

The claim now follows from the density of $\mathscr{S}(\mathbb{R}^2)$ in $L^2(\rho_\ell)$. □

Lemma 4.5.

$$\|\int_0^T Q_r u_r dr\|^2_{H^1(\rho_\ell)} \leq C \int_0^\infty \|u_t\|^2_{L^2(\rho_\ell)} dt.$$

Proof. We again use Parseval's identity and the definition of the Sobolev norms

$$\|f\|_{H^s(\rho)} = \|\langle\nabla\rangle^s f\|_{L^2(\rho)},$$

to compute for $f \in \mathscr{S}(\mathbb{R}^2)$,

$$\begin{aligned}
&\|\int_0^T Q_t u_t dt\|^2_{H^1(\rho_\ell)} \\
=&(2\pi)^4\|\int_0^T dt\, \mathscr{F}(\rho_\ell(1-\Delta)^{\frac{1}{2}}Q_t u_t)\|_{L^2} \\
=&(2\pi)^4\|\int_0^T dt\, \hat{\rho}_\ell * \left(\frac{(|\cdot|^2+1)^{\frac{1}{2}}}{t}e^{-\frac{m^2+|\cdot|^2}{2}}\hat{u}_t\right)\|^2_{L^2} \\
=&(2\pi)^4\|\int_0^T dt\int_{\mathbb{R}^2} dk\hat{\rho}_\ell(\xi-k)\frac{(|k|^2+1)^{\frac{1}{2}}}{t}e^{-\frac{m^2+|k|^2}{2t}}\hat{u}_t(k)\|^2_{L^2(d\xi)}, \\
=&(2\pi)^4\|\int_0^T dt\int_{\mathbb{R}^2} dk\left(\frac{(|k|^2+1)^{\frac{1}{2}}}{t}e^{-\frac{m^2+|k|^2}{2t}}\right)(\hat{u}_t(k)\hat{\rho}_\ell(\xi-k))\|^2_{L^2_\xi},
\end{aligned}$$

Now applying Hölder's inequality in time, we have

$$\begin{aligned}
\leq& C\|\int_{\mathbb{R}^2} dk\left(\int_0^T dt\frac{(|k|^2+1)}{t^2}e^{-\frac{m^2+|k|^2}{t}}\right)^{\frac{1}{2}}\left(\int_0^T dt\,\hat{u}_t^2(k)\hat{\rho}_\ell^2(\xi-k)\right)^{\frac{1}{2}}\|^2_{L^2(d\xi)} \\
\leq& C\sup_k\left(\frac{|k|^2+1}{|k|^2+m^2}\right)\|\int_{\mathbb{R}^2} dk\left(\int_0^T dt\,\hat{u}_t^2(k)\hat{\rho}_\ell^2(\xi-k)\right)^{\frac{1}{2}}\|^2_{L^2(d\xi)} \\
=& C\|\int_{\mathbb{R}^2} dk\left(\int_0^T dt\,\hat{u}_t^2(k)\hat{\rho}_\ell^2(\xi-k)\right)^{\frac{1}{2}}\|^2_{L^2(d\xi)}
\end{aligned}$$

Expanding the L^2-norm in ξ and applying Hölder's inequality now for the integral over k, we arrive at the desired estimate

$$\begin{aligned}
\leq& C\int_0^T dt\iint dk d\xi\,\hat{u}_t^2(k)\hat{\rho}_\ell^2(\xi-k) \\
=& C\int_0^T dt\|\hat{u}_t(k)*\hat{\rho}_\ell\|^2_{L^2} \\
=& C\int_0^T \|u_t\|^2_{L^2(\rho_\ell)}dt.
\end{aligned}$$

For a general $f \in L^2(\rho_\ell)$ the claim follows by approximation with functions in $\mathscr{S}(\mathbb{R}^2)$. $\square$

Lemma 4.6. *Let $|\gamma| < m$, and $w_\gamma(x) = \exp(\gamma |x|)$, the estimate from Lemma 4.4 stays true for $\langle x\rangle^{-\ell}$ replaced by w_γ, that is*

$$\|\int_0^T Q_r u_r dr\|_{H^1(w_\gamma)}^2 \leq C \int_0^\infty \|u_t\|_{L^2(w_\gamma)}^2 dt.$$

Proof. We write $w := w_\gamma$ and split the integral as

$$\|\int_0^T Q_r u_r dr\|_{H^1(w)} \leq \|\int_0^m Q_r u_r dr\|_{H^1(w)} + \|\int_m^T Q_r u_r dr\|_{H^1(w)} =: (\mathrm{I}) + (\mathrm{II}).$$

Estimate for (II). To ensure convergence for large times t, the Lipschitz estimate

$$(w(x) - w(y))w^{-1}(y) \leq |x-y|,$$

will be crucial to compensate divergencies introduced by $\nabla_x k_t^{(\frac{1}{2})}$. With this estimate and Hölder's inequality in time, we have

$$\begin{aligned}
(\mathrm{II})^2 =& \|\int_m^T Q_r u_r\|_{H^1(w)}^2 \\
\leq& \|\int_m^T Q_r u_r w dr\|_{H^1}^2 + \|\int_0^T dt \int_{\mathbb{R}^2} dy (1+\nabla_x) k_t(x,y)(w(y)-w(x)) u_t(y)\|_{L^2(dx)}^2 \\
\leq& C\int_m^T \|u_t\|_{L^2(w)}^2 dt + \|\int_{\mathbb{R}^2} dy \|(1+\nabla_x) k_t^{(\frac{1}{2})}(x,y)\|_{L^2[m,\infty)} |x-y| \\
&\times w(y)\|u_t(y)\|_{L^2[m,\infty)}\|_{L^2(dx)}^2.
\end{aligned}$$

By Young's convolution inequality this can be estimated as

$$\leq C\left(1 + \left(\int_{\mathbb{R}^2} \|(1+\nabla_x) k_t^{(\frac{1}{2})}(x,0)|x|\|_{L_t^2([m,\infty))} dx\right)^2\right) \int_0^T \|u_t\|_{L^2(w)}^2 dt.$$

It remains to check that the norm of the kernel is finite. We first estimate the $L^2([m,\infty))$-norm of the kernel

$$\begin{aligned}
&\|(1+\nabla_x) k_t^{(\frac{1}{2})}(x,y)\|_{L^2([m,\infty)}^2 \\
\leq& C\int_m^\infty \left(-4t^2(x-y)^2 + 1\right) e^{-\frac{m^2}{2t}} e^{-4t|x-y|^2} dt \\
\leq& p(|x-y|) e^{-4m|x-y|^2} |x-y|^{-4} \\
\leq& C e^{-2m|x-y|^2} |x-y|^{-4}
\end{aligned}$$

where p is a polynomial. Thus, the $L^1(\mathbb{R}^2)$-norm can be computed as

$$\begin{aligned}
&\int_{\mathbb{R}^2} \|(1+\nabla_x)k_t^{(\frac{1}{2})}(x,0)\,|x|\|_{L_t^2([m,\infty))}dx \\
\leq & C\int_{\mathbb{R}^2} e^{-2m|x|^2}\,|x|^{-4}\,|x|\,dx \\
\leq & C\int_0^\infty e^{-2mr^2}r^{-1}dr < \infty.
\end{aligned}$$

This shows the estimate on (II).

Estimate on (I). To this end we use the estimate

$$e^{-\frac{m^2}{2t}}e^{-2t|x|^2} \leq e^{-m|x|}, \tag{4.2}$$

for $t \leq m$, combined with the assumption that $|\gamma| < m$. Indeed, applying (4.2) yields

$$\begin{aligned}
(\mathrm{I})^2 = &\|\int_0^m Q_r u_r dr\|_{H^1(w)}^2 \\
=&\|\int_0^m dt\int_{\mathbb{R}^2} dy\left|(1+\nabla_x)k^{(\frac{1}{2})}(x,y)u_t(y)w(x)\right|\|_{L^2(dx)}^2 \\
=&C\|\int_0^m dt\int_{\mathbb{R}^2} dy\left(4t^2\,|x-y|^2+1\right)e^{-\frac{m^2}{2t}}e^{-4t|x-y|^2}e^{\gamma|x|}u_t(y)\|_{L^2(dx)}^2 \\
\leq&C\|\int_0^m dt\int_{\mathbb{R}^2} dy\left(4t^2\,|x-y|^2+1\right)e^{-m|x-y|}e^{\gamma|x|}u_t(y)\|_{L^2(dx)}^2,
\end{aligned}$$

By the triangle inequality $e^{\gamma|x|} \leq e^{|\gamma||x-y|}e^{\gamma|y|}$ and thus

$$\begin{aligned}
\leq&C\|\int_0^m dt\int_{\mathbb{R}^2} dy\left(4t^2\,|x-y|^2+1\right)e^{(\gamma-m)|x-y|}e^{\gamma|y|}u_t(y)\|_{L^2(dx)}^2 \\
\leq&C\|\int_{\mathbb{R}^2} dy\left(4m^2\,|x-y|^2+1\right)e^{(\gamma-m)|x-y|}\int_0^T dt e^{\gamma|y|}u_t(y)\|_{L^2(dx)}^2.
\end{aligned}$$

Proceeding as before with Young's inequality we obtain

$$\begin{aligned}
\leq&C\left(\int_{\mathbb{R}^2} dx\left(4m^2\,|x|^2+1\right)e^{(\gamma-m)|x|}\right)\int_0^T \|u_t\|_{L^2(w)}^2 dt \\
\leq&C\int_0^T \|u_t\|_{L^2(w)}^2 dt,
\end{aligned}$$

which concludes the proof. □

Remark 4.7. While we will always use the more general polynomial weights for our estimates, Lemma 4.6 shows that we could instead also work with exponential weights. The exponential weights would be particularly useful to show e.g. exponential clustering for the limiting measure.

Lemma 4.8. *For any $\delta > 0$,*

$$\|\int_0^T Q_s u_s ds\|_{W^{1,\infty}} \leq C\|\langle s\rangle^{\frac{1}{2}+\delta} u_s\|_{L^\infty([0,T]\times\mathbb{R}^2)}.$$

Proof. For these estimates, we rely on the explicit kernel Lemma 4.3 for Q on L^2 and compute

$$\begin{aligned}
&\|\int_0^T Q_t u_t dt\|_{L^\infty(\mathbb{R}^2)} \\
&= \sup_{x\in\mathbb{R}^2}\left|\int_0^T dt\langle t\rangle^{-\delta}\int_{\mathbb{R}^2} dy k_t^{Q_t}(x,y)\langle t\rangle^{\delta} u_t(y)\right| \\
&= \|\langle t\rangle^{\delta} u_t\|_{L^\infty([0,T]\times\mathbb{R}^2)}\frac{1}{2\pi}\int_0^T dt e^{-\frac{m^2}{2t}}\langle t\rangle^{-\delta}\int_{\mathbb{R}^2} dy e^{-2t|y|^2} \\
&\leq \|\langle t\rangle^{\delta} u_t\|_{L^\infty([0,T]\times\mathbb{R}^2)}\frac{1}{4}\int_0^\infty t^{-1}\langle t\rangle^{-\delta} e^{-\frac{m^2}{2t}} dt \\
&\leq C\|\langle t\rangle^{\delta} u_t\|_{L^\infty([0,T]\times\mathbb{R}^2)} \\
&\leq C\|\langle t\rangle^{\frac{1}{2}+\delta} u_t\|_{L^\infty([0,T]\times\mathbb{R}^2)}.
\end{aligned}$$

In the same way, we can proceed for the derivative with

$$\nabla_x k_t^{(\frac{1}{2})}(x,y) = \frac{-2t}{\pi}(x-y)e^{-2t|x-y|^2} e^{\frac{m^2}{2t}},$$

to obtain

$$\begin{aligned}
&\|\int_0^T \nabla_x Q_t u_t\|_{L^\infty} \\
&= \|\langle t\rangle^{\frac{1}{2}+\delta} u_t\|_{L^\infty([0,T]\times\mathbb{R}^2)}\int_0^T dt\frac{2}{\pi}t^{1-\frac{1}{2}-\delta}e^{-\frac{m^2}{2t}}\int_{\mathbb{R}^2} dx e^{-2t|x|^2}|x| \\
&= \|\langle t\rangle^{\frac{1}{2}+\delta} u_t\|_{L^\infty([0,T]\times\mathbb{R}^2)}\int_0^T dt\frac{2}{\pi}t^{\frac{1}{2}-\delta}e^{-\frac{m^2}{2t}}\int_0^\infty dr e^{-2tr^2}r^2 \\
&= \|\langle t\rangle^{\frac{1}{2}+\delta} u_t\|_{L^\infty([0,T]\times\mathbb{R}^2)}\int_0^\infty \frac{\sqrt{2}}{4\sqrt{\pi}}t^{-\frac{3}{2}}t^{\frac{1}{2}-\delta}e^{-\frac{m^2}{t}} dt \\
&\leq C\|\langle t\rangle^{\frac{1}{2}+\delta} u_t\|_{L^\infty([0,T]\times\mathbb{R}^2)}.
\end{aligned}$$

□

4.1.3 The Martingale Renormalisation

The goal in this section is twofold: First, we want to find the renormalisation associated with the heat-kernel decomposition just introduced. In a second step, we show that with this renormalisation, $V_\infty(W_\infty)$ is a well-defined distribution. From the definition of W_t, its covariance operator on $L^2(\mathbb{R}^2)$ is given by $C_t = \int_0^t Q_s^2 ds$, that is,

$$\mathbf{E}[\langle W_t, f\rangle_{L^2(\mathbb{R}^2)}\langle W_t, g\rangle_{L^2(\mathbb{R}^2)}] = \langle C_t f, g\rangle_{L^2(\mathbb{R}^2)}.$$

Let $V_t(W_t) = \int \alpha_t \cos\beta W_t$ for some α_t to be determined. We compute using Itô's formula Lemma 2.3,

$$dV_t(W_t) = \left(\partial_t \alpha_t \int \cos(\beta W_t) + \frac{\beta^2}{2} k_t(0,0)\alpha_t \int \cos(\beta W_t)\right) dt + \nabla V_t(W_t) Q_t dB_t. \quad (4.3)$$

Choosing $\alpha_t := e^{\frac{\beta^2}{2}\int_0^t k_s(0,0)ds}$, where k_s is the kernel of Q_t^2, the drift vanishes. Now $\nabla V_t(W_t) Q_t$ is bounded and the stochastic integral is a martingale. From Lemma 4.3, we have that $\alpha_t \le C\langle t\rangle^{\frac{\beta}{8\pi}}$. In particular, for $\beta^2 < 4\pi$ this implies the bound $\alpha_t \le C\langle t\rangle^{\frac{1}{2}-\delta}$. We will rely on this fact extensively. Note that by definition of α_t, we also have that $v_t(x) = \alpha_t \sin(\beta x)$ satisfies the Fokker-Planck equation

$$\partial_t v_t = -\frac{1}{2} k_t(0,0)\partial_x^2 v_t. \quad (4.4)$$

Let us fix the following notation:

$$[\![\cos(\beta W_t)]\!] := \alpha_t \cos(\beta W_t) \quad \text{and} \quad [\![\sin(\beta W_t)]\!] := \alpha_t \sin(\beta W_t).$$

Remark 4.9. This definition of course coincides with the usual Wick-exponential or complex multiplicative chaos [37]. For a centred Gaussian random variable ζ on a Hilbert space H and any $h \in H$, it is defined by

$$[\![e^{i\beta\langle\zeta,h\rangle}]\!] := e^{\frac{\beta^2}{2}\mathbf{E}[\langle\zeta,h\rangle^2]} e^{i\beta\langle\zeta,h\rangle}.$$

Since the 2-dimensional free field is log-correlated, we expect polynomial correlations for the complex chaos and the sine-Gordon model (see also Remark 4.12).

We now want to show that the martingale $[\![\cos(\beta W_t)]\!]$ is also uniformly bounded in $L^2(P, H^{-1+\delta}(\rho_\ell))$, which will imply convergence in the same space and almost surely by the martingale convergence theorem. To this end, we roughly follow [3] and [38]. For the notation regarding the Besov spaces and the Littlewood-Paley blocks, we refer to Appendix A.

Proposition 4.10. *For any $p \ge 1$ and $\|\rho\|_{L^p(\mathbb{R}^2)} < \infty$, the Wick-ordered cosine satisfies*

$$\sup_{t\ge 0} \mathbf{E}\left[\|[\![\cos\beta W_t]\!]\|^p_{B_{p,p}^{-\frac{\beta^2}{4\pi}-2\delta}(\rho)} \right] < \infty,$$

and converges in $L^p(P, B_{p,p}^{-\frac{\beta^2}{4\pi}-2\delta}(\rho))$ and almost surely to a limit which we denote by $[\![\cos(\beta W_\infty)]\!]$. The analogous statement holds also for the Wick-ordered sine.

The proof relies on the following deterministic estimates on the quadratic variation.

Lemma 4.11. *The quadratic variation of the martingale $\Delta_i[\![\cos(\beta W_t)(x)]\!]$ satisfies uniformly in x, t for any $\delta > 0$,*

$$|\langle \Delta_i [\![\cos(\beta W_t)(x)]\!]\rangle_t| \leq C_\delta 2^{i\frac{\beta^2}{2\pi}+\delta}.$$

Proof. Using $[\![\cos \beta W_t]\!] = 1 - \int_0^t \beta [\![\sin(\beta W_s)]\!] dW_s$ and the definition of W_s, we can compute the quadratic variation of N^i as

$$\begin{aligned}
&|\langle \Delta_i [\![\cos(\beta W_t)]\!](x)\rangle_t| \\
=&\beta^2 \left| \int_{(\mathbb{R}^2)^2} dy_1 dy_2 \int_0^t \varphi_i(x-y_1)\varphi(x-y_2) \right.\\
&\qquad \left. \times \alpha_s^2 \sin(\beta W_s)(y_1)\sin(\beta W_s)(y_2)\langle dW(y_1), dW(y_2)\rangle_s \right| \\
=&\beta^2 \int_{(\mathbb{R}^2)^2} dy_1 dy_2 \int_0^t ds \left|\varphi_i(x-y_1)\varphi(x-y_2)\alpha_s^2 \sin(\beta W_s)(y_1)\sin(\beta W_s)(y_2) k_s(y_1,y_2)\right| \\
\leq&\beta^2 \int_{(\mathbb{R}^2)^2} dy_1 dy_2 \, |\varphi_i(x-y_1)\varphi(x-y_2)| \int_0^t ds \alpha_s^2 k_s(y_1,y_2).
\end{aligned}$$

We estimate the last integral with the change of variables $\tilde{s} = s\,|y_1 - y_2|^2$ to obtain,

$$\begin{aligned}
\int_0^t ds \alpha_s^2 k_s(y_1,y_2) \leq& \int_0^\infty s^{\frac{\beta^2}{4\pi}} e^{-\frac{s}{4}|y_1-y_2|^2} e^{-\frac{m^2}{s}} \frac{ds}{s} \\
=&|y_1-y_2|^{-\frac{\beta^2}{2\pi}} \int_0^\infty s^{\frac{\beta^2}{4\pi}-1} e^{-\frac{1}{4}s} e^{-\frac{m^2}{s}|x-y|^2} ds \\
\leq& C\,|y_1-y_2|^{-\frac{\beta^2}{2\pi}}.
\end{aligned}$$

Splitting up the integral $\int_{(\mathbb{R}^2)^2} = \int_{|y_1-y_2|<1} + \int_{|y_1-y_2|\geq 1}$ and estimating $\frac{1}{|y_1-y_2|} \leq 1$, this implies

$$\begin{aligned}
&|\langle \Delta_i [\![\cos(\beta W_t)(x)]\!]\rangle_t| \\
\leq& C\beta^2 \int_{(\mathbb{R}^2)^2} dy_1 dy_2 \, |\varphi_i(x-y_1)\varphi_i(x-y_2)|\,|y_1-y_2|^{-\frac{\beta^2}{2\pi}} \\
\leq& C\beta^2 \int_{|y_1-y_2|<1} dy_1 dy_2 \, |\varphi_i(x-y_1)\varphi_i(x-y_2)|\,|y_1-y_2|^{-\frac{\beta^2}{2\pi}} + C\beta^2 \|\varphi_i\|_{L^1}^2.
\end{aligned}$$

Now by Young's inequality, for $\frac{1}{p}+\frac{1}{q}=1$

$$
\begin{aligned}
&\int_{|y_1-y_2|<1} dy_1 dy_2\, |\varphi_i(x-y_1)\varphi_i(x-y_2)|\, |y_1-y_2|^{-\frac{\beta^2}{2\pi}} \\
&\leq \|\varphi_i\|_{L^q}\|\varphi\|_{L^1}\int_{B_1(0)} |y|^{-p\frac{\beta^2}{2\pi}}\, dy \\
&\leq \|\varphi_i\|_{L^q}\|\varphi\|_{L^1}\int_0^1 r^{-p\frac{\beta^2}{2\pi}+1} dr \\
&\leq \|\varphi_i\|_{L^q}\|\varphi\|_{L^1}\left(2-p(\frac{\beta^2}{2\pi})\right)^{-1},
\end{aligned}
$$

provided that $p\in(1,\frac{4\pi}{\beta^2})$. From the estimates on the L^p-norms of the Littlewood-Paley kernels (see (A.1)), for $i\geq 0$,

$$
\|\varphi_i\|_{L^1}\leq C, \quad \|\varphi_i\|_{L^q}\leq C2^{2i\frac{q-1}{q}},
$$

and consequently,

$$
|\langle\Delta_i[\![\cos(\beta W_t)]\!](x)|\rangle_t \leq C\beta^2 2^{2i\frac{q-1}{q}}.
$$

The condition on p implies that $q\geq\frac{1}{1-\frac{\beta^2}{4\pi}}$. Therefore, for q sufficiently close to $\frac{1}{1-\frac{\beta^2}{4\pi}}$, we have

$$
\frac{q-1}{q}\leq\frac{\beta^2}{4\pi}+\delta,
$$

which combined with Lemma 4.1.3 yields the claim. □

Proof of Proposition 4.10. By the definition of the norms we want to estimate

$$
\mathbf{E}\left[\|N\|^p_{B^{-s}_{p,p}(\rho)}\right]=\sum_{i\geq -1}2^{-ips}\,\mathbf{E}\left[\|\Delta_i N\|^p_{L^p(\rho)}\right]. \tag{4.5}
$$

The Burkholder-Davis-Gundy inequalities and the estimate from Lemma 4.11 allow us to compute

$$
\begin{aligned}
\mathbf{E}\left[\|\Delta_i N\|^p_{L^p(\rho)}\right] &= \int_{\mathbb{R}^2}\rho^p(x)\,\mathbf{E}[|\Delta_i[\![\cos(\beta W_t)]\!](x)|^p]dx \\
&\leq \int_{\mathbb{R}^2}\rho^p(x)\,\mathbf{E}[\langle\Delta_i[\![\cos(\beta W_t)]\!](x)\rangle_t]^{\frac{p}{2}}dx \\
&\leq C2^{i\frac{p\beta^2}{4\pi}+\frac{\delta}{2}}\int_{\mathbb{R}^2}\rho^p(x)dx.
\end{aligned}
$$

Hence, (4.5) is finite for

$$
\frac{\beta^2}{4\pi}+\frac{\delta}{2}-s\leq 0 \iff s>\frac{\beta^2}{4\pi}+\frac{\delta}{2}.
$$

□

Remark 4.12. The same computation also shows the polynomial correlations of the complex chaos,

$$\mathbf{E}[[\![\cos(\beta W_t)]\!](x)[\![\cos(\beta W_t)]\!](y)] \le C\beta^2 |x-y|^{-\frac{\beta^2}{2\pi}} .$$

With these definitions, we can make the formal measure from the introduction precise. For the Gaussian measures μ^T with covariance $C_T = \int_0^T Q_s^2 ds$, we define the measures

$$\nu_{\mathrm{SG}}^{T,\xi}(d\varphi) := \frac{\exp\{-V_T^{\xi}(\varphi)\}\mu^T(d\varphi)}{\int \exp\{-V_T^{\xi}(\varphi)\}\mu^T(d\varphi)},$$

where $V_T^{\xi}(\varphi) = \alpha_T \lambda \int \xi \cos(\beta\varphi)$ where $T \in [0,\infty)$ and $|\mathrm{supp}(\xi)| < \infty$.

Remark 4.13. (i) Here we also see that the renormalisation is indeed necessary to obtain a non-trivial limit: Indeed, the estimates on $k_t(0,0)$ show that $\alpha_T \to \infty$ as $T \to \infty$ and thus (4.3) implies $\mathbf{E}[\cos(\beta W_T)] \to 0$ as $T \to \infty$. In other words, this means that the resulting field theory without renormalisation would be free.

(ii) Even though the growth of α_s depends on β^2, we have the bound

$$\alpha_s \le C\langle s\rangle^{\frac{1}{2}-\delta}, \quad \text{for } \beta^2 < 4\pi.$$

We will use this estimate extensively in what follows.

(iii) The regularity of the limit depends on the size of the parameter β^2. For $\beta^2 < 4\pi$, convergence holds in $H^{-1+\delta}(\langle x\rangle^{-\ell})$ for any $\delta > 0$. Noting that H^1 corresponds to the Cameron-Martin space of the free field, this means that we can define the products $\langle h, [\![\cos(\beta W_\infty)]\!]\rangle$ for $h \in H_{\mathrm{CM}}(\mu)$. For $\beta^2 > 4\pi$, it is known that the sine-Gordon measure is singular with respect to the Gaussian free field even in the finite volume and requires additional renormalisation [43, 51].

4.2 A Variational Description

In the spirit of [4], the decomposition of the Gaussian free field in terms of a cylindrical Brownian motion enables a variational description for the Laplace transform of the Gibbs-measures of the form (1.1). For this section, let $\xi \in C_c^\infty(\mathbb{R}^2;[0,1])$ and $V_T^{\xi} \in C_b^2(L^2(\mathbb{R}^2))$.

The starting point for the variational approach is the following characterisation of functionals of Brownian motions due to Boué and Dupuis [13].

Lemma 4.14 (Boué-Dupuis 1999). *Let $W_\cdot = \int_0^\cdot C_t dB_t$ be a Brownian motion with covariance $\mathcal{C}_t = \int_0^t C_t^2 dt$ and define*

$$dX_t(u) = C_t u_t dt + dW_t.$$

If G is a real-valued, bounded and Borel-measurable functional, then with $\mathcal{A} = \mathbb{H}_T^2(L^2(H))$,

$$-\log \mathbf{E}[e^{-G(W)}] = \inf_{u\in\mathcal{A}} \mathbf{E}\left[G(X(u)) + \frac{1}{2}\int_0^T \|u_s\|_{L^2}^2 ds\right]. \tag{4.6}$$

In our specific setting, we are interested in the measures

$$\nu^T(d\varphi) = \Xi_T^{-1} \int \exp -V_T^{\xi,g}(\varphi)\mu^T(d\varphi),$$

and their Laplace transform

$$\int e^{-g(\varphi)}\nu^T(d\varphi) = \Xi_T^{-1}\,\mathbf{E}[\exp(-(V_T^{\xi,g} + g)(\varphi + W_T))],$$

with the partition function $\Xi_T = \mathbf{E}[\exp(-V_T^{\xi,g}(W_T))]$. Hence, we want to control the process

$$X_t(u) = \varphi + \int_0^t Q_s u_s dt + W_t,$$

subject to the cost function

$$J_T^{\xi,g}(u) = \mathbf{E}\left[V_T^{\xi,g}(X_T(u)) + \frac{1}{2}\int_0^T \|u_s\|_{L^2} ds \right],$$

where $u \in \mathcal{A} := \mathbb{H}_T^2(L^2(\Lambda))$ is the control and $V_T^{\xi,g} = V_T^{\xi,g} + g$ for a bounded and continuous perturbation g. Since the functional $V_T^{\xi,g}$ is bounded for $T < \infty$, we can directly apply Lemma 4.14 and arrive at a variational description

$$-\log \mathbf{E}[\exp(-V_T^{\xi,g}(\varphi + W_T))] = \inf_{u\in\mathcal{A}} J_T^{\xi,g}(u) =: \mathcal{V}_T^{\xi,g}. \tag{4.7}$$

As an immediate consequence, we also note the variational characterisation of the Laplace transform

$$\mathcal{W}^{\xi,T}(g) := \int e^{-g(\varphi)}\nu^T(d\varphi) = \mathcal{V}_T^{\xi,g} - \mathcal{V}_T^{\xi},$$

where we dropped the superscript g for $g = 0$. The above holds for any finite T, as V_T^{ξ} is bounded. In the limit $T \to \infty$, we lose this property. Fortunately, a convenient sufficient condition for the formula (4.6) by Hariya and Watanabe [35] and Üstünel [62] is still applicable.

Proposition 4.15. *If G satisfies*

$$\mathbf{E}[|pG(W)| + \exp(-qG(W))] < \infty, \tag{4.8}$$

for Hölder conjugates $\frac{1}{p} + \frac{1}{q} = 1$ or $p = 1 = q$, then the variational formula (4.6) holds.

Proposition 4.16. *The functional $V_\infty^{\xi}(W_\infty) = \int \xi\lambda[\![\cos\beta W_\infty]\!]$ satisfies the condition (4.8) and in particular the variational formula (4.7) also holds for $T = \infty$.*

Proof. Building on the proof of Lemma 4.11, one can show (see Theorem 1.3 in [38]), that the exponential moments of $\int \xi[\![\cos(\beta W_\infty)]\!]$ are finite, that is

$$\mathbf{E}\left[\exp\left\{q\int \xi[\![\cos(\beta W_\infty)]\!]\right\}\right] < \infty,$$

for any $q > 0$ and any smooth compactly supported ξ. The claim now follows from Proposition 4.15. □

Remark 4.17. For $u \in \mathcal{A}$, the expression $V_\infty^\xi(W_\infty + I_\infty(u))$ is well-defined for $\beta^2 < 4\pi$. Indeed, $[\![\cos(\beta W_\infty)]\!] \in H^{-1+\delta}(\rho_\ell)$ and by Lemma 4.5, $I_\infty(u) \in H^1(\rho_\ell)$. Now, by the angle-sum identities, we can write the potential as

$$\begin{aligned}
&V_\infty^\xi(W_\infty + I_\infty(u))\\
:=&\int \xi[\![\cos(W_\infty + I_\infty(u))]\!]\\
:=&\int \xi[\![\cos(\beta W_\infty)]\!]\cos(\beta I_\infty(u)) + [\![\sin(\beta W_\infty)]\!]\sin(\beta I_\infty(u)),
\end{aligned}$$

which is well-defined via the dual pairing of the Sobolev spaces H^{-1} and H^1. We defer the precise verification to Theorem 4.44.

In case the infimum is attained, the following lemma from [5] gives another useful implication of the representation.

Lemma 4.18. *Let $g : \mathscr{S}'(\mathbb{R}^2) \to \mathbb{R}$ be bounded and continuous and $T \in [0, \infty]$. If for $\alpha \in \mathbb{R}$ the variational problem $\inf_{u\in\mathcal{A}} J_T^{\alpha g}$ has a minimiser $u^{\alpha g}$, then the derivative of $\alpha \mapsto \mathcal{V}_T^{\alpha g}$ satisfies*

$$\frac{d}{d\alpha}\mathcal{V}_T^{\alpha g} = \mathbf{E}[g(X_T(u^{\alpha g}))].$$

Proof. By the Boué-Dupuis formula (4.7), $\mathcal{V}_T^{\alpha g} = -\log\int e^{(\alpha g + V_T)(\varphi)}\mu^T(d\varphi)$ and we see that the right-hand side is differentiable in α for bounded functions g. Then, by the optimality of $u^{\alpha g}$,

$$\inf_u J_T^{\alpha g}(u) - \inf_u J_T^{(\alpha-\gamma)g}(u) \geq J_T^{\alpha g}(u^{\alpha g}) - J_T^{(\alpha-\gamma)}(u^{\alpha g}) = \mathbf{E}[\gamma g(X(u^{\alpha g}))].$$

and thus

$$\frac{d}{d\alpha}\mathcal{V}_T^{\alpha g} = \lim_{\gamma\downarrow 0}\gamma^{-1}(\mathcal{V}_T^{\alpha g} - \mathcal{V}_T^{(\alpha-\gamma)g}) \geq \mathbf{E}[g(X_{T,T}(u^{\alpha g}))].$$

In the same way, we obtain the remaining inequality

$$\frac{d}{d\alpha}\mathcal{V}_T^{\alpha g} = \lim_{\gamma\downarrow 0}\gamma^{-1}(\mathcal{V}_T^{(\alpha+\gamma)g} - \mathcal{V}_T^{\alpha g}) \leq \mathbf{E}[g(X_{T,T}(u^{\alpha g}))].$$

□

Remark 4.19. In particular, for $\alpha = 0$ and a minimiser $u^{\xi,T}$ of J_T^{ξ},

$$
\begin{aligned}
\int g(\varphi)\nu^{\xi,T}(d\varphi) &= \frac{d}{d\alpha}\bigg|_{\alpha=0} \int \exp\{(\alpha g + V_T^{\xi})(\varphi)\}\mu^T(d\varphi) \\
&= \frac{d}{d\alpha}\bigg|_{\alpha=0} \mathcal{V}_T^{\alpha g} \\
&= \mathbf{E}[g(X_{T,T}(u^{\xi,T}))],
\end{aligned}
$$

and the law of the optimally controlled process X at the terminal time $t = T \in [0, \infty]$ is given by the approximate measure ν^T. In other words, if there is an optimal control $u^{\xi,T}$ for (4.7), the control system transports the (regularised) Gaussian free field μ^T to the (regularised) measure $\nu^{\xi,T} = \mathrm{Law}(X_T(u^{\xi,T}))$. Intuitively, the optimal control $u^{\xi,T}$ should be a function $\tilde{u}^{\xi,T}$ of the process X so that we expect to obtain an equation for the optimal dynamics

$$X_t = I_t(\tilde{u}^{\xi,T}(X)) + W_t, \quad \text{with } \mathrm{Law}(X_T) = \nu^{\xi,T}.$$

In this sense, the stochastic differential equation for X can be understood as a way to sample from ν^T using the well-understood Gaussian distribution μ^T, very much in the spirit of a theoretical Markov Chain Monte Carlo scheme. This point of view goes back to the idea of *stochastic quantisation* due to Parisi and Wu [58], where the authors advocate the use of stochastic (partial) differential equations to construct the target measure ν on the function space by introducing additional degrees of freedom, here in the form of a fictitious time $t \in \mathbb{R}_+$.

While the entropy $\int_0^T \|u_s\|_{L^2}^2 ds$ is convex, the renormalisation in general spoils convexity of the variational problem as $T \to \infty$ even for small correlations λ. Hence, the existence of a minimiser to (4.7) is not a priori clear. Proving that the control problem does indeed have an optimal control will be one of the goals in the next section.

4.3 An optimal Forward-Backward System

To obtain a more explicit expression from the variational characterisation introduced via the Boué-Dupuis formula, we need to ensure that the infimum is already a minimum. Thanks to the observations in the previous section, identifying the optimal control also provides an explicit characterisation of the approximate measure as the law of the optimally controlled process.

Towards this goal, we interpret the variational problem as a stochastic control problem as in the general setting in Chapter 3. This results in the following simple forward-backward system for a continuous and bounded functional V_T^{ξ},

$$
\begin{cases}
X_t(u) &= \varphi + \int_0^t Q_s u_s ds + \int_0^t Q_s dB_s, \\
Y_{t,T}(u) &= V_T^{\xi}(X_T(u)) + \int_t^T \frac{1}{2}\|u_s\|_{L^2}^2 ds - \int_t^T Z_{s,T}(u) dBs.
\end{cases}
\tag{4.9}
$$

4.3.1 Conventions for BSDEs.

For the reader's convenience, we briefly recall some of the conventions introduced previously in Chapter 2. Given a FBSDE in $\Theta^{t,x} = (X^{t,x}, Y^{t,x}, Z^{t,x})$ on a time interval $[t, T]$ with deterministic coefficients defined by

$$\begin{cases} X_s^{t,x} = x + \int_t^s b(r, \Theta_r^{t,x}) dr + \int_t^s \sigma(r, \Theta_r^{t,x}) dB_r, \\ Y_s^{t,x} = g(X_T^{t,x}) + \int_s^T f(r, \Theta_r^{t,x}) dr + \int_s^T Z_r^{t,x} dB_r, \end{cases}$$

we agree to use the conventions below.

- **Solutions:** Unless indicated otherwise, a *solution* the FBSDE always refers to the *square-integrable* solution in the sense of Definition 2.1.
- **Extension past** T**:** for $s \geq T$ we define $X_s^{t,x} = X_T^{t,x}$ and $(Y_s^{t,x}, Z_s^{t,x}) = (g(X_T^{t,x}), 0)$. This was justified in Proposition 2.15.
- **Extension past** t**:** for $s \leq t$, we define $X_s^{t,x} = x$ and $(Y_s^{t,x}, Z_s^{t,x}) = (Y_t^{t,x}, 0)$. The solution defined in this way is still adapted, as we have seen in Proposition 2.25.
- For $t = 0$, we usually omit the superscript (t, x), that is $X = X^{0,x}$.

Let us also recall Itô's formula for the most important special case of the weighted $L^2(\rho)$-norms from Lemma 2.3. For $dX_t = b_t dt + \sigma_t dt$, we have

$$d\|X_t\|_{L^2(\rho)}^2 = 2\langle X_t, b_t \rangle_{L^2(\rho)} dt + 2\langle X_t, \sigma dB_t \rangle_{L^2(\rho_\ell)} + \|\sigma_t\|_{\mathcal{L}_2(\rho)}^2 dt,$$

where $\|\sigma\|_{\mathcal{L}_2(\rho)}^2 = \mathrm{Tr}_{L_x^2}(\rho^2(x)\sigma\sigma^*)$.

4.3.2 Existence of an Optimal Control

In this section, we introduce three related but different approaches to obtain such a characterisation of the minimiser.

The first option is to understand the relation between the BSDE and the control problem more directly by relaxing the control problem. The relaxed control problem admits an optimal control from which we can construct an optimal control for the original, strong formulation. Alternatively, we can essentially rely on the Hamilton-Jacobi-Bellmann equation and verification theorem Theorem 3.2 to confirm that the usual candidate given by the gradient of the value function is optimal for the control problem. Finally, we also derive a system for the optimally controlled process using a stochastic maximum principle. This point of view is most closely related to the Euler-Lagrange equations and the approach taken in [5]. Unlike the other two approaches, this derivation requires an additional argument for the existence of a strong minimiser (usually in the form of convexity) for the control problem but is more straightforward and does not rely on the boundedness of the potential.

Relaxing the Variational Problem

Following the more general observations in Section 3.2, we derive a weak formulation for the stochastic control problem (4.7). In this relaxed setting, the existence of a minimiser is straightforward. By constructing a strong solution to the associated FBSDE, we can later recover the original, strong formulation of the control problem. We use the notation introduced in Section 3.2.

To relax the variational problem, the key relation between the drift and the diffusion of the controlled process is always satisfied for the control problem given by the Boué-Dupuis formula. With $\theta^u_t = \theta(t,x,u) = u_t$, we have

$$b(t,x,u) = \sigma(t,x)\theta^u_t.$$

The problem, however, is that for $u \in \mathbb{H}^2_T(L^2(\Lambda))$, the process $M^{\theta^u}_t$ is in general only a local martingale. Since this not sufficient to proceed with the Girsanov change of measure argument required for the weak formulation, we introduce the set of bounded admissible controls

$$\mathcal{A}_b := \left\{u \in \mathbb{H}^2_T(L^2(\Lambda)),\ \|u\|_{L^\infty(P\times[0,T]\times\mathbb{R}^2)} < \infty\right\},$$

and focus our attention on the restricted control problem

$$\mathcal{V}^w_{T,b} = \inf_{u\in\mathcal{A}_b} J^w_T(u) \text{ where } J^w_T(u) = \mathbf{E}^{P^u}\left[V^\xi_T(X_T) + \frac{1}{2}\int_0^T \|u_s\|^2_{L^2} ds\right],$$

with the forward process $X_t = \varphi + \int_0^t Q_s dB_s$, and hope to recover the full control problem later. For the bounded controls $u \in \mathcal{A}_b$, the stochastic exponential M^u_t is a martingale and we can proceed as in the general case. Following Section 3.2, this means we want to minimise the family of BSDEs

$$Y^u_t = V^\xi_T(X_T) + \int_t^T \frac{1}{2}\|u_s\|^2_{L^2} + \langle Z^u_s, u_s\rangle_{L^2} ds - \int_t^T Z^u_s dB_s. \tag{4.10}$$

With the Hamiltonians

$$H(s,x,u,z) = \frac{1}{2}\|u\|^2_{L^2} + \langle z,u\rangle_{L^2} \quad \text{and} \quad H^*(s,x,z) = \inf_{u\in L^2} H(s,x,u,z),$$

the optimally controlled BSDE is

$$Y^*_t = V^\xi_T(X_T) + \int_t^T H^*(s,X_s,Z_s)ds - \int_t^T Z^*_s dB_s. \tag{4.11}$$

Here, the infimum in H^* is attained for $u_s = -Z_s$ and in case $u_s \in \mathcal{A}_b$, the optimal BSDE (4.11) becomes

$$Y^*_t = V_T(X_T) - \int_t^T \frac{1}{2}\|Z^*_s\|^2_{L^2} ds - \int_t^T Z^*_s dB_s. \tag{4.12}$$

By Proposition 3.5, if (4.10) has a unique solution and (4.11) has a unique bounded solution, we can conclude

$$Y_0^* = \inf_{u \in \mathcal{A}_b} J^w(u) = J_T^w(-Z^*). \tag{4.13}$$

Equation (4.10) is a BSDE with standard parameters and thus has a unique solution by Theorem 2.2. That (4.12) also has a unique solution is the result of a simple change of variables.

Proposition 4.20. *The BSDE (4.12) has a unique solution $(Y^*, Z^*) \in \mathbb{H}_T^\infty(\mathbb{R}) \times \mathbb{H}_T^2(L^2(\Lambda))$. If the terminal condition satisfies $V_T \in C_b^2(L^2(\mathbb{R}^2))$ then also the martingale part is bounded - more precisely $\|Z^*\|_{L^\infty(P\times[0,T]\times\mathbb{R}^2)} < \infty$ almost surely.*

Proof. Suppose $(Y^*, Z^*) \in \mathbb{H}_T^\infty(\mathbb{R}) \times \mathbb{H}_T^2(L^2(\Lambda))$ satisfies (4.12) and define $y_s = \exp -Y_s^*$. By Itô's formula,

$$y_s = \exp(-V_T(W_T)) - \int_s^T z_r dB_r, \quad \text{where } z_r = y_r Z_r^*.$$

As V_T is bounded, the terminal condition is square-integrable and we see that the BSDE above has a unique square-integrable solution with

$$y_s = \mathbf{E}[\exp(-V_T(W_T))|\mathcal{F}_s] \geq \exp(-\|V_T\|_\infty) > 0. \tag{4.14}$$

But then we can apply Itô's formula to $-\log y_s$ and find that

$$Y_s = -\log y_s, \quad Z_s = \frac{z_s}{y_s},$$

satisfies the original quadratic BSDE with the required integrability. Noting that this argument holds for any solution $(Y^*, Z^*) \in \mathbb{H}_T^\infty(\mathbb{R}) \times \mathbb{H}_T^2(L^2(\Lambda))$, the uniqueness of the solution (y, z) implies uniqueness for (Y^*, Z_s^*). To see that Z^* is bounded if ∇V_T is, we introduce the deterministic function

$$v(t, \varphi) := Y_t^{t,\varphi}, \tag{4.15}$$

where as before the process $Y^{t,\varphi}$ is defined by the FBSDE

$$\begin{cases} X_s^{t,\varphi} = \varphi + \int_t^s dW_r \\ Y_s^{t,\varphi} = V_T(X_T^{t,\varphi}) - \int_s^T \frac{1}{2}\|Z_r^{t,\varphi}\|_{L^2}^2 dr - \int_s^T Z_r^{t,\varphi} dB_r. \end{cases}$$

These considerations imply that $v(t, \varphi) = -\log \mathbf{E}[-\exp(-V_T(X_T^{t,\varphi}))]$, which is differentiable in φ whenever V_T is. In this case,

$$\nabla v(t, \varphi) = -\frac{\nabla y_t}{y_t} = y_t^{-1}\, \mathbf{E}[-\nabla V_T(X_T^{t,\varphi})|\mathcal{F}_t].$$

By assumption, ∇V_T is bounded and we know from (4.14) that y_t is bounded away from 0. This implies thanks to the chain rule [42]

$$Z_s^{t,\varphi} = D_s Y_s^{t,\varphi} = \nabla v(s, X_s^{t,\varphi}) D_s X_s^{t,\varphi} = \nabla v(s, X_s^{t,\varphi}) Q_s,$$

and we conclude that Z^* is bounded uniformly. □

Now Proposition 4.20 also shows $Z^* \in \mathcal{A}_b$, and the relation (4.13) holds. To finish up this argument, we still need to show that the cost functional J is continuous. Recall the notation $\mathcal{A}_S$ and $\tilde{\mathcal{V}}$ from Definition 3.8.

Lemma 4.21. $\tilde{\mathcal{V}} = \inf_{u\in\mathcal{A}_S} J_T(u) = \inf_{u\in\mathcal{A}} J_T(u) = \mathcal{V}$.

Proof. We first show the claim for bounded controls, or more precisely that

$$\inf_{u\in\mathcal{A}_b} J_T(u) = \inf_{u\in\mathcal{A}_S} J_T(u).$$

By definition, $\inf_{u\in\mathcal{A}_b} J_T(u) \le \inf_{u\in\mathbb{A}_S} J_T(u)$, since any simple control is also bounded. If u is bounded, by Lemma 3.10, there is a sequence $u^{(n)}$ of uniformly bounded simple processes approximating u in $L^2([0,T];L^2(\mathbb{R}^2))$, that is

$$\lim_{n\to\infty} \mathbf{E} \int_0^T \|u_s^{(n)} - u_s\|_{L^2}^2 ds = 0.$$

By the estimates on the kernel Lemma 4.4 and Hölder's inequality

$$\begin{aligned}
\|\int_0^T Q_s(u_s^{(n)} - u_s)ds\|_{L^2}^2 \le & \left(\int_0^T \langle s\rangle^{-1} \|u_s^{(n)} - u_s\|_{L^2} ds\right)^2 \\
\le & \int_0^T \langle s\rangle^{-2} ds \int_0^T \|u_s^{(n)} - u_s\|_{L^2}^2 ds \\
\le & C \int_0^T \|u_s^{(n)} - u_s\|_{L^2}^2 ds.
\end{aligned}$$

But then, $J_T(u^{(n)}) \to J_T(u)$ as $n \to \infty$ for any bounded u, which gives the remaining inequality. Finally, approximation for $u \in \mathbb{H}_T^2(H)$ with bounded processes and dominated convergence allows us to remove the restriction to bounded controls. □

In summary, we have derived a lower bound for the value function

$$\mathcal{V}_T := \inf_{u\in\mathcal{A}} J_T(u) = \tilde{\mathcal{V}}_{T,b}^w \ge \mathcal{V}_{T,b}^w = Y_0^*.$$

A simple change of measure shows that this lower bound for $\mathcal{V}$ is already attained.

Proposition 4.22. *For any $T \in [0,\infty)$, the system*

$$\begin{cases} X_{s,T} = \varphi - \int_0^s Q_r^2 \nabla Y_{r,T} dr + \int_0^s Q_r dB_r, \\ Y_{s,T} = V_T^{\xi}(X_{T,T}) + \frac{1}{2}\int_s^T \|Q_r \nabla Y_{r,T}\|_{L^2}^2 dr - \int_t^T \nabla Y_{r,T} Q_r dB_r, \\ \nabla Y_{s,T} = \nabla V_T^{\xi}(X_{T,T}) - \int_s^T \nabla Z_{r,T} dB_r, \end{cases} \tag{4.16}$$

has a unique solution $(X, Y, \nabla Y, \nabla Z)$ and $u_r^T = -Q_r^2 \nabla Y_{r,T}$ is optimal for (4.9). In particular,

$$Y_{0,T} = J_T(u^T) = \mathcal{V}_T = \mathcal{V}_T^w = Y_0^*.$$

Proof. We have seen in Proposition 4.20 that $\nabla Y^{t,\varphi}$ exists. Differentiating the BSDE for $Y^{t,\varphi}$, we obtain (using that $\nabla X^{t,\varphi} = 1$ by Proposition 2.29),

$$\nabla Y_s = \nabla V_T^{\xi}(X_T) - \int_t^T Z_s \nabla Z_s ds + \int_t^T \nabla Z_s dB_s.$$

Since Z is bounded, the exponential

$$M_t^{-Z} = \exp\left(\int_0^t \langle Z_s, dB_s\rangle - \frac{1}{2}\int_0^t \|Z_s\|_{L^2}^2 ds\right),$$

is a true martingale and $\overline{B}_t = B_t + \int_0^t Z_s ds$ is a Brownian motion under $d\overline{P} = M_T^{-Z} dP$. Passing to the $\overline{P}$-Brownian motion $\overline{B}$, and using $Z_s = \nabla Y_s Q_s$, we arrive at the system

$$\begin{cases} X_s = \varphi - \int_0^s Q_r^2 \nabla Y_r dr + \int_0^s Q_r d\overline{B}_r, \\ Y_s = V_T(X_T) + \frac{1}{2}\int_s^T \|Q_r \nabla Y_r\|_{L^2}^2 dr - \int_t^T \nabla Y_r Q_r d\overline{B}_r, \\ \nabla Y_s = \nabla V_T(X_T) - \int_s^T \nabla Z_r d\overline{B}_r. \end{cases}$$

Initially, we only know that there is a $\mathcal{F}^B$-adapted solution. Since $\mathcal{F}^{\overline{B}}$ is a priori smaller than $\mathcal{F}^B$, we can not yet conclude that $u_s^* = -Q_s \nabla Y_s$ is optimal in the strong formulation as it might not be adapted to the driving Brownian motion. Recalling that we showed $\nabla Y_s = \nabla v(s, X_s)$ in Proposition 4.20, the control $u^* = -Q_s \nabla Y_s$ being optimal is equivalent to

$$X_t = \varphi - \int_0^t Q_s^2 \nabla v(s, X_s) ds + \int_0^t Q_s d\overline{B}_s, \tag{4.17}$$

admitting a $\mathcal{F}^{\overline{B}}$ adapted solution. But since

$$\nabla v(t,\varphi) = \frac{\mathbf{E}[\nabla V_T(\varphi + W_T)\exp(-V_T(\varphi + W_T))]}{\mathbf{E}[\exp(-V_T(\varphi + W_T))]},$$

is uniformly Lipschitz in φ, by a standard fixed point argument, there exists a unique strong solution X_t to the SDE (4.17). Thus,

$$Y_0 = J(u^*) = \mathcal{V} \geq \mathcal{V}^w = Y_0^*.$$

By the equivalence of the probability measures, the initial values $Y_0$and Y_0^* coincide P- and $\overline{P}$-almost surely and the claim follows. □

Remark 4.23. In the case of $V_T^{\xi}(\varphi) = \int \xi \lambda \alpha_T \cos(\beta\varphi)$, the trivial Lipschitz constant of ∇v depends on the support of ξ and is therefore not something we can rely on to derive existence and uniqueness when the cut-off is removed, i.e. for $\xi \equiv 1$.

Remark 4.24. In the weak formulation, the non-convex part $V_T^{\xi}(X_T)$ is fixed on the state space. We only control its distribution and the convex Hamiltonian. The crucial ingredients we need for this argument to undo the change of measure are good exponential moment bounds on $V_T(X_T)$ and then also on the control Z to return to the strong formulation.

Verification via the HJB-equation

Instead of understanding the equality of the value functions for the weak and the strong formulation directly, we can also rely on the verification theorem (Theorem 3.2).

In this case, the associated Hamilton-Jacobi-Bellmann equation for the stochastic control problem (4.9) is

$$\begin{cases} \partial_t v(t,\varphi) + \inf_{a\in L^2(\mathbb{R}^2)}\{\frac{1}{2}\operatorname{Tr}(Q_t^2 \operatorname{Hess} v(t,\varphi)) + \langle \nabla v, Q_t a\rangle + \frac{1}{2}\|a\|_{L^2}^2\} = 0, \\ v(T,\varphi) = V_T^\xi(\varphi). \end{cases}$$

If we look for L^2-solutions v, the infimum reduces to a quadratic minimisation problem and the unique minimum is attained at $a^* = -Q_t\nabla v(t,\varphi)$. Hence, the PDE above becomes

$$\begin{cases} \partial_t v(t,\varphi) + \frac{1}{2}\operatorname{Tr}(Q_t^2 \operatorname{Hess} v(t,\varphi)) - \frac{1}{2}\|Q_t\nabla v(t,\varphi)\|_{L^2}^2 = 0, \\ v(T,\varphi) = V_T^\xi(\varphi). \end{cases} \tag{4.18}$$

From the nonlinear Feynman-Kac representation (2.9), we see that this is the parabolic PDE associated with (4.12) and, indeed, we can quickly verify that the function defined by (4.15) is a classical solution to the PDE above.

Lemma 4.25. *The function v defined by* (4.15) *is the unique bounded solution to* (4.18).

Proof. We know from the proof of Proposition 4.20 that (4.15) admits the representation

$$v(t,\varphi) = -\log \mathbf{E}[\exp(-V_T^\xi(\varphi + W_T - W_t))],$$

which immediately implies differentiability in φ with bounded derivatives. To see that v satisfies the PDE, let $(W^{t,\varphi}, Y^{t,\varphi}, Z^{t,\varphi})$ be the solution to

$$\begin{cases} W_s^{t,\varphi} = \varphi + \int_t^s Q_r dB_r \\ Y_S^{t,\varphi} = V_T^\xi(W_T^{t,\varphi}) - \frac{1}{2}\int_s^T \|Z_r^{t,\varphi}\|_{L^2}^2 dr - \int_s^T Z_r^{t,\varphi} dB_r. \end{cases}$$

From Proposition 4.20 we know $Y_s^{t,\varphi} = v(s, W_s^{t,\varphi})$ and $Z_s^{t,\varphi} = \nabla v(s, W_s^{t,\varphi})Q_s$. So by the Markov property,

$$\begin{aligned} v(t+h,\varphi) - v(t,\varphi) &= v(t+h,\varphi) - v(t+h, W_{t+h}^{t,\varphi}) + v(t+h, W_{t+h}^{t,\varphi}) - v(t,\varphi) \\ &= v(t+h,\varphi) - v(t+h, W_{t+h}^{t,\varphi}) + Y_{t+h}^{t,\varphi} - Y_t^{t,\varphi}. \end{aligned}$$

Rewriting this with the help of Itô's formula,

$$\begin{aligned} v(t+h,\varphi) - v(t,\varphi) = &-\frac{1}{2}\int_t^{t+h} \operatorname{Tr}(Q_s^2 \operatorname{Hess} v(s, W_s^{t,\varphi}))ds - \int_t^{t+h} \nabla v(s, W_s^{t,\varphi})Q_s dB_s \\ &+ \frac{1}{2}\int_t^{t+h} \|Z_s^{t,\varphi}\|_{L^2}^2 ds + \int_t^{t+h} Z_s^{t,\varphi} dB_s. \end{aligned}$$

The stochastic integrals cancel, we see that v is differentiable in t and the claim follows by letting $h \to 0$. $\square$

Now that we have a classical solution to the Hamilton-Jacobi-Bellmann equation, we can use the verification theorem (or even the comparison theorem Theorem 2.13) to arrive at the system (4.16) and Proposition 4.22, relying on the Lipschitz continuity of ∇v in the same way as before.

Remark 4.26. For this argument, we rely on the HJB-equation to understand the fact that $\mathcal{V}^w$ and $\mathcal{V}$ define the same value function. In general, the BSDE is more likely to have a solution than the PDE (4.18) due to lower regularity requirements.

Remark 4.27. Since $Y_0^* = \mathcal{V}$ satisfies the HJB-equation, by Remark 3.3 the optimal control obtained in this way is the unique minimiser. As a result, also the measure $\nu_{\mathrm{SG}}^{T,\xi}$ is uniquely determined by Lemma 4.18.

Remark 4.28. The equation (4.18) is also known as the Polchinski equation in the physics literature. The nonlinear PDE can be understood as a continuous adaptation to Wilson's renormalisation group [63] and was first introduced by Polchinski in [60]. Since its introduction, it has played an important role in the development of constructive quantum field theory. For two related applications in the context of the sine-Gordon EQFT, we refer to [6, 17]. A review of the history and its influence on some more recent developments in quantum field theory can be found in [41].

A Stochastic Maximum Principle

An alternative approach to the stochastic control problem is via the stochastic maximum principle, which provides necessary conditions any optimal control must satisfy. This approach is more general and does not rely the boundedness of V_T directly, however, this comes at a cost. Comparable to the indirect method to variational problems, it shares the disadvantage of not providing sufficient conditions for optimality. Obtaining existence requires an additional argument, usually by relaxing the optimisation problem. The necessary condition is derived in the usual way: We perturb the control $u \in \mathbb{H}_T^2(L^2(\Lambda))$ by some variation $\varepsilon\delta u$, and differentiate with respect to ε. This procedure leads to a first-order condition for the optimal control and in our case, this uniquely determines the control.

Existence of an Optimal Control.

It is not a priori clear that an optimal control in the space $\mathcal{A} = \mathbb{H}^2(L^2(\mathbb{R}^2))$ can be found. By completing the space of optimal controls in a suitable (weak) topology, Barashkov and Gubinelli show in [5] that under suitable admissibility conditions on the cost functional J, most notably strong coercivity and lower-semicontinuity (see Definition 6 and Lemma 7 in [5]), optimal controls to the relaxed variational problem exist. As this is not our main focus, we refer to [5] for the details. At least when ξ is compactly supported, we can readily verify that the conditions of Lemma 7 in [5] are satisfied for the sine-Gordon model. Again, for the sake of brevity, we do not include a precise verification here.

Denoting the completion of $\mathcal{A}$ defined in this way by $\check{\mathcal{A}}$ and the corresponding cost functional by $\check{J}$, the value functions still coincide

$$\inf_{u\in\mathbb{A}} J(u) = \inf_{u\in\check{\mathcal{A}}} \check{J}(u) = \min_{u\in\check{\mathcal{A}}} \check{J}(u) =: \check{\mathcal{V}}.$$

However, the infimum might not be attained in $\mathcal{A}$ and we, therefore, have to look for more general minimisers in $\check{\mathcal{A}}$. The main point is that in contrast to $\mathcal{A}$, the completion $\check{\mathcal{A}}$ is not restricted to controls adapted to the Brownian motion B.

A Necessary Condition.

Since the filtration is now no longer generated by the Brownian motion itself, we are in the situation of the generalised BSDEs introduced in section 2.5. In other words, the system is given by

$$\begin{cases} X_t(u) & = \varphi + \int_0^t Q_s u_s ds + \int_0^t Q_s dB_s, \\ Y_{t,T}(u) & = V_T^{\xi}(X_T(u)) + \int_t^T \frac{1}{2}\|u_s\|_{L^2}^2 ds - \int_t^T Z_{s,T}(u) dBs + \int_t^T dM_s. \end{cases}$$

for an orthogonal martingale M. For a variation $\delta u_s \in \overline{\mathcal{A}}(u)$, it follows the stability of BSDEs, Proposition 2.18 that the solution $\Theta(u+\varepsilon\delta u)$ is differentiable in ε. Upon differentiating and evaluating at $\varepsilon = 0$ we obtain,

$$\begin{cases} \nabla_\varepsilon X_t^{u,\delta u} = \int_0^t Q_r \delta u_r dr \\ \nabla_\varepsilon Y_{t,T}^{u,\delta u} = \mathbf{E}\Big[\nabla V_T^{\xi}(X_T(u))\nabla_\varepsilon X_T^{u,\delta u} + \int_t^T u_s \delta u_s ds\Big]. \end{cases}$$

Thus, after inserting $\nabla_\varepsilon X$ into the backward equation,

$$\begin{aligned} \nabla_\varepsilon Y_{0,T}^{u,\delta u} &= \mathbf{E}\int_0^T (\nabla V_T^{\xi}(X_T(u))Q_s + u_s)\delta u_s ds \\ &= \mathbf{E}\int_0^T \Big(\mathbf{E}\Big[\nabla V_T^{\xi}(X_T(u))Q_s|\mathcal{F}_s\Big] + u_s\Big)\delta u_s ds. \end{aligned} \tag{4.19}$$

If the control u^* is optimal, the cost difference satisfies for any direction δu,

$$Y_{0,T}(u^* + \varepsilon\delta u) - Y_{0,T}(u^*) \geq 0, \quad P\text{-almost surely.}$$

First, this implies that the gradient $\nabla_\varepsilon Y_0^{u^*,\delta u}$ must be nonnegative. Moreover, since (4.19) holds for all directions δu and as

$$\nabla Y_s(u) := \mathbf{E}\Big[\nabla V_T^{\xi}(X_T(u))Q_s|\mathcal{F}_s\Big],$$

does not depend on the direction δu, by the fundamental lemma of calculus of variations, we arrive at the first order condition for optimality

$$u_s + Q_s\nabla Y_s(u) = 0 \iff u_s = u_s^* = -Q_s\nabla Y_s(u^*). \tag{4.20}$$

This condition is necessary and uniquely determines the optimal control u^*. Inserting the optimal control back into the system yields the optimally controlled GFBSDE

$$\begin{cases} X_{t,T} &= \varphi - \int_0^t Q_s^2 \nabla Y_{s,T} ds + \int_0^t Q_s dB_s, \\ Y_{t,T} &= V_T^{\xi}(X_{T,T}) + \int_t^T \frac{1}{2} \|Q_s \nabla Y_{s,T}\|_{L^2}^2 ds - \int_t^T Z_{s,T} dB_s - \int_t^T dM_s, \\ \nabla Y_{t,T} &= \nabla V_T^{\xi}(X_{T,T}) - \int_t^T \nabla Z_{s,T} dB_s - \int_t^T d(\nabla M_s), \end{cases} \tag{4.21}$$

where the martingales M and ∇M are orthogonal to the Brownian motion B.

Remark 4.29. (i) Following the same steps for a more general controlled FBSDE will yield the general stochastic maximum principle for convex control domains $\mathcal{A}$.

(ii) The argument still holds for $T = \infty$ and as a result shows that any optimal control u^* must satisfy (4.20). In other words, if the FBSDE has a unique solution, there is a unique optimal control u.

(iii) Any solution to (4.22) below also solves (4.21) with $M = 0$, thus we only consider the system with $M = 0$ going forward. This means in particular that the optimal control to relaxed variational problem has a version adapted to the Brownian motion.

4.4 Uniform Bounds on the Control

In this section, we want to return to the perturbed sine-Gordon interaction with a fixed smooth cut-off $\xi \in C_c^\infty(\mathbb{R}^2; [0,1])$ given by

$$V_T^{\xi,g}(\varphi) = g(\varphi) + \int \xi \alpha_T \cos(\beta\varphi),$$

for a bounded and continuous functional g and derive uniform bounds on the optimally controlled FBSDE in this setting. Recall that we derived the system for the optimally controlled FBSDE

$$\begin{cases} X_{t,T} &= \varphi - \int_0^t Q_s^2 \nabla Y_{s,T} ds + \int_0^t Q_s dB_s, \\ Y_{t,T} &= V_T^{\xi,g}(X_{T,T}) + \int_t^T \frac{1}{2} \|Q_s \nabla Y_{s,T}\|_{L^2}^2 ds - \int_t^T \nabla Y_{s,T} Q_s dB_s \\ \nabla Y_{t,T} &= \nabla V_T^{\xi,g}(X_{T,T}) - \int_t^T \nabla Z_{s,T} dB_s. \end{cases} \tag{4.22}$$

We have also seen that for a compactly supported cut-off and $T < \infty$, this system has a unique solution, which determines the unique optimal control and that

$$Y_{0,T} = \mathcal{V}_T^w = \mathcal{V}_T.$$

Let us make some additional remarks on the structure of (4.22). As a first simple but important observation we note that the dynamics are captured entirely by $(X, \nabla Y)$.

Provided we have good bounds on ∇Y, we can pass to the limits $\xi \to 1$ and $T \to \infty$. Moreover, an equivalent formulation of the system is given by

$$\begin{cases} X_{t,T} &= \varphi - \int_0^t \mathbf{E}[\nabla V_T^{\xi,g}(X_{T,T})|\mathcal{F}_s]ds + W_t, \\ Y_{t,T} &= V_T^{\xi,g}(X_{T,T}) + \int_t^T \frac{1}{2}\|Q_s \mathbf{E}[\nabla V_T^{\xi,g}(X_{T,T})|\mathcal{F}_s]\|_{L^2}^2 ds - \int_t^T Z_s dB_s, \end{cases}$$

where the coefficients of the forward equation now depend on the *distribution of* X. Solving this SDE in X via a fixed point iteration is equivalent to solving the forward-backward system (4.22).

Another key observation is the fact that the equations in $X, \nabla Y$ do not depend on the potential V_T^ξ but only on its gradient ∇V_T^ξ. Even though the V_T is ill-defined when the cut-off is removed, the gradient $\nabla V_T(\varphi) = -\alpha_T \beta\lambda \sin(\beta\varphi)$ and thus the processes $(X, \nabla Y)$ can also be studied on $\mathbb{R}^2$. We can of course still not expect ∇V_T to be in $L^2(\mathbb{R}^2)$, but passing to appropriate weighted spaces, we can control the weighted $L^2(\rho_\ell)$-norms of $(X, \nabla Y)$ uniformly in the cut-off ξ.

Finally, while the terminal condition $\nabla V_T^{\xi,g}(X_{T,T})$ is bounded for any $T < \infty$, this bound degenerates as $T \to \infty$. In the same way, the trivial Lipschitz constant for $\nabla v(t, \varphi)$ as defined in (4.15) blows up as $\xi \to 1$.

Conventions. We recall that the generic positive constant C, may depend on β and ℓ but is independent of T and λ. For $\ell < 0$, the constant C is also independent of ξ. Furthermore, we will consider small perturbations $g \in C_b^2(L^2(\rho_\ell)) \cap C_b^2(H^{-\delta}(\rho_\ell))$ of the terminal condition which satisfy a Lipschitz condition with

$$\mathrm{Lip}_\ell(\nabla g) := \sup_{\phi,\psi \in L^2(\rho_\ell)} \frac{\|\nabla g(\psi) - \nabla g(\phi)\|_{L^2(\rho_\ell)}}{\|\psi - \phi\|_{L^2(\rho_\ell)}} < \lambda.$$

Here, the space $C_b^2(L^2(\rho_\ell)) \cap C_b^2(H^{-\delta}(\rho_\ell))$ is the space of functions in $C_b^2(L^2(\rho_\ell))$ which extend continuously to $H^{-\delta}(\rho_\ell)$. Let us also fix the notation

$$|\nabla g|_\ell := \sup_{\varphi \in L^2(\rho_\ell)} \|\nabla g(\varphi)\|_{L^2(\rho_\ell)}.$$

4.4.1 Passing to the Remainder

By the construction of the potential V_t, the process $\{\nabla V_t(W_t), t \geq 0\}$ is a martingale. Seeing that the controlled process $X_t = \varphi - \int_0^t Q_r u_r dr + W_t$ is essentially a small perturbation of the Q-Brownian motion W_t given by a random shift, we consider the Ansatz

$$\mathbf{E}[\nabla V_T(X_{T,T})|\mathcal{F}_t] = \nabla V_t(X_{t,T}) + R_{t,T},$$

and want to control the remainder $R_{t,T}$ uniformly in T. These considerations lead to the system

$$\begin{cases} X_{t,T} = \varphi - \int_0^t Q_s^2(\nabla V_s(X_{s,T}) + R_{s,T})ds + W_t \\ R_{t,T} = \int_t^T d(\nabla V_T^\xi(X_{s,T})) - \int_t^T \nabla Z_{s,T} dB_s, \end{cases}$$

which we want to rewrite with Itô's formula (c.f. Lemma 2.3). Recall that

$$\nabla V_t(\varphi) = -\lambda\beta\alpha_t \sin(\beta\varphi) = -\lambda\beta\, \mathbf{E}[[\![\sin(\beta(\varphi + W_t))]\!]] = (\nabla V_t * \mu^t)(\varphi),$$

satisfies the Fokker-Planck equation (4.4) for C_t and, hence, the differential reduces to

$$d([\![\sin(\beta X_{t,T})]\!]) = \nabla[\![\sin(\beta X_{t,T})]\!] dX_{t,T}.$$

Allowing again a small perturbation g in the terminal condition, we obtain the system

$$\begin{cases} X_{t,T} = \varphi - \int_0^t Q_s^2(\nabla V_s^{\xi,g}(X_{s,T}) + R_{s,T})ds + W_t \\ R_{t,T} = \nabla g(X_{T,T}) + \int_t^T h^\xi(s, X_{t,T}, R_{t,T})ds - \int_t^T \tilde{Z}_{s,T} dB_s, \end{cases} \tag{4.23}$$

where $\tilde{Z}_{s,T} = \nabla Z_{s,T} - \beta^2\lambda\xi[\![\cos\beta X_{s,T}]\!]Q_s$ and

$$h^\xi(s,x,r) := \beta^2\lambda\xi[\![\cos(\beta x)]\!]_s Q_s^2(\nabla V_s^{\xi,g}(x) + r). \tag{4.24}$$

In contrast to the original BSDE, the terminal condition ∇g is now bounded uniformly in T, at the cost of including the additional drift h. In other words, the data g, h of the BSDE in $X, R, \tilde{Z}$ no longer depends on the terminal time. We want to leverage this to obtain uniform bounds on the solution.

4.4.2 Well-posedness

On a finite volume, we have already seen that there is a unique solution to the FBSDE using the (regular) decoupling field $v(t,\varphi) = -\log \mathbf{E}[\exp(-V_T^{\xi,g}(\varphi + W_t))]$. The argument however relied on trivial bounds depending on T and ξ and is not available for $\xi \equiv 1$. In this section, we show existence and uniqueness for the FBSDE in (X,R) on $\mathbb{R}^2$ with small correlations λ, without relying on $V_T^\xi(\varphi) = \int \xi\alpha_T\lambda\cos(\beta\varphi)$ being well-defined. For notational convenience, we surpress the dependence on the cut-off in $\Theta_T^\xi = \Theta_T = (X_{\cdot,T}, R_{\cdot,T}, \tilde{Z}_{\cdot,T})$. We always assume that either

- $\ell < -1$, which ensures $\rho_\ell^2(x)dx$ is a finite measure on $\mathbb{R}^2$, or
- ξ has compact support.

Note that this already includes the infinite volume case $\xi \equiv 1$ in the weighted L^2-space.

Theorem 4.30. *Fix any $T \in [0,\infty)$. If $\lambda > 0$ is sufficiently small, the FBSDE (4.23) has a unique solution*

$$\Theta_T = (X_{t,T}, R_{t,T}, \tilde{Z}_{t,T}) \in \mathbb{H}_T^\infty(H^{-\delta}(\rho_\ell)) \times \mathbb{H}_T^\infty(H^{-\delta}(\rho_\ell)) \times \mathbb{H}_T^2(\mathcal{L}_{2,\ell}(L^2(\mathbb{R}^2))).$$

The proof is based on a fixed point argument and relies on estimates of the spread of solutions to the FBSDE which are also useful beyond this proof. Towards this goal, given $r \in \mathbb{H}_T^2(L^2(\rho_\ell))$, we introduce the decoupled FBSDE

$$\begin{cases} X_{t,T}^r = \varphi - \int_0^t Q_s^2(\nabla V_s^{\xi,g}(X_{s,T}) + r_s)ds + W_t, \\ R_{t,T}^r = \nabla g(X_{T,T}^r) + \int_t^T h^\xi(s, X_{s,T}^r, r_s)ds + \int_t^T \tilde{Z}_s^r dB_s. \end{cases} \tag{4.25}$$

For the decoupled system, the existence of a solution follows from the general framework introduced in Chapter 2 and the regularity of the coefficients.

Lemma 4.31. *Let $b(s,x,r) = -Q_s^2(\nabla V_s^{\xi,g}(x) + r)$ be the drift of the forward equation and h as defined by (4.24). Then, b and h are Lipschitz in x and r as functions $L^2(\rho_\ell) \to L^2(\rho_\ell)$, uniformly on $[0,T]$. In particular, for any $r \in \mathbb{H}_T^\infty(L^2(\rho_\ell))$ and $T \in [0,\infty)$, the system (4.25) has a unique solution $(X^r, R^r, \tilde{Z}^r) \in \mathbb{H}_T^\infty(L^2(\rho_\ell)) \times \mathbb{H}_T^\infty(L^2(\rho_\ell)) \times \mathbb{H}_T^2(\mathcal{L}_{2,\ell}(L^2(\mathbb{R}^2)))$.*

Proof. First, we note that, by Lemma 4.4, Q_s is a bounded operator on $L^2(\rho_\ell)$. Moreover,

$$\|Q_s\alpha_s \sin(\beta\varphi^1) - Q_s\alpha_s\sin(\beta\varphi^2)\|_{L^2(\rho_\ell)} \leq C\langle s\rangle^{-\frac{1}{2}-\delta}\|\varphi^1 - \varphi^2\|_{L^2(\rho_\ell)}, \quad s \in [0,T]$$

and in the same way for the renormalised cosine. This implies that f and b are Lipschitz in x, and by the boundedness also in r.

To see that there is a unique solution, observe that for $r \in \mathbb{H}_T^\infty(L^2(\rho_\ell))$, the function $b(s,\cdot,r_s)$ is adapted, bounded and uniformly Lipschitz. By a standard fixed point argument, the forward equation has a unique solution $X_{t,T}^r$. In the same way, the generator $h(s, X_{s,T}^r, r_s)$ is adapted, bounded in $L^2(\rho_\ell)$ and trivially Lipschitz in R with constant 0. Thus, $h(s, X_{s,T}^r, 0) \in \mathbb{H}_T^2(L^2(\rho_\ell))$ and $(0,f)$ are standard parameters. By Theorem 2.2, there is a unique solution $(R_{t,T}^r, \tilde{Z}_{t,T}^r)$ to (4.25). The bounds in $\mathbb{H}_T^\infty(L^2(\rho_\ell))$ follow from the integrability of b and h and Proposition 2.9. □

Proposition 4.32. *Let Θ^{r_i} be the solution to (4.25) with $r_i \in \mathbb{H}_T^\infty(L^2(\rho_\ell))$ for $i = 1,2$, and denote their spread by $\delta\Theta = (\delta X, \delta R, \delta Z)$. Then, there is a constant $C > 0$, not depending on $T \in [0,\infty)$, such that for any $t \in [0,T]$,*

$$\|\delta X_t\|_{L^2(\rho_\ell)} \leq Ce^{-\frac{2m}{t}} \sup_{s\in[0,t]} \|\delta r_s\|_{L^2(\rho_\ell)},$$

and

$$\mathbf{E}\left[\sup_{t\in[0,T)} \|\delta R_t\|_{L^2(\rho_\ell)}^2 + \int_0^T \|\delta Z_s\|_{\mathcal{L}_{2,\ell}}^2 ds\right] \leq (\lambda^2 + \mathrm{Lip}_\ell(\nabla g))Ce^{-\frac{4m}{T}}\,\mathbf{E}\left[\sup_{t\in[0,T]} \|\delta r_r\|_{L^2(\rho_\ell)}^2\right].$$

Before we proceed with these estimates, let us first derive Theorem 4.30.

Proof of Theorem 4.30. Existence for $T \in [0,\infty)$. For $\mathrm{Lip}_\ell(\nabla g) < \lambda$ the constant for any $T \in [0,\infty)$ on the right-handside of Proposition 4.32 reduces to $C\lambda^2$. Thus, if λ is sufficiently small, the map $\Phi : r \mapsto R^r$ is a contraction on $\mathbb{H}_T^\infty(L^2(\rho_\ell))$ and there is a unique fixed point $\overline{R}$. Denoting by (X,R) the continuous versions of the solutions to (4.25) $(\overline{X}^{\overline{R}}, \overline{R}^{\overline{R}})$, which by Lemma 4.31 exist for $T < \infty$, and letting $Z = Z^R$, we see that (X,R,Z) is a solution to (4.23).

Uniqueness for $T \in [0,\infty]$. For any solution to (4.23) on $[0,\infty)$, the backward component R is a fixed point of Φ. Since Φ has a unique fixed point, uniqueness must also hold for the solution (X,R,Z) to (4.23). □

Remark 4.33. By passing to $H^{-\delta}(\rho_\ell)$ instead of $L^2(\rho_\ell)$, we could show existence and uniqueness for (4.25) for $T = \infty$. Combined with the estimates in Proposition 4.32, this is sufficient to show existence on $[0, \infty]$ for λ small. We prefer to stay in $L^2(\rho_\ell)$ as this later shows that $R_{t,\infty} \in L^2(\rho_\ell)$ for any $t < \infty$ which in turn will yield converngence of the drift in $L^2(P, H^1(\rho_\ell))$.

Remark 4.34. Comparing these estimates to Proposition 2.9 and Lemma 2.33, we now see why the local solution for a coupled FBSDE may not be extended to any time interval in the same way as for decoupled systems: While we still obtain a contraction mapping for small time intervals (for any λ), the size of the time interval can degenerate in the coupled case because the Lipschitz constant of the terminal condition now affects the stepsize.

Proof of Proposition 4.32. The difference satisfies the BSDE

$$\begin{cases} \delta X_t = -\int_0^t Q_s^2(\nabla V_T^{\xi,g}(X_{s,T}^{r_1}) - \nabla V_T^{\xi,g}(X_{s,T}^{r_2}))ds + \int_0^t Q_s^2 \delta r_s ds, \\ \delta R_t = \delta \nabla g + \int_t^T (h(s, X_{s,T}^{r_1}, r_1) - h(s, X_{s,T}^{r_2}, r_2))ds - \int_t^T \delta \tilde{Z}_s dB_s, \end{cases}$$

where $\delta \nabla g = \nabla g(X_{T,T}^{r_1}) - \nabla g(X_{T,T}^{r_2})$. For the estimate on δX_t, we have by the Lipschitz continuity of the coefficients and the estimates on the kernel $\|Q_t f\|_{L^2(\rho_\ell)} \lesssim \langle t \rangle^{-1} \|f\|_{L^2(\rho_\ell)}$,

$$\begin{aligned} &\|\delta X_t\|_{L^2(\rho_\ell)} \\ \leq& \lambda C \int_0^t \langle s \rangle^{-2} \|\nabla V_s^{\xi,g}(X_{s,T}^{r_1}) - \nabla V_s^{\xi,g}(X_{s,T}^{r_2})\|_{L^2(\rho_\ell)} ds + \| \int_0^t Q_s^2 \delta r_s \|_{L^2(\rho_\ell) ds} \\ \leq& \lambda C \int_0^t \langle s \rangle^{-2} \alpha_s \|\delta X_s\|_{L^2(\rho_\ell)} ds + \| \int_0^t Q_s^2 \delta r_s ds \|_{L^2(\rho_\ell)}. \end{aligned}$$

Using that $\alpha_s \leq C \langle s \rangle^{\frac{\beta^2}{8\pi}}$, we see that $\langle s \rangle^{-2} \alpha_s$ is integrable on $\mathbb{R}_+$ for any $\beta^2 < 8\pi$. Thus, by Gronwalls lemma and Lemma 4.4,

$$\begin{aligned} \|\delta X_t\|_{L^2(\rho_\ell)} \leq& C \int_0^t \|Q_s^2 \delta r_s\|_{L^2(\rho_\ell)} ds \\ \leq& \int_0^t s^{-2} e^{-\frac{2m}{s}} \|\delta r_s\|_{L^2(\rho_\ell)} ds \\ \leq& \frac{C}{2m} e^{-\frac{2m}{t}} \sup_{s \leq t} \|\delta r_s\|_{L^2(\rho_\ell)}. \end{aligned}$$

To show the estimates on the remainder R, we proceed in the usual way by applying Itô's formula to obtain

$$\begin{aligned} &\|\delta R_t\|_{L^2(\rho_\ell)}^2 + \int_t^T \|\delta Z_s\|_{\mathcal{L}_{2,\ell}}^2 ds \\ =& \|\delta \nabla g\|_{L^2(\rho_\ell)}^2 + \int_t^T 2 \langle \delta R_s, h_s^1(r_1) - h_s^2(r_2) \rangle_{L^2(\rho_\ell)} ds - \int_t^T 2 \langle \delta R_s, \delta Z_s dB_s \rangle_{L^2(\rho_\ell)}, \end{aligned} \tag{4.26}$$

where $h_s^i(r) = h(s, X_{s,T}^{r_i}, r_s)$. By Lemma 2.7, the stochastic integral is a martingale and vanishes upon taking expectation. Hence,

$$\mathbf{E}\int_0^T \|Z_s\|_{\mathcal{L}_{2,\ell}}^2 ds \le \mathbf{E}\|\delta\nabla g\|_{L^2(\rho_\ell)}^2 + \mathbf{E}\left[\sup_{t\in[0,T]}\int_t^T 2\langle \delta R_s, h_s^1(r_1) - h_s^2(r_2)\rangle_{L^2(\rho_\ell)} ds\right].$$

The same argument as in the proof of Theorem 2.2, using the Burkholder-Davis-Gundy inequality, shows that

$$\mathbf{E}\left[\sup_{t\in[0,T]}\int_t^T 2\langle \delta R_s, \delta Z_s dB_s\rangle_{L^2(\rho_\ell)}\right] \le \frac{1}{4}\mathbf{E}\left[\sup_{t\in[0,T]}\|\delta R_t\|_{L^2(\rho_\ell)}^2\right] + 4c\,\mathbf{E}\int_0^T \|\delta Z_s\|_{\mathcal{L}_{2,\ell}}^2 ds.$$

Thus, since all terms are finite by Lemma 2.7, taking the supremum and expectation in (4.26) yields after rearranging,

$$\begin{aligned}
&\mathbf{E}\left[\sup_{t\in[0,T]}\|\delta R_t\|_{L^2(\rho_\ell)}^2 + \int_0^T \|\delta Z_s\|_{\mathcal{L}_{2,\ell}}^2 ds\right] \\
\le& C\,\mathbf{E}\|\delta\nabla g\|_{L^2(\rho_\ell)}^2 + C\,\mathbf{E}\left[\sup_{t\in[0,T]}\int_t^T 2\langle \delta R_s, h_s^1(r_1) - h_s^2(r_2)\rangle_{L^2(\rho_\ell)} ds\right].
\end{aligned} \tag{4.27}$$

The terminal conditions can be estimated from the Lipschitz condition on ∇g and the estimate on δX just derived to obtain,

$$\|\delta\nabla g\|_{L^2(\rho_\ell)} \le L\|\delta X_T\|_{L^2(\rho_\ell)} \le CLe^{\frac{-2m}{T}} \sup_{s\in[0,T]}\|\delta r_s\|_{L^2(\rho_\ell)}.$$

Moreover, for the integral on the right-hand side, note

$$\begin{aligned}
&2\langle \delta R_s, h_s^1(r_1) - h_s^2(r_2)\rangle_{L^2(\rho_\ell)} \\
=&2\langle \delta R_s, h_s^1(r_1) - h_s^1(r_2)\rangle_{L^2(\rho_\ell)} + 2\langle \delta R_s, h_s^1(r_2) - h_s^2(r_2)\rangle_{L^2(\rho_\ell)} \\
\le&\varepsilon_s\|\delta R_s\|_{L^2(\rho_\ell)}^2 + 2\varepsilon_s^{-1}\|\delta h_s^1\|_{L^2(\rho_\ell)}^2 + 2\varepsilon_s^{-1}\|\delta_2 h_s\|_{L^2(\rho_\ell)}^2,
\end{aligned}$$

where $\delta h_s^1 = h_s^1(r_1) - h_s^1(r_2)$ and $\delta_2 h_s = h_s^1(r_2) - h_s^2(r_2)$. Let $\varepsilon_s = K\langle s\rangle^{-1-\delta}$ for some constant K to be chosen later. Then, using the bounds on $[\![\cos(\beta\delta\varphi)]\!]_s \lesssim \langle s\rangle^{\frac{1}{2}-\delta}$ and the estimates from Lemma 4.4

$$\begin{aligned}
\mathbf{E}\int_0^T \varepsilon_s^{-1}\|\delta h_s^1\|_{L^2(\rho_\ell)}^2 ds &= \mathbf{E}\,\beta^4\lambda^2\int_0^T \varepsilon_s^{-1}\|Q_s^2[\![\cos\beta X_s^{r_1}]\!]\delta r_s\|_{L^2(\rho_\ell)}^2 ds \\
&\le C\lambda^2K^{-1}\,\mathbf{E}\left[\sup_{t\in[0,T]}\|\delta r_t\|_{L^2(\rho_\ell)}^2\right]\int_0^\infty s^{1+\delta}s^{-4}e^{-\frac{4m}{s}}s^{1-2\delta} ds \\
&\le C\lambda^2K^{-1}e^{-\frac{4m}{T}}\,\mathbf{E}\left[\sup_{t\in[0,T]}\|\delta r_t\|_{L^2(\rho_\ell)}^2\right].
\end{aligned}$$

Likewise, due to the Lipschitz continuity of $\nabla V^{\xi,g}$, combined with the estimate on $\|\delta X_s\|_{L^2(\rho_\ell)}$,

$$
\begin{aligned}
&\mathbf{E}\int_0^T \varepsilon_s^{-1}\|\delta_2 h_s\|^2_{L^2(\rho_\ell)}ds \\
\leq& C\lambda^4\,\mathbf{E}\int_0^T s^{1+\delta}s^{-4}e^{-\frac{4m}{s}}\alpha_s^2\|\nabla V_s^{\xi,g}(X_s^{r_1})-\nabla V_s^{\xi,g}(X_s^{r_2})\|^2_{L^2(\rho_\ell)}ds \\
\leq& C\lambda^4K^{-1}\,\mathbf{E}\int_0^T s^{1+\delta}s^{-4}e^{-\frac{4m}{s}}\alpha_s^4\|\delta X_s\|^2_{L^2(\rho_\ell)}ds \\
\leq& C\lambda^4K^{-1}\,\mathbf{E}\left[\sup_{t\in[0,T]}\|\delta r_t\|^2_{L^2(\rho_\ell)}\right]\int_0^T s^{-4}s^{2-4\delta}s^{1+\delta}e^{-\frac{4m}{s}}ds \\
\leq& C\lambda^4K^{-1}e^{-\frac{4m}{T}}\,\mathbf{E}\left[\sup_{t\in[0,T]}\|\delta r_t\|^2_{L^2(\rho_\ell)}\right].
\end{aligned}
$$

Thanks to the integrability of ε_s on $[0,\infty)$, the remaining term can be estimated as

$$
\mathbf{E}\int_0^T \varepsilon_s\|\delta R_s\|^2_{L^2(\rho_\ell)}ds \leq K\int_0^\infty \langle s\rangle^{-1-\delta}ds\,\mathbf{E}\left[\sup_{t\in[0,T]}\|\delta R_t\|^2_{L^2(\rho_\ell)}\right] \leq KC\,\mathbf{E}\left[\sup_{t\in[0,T]}\|\delta R_t\|^2_{L^2(\rho_\ell)}\right].
$$

Combined, using the estimates in (4.27), we obtain

$$
\begin{aligned}
&\mathbf{E}\left[\sup_{t\in[0,T]}\|\delta R_t\|^2_{L^2(\rho_\ell)}+\int_0^T\|\delta Z_s\|^2_{\mathcal{L}_{2,\ell}}ds\right] \\
\leq&(CL^2+C\lambda^2K^{-1})e^{-\frac{4m}{T}}\,\mathbf{E}\left[\sup_{t\in[0,T]}\|\delta r_r\|^2_{L^2(\rho_\ell)}\right]+CK\,\mathbf{E}\left[\sup_{t\in[0,T]}\|\delta R_t\|^2_{L^2(\rho_\ell)}\right].
\end{aligned}
$$

Choosing $K=\frac{1}{2C}$, we obtain the claim after rearranging. □

4.4.3 A priori Estimates

The following representation for the remainder will be helpful.

Proposition 4.35. *The remainder admits the representation*

$$
R_{t,T}=\mathbf{E}\left[\Gamma_T^t\nabla g(X_{T,T})+\int_t^T\Gamma_s^t\phi_s ds\Big|\mathcal{F}_t\right],
$$

where $\Gamma_s^t=\exp\left\{\int_t^s\beta^2\lambda[\![\cos\beta X_{u,T}]\!]Q_u^2du\right\}$ *is the solution to the adjoint forward equation and* $\phi_s=\beta^2\lambda[\![\cos\beta X_{s,T}]\!]Q_s^2V_s^{\xi,g}(X_{s,T})$ *is the inhomogeneous part of the backward equation.*

Proof. The same argument as in Proposition 2.12 shows that the representation holds, if Γ_s^t is a bounded operator $L^2(\rho_\ell)\to L^2(\rho_\ell)$ and $\phi\in\mathbb{H}_T^2(L^2(\rho_\ell))$. Since $X_{s,T}$ is adapted, so is φ. Moreover, it holds

$$
\|\phi_s\|_{L^2(\rho_\ell)}\leq C\lambda^2\langle s\rangle^{\frac{\beta^2}{4\pi}-2}.
$$

Hence, $\phi \in \mathbb{H}^2_T(L^2(\rho_\ell))$ for $\beta^2 < 4\pi$. For $\gamma_r := \beta^2\lambda[\![\cos\beta X_{r,T}]\!]Q_r^2$, we compute

$$\|\gamma_r\|_{L^2(\rho_\ell)\to L^2(\rho_\ell)} \le C\|[\![\cos\beta X_{r,T}]\!]Q_r^2\|_\infty \le C\alpha_r\langle r\rangle^{-2} \le C\langle r\rangle^{\frac{\beta^2}{8\pi}-2},$$

for any $r \in [0,\infty)$. Since the right hand-side is integrable in time, $\int_s^t \gamma_r dr$ is a bounded linear operator and thus the same is true for its exponential Γ_s^t. □

Proposition 4.36. *There is a constant C, such that*

$$\|R_{t,T}\|_{L^\infty} \le |\nabla g|_\ell + C\lambda^2\langle t\rangle^{-2\delta},$$

and

$$\|R_{t,T}\|_{L^2(\rho_\ell)} \le |\nabla g|_\ell + C\lambda^2\langle t\rangle^{-2\delta}.$$

Moreover, for any $t \in [0,T]$

$$\mathbf{E}\left[\|R_{t,T}\|^2_{L^2(\rho_\ell)} + \int_t^T \|\tilde{Z}_{s,T}\|^2_{\mathcal{L}_2(\rho_\ell)} ds\right] \le |\nabla g|_\ell + C\lambda^4\langle t\rangle^{-4\delta},$$

and in particular, the remainder R is bounded uniformly in T and ξ.

Proof. In the notation from Proposition 4.36, the first two estimates are direct consequences of the aforementioned proposition, the boundedness of Γ and the integrability of $\int_t^T \|\Gamma_s^t\varphi_s\|^2_{L^\infty} ds \le C\langle t\rangle^{-2\delta}$. To transfer the bound to the martingale part, we apply Itô's formula

$$\begin{aligned}
&\mathbf{E}[\|R_{t,T}\|^2_{L^2(\rho_\ell)}] + \mathbf{E}\int_t^T \|Z_{s,T}\|^2_{\mathcal{L}_{2,\ell}} ds \\
=&\|\nabla g(X_T)\|^2_{L^2(\rho_\ell)} + \mathbf{E}\int_t^T 2\langle R_{s,T}, h(s, X_{s,T}, R_{s,T})\rangle_{L^2(\rho_\ell)} ds \\
\le&\|\nabla g(X_T)\|^2_{L^2(\rho_\ell)} + \mathbf{E}\int_t^T \varepsilon_s\|R_{s,T}\|^2_{L^2(\rho_\ell)} + \varepsilon_s^{-1}\|h(s, X_{s,T}, R_{s,T})\|^2_{L^2(\rho_\ell)} ds.
\end{aligned}$$

Then, suppressing the arguments in h, using the same estimates we already used

$$\|h_s\|^2_{L^2(\rho_\ell)} \le C\lambda^4\langle s\rangle^{-4}\alpha_s^4 + C\lambda^2\langle s\rangle^{-4}\alpha_s^2\|R_{s,T}\|^2_{L^2(\rho_\ell)}.$$

Inserting the bound we just derived and letting $\varepsilon_s = \langle s\rangle^{-1}$, the claim follows. □

4.4.4 Dependence on the Ultraviolet cut-off T

With the uniform bound derived in the previous section, the standard a priori estimates allow us to study the dependence on the cut-off T.

Proposition 4.37. *Let $T_1 \le T_2$ and denote the spread between the solutions to (4.23) with terminal times T_1, T_2 by $\delta_T \Theta = \Theta^{T_1} - \Theta^{T_2}$. For $\lambda > 0$ sufficiently small, we have*

$$\mathbf{E}\left[\sup_{t\ge 0}\|\delta_T R_t\|^2_{L^2(\rho_\ell)} + \int_0^{T_2}\|\delta_T Z_s\|^2_{\mathcal{L}_{2,\ell}}ds\right] \le C\lambda^4\langle T_1\rangle^{-4\delta}.$$

Proof. Since $R_{T_1,T_1} = \nabla g(X_{T_1,T_1})$, the difference statisfies the BSDE

$$\delta_T R_t = \nabla g(X_{T_1,T_1}) - \nabla g(X_{T_2,T_2}) + \int_t^{T_2}(h_s^2 - h_s^1)ds - \int_t^{T_2}\delta_T Z_s dB_s, \quad t\in[0,T_2],$$

where

$$h_s^i = \mathbb{1}_{\{s\le T_i\}}\beta^2\lambda[\![\cos(\beta X_{s,T_i})]\!]Q_s^2(\nabla V_s^{\xi,g}(X_{s,T_i}) + R_{s,T_i}),$$

denotes the generators for the BSDEs on $[0,T_i]$ for $i=1,2$. Applying the usual procedure, using the bound from Proposition 4.36, we have

$$\begin{aligned}
&\mathbf{E}\left[\sup_s\|\delta_T R_s\|^2_{L^2(\rho_\ell)} + \int_0^{T_2}\|\delta_T Z_s\|^2_{L^2(\rho_\ell)}ds\right] \qquad (4.28)\\
&\le \mathbf{E}[\|\nabla g(X_{T_1,T_1}) - \nabla g(X_{T_2,T_2})\|^2_{L^2(\rho_\ell)}] + \mathbf{E}\int_0^{T_1}2\langle\delta_T R_s, h_s^2 - h_s^1\rangle_{L^2(\rho_\ell)}ds\\
&\quad + \mathbf{E}\int_{T_1}^{T_2}2\langle\delta_T R_s, h_s^2\rangle_{L^2(\rho_\ell)}ds.
\end{aligned}$$

The last integral over $[T_1,T_2]$ can be bounded in the same way as before in the proof of Proposition 4.36,

$$\mathbf{E}\int_{T_1}^{T_2}2\langle\delta_T R_s, h_s^2\rangle_{L^2(\rho_\ell)}ds \le C\lambda^4\langle T_1\rangle^{-4\delta}.$$

For the integral over $[0,T_1]$, we again estimate the difference between the generators h_s^i in terms of $\|\delta_T R_s\|^2_{L^2(\rho_\ell)}$. To this end note that, for $s\le T_1$,

$$\begin{aligned}
&\|h_s^2 - h_s^1\|_{L^2(\rho_\ell)}\\
\le&\beta^6\lambda^4\|[\![\cos(\beta X_{s,T_1})]\!]Q_s^2\nabla V_s^{\xi,g}(X_{s,T_1})\delta_T R_s\|_{L^2(\rho_\ell)}\\
&+\beta^6\lambda^4\|Q_s^2(\nabla V_s^{\xi,g}(X_{s,T_2}) - \nabla V_s^{\xi,g}(X_{s,T_1}))\|_{L^2(\rho_\ell)}\\
&+\beta^4\lambda^2\|([\![\cos(\beta X_{s,T_2})]\!] - [\![\cos(\beta X_{s,T_1})]\!])Q_s^2 R_{s,T_1}\|_{L^2(\rho_\ell)}\\
=:&\text{(I)} + \text{(II)} + \text{(III)}.
\end{aligned}$$

Using the boundedness of $\|V_s^{\xi,g}(\varphi)\|_{L^2(\rho_\ell)} \le C\langle s\rangle^{\frac{1}{2}-\delta}$, combined with the bound on $\delta_T X = X^{R_\cdot,T_2} - X^{R_\cdot,T_1}$ from Proposition 4.32, the first two terms can be estimated in the same way as before,

$$\mathbf{E}\,(\text{I})^2 \le \lambda^4 C\langle s\rangle^{-2-4\delta}\,\mathbf{E}\|\delta_T R_s\|^2_{L^2(\rho_\ell)},$$

and

$$\mathbf{E}\,(\mathrm{II})^2 \leq \lambda^4 C \langle s\rangle^{-2-4\delta}\,\mathbf{E}\left[\sup_{\ell}\|\delta_T R_\ell\|^2_{L^2(\rho_\ell)}\right].$$

For the third term, we also use the uniform bounds from Proposition 4.36 and argue in the same way that

$$\begin{aligned}\mathbf{E}\,(\mathrm{III})^2 &\leq \mathbf{E}\Big[\lambda^2 C\|Q_s^2([\![\cos\beta X_{s,T_2}]\!]-[\![\cos\beta X_{s,T_1}]\!])\|^2_{L^2(\rho_\ell)}\|R_{s,T_1}\|^2_{L^\infty}\Big]\\ &\leq \lambda^2 C\langle s\rangle^{-3}\langle s\rangle^{-2\delta}\,\mathbf{E}\|[\![\cos\beta X_{s,T_2}]\!]-[\![\cos\beta X_{s,T_1}]\!]\|^2_{L^2(\rho_\ell)}\\ &\leq \lambda^2 C\langle s\rangle^{-4-2\delta}\langle s\rangle^{-2\delta}\,\mathbf{E}\|X_{s,T_1}-X_{s,T_2}\|^2_{L^2(\rho_\ell)}\\ &\leq \lambda^2 C\langle s\rangle^{-4-2\delta}\langle s\rangle^{-2\delta}\,\mathbf{E}\left[\sup_{\ell}\|\delta_T R_\ell\|^2_{L^2(\rho_\ell)}\right].\end{aligned}$$

Using these bounds and the Lipschitz continuity of ∇g and X^r in (4.28), we have

$$\begin{aligned}&\mathbf{E}\left[\sup_t\|\delta_T R_t\|^2_{L^2(\rho_\ell)}+\int_0^{T_2}\|\delta_T Z_s\|^2_{\mathcal{L}_{2,\ell}}ds\right]\\ \leq&\lambda^4 C\langle T_1\rangle^{-4\delta}+\left(C\lambda^4\int_0^\infty c(s)ds+\lambda^2 C\right)\mathbf{E}\left[\sup_t\|\delta_T R_t\|^2_{L^2(\rho_\ell)}\right],\end{aligned}$$

for some positive $c(s)$ with $\int_0^\infty c(s)ds<\infty$. Choosing $\lambda>0$ small enough we obtain the desired bound after rearranging. □

4.4.5 Dependence on the Infrared cut-off ξ

While the bounds in the previous section are sufficient to show tightness and existence of a solution on the infinite volume $\mathbb{R}^2$, we can also show convergence $\nu_{\mathrm{SG}}^{T,\xi}\to\nu_{\mathrm{SG}}$ for the entire sequence in the infinite volume limit to a *unique* limit, at least if $\lambda>0$ is small. This will allow us to reconnect the FBSDE on $\mathbb{R}^2$ to the variational problem and, consequently, the approximate measures ν_{SG}^{ξ} to ν_{SG}. Recall again the definition

$$V_T^\xi(\varphi)=\lambda\int\xi\alpha_T\cos(\beta\varphi),$$

for a smooth cut-off ξ. To study the dependence on the cut-off, we introduce a system for the spread $\delta_\xi\Theta=\Theta_T^{\xi_2}-\Theta_T^{\xi_1}$ between two solutions with terminal conditions $V_T^{\xi_1,g}$ and $V_T^{\xi_2,g}$ sastisfying the equation,

$$\begin{cases}\delta_\xi X_t &=-\int_0^t(Q_s^2(\xi_2\nabla V_s^{\xi_2,g}(X_s^{\xi_2})-\xi_1\nabla V_s^{\xi_1,g}(X_s^{\xi_1}))+Q_s^2\delta_\xi R_s)ds\\ \delta_\xi R_t &=\nabla g(X_T^{\xi_1})-\nabla g(X_T^{\xi_2})+\int_t^T h_s^{\xi_1}(X_s^{\xi_1},R_s^{\xi_1})-h_s^{\xi_2}(X_s^{\xi_2},R_s^{\xi_2})ds-\int_t^T\delta_\xi\tilde{Z}_s dB_s.\end{cases}$$

Proposition 4.38. *For $\lambda > 0$ sufficiently small, there is a constant C such that*

$$\mathbf{E}\left[\sup_{s\in[t,T]} \|\delta_\xi R_s\|^2_{L^2(\rho_\ell)} + \int_t^T \|\delta_\xi \tilde{Z}_s\|^2_{\mathcal{L}_{2,\ell}} ds\right] \le C\langle t\rangle^{-2\delta} \|\xi_1 - \xi_2\|^2_{L^2(\rho_\ell)}.$$

Proof. In the same way as before, this estimate comes down to estimating

$$\|\nabla g(X_T^{\xi_1}) - \nabla g(X_T^{\xi_2})\|_{L^2(\rho_\ell)} \le C\lambda \sup_t \|\delta_\xi R_t\|_{L^2(\rho_\ell)},$$

and the spread in the generator h. Splitting the difference in the same way as in the proof of Proposition 4.37, we see that

$$\begin{aligned}
&\|\xi_1 \alpha_s \cos(\beta X_s^{\xi_1}) Q_s^2 R_s^{\xi_1} - \xi_1 \alpha_s \cos(\beta X_s^{\xi_1}) Q_s^2 R_s^{\xi_1}\|_{L^2(\rho_\ell)} \\
&\le \|\xi_1(\alpha_s \cos(\beta X_s^{\xi_1}) Q_s^2 R_s^{\xi_1} - \alpha_s \cos(\beta X_s^{\xi_2}) Q_s^2 R_s^{\xi_2})\|_{L^2(\rho_\ell)} \\
&\quad + \|(\xi_1 - \xi_2)\alpha_s \cos(\beta X_s^{\xi_2}) Q_s^2 R^{\xi_2}\|_{L^2(\rho_\ell)} \\
&\le \|\xi_1 \alpha_s \cos \beta X^{\xi_1} Q_s^2 \delta_\xi R_s\|_{L^2(\rho_\ell)} \\
&\quad + \|\xi_1 \alpha_s (\cos \beta X_s^{\xi_1} - \alpha_s \cos \beta X^{\xi_2}) Q_s^2 R_s^{\xi_2}\|_{L^2(\rho_\ell)} \\
&\quad + \|\alpha_s \cos(\beta X_s^{\xi_2}) Q_s^2 R^{\xi_2}\|_{L^\infty} \|\xi_1 - \xi_2\|_{L^2(\rho_\ell)} \\
&\le C\langle s\rangle^{-\frac{3}{2}-\delta} \|\delta_\xi R_s\|_{L^2(\rho_\ell)} + C\langle s\rangle^{-\frac{3}{2}-2\delta} \|\delta R_s\|_{L^2(\rho_\ell)} + C\langle s\rangle^{-\frac{3}{2}-2\delta} \|\xi_1 - \xi_2\|_{L^2(\rho_\ell)}.
\end{aligned}$$

The other term in h can be treated in the same way and we obtain

$$\begin{aligned}
\mathbf{E}\|h_s^{\xi_1}(X_s^{\xi_1}, R_s^{\xi_1}) - h_s^{\xi_2}(X_s^{\xi_2}, R_s^{\xi_2})\|^2_{L^2(\rho_\ell)} \le \lambda^2 C\langle s\rangle^{-2-4\delta}\, \mathbf{E}\left[\sup_s \|\delta_\xi R_s\|^2_{L^2(\rho_\ell)}\right] \\
+ C\lambda^2 \langle s\rangle^{-2-4\delta} \|1_{\xi_1} - 1_{\xi_2}\|^2_{L^2(\rho_\ell)}
\end{aligned}$$

Therefore, by arguments already used,

$$\begin{aligned}
&\mathbf{E}\left[\sup_{t\le T} \|\delta_\xi R_t\|^2_{L^2(\rho_\ell)(A)} + \int_0^T \|\delta_\xi \tilde{Z}_s\|^2_{\mathcal{L}^2(\rho_\ell)} ds\right] \\
\le& \lambda^2 C\, \mathbf{E}[\sup_{t\le T} \|\delta_\xi R_t\|^2_{L^2(\rho_\ell)}] + C\lambda^2 \|\xi_1 - \xi_2\|^2_{L^2(\rho_\ell)}.
\end{aligned}$$

For $\lambda > 0$ small, the result follows after rearranging. □

4.4.6 Dependence on the Perturbation g

To control the Laplace transform, it will be useful to quantify the dependence on the perturbation g. We fix the terminal time $T > 0$ and the smooth cut-off ξ. Following the same ideas as before, we introduce a system for the spread $\delta_g \Theta_T = \Theta_T^g - \Theta_T$ between the solution to (4.23) for $V_T^{\xi,g} = V_T^\xi + g$ and $V_T^0 = V_T$. We stress the fact that the following result also holds for $\ell > 0$: The effect of a perturbation ∇g which is concentrated on a compact set, only has a localised effect on the solution R and, thus, the optimal control u.

Proposition 4.39. *For the spread of the solution, it holds for any ℓ,*

$$\mathbf{E}\left[\sup_{t\in[0,T]}\|\delta_g R_t\|^2_{L^2(\rho_\ell)}+\int_0^T\|\delta_g Z_s\|^2_{\mathcal{L}_2(\rho_\ell)}ds\right]\leq C\,|\nabla g|^2_\ell\,.$$

Proof. The difference $\delta_g\Theta$ satisfies the FBSDE

$$\begin{cases}\delta_g X_t &= \int_0^t Q_s^2(\nabla V_T^{\xi,g}(X_{s,T}^g)-\nabla V_T^{\xi}(X_{s,T})+\delta_g R_s)ds,\\ \delta_g R_t &= \nabla g(X_{T,T}^g)+\int_t^{T_2}(h_s^g-h_s^0)ds-\int_t^{T_2}\delta_g Z_s dB_s,\end{cases}$$

where

$$h_s^f=h(s,X_{s,T}^f,R_{s,T}^f).$$

For X, the same argument as in Proposition 4.32, applies and

$$\|\delta_g X_{s,T}\|_{L^2(\rho_\ell)}\leq\lambda C\int_0^t\langle s\rangle^{-2}\|\nabla V_s^{\xi,g}(X_{s,T}^g)-\nabla V_s(X_{s,T})\|_{L^2(\rho_\ell)}ds+\int_0^t\|Q_s^2\delta_g R_s\|_{L^2(\rho_\ell)}ds.$$

By the definition of $V^{\xi,g}$ and the Lipschitz continuity,

$$\|\nabla V_s^{\xi,g}(X_{s,T}^g)-\nabla V_s(X_{s,T})\|_{L^2(\rho_\ell)}\leq\|\nabla g(X_{s,T}^g)\|_{L^2(\rho_\ell)}+C\alpha_s\|\delta_g X_s\|_{L^2(\rho_\ell)}.$$

Thus, by Gronwalls inequality,

$$\|\delta_g X_t\|_{L^2(\rho_\ell)}\leq C(|\nabla g|_\ell+\int_0^t\|Q_s^2\delta_g R_s\|_{L^2(\rho_\ell)}).\tag{4.29}$$

For the difference in the backward equation, we split up h^g-h^0 as

$$\begin{aligned}&\|h_s^g-h_s^0\|_{L^2(\rho_\ell)}\\ \leq&\beta^2\lambda\|[\![\cos(\beta X_{s,T}^g)]\!]Q_s^2(\nabla V_s^{\xi,g}(X_{s,T}^g)-\nabla V_s(X_{s,T}))\|_{L^2(\rho_\ell)}\\ &+\beta^2\lambda\|([\![\cos(\beta X_{s,T}^g)]\!]-[\![\cos(\beta X_{s,T})]\!])Q_s^2\nabla V_s(X_{s,T})\|_{L^2(\rho_\ell)}\\ &+\beta^2\lambda\|[\![\cos(\beta X_{s,T}^g)]\!]Q_s^2\delta_g R_s\|_{L^2(\rho_\ell)}+\beta^2\lambda\|([\![\cos(\beta X_{s,T}^g)]\!]-[\![\cos(\beta X_{s,T}^g)]\!])Q_s^2 R_s\|_{L^2(\rho_\ell)}.\end{aligned}$$

In the same way as before (as in the proof of Proposition 4.37), we use the Lipschitz continuity, the bounds on α_s and the uniform estimates on R from Proposition 4.36 to obtain

$$\begin{aligned}\|h_s^g-h_s^0\|_{L^2(\rho_\ell)}\leq&\,\beta^2\lambda\alpha_s\langle s\rangle^{-2}|\nabla g|_\ell+\beta^2\lambda\alpha_s\langle s\rangle^{-2}\|\delta_g R_s\|_{L^{2,\ell}}\\ &+(\beta^3\lambda^2\alpha_s^2\langle s\rangle^{-2}C+\beta^2\lambda\alpha_s\langle s\rangle^{-2})\|\delta_g X_s\|_{L^2(\rho_\ell)}\\ \leq&\,C\lambda\alpha_s^2\langle s\rangle^{-2}|\nabla g|_\ell+C\lambda\alpha_s^2\langle s\rangle^{-2}\sup_{s\in[0,T]}\|\delta_g R_s\|_{L^2(\rho_\ell)}.\end{aligned}$$

where we also applied (4.29) in the second step. The usual a priori estimate with $\varepsilon_s = \langle s\rangle^{-1-2\delta}$ as in (4.27) then allows us to conclude

$$\begin{aligned}
&\mathbf{E}\left[\sup_{t\in[0,T]}\|\delta_g R_t\|^2_{L^2(\rho_\ell)}+\int_0^T\|\delta_g Z_s\|^2_{\mathcal{L}_2(\rho_\ell)}ds\right]\\
\leq& C\,|\nabla g|^2_\ell + C\,\mathbf{E}\left[\int_0^T(\varepsilon_s\|\delta_g R_s\|^2_{L^2(\rho_\ell)}+\varepsilon_s^{-1}\|h^g_s-h^0_s\|^2_{L^2(\rho_\ell)})ds\right]\\
\leq& C\,|\nabla g|^2_\ell + \lambda^2 C\,\mathbf{E}\left[\sup_{s\in[0,T]}\|\delta_g R_s\|^2_{L^2(\rho_\ell)}\right].
\end{aligned}$$

Choosing $\lambda > 0$ sufficiently small, the claim follows after rearranging. □

4.5 Removing the cut-off

The estimate we derived in the previous section now allows us to remove both the ultraviolet and the infrared cut-off. Since all estimates are uniform in T and ξ, the order in which these limits are applied does not matter.

4.5.1 The Ultraviolet Limit

In this section, we apply the bounds we derived to conclude convergence for the control as $T\to\infty$. We always assume that $\lambda > 0$ is small enough and fix the cut-off ξ. We leave the dependency implicit whenever possible to simplify the notation. Since the bounds derived in the previous section are uniform in ξ, provided $\ell < -1$, this does not lead to ambiguities.

Let us recall that the optimal control u is given by

$$u_{t,T} = u^{\xi,g}_{t,T} = Q_t\nabla Y^{\xi,g}_{t,T} = Q_t(\nabla V^{\xi,g}_t(X^{\xi,g}_{t,T}) + R^{\xi,g}_{t,T}).$$

Lemma 4.40. *The sequence $\{R_{t,T}, t\geq 0\}_T$ of solutions to the FBSDE (4.23) converges in $\mathbb{H}^\infty_T(L^2(\rho_\ell))$ to a limit $\{\overline{R}_t,\ t\in[0,\infty)\}$. The convergence is uniform in ξ and the limit is bounded uniformly in t and ξ, with*

$$\|\overline{R}_t\|^2_{L^\infty}+\|\overline{R}_t\|^2_{L^2(\rho_\ell)} \leq |\nabla g|^2_\ell + C\lambda^4\langle t\rangle^{-4\delta} < \infty.$$

Moreover, using the notation $I_t(u) = \int_0^t Q_s u_s ds$, we have as $T\to\infty$,

$$I_T(Q.R_{\cdot,T}) \to I_\infty(Q\overline{R}) \in L^\infty(P\times\mathbb{R}^2) \text{ in } L^2(P, H^1(\rho_\ell)).$$

Proof. By Proposition 4.37, the sequence $\{R_{t,T}, t\geq 0\}_T$ is Cauchy in $\mathbb{H}^\infty(L^2(\rho_\ell))$ and thus converges to a limit which we denote by $\overline{R}$. The bounds now follow immediately from Proposition 4.36. The boundedness of $I_\infty(Q\overline{R})$ follows from the bound on $\overline{R}$ and Lemma 4.4. Finally, the convergence in $L^2(P, H^1(\rho_\ell))$ follows from Lemma 4.5, since

$$\begin{aligned}\mathbf{E}\|\int_0^\infty Q_s^2(R_{s,T}-\overline{R}_s)ds\|^2_{H^1(\rho_\ell)} &\le \mathbf{E}\int_0^\infty \|Q_s(R_{s,T}-\overline{R}_s)\|^2_{L^2(\rho_\ell)}ds\\ &\le \mathbf{E}\int_0^\infty \langle s\rangle^{-2+\delta}\|R_{s,T}-\overline{R}_s\|^2_{L^2(\rho_\ell)}ds\\ &\le C\lambda^4\langle T\rangle^{-2\delta}.\end{aligned}$$

□

Theorem 4.41. *The process $\overline{X}_t = X_t^{\overline{R}}$, where $\overline{R}$ is the limit introduced in Lemma 4.40, converges in $L^2(P, H^{-\delta}(\rho_\ell))$ to a limit*

$$\overline{X}_\infty = \mathcal{I}_\infty + W_\infty,$$

where $W_\infty \sim \mu$ is a realisation of the Gaussian free field on $H^{-\delta}(\rho_\ell)$ and $I_\infty \in L^2(P, H^1(\rho_\ell))$. Moreover, $(\overline{X}, \overline{R}, \overline{Z})$ is the unique solution to (4.23) at $T = \infty$.

Proof. The convergence of $W_t \to W_\infty$ in $L^2(P;\ H^{-\delta}(\rho_\ell))$ with $\mathrm{Law}(W_\infty) = \mu$ follows from Proposition 4.1. For the drift, it only remains to show convergence of the integral $\int_0^\infty Q_s^2 V_s^{\xi,g}(\overline{X}_s)ds$. Thanks to the boundedness of $V_s^{\xi,g}$ in $L^2(\rho_\ell)$, the integral exists and by Lemma 4.5,

$$\begin{aligned}&\mathbf{E}\|\int_0^\infty Q_s^2\nabla V_s^{\xi,g}(\overline{X}_s)ds\|^2_{H^1(\rho_\ell)}\\ \le\ &\mathbf{E}\int_0^\infty \|Q_s\nabla V_s^{\xi,g}(\overline{X}_s)\|^2_{L^2(\rho_\ell)}ds\\ \le\ &C\int_0^\infty \langle s\rangle^{-2}\langle s\rangle^{1-2\delta}ds < \infty.\end{aligned}$$

Thus $\overline{X}_\infty = \mathcal{I}_\infty + W_\infty$ with the claimed regularity. The pair $(\overline{X}, \overline{R}, \overline{Z})$ satisfies (4.23) by construction and uniqueness was already shown in Theorem 4.30. □

Let us also note the convergence of the optimal control for future reference.

Lemma 4.42. *There is a control $\overline{u}^\xi \in \mathbb{H}^2(L^2(\rho_\ell))$ such that*

$$\lim_{T\to\infty}\mathbf{E}\int_0^\infty \|\overline{u}_t^\xi - u_{t,T}^\xi\|^2_{L^2(\rho_\ell)} = 0.$$

Moreover, $\|\langle t\rangle^{\frac{1}{2}+\delta}\overline{u}_t^\xi\|_{L^\infty(\mathbb{R}_+\times\mathbb{R}^2)} < \infty$.

Lemma 4.43. *There is a constant C such that*

$$|\mathcal{V}_T(\varphi)| = \left|Y_{0,T}^{0,\varphi}\right| \le \lambda\,|\mathrm{supp}(\xi)|\,C.$$

Proof. To obtain a uniform bound on Y, we rely on the same idea as for ∇Y and rewrite

$$Y_{t,T} = V_t(X_{t,T}) + \int_t^T \langle \nabla V_s(X_{s,T}), dX_{s,T}\rangle_{L^2} + \frac{1}{2}\int_t^T \|Q_s \nabla Y_{s,T}\|_{L^2}^2 ds - \int_t^T \nabla Y_{s,T} Q_s dB_s.$$

Here we again used that the Wick-ordered cosine satisfies the Fokker-Planck equation (4.4). We compute,

$$\int_t^T \langle \nabla V_s(X_{s,T}), d(X_{s,T})\rangle_{L^2} ds = \lambda\beta \int_s^T \int \xi [\![\sin(\beta X_{s,T})]\!] Q_s^2 \nabla Y_{s,T} ds + \int_s^T \int \xi\lambda\beta [\![\sin(\beta X_{s,T})]\!] Q_s dB_s.$$

The stochastic integral is a martingale since

$$\mathbf{E}\int_t^T \int \|\xi [\![\sin(\beta X_{t,T})]\!] Q_s\|_{L^2}^2 ds \leq C \int_t^T \langle s\rangle^{-1-2\delta} |\mathrm{supp}(\xi)|^2 ds < \infty.$$

Then, letting $t = 0$ and taking expectation, we see that

$$Y_{0,T} = \mathbf{E}\left[\lambda |\mathrm{supp}(\xi)| + \int_0^T \lambda\beta \int [\![\sin(\beta X_{s,T})]\!] Q_s^2 \nabla Y_{s,T} ds + \frac{1}{2}\int_0^T \|Q_s \nabla Y_{s,T}\|_{L^2}^2 ds\right].$$

Since we have already seen that $\int_0^T \|Q_s \nabla Y_{s,T}\|_{L^2}^2 ds$ is uniformly bounded, it only remains to take care of the term in the middle. By Hölder's inequality, the estimate on α and Lemma 4.4,

$$\left|\int \xi [\![\sin(\beta X_{s,T})]\!] Q_s^2 \nabla Y_{s,T}\right| \leq \|\xi [\![\sin(\beta X_{s,T})]\!]\|_{L^2} \|Q_s^2 \nabla Y_{s,T}\|_{L^2} \leq |\mathrm{supp}(\xi)| \langle s\rangle^{-1-2\delta}.$$

As this is integrable over $\mathbb{R}_+$, the claim follows. □

By undoing the change of variables, the convergence of the remainder R also means there is a unique solution to (4.22) at $T = \infty$.

Theorem 4.44. *There is a unique solution* $(\overline{X}^\xi, \nabla\overline{Y}^\xi, \nabla\overline{Z}^\xi)$ *to* (4.22) *at* $T = \infty$. *More precisely for any* $0 \leq t \leq T < \infty$, *the triple satisfies*

$$\begin{cases} \overline{X}_t^\xi &= \varphi - \int_0^t Q_s^2 \nabla\overline{Y}_s^\xi ds + \int_0^t Q_s dB_s, \\ \nabla\overline{Y}_t^\xi &= \nabla\overline{Y}_T^\xi - \int_t^T \nabla\overline{Z}_s^\xi dB_s. \end{cases}$$

If $|\mathrm{supp}(\xi)| < \infty$, *the pair* $(\overline{X}^\xi, \overline{u}^\xi)$ *is optimal for the variational problem*

$$\mathcal{V}_\infty^\xi = \inf_{u\in\mathcal{A}} J_\infty^\xi(u) = \inf_{u\in\mathcal{A}} \left\{ V_\infty^\xi(W_\infty + I_\infty(u)) + \frac{1}{2}\int_0^\infty \|u_s\|_{L^2(\Lambda)}^2 ds \right\},$$

where again $I_t(u) = \int_0^t Q_s u_s ds$ *and*

$$\begin{aligned} V_\infty^\xi(I_\infty(u)+W_\infty) &= \int \xi [\![\cos(\beta(W_\infty + I_\infty(u)))]\!] \\ &= \int \xi [\![\cos(\beta W_\infty)]\!](\cos(\beta I_\infty(u)) + [\![\sin(\beta W_\infty)]\!] \sin(\beta I_\infty(u))). \end{aligned}$$

In particular, $\mathrm{Law}(\overline{X}_\infty^\xi) = \mathrm{Law}(W_\infty + \mathcal{I}_\infty^\xi) = \nu_{SG}^\xi$.

Proof. By definition of the remainder,

$$\begin{cases} \nabla \overline{Y}_t^\xi = \nabla V_t^\xi(\overline{X}_t^\xi) + \overline{R}_t^\xi \\ \nabla \overline{Z}_t^\xi = \overline{Z}_t^\xi + \beta^2 \lambda^\xi [\![\cos \beta \overline{X}_t^\xi]\!] Q_t. \end{cases}$$

Hence, it follows directly from the definition of R, $\overline{R}$ and $\overline{X}$ and Theorem 4.41 that the triple $(\overline{X}^\xi, \nabla \overline{Y}^\xi, \nabla \overline{Z}^\xi)$ satisfies (4.22). If $\overline{X}^\xi$ is optimal, the statement about the law follows from Lemma 4.18. It remains to check that for a compactly supported cut-off ξ with $\mathrm{supp}(\xi) \subset \Lambda$ bounded, we also have convergence for

$$\overline{Y}_t^\xi = \overline{Y}_T^\xi + \int_t^T \frac{1}{2} \|Q_s \nabla \overline{Y}_s^\xi\|_{L^2}^2 ds - \int_t^T \nabla \overline{Y}_s^\xi Q_s dB_s,$$

and that $\overline{X}^\xi$ is the optimally controlled process with $\overline{Y}_0^\xi = \mathcal{V}_\infty^\xi$.

Convergence of the terminal condition. Let us start by rewriting $\nabla V_T^\xi(\mathcal{I}_T + W_T)$ as

$$\begin{aligned} [\![\sin(\beta X_{T,T}^\xi)]\!] &= [\![\sin(\beta W_T + \beta \mathcal{I}_T^\xi)]\!] \\ &= [\![\cos(\beta W_T)]\!] \cos(\beta \mathcal{I}_T^\xi) + [\![\sin(\beta W_T)]\!] \sin(\beta \mathcal{I}_T^\xi). \end{aligned}$$

Thus, splitting up the integral,

$$\begin{aligned} &\mathbf{E} \int \xi \Big([\![\cos(\beta W_T)]\!] \cos(\beta \mathcal{I}_T^\xi) - [\![\cos(\beta W_\infty)]\!] \cos(\beta \mathcal{I}_\infty^\xi) \Big) \\ =&\mathbf{E} \int \xi \Big([\![\cos(\beta W_T)]\!] \{\cos(\beta \mathcal{I}_T^\xi) - \cos(\beta \mathcal{I}_\infty^\xi)\} \\ &\quad + \{[\![\cos(\beta W_T)]\!] - [\![\cos(\beta W_\infty)]\!]\} \cos(\beta \mathcal{I}_\infty^\xi) \Big). \end{aligned}$$

By the dual pairing of $H^{-1+\delta}$ and $H^{1-\delta}$,

$$\begin{aligned} &\mathbf{E} \int \xi [\![\cos(\beta W_T)]\!] \{\cos(\beta \mathcal{I}_T^\xi) - \cos(\beta \mathcal{I}_\infty^\xi)\} \\ \leq &\mathbf{E} \Big[\|[\![\cos(\beta W_T)]\!]\|_{H^{-1+\delta}(\Lambda)} \|\cos(\beta \mathcal{I}_T^\xi) - \cos(\beta \mathcal{I}_\infty^\xi)\|_{H^1(\Lambda)} \Big] \\ \leq &\mathbf{E} \Big[\|[\![\cos(\beta W_T)]\!]\|_{H^{-1+\delta}(\Lambda)}^2 \Big]^{\frac{1}{2}} \mathbf{E} \Big[\|\cos(\beta \mathcal{I}_T^\xi) - \cos(\beta \mathcal{I}_\infty^\xi)\|_{H^1(\Lambda)}^2 \Big]^{\frac{1}{2}}. \end{aligned}$$

Thanks to the boundedness and continuity of the sine and cosine, the convergence $\mathcal{I}_T^\xi \to \mathcal{I}_\infty^\xi$ in $L^2(P, H^1(\Lambda))$ is inherited for $\cos(\beta \mathcal{I}_T^\xi)$ and $\sin(\beta \mathcal{I}_T)$. By Proposition 4.10, the renormalised cosine and sine are martingales, convergent in $L^2(P, H^{-1+\delta}(\Lambda))$ provided $\beta^2 < 4\pi$. In other words, the first term is bounded uniformly in T, and the second term approaches 0 as $T \to 0$. It follows in the same way that

$$\mathbf{E}\int \xi\{[\![\cos(\beta W_T)]\!] - [\![\cos(\beta W_\infty)]\!]\}\cos(\beta \mathcal{I}_\infty^\xi) \to 0, \quad (T \to \infty).$$

Of course, the exact same argument shows convergence for the cosine replaced by sine. Hence,

$$\lim_{T\to\infty} \mathbf{E}\left[V_T^\xi(\overline{X}_T^\xi) - V_\infty^\xi(\overline{X}_\infty^\xi)\right] = 0,$$

as required.

Backward equation at $T = \infty$. We have already seen that the terminal condition for $\overline{Y}^\xi$ converges to V_∞^ξ. It remains to check convergence for the drift of ∇Y^ξ. Recalling that $\nabla \overline{Y}_t^\xi = \nabla V_t^\xi(X_t^\xi) + \overline{R}_t^\xi$ we can use the estimates on $\overline{R}_t^\xi$, the estimates on the renormalisation $\alpha_t \lesssim \langle t \rangle^{\frac{1}{2}-\delta}$ and Lemma 4.4 to see that

$$\mathbf{E}\int_0^\infty \|Q_r \nabla Y_r^\xi\|_{L^2(\Lambda)}^2 dr \leq \int_0^t \langle r \rangle^{-2} \|\nabla V_t^\xi(\overline{X}_t) + \overline{R}_t^\xi\|_{L^2(\Lambda)}^2 dr \leq \int_0^\infty \langle r \rangle^{-2} \langle r \rangle^{1-2\delta} dt < \infty.$$

Thus, as $T \to \infty$,

$$\begin{aligned}\overline{Y}_t^\xi &= \overline{Y}_T^\xi + \int_t^T \|Q_r \nabla Y_r^\xi\|_{L^2(\Lambda)}^2 dr + \int_t^T \nabla Y_r^\xi Q_r dB_r \\ &\to V_\infty^\xi(\overline{X}_\infty^\xi) + \int_t^\infty \|Q_r \nabla Y_r^\xi\|_{L^2(\Lambda)}^2 dr + \int_t^\infty \nabla Y_r^\xi Q_r dB_r.\end{aligned}$$

Optimality. We show that the limit still coincides with the value function at $T = \infty$. Let $X_{t,T}^\xi(u)$ be the solution to (4.9) with control u. In the same way as before we can show for any $u \in \mathbb{H}^2(L^2(\Lambda))$,

$$\mathbf{E}[V_T^\xi(X_{T,T}^\xi(u)) - V_\infty^\xi(X_{\infty,\infty}^\xi(u))] \to 0, \quad (T \to \infty),$$

and use monotone convergence for the quadratic term to conclude,

$$\lim_{T\to\infty} J_T^\xi(u) = J_\infty^\xi(u),$$

for any fixed control u. Hence, since $u^{\xi,T}$ is optimal for J_T^ξ,

$$J_\infty^\xi(u) = \lim_{T\to\infty} J_T^\xi(u) \geq \liminf_{T\to\infty} \inf_{u\in\mathcal{A}} J_T^\xi(u) = \liminf_{T\to\infty} J_T^\xi(u^{\xi,T}).$$

On the other hand, for the sequence of optimal controls $u^{\xi,T}$, we have by Fatou's lemma,

$$\liminf_{T\to\infty} \mathbf{E}\int_0^T \|u_s^{\xi,T}\|_{L^2}^2 ds \geq \mathbf{E}\int_0^\infty \|\overline{u}_s^\xi\|_{L^2}^2 ds.$$

Since $u^{\xi,T}$ minimises J_T^ξ for any finite T, this implies with the convergence of $V_T^\xi(X_T^\xi(u))$,

$$\liminf_{T\to\infty} J_T^\xi(u^{\xi,T}) \geq J_\infty^\xi(\overline{u}^\xi).$$

Combining both estimates

$$J_\infty^\xi(\overline{u}^\xi) \leq \liminf_{T\to\infty} J_T^\xi(u^{\xi,T}) \leq \inf_{u\in\mathcal{A}} J_\infty^\xi(u).$$

As $\overline{u}^\xi \in \mathbb{H}^2(L^2(\Lambda))$ is admissible, this implies $\overline{Y}_\infty^\xi = J_\infty^\xi(\overline{u}^\xi) = \inf_{u\in\mathcal{A}} J_\infty^\xi(u)$. □

Remark 4.45. By Remark 3.3 the condition (4.20) is necessary and uniquely determines the optimal control. In the proof above, we have just shown that for weak interactions λ there is a unique solution to (4.20). Therefore, $\overline{X}_\infty$ is the unique optimally controlled process and by Lemma 4.18 also the measure $\nu_{\mathrm{SG}}^\xi = \mathrm{Law}(\overline{X}_\infty^\xi)$ is unique.

4.5.2 The Infinite Volume Limit

For small interactions $\lambda > 0$, we show weak convergence to a unique measure on $\mathbb{R}^2$. We assume that $\ell > 1$.

Theorem 4.46. *For a smooth cut-off ξ, let $\overline{u}_s^\xi = Q_s\nabla\overline{Y}_s^\xi$ be the optimal control for the control problem on* $\mathrm{supp}(\xi)$*. Then, For $\lambda > 0$ sufficiently small there is a $\overline{u}_t \in \mathbb{H}^2(L^2(\rho_\ell))$ such that*

$$\lim_{\xi\to 1} \mathbf{E}\int_0^\infty \|\overline{u}_t - \overline{u}_t^\xi\|_{L^2(\rho_{-\ell})}^2 dt = 0.$$

Moreover, $\|\overline{u}_t\langle t\rangle^{\frac{1}{2}+\delta}\|_{L^\infty(\mathbb{R}_+\times\mathbb{R}^2)} < \infty$.

Proof. By definition $\overline{u}_s^\xi = Q_s(\nabla V_s^\xi(\overline{X}_s^\xi) + \overline{R}_s^\xi)$. Proposition 4.38 implies

$$\mathbf{E}\int_0^\infty \|Q_s(\overline{R}_s - \overline{R}_s^\xi)\|_{L^2(\rho_\ell)}^2 ds \leq \mathbf{E}\int_0^\infty \langle s\rangle^{-2-\delta}\|1-\xi\|_{L^2(\rho_{-\ell})}^2 ds \to 0.$$

Thanks to Proposition 4.32, we also see that $\overline{X}_s^\xi \to \overline{X}_s$ in $L^2(P, L^2(\rho_{-\ell}))$. Hence, with the continuity of ∇V_s and the estimates on α_s,

$$\begin{aligned}
&\mathbf{E}\int_0^\infty \|Q_s(\nabla V_s^\xi(\overline{X}_s^\xi) - \nabla V_s(\overline{X}_s^\xi))\|_{L^2(\rho_{-\ell})}^2 ds \\
\leq &\mathbf{E}\int_0^\infty \langle s\rangle^{-2}\alpha_s^2(\|\overline{X}_s^\xi - \overline{X}_s\|_{L^2(\rho_{-\ell})}^2 + \|(1-\xi)\nabla V_s(\overline{X}_s)\|_{L^2(\rho_{-\ell})}^2) ds \\
\leq &\|1-\xi\|_{L^2(\rho_{-\ell})}^2 \int_0^\infty \langle s\rangle^{-1-2\delta} ds \to 0, \quad \text{as} \quad \xi \to 1,
\end{aligned}$$

which allows us to conclude the desired convergence. The moreover part follows from the bounds on V_t and $\overline{R}$, since uniformly in t,

$$\|\overline{u}_t\|_{L^\infty(\mathbb{R}^2)} \le t^{-1}\|V_t(\overline{X}_t)+\overline{R}_t\|_{L^\infty(\mathbb{R}^2)} \le C\langle t\rangle^{-\frac{1}{2}-\delta}.$$

□

Corollary 4.47. *For a smooth cut-off ξ, let $\mathcal{I}_t^\xi = \int_0^t Q_s \overline{u}_s^\xi ds$. There is a unique $\mathcal{I}_\infty = I_\infty(\overline{u}) \in L^\infty(P \times \mathbb{R}^2)$ such that*

$$\lim_{\xi\to 1} \mathbf{E}\|\mathcal{I}_\infty - \mathcal{I}_\infty^\xi\|^2_{H^1(\rho_{-\ell})} \to 0,$$

and thus $\overline{X}_\infty^\xi \to \overline{X}_\infty$ in $L^2(P, H^{-\delta}(\rho_{-\ell}))$. In particular, the family $\{\nu_{SG}^\xi\}_\xi$ has a unique weak limit $\nu_{SG} = \mathrm{Law}(\overline{X}_\infty)$ on $H^{-\delta}(\rho_{-\ell})$.

Proof. The convergence of $\mathcal{I}^\xi$ is a direct consequence of Theorem 4.46 and Lemma 4.5. This in turn also implies the convergence of X_∞^ξ in $L^2(P, H^{-\delta}(\rho_{-\ell}))$ by Proposition 4.10. If $|\mathrm{supp}(\xi)| < \infty$, we know from Theorem 4.41 that for any bounded and continuous function $f : H^{-\delta}(\rho_{-\ell}) \to \mathbb{R}$

$$\nu_{\mathrm{SG}}^\xi(f) = \mathbf{E} f(\overline{X}_\infty^\xi) \to \mathbf{E} f(\overline{X}_\infty), \quad (\xi \to 1).$$

In other words, the measure $\nu_{\mathrm{SG}} := \mathrm{Law}(\overline{X}_\infty)$ is the unique weak limit of ν_{SG}^ξ. □

Remark 4.48. While with our estimates, the convergence of the stochastic processes to a unique limit requires λ to be sufficiently small, we still have tightness of $\{\nu_{\mathrm{SG}}^\xi\}_\xi$. Indeed, by Lemma 4.42, $\sup_\xi \|\langle t\rangle^{\frac{1}{2}+\delta}\overline{u}_t^\xi\|_{L^\infty(\mathbb{R}_+\times\mathbb{R}^2)} < \infty$ and thus

$$\mathcal{I}_T^\xi = \int_0^T Q_s^2 \overline{u}_{s,T}^\xi ds,$$

is uniformly bounded in $L^\infty(P; W^{1,\infty}(\mathbb{R}^2))$ and $L^2(P; H^1(\rho_\ell))$ by Lemma 4.8 and respectively Lemma 4.5. Thus, there is a weakly convergent subsequence of $\{\mathcal{I}_\infty^\xi\}$ with an accumulation point which we denote by $\mathcal{I}_\infty$. This means we can still interpret the shifted free field $\mathcal{I}_\infty + W_\infty$ as a realisation of the sine-Gordon measure on $H^{-\delta}(\rho_{-\ell})$. We can, however, no longer deduce uniqueness for the limit from our estimates.

Thanks to the uniform boundedness of $\{\mathcal{I}_\infty^\xi\}_\xi$ in $W^{1,\infty}$, we also immediately get that any accumulation point of $\{\nu_{\mathrm{SG}}^\xi\}_\xi$ has Gaussian tails.

Corollary 4.49. *Let $\mathcal{I}_\infty$ be an accumulation point of $\{\nu_{SG}^\xi\}$. For any $\delta > 0$, there is a $\gamma > 0$ such that*

$$\int e^{\gamma\|\varphi\|^2_{H^{-\delta}(\rho_{-\ell})}} \nu_{SG}(d\varphi) = \mathbf{E}[\exp(\|\mathcal{I}_\infty + W_\infty\|^2_{H^{-\delta}(\rho_{-\ell})})] < \infty.$$

4.6 Variational Description on $\mathbb{R}^2$

Even though the value function is not meaningful on the infinite volume, we can derive a variational principle for the Laplace transform from Proposition 4.39. The section follows [3] modulo minor changes in some estimates and some additional details.

The crucial ingredient is the following locality property of the optimal control.

Lemma 4.50. *Let $u^{g,\xi}$ be optimal for g and denote as before by $\overline{u}^{\xi} = u^{0,\xi}$ the optimal control for the unperturbed variational problem. Then,*

$$\mathbf{E}\int_0^\infty \|u_s^{g,\xi} - \overline{u}_s^{\xi}\|^2_{L^2(\langle\rho\rangle^n)} ds \leq |\nabla g|_n^2.$$

Proof. The proof follows from Proposition 4.39 in complete analogy to the proof of Theorem 4.46 from Proposition 4.38. □

Remark 4.51. Thanks to Lemma 4.6, the statement of Lemma 4.50 and consequently also Theorem 4.52 applies when the polynomial weight is replaced by an exponential weight. In particular, the effect of a perturbation g whose gradient decays exponentially/polynomially in space on the control is also localised exponentially/polynomially.

Theorem 4.52. *For a sufficiently large constant $C > 0$ depending only on $|\nabla g|_n$ where $n > 2$, define*

$$\mathcal{A}(g) := \{v \in \mathbb{H}^2(L^2(\rho_n));\ \mathbf{E}\int_0^\infty \|v_t\|^2_{L^2(\rho_n)} dt < C\}.$$

With the notation $\mathcal{W}^{\xi}(g) = \mathcal{V}^{\xi,g} - \mathcal{V}^{\xi,0}$ for the Laplace tranform of ν_{SG}^{ξ}, we have

$$\lim_{\xi\to 1} \mathcal{W}^{\xi}(g) = \inf_{v\in\mathcal{A}(g)} \overline{J}^g(v),$$

where the cost functional is defined as

$$\begin{aligned}\overline{J}^g(v) = &\mathbf{E}[g(W_\infty + \mathcal{I}_\infty + I_\infty(v))\\ &+ \lambda\int ([\![\cos(\beta(W_\infty + \mathcal{I}_\infty + I_\infty(v)))]\!] - [\![\cos(\beta(W_\infty + \mathcal{I}_\infty))]\!])\\ &+ \int_0^\infty \langle \overline{u_t}, v_t\rangle_{L^2} dt + \frac{1}{2}\int_0^\infty \|v_t\|^2_{L^2} dt].\end{aligned}$$

Here, we again use the notation $\mathcal{I}_\infty = \int_0^\infty Q_s \overline{u}_s ds$ and $\overline{u}_s = Q_s \nabla \overline{Y}_s$ for the optimal control on $\mathbb{R}^2$ from Corollary 4.47.

The following Lipschitz estimate for bounded functions from [3] will be useful.

Lemma 4.53. *For $\|f^1\|_{W^{1,\infty}} + \|f^2\|_{W^{1,\infty}} \leq M$ and $h \in H^1(\rho_\ell)$, we have for any $\varepsilon \in (0,1)$,*

$$\begin{aligned}&\|(\cos(f^1 + h) - \cos(f^2 + h))h\|_{W^{1,1}(\rho_\ell)}\\ \leq &C_M(\|f^1 - f^2\|_{H^1(\rho_{-\ell})}\|h\|_{H^{1,}(\rho_{2\ell})} + \|f^1 - f^2\|^{\varepsilon}_{L^2(\rho_{-\ell})}\|h\|_{H^1(\rho_{2\ell})}).\end{aligned}$$

The analogous statement holds for the sine.

Proof. By the Lipschitz continuity, we have

$$\|(\cos(f^1+h)-\cos(f^2+h))h\|_{L^1(\rho_\ell)} \le \|f^1-f^2\|_{L^2(\rho_{-\ell})}\|h\|_{L^2(\rho_{2\ell})}.$$

For the derivative, we compute

$$\begin{aligned}
&\nabla((\cos(f^1+h)-\cos(f^2+h))h)\\
&=(\cos(f^1+h)-\cos(f^2+h))\nabla h+(\sin(f^1+h)(\nabla f^1+\nabla h)-\sin(f^2+h)(\nabla f^2+\nabla h))h\\
&=(\cos(f^1+h)-\cos(f^2+h))\nabla h+(\sin(f^1+h)-\sin(f^2+h))h\nabla f^2\\
&\quad+\sin(f^1+h)(\nabla f^1-\nabla f^2)h+(\sin(f^1+h)-\sin(f^2+h))h\nabla h.
\end{aligned}$$

The first term can be estimated as before,

$$\|(\cos(f^1+h)-\cos(f^2+h))\nabla h\|_{L^1(\rho_\ell)} \le \|f^1-f^2\|_{L^2(\rho_{-\ell})}\|\nabla h\|_{L^2(\rho_{2\ell})}.$$

Using the boundedness and Lipschitz continuity of the cosine,

$$\begin{aligned}
&\|(\sin(f^1+h)-\sin(f^2+h))h\nabla f^2\|_{L^1(\rho_\ell)} \le \|f^1-f^2\|_{L^2(\rho_{-\ell})}\|h\|_{L^2(\rho_{2\ell})}\|\nabla f^1\|_{L^\infty},\\
&\|\sin(f^1+h)(\nabla f^1-\nabla f^2)h\|_{L^1(\rho_\ell)} \le \|\nabla f^1-\nabla f^2\|_{L^2(\rho_{-\ell})}\|h\|_{L^2(\rho_{2\ell})},\\
&\|(\sin(f^1+h)-\sin(f^2+h))h\nabla h\|_{L^1(\rho_\ell)} \le \|f^1-f^2\|_{L^{p,}(\rho_{-\ell})}\|h\|_{L^q}\|\nabla h\|_{L^2(\rho_{2\ell})},
\end{aligned}$$

where we also used Hölder's inequality with $\frac{1}{p}+\frac{1}{q}=\frac{1}{2}$ for the last estimate. Interpolating between L^∞ and $L^2(\rho_{-\ell})$, we have for $\frac{1}{p}=\frac{\varepsilon}{2}$

$$\|f^1-f^2\|_{L^{p,}(\rho_{-\ell})} \le \|f^1-f^2\|_{L^\infty}^{1-\varepsilon}\|f^1-f^2\|_{L^2(\rho_{-\ell})}^{\varepsilon}.$$

To deal with the last L^q-norm, we use the Sobolev embedding $W^{1,2}\hookrightarrow L^q$ with $\frac{1}{q}=\frac{1}{2}-\frac{\varepsilon}{2}\in(0,\frac{1}{2})$,

$$\|h\|_{L^q} \le \|h\|_{H^1(\rho_{2\ell})}.$$

Combined with the assumption $\|f^1\|_{W^{1,\infty}}+\|f^2\|_{W^{1,\infty}}<M$, this yields the claim. □

Proof of Theorem 4.52. First, we note that with the optimal control $\overline{u}_s^\xi=Q_s\nabla\overline{Y}_s^\xi$,

$$\mathcal{V}^{\xi,g}-\mathcal{V}^{\xi,0}=\inf_{u\in\mathcal{A}}(J^{\xi,g}(u)-J^{\xi,0}(\overline{u}^\xi)).$$

Introducing the change of variables $v:=u-\overline{u}^\xi$, we define

$$\overline{J}^{\xi,g}(v):=J^{\xi,g}(v+\overline{u}^\xi)-J^{\xi,g}(\overline{u}^\xi).$$

Restriction to $\mathcal{A}(g)$. We claim that the minimiser $\overline{u}^{\xi,g}$ of $J^{\xi,g}(v+\overline{u})$ is already contained in $\mathcal{A}(g)$ provided that $C\gtrsim|\nabla g|_n^2$, from which the claim follows. Indeed, by Lemma 4.50, the minimiser v^* for $\overline{J}^\xi$ satisfies uniformly in ξ

$$\mathbf{E}\int_0^\infty\|v_t^*\|_{L^2(\rho_n)}^2dt=\mathbf{E}\int_0^\infty\|\overline{u}_t^{\xi,g}-\overline{u}_t^{\xi,0}\|_{L^2(\rho_n)}^2dt<C|\nabla g|_n^2.$$

Convergence of $\overline{J}^{\xi,g}$. We show instead that $\overline{J}^{\xi,g} \to \overline{J}^{g}$ uniformly on $\mathcal{A}(g)$, term by term, which then implies the claim. We need to control

$$\begin{aligned}
&\overline{J}^{g}(v)-\overline{J}^{\xi,g}(v)\\
&=\mathbf{E}\Big[g(W_\infty-I_\infty(v)-\mathcal{I}_\infty)-g(W_\infty-I_\infty(v)-\mathcal{I}_\infty^{\xi})\\
&\quad+\lambda\int \xi[\![\cos(\beta W_\infty)]\!]\big\{(\cos(\beta(I_\infty(v)+\mathcal{I}_\infty))-\cos(\beta(I_\infty(v)+\mathcal{I}_\infty^{\xi})))\\
&\qquad+\cos(\beta\mathcal{I}_\infty^{\xi})-\cos(\beta\mathcal{I}_\infty)\big\}\\
&\quad+\lambda\int \xi[\![\sin(\beta W_\infty)]\!]\big\{(\sin(\beta(I_\infty(v)+\mathcal{I}_\infty))-\sin(\beta(I_\infty(v)+\mathcal{I}_\infty^{\xi})))\\
&\qquad+\sin(\beta\mathcal{I}_\infty^{\xi})-\sin(\beta\mathcal{I}_\infty)\big\}\\
&\quad+\lambda\int(1-\xi)[\![\cos(\beta W_\infty)]\!]\{\cos(\beta(I_\infty(v)+\mathcal{I}_\infty))-\cos(\beta(\mathcal{I}_\infty))\}\\
&\quad+\lambda\int(1-\xi)[\![\sin(\beta W_\infty)]\!]\{\sin(\beta(I_\infty(v)+\mathcal{I}_\infty))-\sin(\beta(\mathcal{I}_\infty))\}\\
&\quad+\int_0^\infty\langle\overline{u},v_t\rangle_{L^2}dt-\frac{1}{2}\int_0^\infty\|v_t\|_{L^2}^2+\|\overline{u}^{\xi}\|_{L^2}^2-\|\overline{u}_t^{\xi}-v_t\|_{L^2}^2dt\Big].
\end{aligned}$$

For the first term, note that for any $\varphi,\psi\in L^2(\rho_{-\ell})$,

$$|g(\varphi-\psi)|=\left|\int\nabla g((1-\theta)\varphi+\theta\psi)(\varphi-\psi)d\theta\right|\le|\nabla g|_\ell\,\|\varphi-\psi\|_{L^2(\rho_{-\ell})}.$$

Hence,

$$\begin{aligned}
&\mathbf{E}\left|g(W_\infty+I_\infty(v)+\mathcal{I}_\infty)-g(W_\infty+I_\infty(v)+\mathcal{I}_\infty^{\xi})\right|\\
&\le\mathbf{E}\big[|\nabla g|_n^2\big]^{\frac{1}{2}}\,\mathbf{E}\Big[\|\mathcal{I}_\infty-\mathcal{I}_\infty^{\xi}\|_{L^2(\rho_{-n})}^2\Big]^{\frac{1}{2}},
\end{aligned}$$

which converges to 0 uniformly on $\mathcal{A}(g)$ by Corollary 4.47.

For the integrals involving the cosine, we know that $[\![\cos(\beta W_\infty)]\!]$ converges in $L^p(P;B^{-1+\delta}_{p,p}(\rho_{-p\ell}))$ for any $p\ge2$ and $\ell>1$ by Proposition 4.10. Moreover, by Proposition A.4 and Hölder's inequality with $\frac{1}{p}+\frac{1}{q}=1$

$$\begin{aligned}
&\mathbf{E}\int[\![\cos(\beta W_\infty)]\!]\{\cos(\beta(I_\infty(v)+\mathcal{I}_\infty))-\cos(\beta(I_\infty+\mathcal{I}_\infty^{\xi}))+\cos(\beta I_\infty^{\xi})-\cos(\beta\mathcal{I}_\infty)\}\\
&\le\mathbf{E}\Big[\|[\![\cos(\beta W_\infty)]\!]\|^p_{B^{-1+\delta}_{p,p}(\langle x\rangle^{-p\ell})}\Big]^{\frac{1}{p}}\\
&\quad\times\mathbf{E}\Big[\|\cos(\beta(I_\infty(v)+\mathcal{I}_\infty))-\cos(\beta(I_\infty+\mathcal{I}_\infty^{\xi}))+\cos(\beta I_\infty^{\xi})-\cos(\beta\mathcal{I}_\infty)\|^q_{B^{1-\delta}_{q,q}(\langle x\rangle^{q\ell})}\Big]^{\frac{1}{q}}\\
&=:\mathbf{E}\Big[\|[\![\cos(\beta W_\infty)]\!]\|^p_{B^{-1+\delta}_{p,p}(\langle x\rangle^{-p\ell})}\Big]^{\frac{1}{p}}\times(\mathrm{I}).
\end{aligned}$$

Since the first factor is bounded, this means we want to show convergence for (I) in $B^{1-\delta}_{1,1}(\rho_{\frac{q}{2}\ell})$ for some $q,\ell > 1$ to be chosen later. To this end, we interpolate between $B^{1-\frac{\delta}{2}}_{q,q} = W^{1-\frac{\delta}{2},q}$ and L^∞ using the Gagliardo-Nirenberg inequality, Lemma A.2,

$$\begin{aligned}
&\|\cos(\beta(I_\infty(v)+\mathcal{I}_\infty)) - \cos(\beta(I_\infty(v)+\mathcal{I}^\xi_\infty)) + \cos(\beta\mathcal{I}^\xi_\infty) - \cos(\beta\mathcal{I}_\infty)\|_{B^{1-\delta}_{q,q}(\langle x\rangle^{q\ell})} \\
\leq& C\int_0^1 \|\{\sin(\beta(tI_\infty(v)+\mathcal{I}_\infty)) - \sin(\beta(tI_\infty(v)-\mathcal{I}^\xi_\infty))\}I_\infty(v)\|^{\frac{1}{q}}_{B^{1-\frac{\delta}{2}}_{1,1}(\langle x\rangle^{q\ell})} dt \\
\leq& C\int_0^1 \|\{\sin(\beta(tI_\infty(v)+\mathcal{I}_\infty)) - \sin(\beta(tI_\infty(v)-\mathcal{I}^\xi_\infty))\}I_\infty(v)\|^{\frac{1}{q}}_{W^{1,1}(\langle x\rangle^{q\ell})} dt \\
=:& C \times (\mathrm{II}),
\end{aligned}$$

where $\frac{1}{q} = \frac{1-\delta}{1-\frac{\delta}{2}} < 1$. Here, we also used that $W^{s_2,p} \subset W^{s_1,p}$ for $s_1 < s_2$ and that $B^s_{q,q} \simeq W^{s,q}$ for $s \notin \mathbb{Z}$. Due to Lemma 4.8, the bound on the renormalisation $\alpha_t \leq \langle t\rangle^{\frac{1}{2}-\delta}$ and Lemma 4.42, the hypothesis

$$\|\mathcal{I}_\infty\|_{W^{1,\infty}} + \|\mathcal{I}^\xi_\infty\|_{W^{1,\infty}} \leq \|\overline{u}^\xi_t \langle t\rangle^{\frac{1}{2}+\delta}\|_{L^\infty(\mathbb{R}_+\times\mathbb{R}^2)} + \|\overline{u}_t \langle t\rangle^{\frac{1}{2}+\delta}\|_{L^\infty(\mathbb{R}_+\times\mathbb{R}^2)} < \infty,$$

of Lemma 4.53 is satisfied. Hence, for $\varepsilon \in (0,1)$ to be chosen later

$$\begin{aligned}
&\mathbf{E}[(\mathrm{II})^q] \\
\leq& \mathbf{E}\Big[\|\mathcal{I}_\infty - \mathcal{I}^\xi_\infty\|_{H^1(\rho_{-q\ell})}\|I_\infty(v)\|_{H^{1,}(\rho_{2q\ell})} + \|\mathcal{I}_\infty - \mathcal{I}^\xi_\infty\|^\varepsilon_{L^2(\rho_{-q\ell})}\|I_\infty(v)\|_{H^1(\rho_{2q\ell})}\Big] \\
\leq& \mathbf{E}\Big[\|\mathcal{I}_\infty - \mathcal{I}^\xi_\infty\|^2_{H^1(\rho_{-q\ell})}\Big]^{\frac{1}{2}} \mathbf{E}\Big[\|I_\infty(v)\|^2_{H^{1,}(\rho_{2q\ell})}\Big]^{\frac{1}{2}} \\
&\quad + \mathbf{E}\Big[\|\mathcal{I}_\infty - \mathcal{I}^\xi_\infty\|^{\varepsilon 2}_{L^2(\rho_{-q\ell})}\Big]^{\frac{1}{2}} \mathbf{E}\Big[\|I_\infty(v)\|^2_{H^1(\rho_{2q\ell})}\Big]^{\frac{1}{2}}.
\end{aligned}$$

For δ small enough, $q = \frac{1-\delta}{1-\frac{\delta}{2}}$ is sufficiently close to 1, we can choose $\ell := \frac{n}{2q} > 1$. Thus, since $v \in \mathcal{A}(g)$, we see that

$$\mathbf{E}\Big[\|I_\infty(v)\|^2_{H^1(\rho_n)}\Big]^{\frac{1}{2}},$$

is bounded by Lemma 4.5. For $\varepsilon \geq \frac{1}{2}$, we have $\frac{2\varepsilon}{q} \in (1,2)$ and again with $q\ell = \frac{n}{2} > 1$,

$$\mathbf{E}\Big[\|\mathcal{I}_\infty - \mathcal{I}^\xi_\infty\|^{2\varepsilon}_{L^2(\rho_{-q\ell})}\Big] \to 0, \text{ as } \xi \to 1,$$

by Corollary 4.47. Combined, this shows (uniform) convergence of (I). Of course, the same argument applies to the analogous term with the cosine replaced by the sine.

For the terms

$$\int (1-\xi)[\![\cos(\beta W_\infty)]\!](\cos(\beta(I_\infty(v)+\mathcal{I}_\infty)) - \cos(\beta\mathcal{I}_\infty)),$$

the convergence will follow from the weighted estimate

$$\|(1-\xi)f\|_{L^1(\langle x\rangle^{-\ell})} \leq \sup_{|x|\geq N} \langle x\rangle^{-\frac{\ell}{2}} \|f\|_{L^1(\langle x\rangle^{-\frac{\ell}{2}})} \leq N^{-\frac{\ell}{2}} \|f\|_{L^1(\langle x\rangle^{-\frac{\ell}{2}})},$$

for $\operatorname{supp}(1-\xi) \subset B_N^c(0)$. Indeed, the same arguments as before yield for $q = \frac{1-\delta}{1-\frac{\delta}{2}}$

$$\begin{aligned}
&\mathbf{E}\left[\int (1-\xi)[\![\cos(\beta W_\infty)]\!](\cos(\beta(I_\infty(v)+\mathcal{I}_\infty)) - \cos(\beta\mathcal{I}_\infty))\right] \\
\leq & C\,\mathbf{E}\left[\|[\![\cos(\beta W_\infty)]\!]\|^p_{B^{-1+\delta}_{p,p}(\langle x\rangle^{-p\ell})}\right]^{\frac{1}{p}} \\
&\times \mathbf{E}\left[\|(1-\xi)\cos(\beta(I_\infty(v)+\mathcal{I}_\infty)) - \cos(\beta\mathcal{I}_\infty)\|^q_{B^{1-\delta}_{q,q}(\langle x\rangle^{q\ell})}\right]^{\frac{1}{q}}.
\end{aligned}$$

The first factor is again bounded. For the second term, we estimate

$$\begin{aligned}
&\mathbf{E}\left[\|(1-\xi)\cos(\beta(I_\infty(v)+\mathcal{I}_\infty)) - \cos(\beta\mathcal{I}_\infty)\|^q_{B^{1-\delta}_{q,q}(\langle x\rangle^{q\ell})}\right]^{\frac{1}{q}} \\
\leq & \mathbf{E}\left[\int_0^1 \|(1-\xi)\sin(\beta(tI_\infty(v)+\mathcal{I}_\infty))I_\infty(v)\|_{W^{1,1}(\langle x\rangle^{q\ell})} dt\right]^{\frac{1}{q}}.
\end{aligned}$$

Now, this norm can be estimated similarly to Lemma 4.53,

$$\begin{aligned}
&\|(1-\xi)\sin(\beta(I_\infty(v)+\mathcal{I}_\infty))I_\infty(v)\|_{L^1(\langle x\rangle^{q\ell})} \\
\leq & \|(1-\xi)\sin(\beta(I_\infty(v)+\mathcal{I}_\infty))\|_{L^2(\langle x\rangle^{-q\ell})}\|I_\infty(v)\|_{L^2(\langle x\rangle^{2q\ell})} \\
\leq & N^{-q\ell}\|I_\infty(v)\|_{L^2(\langle x\rangle^{2q\ell})}.
\end{aligned}$$

It remains to estimate the derivative

$$\begin{aligned}
&\nabla(1-\xi)(\sin(\beta(I_\infty(v)+\mathcal{I}_\infty))I_\infty(v)) \\
=&(1-\nabla\xi)(\sin(\beta(I_\infty(v)+\mathcal{I}_\infty))I_\infty(v)) + (1-\xi)\sin(\beta(I_\infty(v)+\mathcal{I}_\infty))\nabla I_\infty(v) \\
&+(1-\xi)\cos(\beta(I_\infty(v)+\mathcal{I}_\infty))\beta(\nabla I_\infty(v)+\nabla\mathcal{I}_\infty)I_\infty(v).
\end{aligned}$$

The first term can be estimated in the same way as before

$$\|(1-\nabla\xi)(\sin(\beta(I_\infty(v)+\mathcal{I}_\infty))I_\infty(v))\|_{L^1(\langle x\rangle^{q\ell})} \leq N^{-4q\ell}\|I_\infty(v)\|_{L^2(\langle x\rangle^{2q\ell})}.$$

Thanks to the boundedness of the sine and cosine,

$$\begin{aligned}
&\|(1-\xi)\cos(\beta(I_\infty(v)+\mathcal{I}_\infty))\beta(\nabla I_\infty(v)+\nabla\mathcal{I}_\infty)I_\infty(v)\|_{L^1(\langle x\rangle^{q\ell})} \\
\leq & CN^{-2q\ell}\|(\nabla I_\infty(v)+\nabla\mathcal{I}_\infty)\|_{L^2(\langle x\rangle^{-2q\ell})}\|I_\infty(v)\|_{L^2(\langle x\rangle^{2q\ell})} \\
\leq & CN^{-2q\ell}\|(\nabla I_\infty(v)+\nabla\mathcal{I}_\infty)\|_{L^2(\langle x\rangle^{-2q\ell})}\|I_\infty(v)\|_{H^1(\langle x\rangle^{2q\ell})},
\end{aligned}$$

and

$$
\begin{aligned}
&\|(1-\xi)\sin(\beta(I_\infty(v)+\mathcal{I}_\infty))\nabla I_\infty(v)\|_{L^1(\langle x\rangle^{q\ell})}\\
\leq & N^{-4q\ell}\|\nabla I_\infty(v)\|_{L^2(\langle x\rangle^{2q\ell})}\\
\leq & N^{-4q\ell}\|\nabla I_\infty(v)\|_{H^1(\langle x\rangle^{2q\ell})}.
\end{aligned}
$$

We have already argued that all norms above are finite provided that $\ell = \frac{n}{2q} > 1$. For q sufficiently close to 1, as $\xi \to 1$ and consequently $N \to \infty$, also this term goes to 0. Since all bounds are uniform in $v \in \mathcal{A}(g)$, the convergence is again uniform.

To finish the proof, we note that the quadratic term can be reduced to

$$
\|v_t\|_{L^2}^2 + \|\overline{u}^\xi\|_{L^2}^2 - \|\overline{u}_t^\xi - v_t\|_{L^2}^2 = 2\langle \overline{u}_t^\xi, v_t\rangle_{L^2},
$$

and by Hölder's inequality

$$
\mathbf{E}\int_0^\infty \langle v_t, \overline{u}_t - \overline{u}_t^\xi\rangle_{L^2} dt \leq \mathbf{E}\left[\int_0^\infty \|v_t\|_{L^2(\rho_\ell)}^2 dt\right]^{\frac{1}{2}} \mathbf{E}\left[\int_0^\infty \|\overline{u}_t - \overline{u}_t^\xi\|_{L^2(\rho_{-\ell})}^2 dt\right]^{\frac{1}{2}}.
$$

This again converges to 0 uniformly on $\mathcal{A}(g)$ by Theorem 4.46, which concludes the proof. □

4.7 Non-Gaussianity

As a consequence of the estimates we derived earlier, we can closely follow [2] to show that the limit is not Gaussian.

Indeed, we know that for $\ell < -1$ the measure ν_{SG} is supported on $H^{-1}(\rho_\ell)$. Suppose toward a contradiction that ν_{SG} is a Gaussian measure on this Hilbert space. Then, we know from Lemma B.2 and Definition B.4 that there is a $m \in H^{-1}(\rho_\ell)$ and a Hilbert space $H_{CM}(\nu_{\mathrm{SG}}) \subset H^{-1}(\rho_\ell)$ such that for any $\psi \in H^{-1}(\rho_\ell)$,

$$
\log\int \exp(-\langle\varphi,\psi\rangle)\,\nu(d\varphi) = \frac{1}{2}\|\psi\|_{H_{CM}(\nu_{\mathrm{SG}})}^2 + \langle m,\psi\rangle_{H^{-1}(\rho_\ell)}.
$$

For $\psi \in C_c^\infty(\mathbb{R}^2)$, defining $g(\varphi) = \langle\psi,\varphi\rangle$, it follows that $\nabla g = \psi$ satisfies the assumptions of Theorem 4.52. From the variational description of ν_{SG}, we can then rewrite the left-hand side as the limit of the approximate measures $\nu_{\mathrm{SG}}^{\xi,T}$ to obtain

$$
\begin{aligned}
&\log\int \exp(-\langle\varphi,\psi\rangle)\,\nu_{\mathrm{SG}}(d\varphi)\\
=&\lim_{\substack{T\to\infty\\ \xi\to 1}} \log\left(\Xi_{T,\xi}^{-1}\int \exp\left(-\langle\psi, C_T\varphi\rangle_{H^{-1}(\rho_\ell)} - V_T^\xi(C_T\varphi)\right)\mu(d\varphi)\right).
\end{aligned}
$$

By the Cameron-Martin formula (Theorem B.5), the shifted Gaussian measure $\mu(d(\varphi - (-\Delta+m^2)^{-1}C_T\psi))$ is absolutely continuous with density

$$\mu(d\varphi) = \exp(\langle C_T\varphi,\psi\rangle_{H^{-1}(\rho_\ell)} - \tfrac{1}{2}\langle C_T\psi,(-\Delta+m^2)^{-1}C_T\psi\rangle) \\ \times \mu(d(\varphi-(-\Delta+m^2)^{-1}C_T\psi)).$$

Hence, applying this shift in the previous equation,

$$\begin{aligned}
&\log\int \exp\big(-\langle\psi,\varphi\rangle_{H^{-1}(\rho_\ell)}\big)\,\nu_{\mathrm{SG}}(d\varphi)\\
&= \lim_{\substack{T\to\infty\\ \xi\to 1}} \log\Big(\tfrac{1}{2}\langle C_T\psi,(-\Delta+m^2)^{-1}C_T\psi\rangle_{H^{-1}(\rho_\ell)}\\
&\qquad\times \Xi_{T,\xi}^{-1}\int \exp\Big(-V_T^\xi\big(C_T\varphi+(-\Delta+m^2)^{-1}C_T\psi\big)\Big)\mu(d\varphi)\Big)\\
&= \lim_{\substack{T\to\infty\\ \xi\to 1}} \tfrac{1}{2}\langle C_T\psi,(-\Delta+m^2)^{-1}C_T\psi\rangle + \mathcal{V}_T^\xi((-\Delta+m^2)\psi) - \mathcal{V}_T^\xi(0).
\end{aligned}$$

As the first part is quadratic in ψ, showing that $\varphi \mapsto \mathcal{V}_T^\xi(\varphi)$ for $\varphi \in H_{\mathrm{CM}}$ cannot be quadratic yields the desired contradiction. Equivalently, we can show that $\nabla\mathcal{V}_T^\xi$ is not linear. Recall that by definition of R,

$$\nabla\mathcal{V}_T^\xi = \nabla Y_{0,T}^\xi = \nabla V_0^\xi(X_{0,T}^\xi) + R_{0,T}^\xi, \tag{4.30}$$

with

$$\nabla V_0^\xi(X_{0,T}^\xi) = \lambda\alpha_0\sin(\beta\varphi).$$

It is clear that the sine is non-linear and to show that the property is preserved in the limit it is enough to find specific test functions in H_{CM}. For example, letting $\varphi,\psi \in C_c^\infty(\mathbb{R}^2)$ with $\mathbb{1}_{\{|x|<1\}}\,\varphi(x) = \frac{\pi}{2}$ and $\mathbb{1}_{\{|x|<1\}}\,\psi(x) = \frac{\pi}{4}$, we compute for any $x \in B_1(0)$,

$$\sin(\varphi(x)+\psi(x)) + \sin(\varphi(x)-\psi(x)) = \sin(\frac{3\pi}{4}) + \sin(\frac{\pi}{4}) = \sqrt{2} \neq 2 = 2\sin(\varphi(x)).$$

Thus, by (4.30) and Proposition 4.36,

$$(\nabla\mathcal{V}_T^\xi(\psi+\varphi) + \nabla\mathcal{V}_T^\xi(\psi-\varphi) - 2\nabla\mathcal{V}_T^\xi(\psi))(x) \geq \lambda(\sqrt{2}-2) - C\lambda^2 \geq \frac{\lambda}{2}(\sqrt{2}-2),$$

for λ small enough. But then by Theorem 4.52, the same estimate applies for the limit $T\to\infty$, $\xi\to 1$. In other words, we arrive at the desired contradiction once we establish that $C_c^\infty(\mathbb{R}^2) \subset H_{\mathrm{CM}(\nu_{\mathrm{SG}})}$. This inclusion is a consequence of the computations above and the estimates on $\mathcal{V}_T^\xi$. Indeed, for $\psi \in C_c^\infty(\mathbb{R}^2)$,

$$\begin{aligned}
\frac{1}{2}\|\psi\|^2_{H_{CM}(\nu_{\mathrm{SG}})} &\leq \liminf_{\substack{T\to\infty\\ \xi\to 1}} \log\int \exp\left(-\langle\psi,\varphi\rangle\right)\nu_{\mathrm{SG}}(d\varphi) - \langle m,\psi\rangle_{H^{-1}(\rho_\ell)}\\
&\leq \sup_{T,\xi}\sup_{\varphi\in L^2}\|\nabla\mathcal{V}_T^\xi(\varphi)\|_{L^\infty}\|(-\Delta+m^2)^{-1}C_T\psi\|_{L^1}\\
&\quad + \|m\|_{H^{-1}(\rho_\ell)}\|\psi\|_{H^{-1}(\rho_\ell)} + \frac{1}{2}\|C_T\psi\|^2_{H_{CM}(\mu)},
\end{aligned}$$

which is finite by Proposition 4.36 since $\nabla\mathcal{V}_T^\xi = \nabla Y_{0,T}^\xi$.

Bibliography

[1] Hajer Bahouri, Jean-Yves Chemin, and Raphaël Danchin. *Fourier analysis and nonlinear partial differential equations*. Vol. 343. Springer, 2011.

[2] Nikolay Barashkov. "A stochastic control approach to Sine Gordon EQFT". arXiv preprint arXiv:2203.06626. 2022.

[3] Nikolay Barashkov. "A variational Approach to Gibbs measures on function spaces". PhD thesis. Universität Bonn, 2021.

[4] Nikolay Barashkov and Massimiliano Gubinelli. "A variational method for Φ^4_3". In: *Duke Mathematical Journal* 169.17 (2020), pp. 3339–3415.

[5] Nikolay Barashkov and Massimiliano Gubinelli. "On the variational method for Euclidean quantum fields in infinite volume". arXiv preprint arXiv:2112.05562. 2021.

[6] Roland Bauerschmidt and Thierry Bodineau. "Log-Sobolev Inequality for the Continuum Sine-Gordon Model". In: *Communications on Pure and Applied Mathematics* 74.10 (2021), pp. 2064–2113.

[7] G Benfatto, G Gallavotti, and Francesco Nicolo. "On the massive sine-Gordon equation in the first few regions of collapse". In: *Communications in Mathematical Physics* 83.3 (1982), pp. 387–410.

[8] G Benfatto et al. "Some probabilistic techniques in field theory". In: *Communications in Mathematical Physics* 59.2 (1978), pp. 143–166.

[9] G Benfatto et al. "Ultraviolet stability in Euclidean scalar field theories". In: *Communications in Mathematical Physics* 71.2 (1980), pp. 95–130.

[10] Jean-Michel Bismut. "An introductory approach to duality in optimal stochastic control". In: *SIAM review* 20.1 (1978), pp. 62–78.

[11] Jean-Michel Bismut. "Conjugate convex functions in optimal stochastic control". In: *Journal of Mathematical Analysis and Applications* 44.2 (1973), pp. 384–404.

[12] Vladimir Igorevich Bogachev. *Gaussian measures*. 62. American Mathematical Soc., 1998.

[13] Michelle Boué and Paul Dupuis. "A variational representation for certain functionals of Brownian motion". In: *The Annals of Probability* 26.4 (1998), pp. 1641–1659.

[14] Haïm Brezis and Petru Mironescu. "Gagliardo–Nirenberg inequalities and non-inequalities: The full story". In: *Annales de l'Institut Henri Poincaré C, Analyse non linéaire*. Vol. 35. 5. Elsevier. 2018, pp. 1355–1376.

[15] Philippe Briand and Fulvia Confortola. "Quadratic BSDEs with random terminal time and elliptic PDEs in infinite dimension". In: *Electronic Journal of Probability* 13 (2008), pp. 1529–1561.

[16] David C Brydges, Jürg Fröhlich, and Alan D Sokal. "A new proof of the existence and nontriviality of the continuum ϕ 2 4 and ϕ 3 4 quantum field theories". In: *Communications in mathematical physics* 91.2 (1983), pp. 141–186.

[17] David C Brydges and T Kennedy. "Mayer expansions and the Hamilton-Jacobi equation". In: *Journal of Statistical Physics* 48.1 (1987), pp. 19–49.

[18] Samuel N Cohen and Robert J Elliott. "Existence, uniqueness and comparisons for BSDEs in general spaces". In: *The Annals of Probability* 40.5 (2012), pp. 2264–2297.

[19] Ruth F Curtain and Peter L Falb. "Ito's lemma in infinite dimensions". In: *Journal of mathematical analysis and applications* 31.2 (1970), pp. 434–448.

[20] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic equations in infinite dimensions*. Cambridge university press, 2014.

[21] Poul H Damgaard and Helmuth Hüffel. "Stochastic quantization". In: *Physics Reports* 152.5-6 (1987), pp. 227–398.

[22] Richard WR Darling and Etienne Pardoux. "Backwards SDE with random terminal time and applications to semilinear elliptic PDE". In: *The Annals of Probability* 25.3 (1997), pp. 1135–1159.

[23] Mark HA Davis and Pravin Varaiya. "The multiplicity of an increasing family of σ-fields". In: *The Annals of Probability* (1974), pp. 958–963.

[24] J Dimock and TR Hurd. "Construction of the two-dimensional Sine-Gordon model for $\beta < 8\pi$". In: *Communications in mathematical physics* 156.3 (1993), pp. 547–580.

[25] Jonathan Dimock and TR Hurd. "Sine-Gordon revisited". In: *Annales Henri Poincaré*. Vol. 1. 3. Springer. 2000, pp. 499–541.

[26] Nicole El Karoui and SJ Huang. "backward stochastic differential equations". In: *Pitman Research Notes in Mathematics Series* (1997).

[27] Nicole El Karoui, Shige Peng, and Marie Claire Quenez. "Backward stochastic differential equations in finance". In: *Mathematical finance* 7.1 (1997), pp. 1–71.

[28] Giorgio Fabbri, Fausto Gozzi, and Andrzej Świkech. *Stochastic Optimal Control in Infinite Dimension*. Springer International Publishing, 2017.

[29] Joel S Feldman and Konrad Osterwalder. "The Wightman axioms and the mass gap for weakly coupled $(\phi^4)_3$ quantum field theories". In: *Annals of Physics* 97.1 (1976), pp. 80–135.

[30] James Glimm and Arthur Jaffe. "Positivity of the φ_3^4 Hamiltonian". In: *Fort-schritte der Physik* 21.7 (1973), pp. 327–376.

[31] James Glimm and Arthur Jaffe. *Quantum physics: a functional integral point of view*. Springer Science & Business Media, 2012.

[32] Massimiliano Gubinelli and Martina Hofmanová. "A PDE Construction of the Euclidean Φ_3^4 Quantum Field Theory". In: *Communications in Mathematical Physics* 384.1 (2021), pp. 1–75.

[33] Massimiliano Gubinelli, Peter Imkeller, and Nicolas Perkowski. "Paracontrolled distributions and singular PDEs". In: *Forum Math. Pi* 3 (2015), p. 75.

[34] Martin Hairer. "A theory of regularity structures". In: *Invent. Math.* 198.2 (2014), pp. 269–504.

[35] Yuu Hariya and Sou Watanabe. "The Boué–Dupuis formula and the exponential hypercontractivity in the Gaussian space". In: *arXiv preprint-arXiv:2110.14852* (2021).

[36] Nobuyuki Ikeda and Shinzo Watanabe. "A comparison theorem for solutions of stochastic differential equations and its applications". In: *Osaka Journal of Mathematics* 14.3 (1977), pp. 619–633.

[37] Svante Janson. *Gaussian Hilbert Spaces*. 129. Cambridge university press, 1997.

[38] Janne Junnila, Eero Saksman, and Christian Webb. "Imaginary multiplicative chaos: Moments, regularity and connections to the Ising model". In: *The Annals of Applied Probability* 30.5 (2020), pp. 2099–2164.

[39] Ioannis Karatzas and Steven Shreve. *Brownian Motion and Stochastic Calculus*. Springer New York, 1998.

[40] Magdalena Kobylanski. “Backward stochastic differential equations and partial differential equations with quadratic growth”. In: *the Annals of Probability* 28.2 (2000), pp. 558–602.

[41] Christoph Kopper. “Renormalization theory based on flow equations”. In: *Rigorous quantum field theory*. Springer, 2007, pp. 161–174.

[42] Raphael Kruse. “A Short Review of the Malliavin Calculus in Hilbert Spaces”. In: *Strong and Weak Approximation of Semilinear Stochastic Evolution Equations* (2014), pp. 85–108.

[43] Hubert Lacoin, Rémi Rhodes, and Vincent Vargas. “A probabilistic approach of ultraviolet renormalisation in the boundary Sine-Gordon model”. In: *arXiv preprint arXiv:1903.01394* (2019).

[44] Jean-Pierre Lepeltier and Jaime San Martin. “Backward stochastic differential equations with continuous coefficient”. In: *Statistics & Probability Letters* 32.4 (1997), pp. 425–430.

[45] Wei Liu and Michael Röckner. *Stochastic partial differential equations: an introduction*. Springer, 2015.

[46] Jin Ma et al. “On well-posedness of forward–backward SDEs—A unified approach”. In: *The Annals of Applied Probability* 25.4 (2015), pp. 2168–2214.

[47] J Magnen and Roland Seneor. “The infinite volume limit of the φ_3^4 model”. In: *Annales de l’IHP Physique théorique*. Vol. 24. 2. 1976, pp. 95–159.

[48] Carlo Marinelli and Michael Röckner. “On the maximal inequalities of Burk-holder, Davis and Gundy”. In: *Expositiones Mathematicae* 34.1 (2016), pp. 1–26.

[49] Edward Nelson. “A quartic interaction in two dimensions”. In: *Mathematical Theory of Elementary Particles, Proc. Conf., Dedham, Mass., 1965*. MIT press. 1966, pp. 69–73.

[50] Edward Nelson. “The free Markoff field”. In: *Journal of Functional Analysis* 12.2 (1973), pp. 211–227.

[51] F Nicolo, J Renn, and A Steinmann. “On the massive sine-Gordon equation in all regions of collapse”. In: *Communications in mathematical physics* 105.2 (1986), pp. 291–326.

[52] David Nualart. *The Malliavin calculus and related topics*. Vol. 1995. Springer, 2006.

[53] Konrad Osterwalder and Robert Schrader. “Axioms for Euclidean Green’s functions”. In: *Communications in mathematical physics* 31.2 (1973).

[54] Konrad Osterwalder and Robert Schrader. “Axioms for Euclidean Green’s functions II”. In: *Communications in Mathematical Physics* 42.3 (1975).

[55] Étienne Pardoux. “Backward Stochastic Differential Equations and Viscosity Solutions of Systems of Semilinear Parabolic and Elliptic PDEs of Second Order”. In: *Stochastic Analysis and Related Topics VI*. Birkhäuser Boston, 1998, pp. 79–127.

[56] Étienne Pardoux and Shige Peng. “Adapted solution of a backward stochastic differential equation”. In: *Systems & Control Letters* 14.1 (1990), pp. 55–61.

[57] Étienne Pardoux and Shige Peng. “Backward stochastic differential equations and quasilinear parabolic partial differential equations”. In: *Stochastic Partial Differential Equations and Their Applications*. Springer-Verlag, 1992, pp. 200–217.

[58] Georgio Parisi, Yong Shi Wu, et al. “Perturbation theory without gauge fixing”. In: *Sci. Sin* 24.4 (1981), pp. 483–496.

[59] Triet Pham and Jianfeng Zhang. “Some norm estimates for semimartingales”. In: *Electronic Journal of Probability* 18 (2013), pp. 1–25.

[60] Joseph Polchinski. “Renormalization and effective Lagrangians”. In: *Nuclear Physics B* 231.2 (1984), pp. 269–295.

[61] Giuseppe Da Prato and Arnaud Debussche. “Strong solutions to the stochastic quantization equations”. In: *The Annals of Probability* 31.4 (2003), pp. 1900–1916.

[62] Ali Süleyman Üstünel. “Variational calculation of Laplace transforms via entropy on Wiener space and applications”. In: *Journal of Functional Analysis* 267.8 (2014), pp. 3058–3083.

[63] Kenneth G Wilson. “Renormalization group and critical phenomena. I. Renormalization group and the Kadanoff scaling picture”. In: *Physical review B* 4.9 (1971), p. 3174.

[64] Jiongmin Yong and Xun Yu Zhou. *Stochastic controls: Hamiltonian systems and HJB equations*. Vol. 43. Springer Science & Business Media, 1999.

[65] Jianfeng Zhang. *Backward Stochastic Differential Equations*. Springer New York, 2017.

A. Analytic Tools

We briefly introduce some tools from Fourier analysis, which we use mainly to obtain convergence for the adjusted potential and some bounds on the kernel. For the sake of brevity, proofs are omitted, but full proofs for the statements can be found in [1, Chapter 2] and [33].

Definition A.1. For $f \in \mathscr{S}(\mathbb{R}^d)$, we define its Fourier transform

$$\mathscr{F}(f)(\xi) = \hat{f}(\xi) = \int_{\mathbb{R}^d} e^{i\langle x,\xi\rangle} f(x)dx,$$

with inverse

$$\mathscr{F}^{-1}(\hat{f})(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i\langle x,\xi\rangle} \hat{f}(\xi)d\xi.$$

With this scaling, $\|\mathscr{F}(f)\|_{L^2}^2 = (2\pi)^d \|f\|_{L^2}^2$.

A.1 Littlewood-Paley decomposition

Motivated the observation that derivatives act almost like dilations in the L^p-norms on distributions whose Fourier transform is supported on an annulus, the Littlewood-Paley decomposition is a dyadic localization in the Fourier-space for tempered distributions on $\mathbb{R}^d$.

This dilation property is captured by estimates like the Bernstein lemma (see e.g. [1, Lemma 2.1]). To leverage this fact also for distributions that are not localized in the frequency space, we introduce a dyadic partition of unity. Let $\chi, \tilde{\varphi} : \mathbb{R}^d \to \mathbb{R}_+$ be smooth and radial functions such that for some $R > 0$,

- $\mathrm{supp}(\chi) \subset B(0,R)$ and $\mathrm{supp}(\tilde{\varphi}) \subset B(0,2R) \setminus B(0,R)$,
- $\tilde{\varphi} \leq 1$ and $\chi(\xi) + \sum_{i\geq 0} \tilde{\varphi}(2^{-i}\xi) = 1$ for any $\xi \in \mathbb{R}^d$.
- $\mathrm{supp}(2^{-i}) \cap \mathrm{supp}(2^{-j}) = \emptyset$ for $|i-j| > 1$.

Then, denoting $\tilde{\varphi}_{-1} = \chi$, the Littlewood-Paley blocks are defined as

$$\Delta_i = \tilde{\varphi}_i(D).$$

For a tempered distribution $u \in \mathscr{S}'(\mathbb{R}^d)$, we define the decomposition

$$u = \sum_{i\geq -1} \Delta_i u.$$

While u is only a distribution in general, $\Delta_j u$ has compact support in Fourier space and thus is a function. Denoting by $\varphi_i = \mathscr{F}^{-1}(\tilde{\varphi}_i)$ the kernel of the Paley-Littlewood-block, we see that

$$\varphi_i(x) = 2^{id}\varphi_0(2^i x) \text{ and } \Delta_i u = \varphi_i * u.$$

Thus, $\|\varphi_i\|_{L^1(\langle x\rangle^{-n})} \leq C$ uniformly in i and

$$\|\varphi_i\|^p_{L^p(\langle x\rangle^{-n})} \leq C 2^{id(p-1)}. \tag{A.1}$$

With this notation, we define the norms

$$\|u\|^q_{B^s_{p,q}(w)} = \sum_{j\geq -1} 2^{sjq}\|\Delta_j u\|^q_{L^p(w)},$$

with the usual modifications for $p, q = \infty$ and $B^s_{p,q}(\mathbb{R}^d) := \{u \in \mathscr{S}'(\mathbb{R}^d) : \|u\|_{B^s_{p,q}} < \infty\}$. One can show that the spaces defined in this way are Banach spaces and that we can identify $H^s \simeq B^s_{2,2}$ for any $s \in \mathbb{R}$. For the cases $p \neq 2$, we still have $B^s_{p,p} \simeq W^{s,p}$, if $s > 0$ is not an integer. Another reason Besov spaces are very convenient is that they are the natural interpolation spaces for $W^{s,p}$.

Lemma A.2 (Gagliardo-Nirenberg inequality). *Let $s_1 \leq s \leq s_2$ and $1 \leq p_1, p_2, p \leq \infty$ with*

$$s = \theta s_1 + (1-\theta)s_2 \quad \frac{1}{p} = \frac{\theta}{p_1} + \frac{1-\theta}{p_2}, \quad \theta \in (0,1).$$

For any $f \in B^{s_1}_{p_1,p_1}(w_1) \cap B^{s_2}_{p_2,p_2}(w_2)$ it holds that

$$\|f\|_{B^s_{p,p}(w)} \leq C\|f\|^\theta_{B^{s_1}_{p_1,p_1}(w_1)}\|f\|^{1-\theta}_{B^{s_2}_{p_2,p_2}(w_2)},$$

where $w = w_1^{\frac{p}{p_1}} w_2^{\frac{p}{p_2}}$. The result stays true for the Sobolev spaces $W^{s,p}$ unless $p_2 = 1, s_2 \geq 1$ is an integer and $s_1 - \frac{1}{p_1} \geq s_2 - \frac{1}{p_2}$.

Proof. See [14] Proposition 5.7 for the Besov space version and Theorem 1 for the Sobolev case. □

Bony paraproduct

For two tempered distributions u, v, we may formally decompose their product as

$$uf = \sum_{i,j\geq -1} \Delta_i u \Delta_j f = T_u f + R(u,f) + T_f u,$$

where with $S_i g = \sum_{i'<i} \Delta_{i'}$ we define

$$T_u f = \sum_{i\geq -1} S_i u \Delta_i f \text{ and } R(u,f) = \sum_{i\geq -1}\sum_{|i-i'|\leq 1} \Delta_i u \Delta_{i'} f.$$

The *paraproducts* $T_u v$ (respectively $T_v u$) are locally finite in the Fourier space and as such always well-defined with regularity no better than u (resprectively v). The *resonant* term $R(u,v)$ is not always well-defined, unless the sum of the regularities of u and v is positive.

Theorem A.3. *[1, Theorem 2.85] Let $f \in B^{s_1}_{p_1,q_1}(w_1)$, $u \in B^{s_2}_{p_2,q_2}(w_2)$ for some $s_2 + s_1 > 0$ with $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} \leq 1$, $\frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2} \leq 1$ and $w = w_1^{\frac{p}{p_1}} w_2^{\frac{p}{p_2}}$. Then, the product*

$$fu = T_u f + R(u,f) + T_f u,$$

is a well-defined, continuous bilinear map with

$$\|fu\|_{B^{s_1 \wedge s_2}_{p,q}(w)} \lesssim \|f\|_{B^{s_1}_{p_1,q_1}(w_1)} \|u\|_{B^{s_2}_{p_2,q_2}(w_2)}.$$

Proposition A.4. *[1, Proposition 2.76] For $(p,r) \in ([1,\infty]^2)$ and $s \in \mathbb{R}$, the map*

$$\langle \cdot, \cdot \rangle : B^s_{p,r} \times B^{-s}_{p',r'} \to \mathbb{R},\ (f,u) \mapsto \sum_{|j-j'|\leq 1} \langle \Delta_j f, \Delta_j u \rangle,$$

is continuous and bilinear with

$$|\langle f,u \rangle| \leq C \|f\|_{B^s_{p_1,q_1}(w_1)} \|u\|_{B^{-s}_{p_2,q_2}(w_2)}.$$

Here, $\langle \Delta_j f, \Delta_j u \rangle$ is the dual pairing of $(\mathscr{S}'(\mathbb{R}^d), \mathscr{S}(\mathbb{R}^d))$, the weights w_1, w_2 are conjugate weights (i.e. $1 = w = w_1^{\frac{1}{p_1}} w_2^{\frac{1}{p_2}}$), and $\frac{1}{q_1} + \frac{1}{q_2} = 1$, $\frac{1}{p_1} + \frac{1}{p_2} = 1$ are Hölder conjugates.

B. Gaussian measures

We recall some important basic properties of Gaussian measures, mostly without proof. For full proofs of the statements and additional background, we refer to [12] as a general reference and to [37] for a more probabilistic and combinatorial point of view.

Definition B.1. Let $\mathcal{X}$ be a locally convex space equipped with a Radon measure μ. We say that μ is Gaussian if for any finite family $x_1^*, x_2^*, \dots, x_k^* \in \mathcal{X}^*$, the push forward $\mu_{(x_1^*, x_2^*, \dots, x_k^*)}$ is a Gaussian measure on $\mathbb{R}^k$.

The restriction on Radon measures is mainly to ensure measures are uniquely determined by their finite-dimensional distribution. For our purposes, the following characterisation of Gaussian measures will be useful later. Recall that the Fourier transform of a Radon measure μ is defined via duality as

$$\mathscr{F}(\mu)(x^*) = \int_{\mathcal{X}} \exp(ix^*(x))\mu(dx), \quad \text{for } x^* \in \mathcal{X}^*.$$

Lemma B.2. *A measure μ on a locally convex space $\mathcal{X}$ is Gaussian if and only if its Fourier transform is of the form*

$$\mathscr{F}(\mu)(x^*) = \exp\left(im(x^*) - \frac{1}{2}B(x^*, x^*)\right),$$

where $m : \mathcal{X}^* \to \mathbb{R}$ is a continuous functional and B is a symmetric, continuous and nonnegative bilinear form on $\mathcal{X}^*$.

Proof. See e.g. [12, Theorem 2.2.4] □

In the notation of Lemma B.2, we call m the *mean* and B the *covariance* of the Gaussian measure. If $H = \mathcal{X}$ is also a Hilbert space, we can identify $m \in H^{**} \simeq H$ thanks to reflexivity. For $m = 0$, we also say that the Gaussian is *centred*. Furthermore, in the Hilbert space setting, the bilinear form may be written as $B(x^*, x^*) = \langle S^* x^*, x^* \rangle$ for a self-adjoint, nonnegative, trace-class operator S^* on H^*. By the Riesz representation theorem, we find a unique self-adjoint, nonnegative trace-class operator S on H such that

$$B(x^*, y^*) = \langle Sx, y \rangle, \quad x, y \in H,$$

where $x^* = \langle x, \cdot \rangle \in H^*$. In this setting, we also have the following existence result.

Lemma B.3. *There is a Gaussian measure on H with covariance S and mean m if and only if S is a positive self-adjoint trace-class operator.*

Proof. See e.g. [12, Theorem 2.3.1] □

The central Gaussian measure for our purposes is the Gaussian free field with mass $m > 0$, the centred Gaussian measure with covariance $(m^2 - \Delta)^{-1}$. More precisely, we say $\Psi_m : \Omega \to \mathscr{S}'(\mathbb{R}^d)$ is a realisation of the Gaussian field with mass $m > 0$, if

$$\mathbf{E}[\langle \Psi_m, f\rangle \langle \Psi_m, g\rangle] = \langle (-\Delta + m^2)^{-1} f, g\rangle_{L^2(\mathbb{R}^d)} = \int_{\mathbb{R}^d} \frac{\mathscr{F}(f)\overline{\mathscr{F}(g)}(k)}{|k|^2 + m^2} dk.$$

Definition B.4 (Cameron Martin space). For a centred Gaussian measure μ on a Hilbert space H with covariance S. We call

$$H_{\mathrm{CM}}(\mu) := \mathrm{range}(S^{\frac{1}{2}}) \subset H,$$

the *Cameron-Martin space of* μ. Equipped with the inner product

$$\langle x, y\rangle_{H_{CM}} := \langle S^{-\frac{1}{2}} x, S^{-\frac{1}{2}} y\rangle_H,$$

the space is a Hilbert space and $x \in H_{CM}(\mu)$ if and only if $\|x\|_{H_{CM}(\mu)} < \infty$.

Theorem B.5 (Cameron Martin Theorem). *Let μ be a Gaussian measure on H with covariance S. For any $h \in H$, the measure $\mu_h := \mu(\cdot - S^{\frac{1}{2}} h)$ is equivalent to μ with Radon-Nikodym derivative*

$$\rho_h(x) = \exp(\langle h, x\rangle + \frac{1}{2}\|S^{\frac{1}{2}} h\|^2_{H_{CM}}).$$

Proof. See e.g. [12, Corollary 2.4.3] □

As in the finite-dimensional case, the convergence of Gaussian measures comes down to the convergence of their covariances.

Lemma B.6. *A sequence of centred Gaussian measures μ_n with covariance C_n converges weakly to a centred Gaussian measure μ with covariance C if and only if $\sqrt{C_n}$ converges to $\sqrt{C}$ in the Hilbert-Schmidt norm.*

Proof. See e.g. [12, Proposition 3.8.12]. □