An introduction to
singular stochastic PDEs:
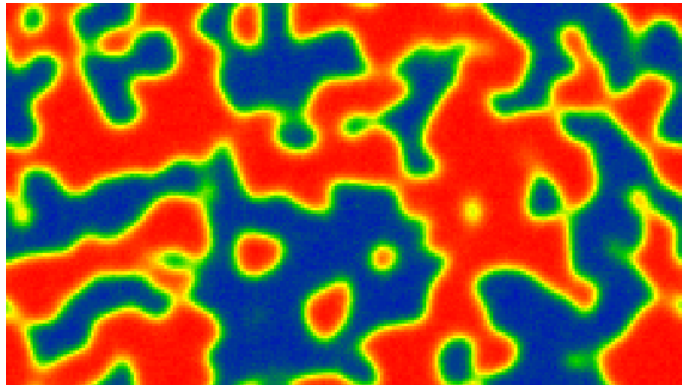# Allen–Cahn equations, metastability
and regularity structures

Nils Berglund

Institut Denis Poisson – UMR 7013

Université d'Orléans, Université de Tours, CNRS

Lecture notes

Sarajevo Stochastic Analysis Winter School 2019

— Version of September 12, 2019 —

Contents

Preface

These notes have been prepared for a series of lectures given at the Sarajevo Stochastic Analysis Winter School, from January 28 to February 1, 2019. There already exist several excellent lecture notes and reviews on the subject, such as [Hai09] on (non-singular) stochastic PDEs, and [Hai15, CW17] on singular stochastic PDEs and regularity structures. The present notes have two main specificities. The first one is that they focus on a particular example, the Allen–Cahn equation, which allows to introduce several of the difficulties of the theory in a gradual way, by increasing the space dimension step by step. The hope is that while this limits the generality of the theory presented, this limitation is more that made up by a gain in clarity. The second specific aspect of these notes is that they go beyond existence and uniqueness of solutions, by covering a few recent results on convergence to equilibrium and metastability in these system.

The author wishes to thank the organisers of the Winter School, Frank Proske, Abdol-Reza Mansouri and Oussama Amine, for the invitation that provided the motivation to compile these notes; his co-workers on SPDEs, Ajay Chandra, Giacomo Di Gesù, Barbara Gentz, Christian Kuehn and Hendrik Weber, who greatly contributed to some results presented here; Yvain Bruned for remarks on an early version of the notes; and Martin Hairer and Lorenzo Zambotti for patiently explaining some subtleties of the theory of regularity structures.

Thanks are also due to the Isaac Newton Institute in Cambridge, as well as David Brydges, Alessandro Giuliani, Massimiliano Gubinelli, Antti Kupiainen, Hendrik Weber and Lorenzo Zambotti for organising the semester "Scaling limits, rough paths, quantum field theory". In particular the tutorial lectures given during the opening workshop of the semester proved of great help in preparing these notes.

Finally, thanks are due to Adrian Martini, for pointing out some typos in the original version of these notes.

Chapter 1

A system of interacting diffusions

In this chapter, we will analyse the dynamics of a system of $N \geqslant 2$ coupled stochastic differential equations (SDEs), which will converge, in a certain sense, to a stochastic Allen–Cahn PDE. This will serve two purposes: firstly, it will introduce a natural physical model that motivates the use of stochastic PDEs, and secondly, it will allow us to recapitulate some notions from the theory of SDEs that will be important in the infinite-dimensional case.

The system of interacting diffusions is given by

$$\mathrm{d}y_t^i = [y_t^i - (y_t^i)^3]\,\mathrm{d}t + \frac{\gamma}{2}[y_t^{i+1} - 2y_t^i + y_t^{i-1}]\,\mathrm{d}t + \sqrt{2\varepsilon}\,\mathrm{d}W_t^i, \qquad i = 1,\dots,N, \tag{1.0.1}$$

in the following setting.

- The real variables $y^1,\dots,y^N \in \mathbb{R}$ can be thought of as the positions of N particles. We use periodic boundary conditions, that is, $y^0 = y^N$ and $y^{N+1} = y^1$. This can also be indicated by considering i as an element of the cyclic group $\Lambda = \mathbb{Z}/N\mathbb{Z}$.
- The term $[y_t^i - (y_t^i)^3]\,\mathrm{d}t$ tends to push each particle towards $+1$ if $y_t^i > 0$ and towards -1 if $y_t^i < 0$, that is, we have a bistable local dynamics.
- The term $\frac{\gamma}{2}[y_t^{i+1} - 2y_t^i + y_t^{i-1}]\,\mathrm{d}t$ is a discretised Laplacian interaction, which describes a ferromagnetic nearest-neighbour coupling pushing the y_t^i towards each other (if $\gamma > 0$).
- The W_t^i are independent standard Wiener processes on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, modelling thermal noise.

The system (1.0.1) can be thought of as a version of the Ising model with continuous spins. We are mainly interested in its long-time dynamics for large values of N and γ. The parameter $\varepsilon \geqslant 0$ will be either small or of order 1.

1.1 Deterministic dynamics

In the deterministic case $\varepsilon = 0$, the system (1.0.1) reduces to the system of ordinary differential equations (ODEs)

$$\dot{y}^i = y^i - (y^i)^3 + \frac{\gamma}{2}[y^{i+1} - 2y^i + y^{i-1}]. \tag{1.1.1}$$

A first observation is that (1.1.1) has gradient form

$$\dot{y} = -\nabla V(y),$$

where V is the quartic potential given by

$$V(y) := \sum_{i\in\Lambda} U(y^i) + \frac{\gamma}{4}\sum_{i\in\Lambda}(y^{i+1} - y^i)^2, \qquad U(\xi) = \frac{1}{4}(\xi^2 - 1)^2. \tag{1.1.2}$$

The stationary points of (1.1.1) are thus the critical points of V. Note that V is invariant under the group G generated by cyclic permutations of the y^i, the reflection $y^i \mapsto y^{N+1-i}$, and the point symmetry $y \mapsto -y$.

The simplest situation arises for $\gamma = 0$: then the set of all critical points of V is $\{-1, 0, 1\}^\Lambda$ and has cardinality 3^N. The critical points in $\{-1, 1\}^\Lambda$ are all local minima of V, and thus describe 2^N stable stationary points of (1.1.1). In fact, one can show [BFG07a, Proposition 2.1] that this situation perturbs to small positive γ, in the sense that there are still 3^N stationary points, 2^N of which are stable, for all γ up to a critical value $\gamma_0(N)$ which is larger than $\frac{1}{4}$ for all N.

We are more interested, however, in the case of large γ of the order N^2. Indeed, this is the natural scaling in which the discrete Laplacian in (1.1.1) approaches the continuous Laplacian.

Proposition 1.1.1. *Let*

$$\gamma_1(N) := \frac{1}{2\sin^2(\pi/N)} = \frac{N^2}{2\pi^2}\Big[1 + \mathcal{O}\Big(\frac{1}{N^2}\Big)\Big].$$

Then the potential V admits exactly 3 stationary points 0 and $\pm(1,\ldots,1)$ if and only if $\gamma \geqslant \gamma_1(N)$. The points $\pm(1,\ldots,1)$ are always local minima of V, while 0 is a saddle with exactly one linearly unstable direction if and only if $\gamma > \gamma_1(N)$.

PROOF: First note that the discrete Laplacian is represented by a Toeplitz matrix (its entries are constant on diagonals), whose eigenvalues can be computed explicitly by discrete Fourier transform. They are given by $-\lambda_k^N$ where

$$\lambda_k^N = \lambda_{N-k}^N := 2\sin^2\Big(\frac{k\pi}{N}\Big), \qquad k = 0,\ldots,N-1 .$$

In particular, $\lambda_0^N = 0$ and $\gamma_1(N) = 1/(\lambda_1^N - \lambda_0^N)$ is the inverse of the spectral gap. Consider now the function

$$W(y) := \frac{1}{2}\sum_{i\in\Lambda}(y^{i+1} - y^i)^2 .$$

Note that W vanishes on the diagonal $\{y^1 = \cdots = y^N\}$ and is otherwise positive. An easy computation (cf. [BFG07a, Proposition 3.1]) shows that

$$\frac{\mathrm{d}}{\mathrm{d}t}W(y_t) \leqslant 2\Big(1 - \frac{\gamma}{\gamma_1(N)}\Big)W(y_t) - \frac{1}{N}W(y_t)^2 .$$

Thus when $\gamma \geqslant \gamma_1(N)$, $t \mapsto W(y_t)$ decreases for y_t outside the diagonal, i.e., W is a Lyapunov function. Therefore, when $\gamma \geqslant \gamma_1(N)$ all orbits of (1.1.1) will converge to the diagonal. The only stationary points on the diagonal are 0 and $\pm(1,\ldots,1)$. The eigenvalues of the linearisation of (1.1.1) at 0 are given by $-\mu_k^N$ where

$$\mu_k^N := -1 + \gamma\lambda_k^N , \qquad k = 0,\ldots,N-1 . \tag{1.1.3}$$

If $\gamma > \gamma_1(N)$, then only $-\mu_0^N$ is positive, while for $\gamma < \gamma_1(N)$, there are at least three positive eigenvalues (two if $N = 2$). Since $V(y)$ is bounded below and goes to infinity as $\|y\| \to \infty$, there must be critical points outside the diagonal in the latter case. Finally, the eigenvalues of the linearisation at $\pm(1,\ldots,1)$ are given by $-v_k^N$ where

$$v_k^N := 2 + \gamma\lambda_k^N , \qquad k = 0,\ldots,N-1 . \tag{1.1.4}$$

These are always negative, showing that these points are always stable (they are local minima of V). $\square$

Exercise 1.1.2. Show that the potential satisfies the lower bound

$$V(y_0 + y_\perp) \geqslant V(y_0) + \frac{1}{2}\Big(\frac{\gamma}{\gamma_1(N)} - 1\Big)\|y_\perp\|^2 \tag{1.1.5}$$

whenever y_0 belongs to the diagonal $\{y^1 = \cdots = y^N\}$ and $y_\perp$ is orthogonal to the diagonal. ♣

This result shows that if $\gamma > \gamma_1(N)$, then V is a double-well potential, with a saddle having one unstable direction located at the origin, and the two minima located at $\pm(1,\ldots,1)$. The unstable manifolds of the saddle are subsets of the diagonal, and connect 0 to the two local minima. This is the case we will mainly be interested in in what follows.
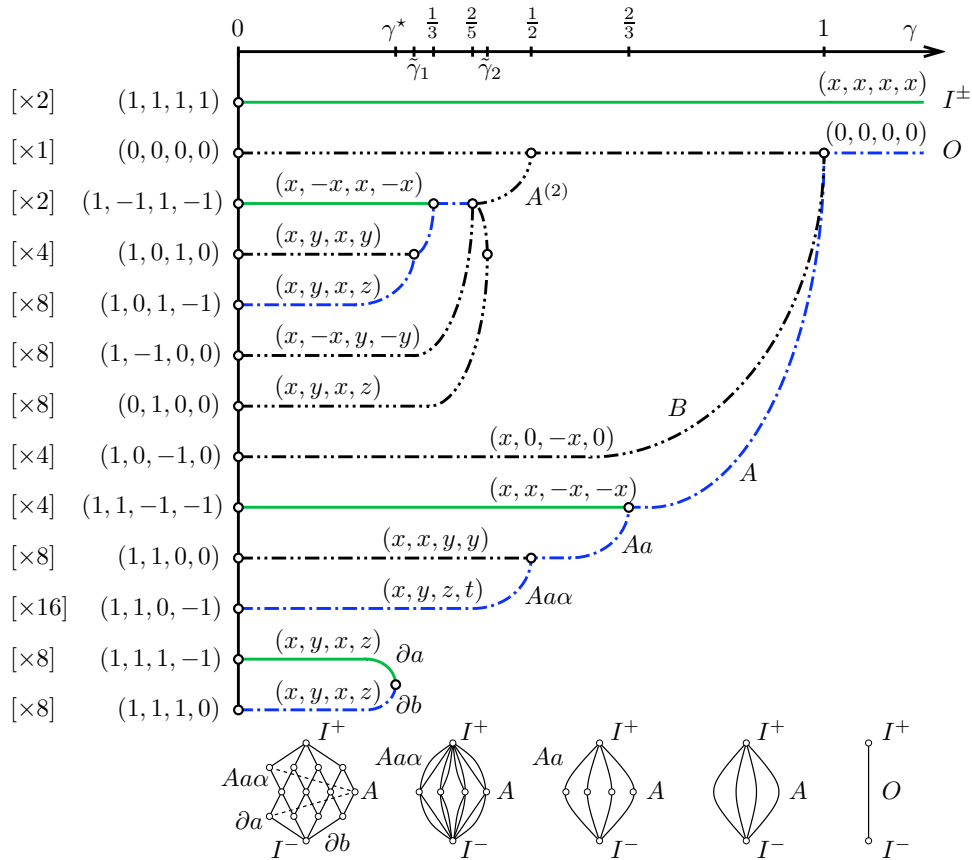
Figure 1.1 – Schematic bifurcation diagram for (1.1.1) in the case $N = 4$. Each branch in this diagram represents a family of stationary points globally invariant under the symmetry group G, with only one representative of the family displayed on the left side of the diagram. Green full lines represent stable stationary points (local minima of the potential V), while broken lines with k dots represent stationary points with k unstable directions (saddles of Morse index k of V).

As γ decreases below $\gamma_1(N)$, the situation becomes much more involved, and is not understood in full detail. What is known is that at $\gamma_1(N)$, the system undergoes a G-symmetric analogue of a pitchfork bifurcation, at which the origin becomes unstable in more than one direction, while expelling a multiple of N new stationary points with lower symmetry. As γ decreases further, these points undergo further symmetry-breaking bifurcations, as illustrated by Figure 1.1 in the case $N = 4$. However, all these further bifurcation occur for $\gamma = \mathcal{O}(N^2)$. In other words, for any $\tilde\gamma > 0$, the number of stationary points is at most of order N when $\gamma > \tilde\gamma N^2$ and N is sufficiently large.

Exercise 1.1.3. In the case $N = 2$, find all stationary points of the potential V and determine their stability as a function of $\gamma \geqslant 0$. **Hint:** Express the potential V in variables $z_\pm = y_1 \pm y_2$.   ♣

1.2   Existence and uniqueness of solutions

We now return to the SDE (1.0.1) for $\varepsilon > 0$. We can write it in compact form as

$$\mathrm{d}y_t = -\nabla V(y_t)\,\mathrm{d}t + \sqrt{2\varepsilon}\,\mathrm{d}W_t\,. \tag{1.2.1}$$

We first note that $y \mapsto \nabla V(y)$ is continuous (and therefore measurable) and locally (but not globally) Lipschitz. We can thus apply a standard result, such as [Øks85, Theorem 5.2.1] to the process $y_{t \wedge \tau}$ where τ is the first-exit time from a large ball to obtain local existence and uniqueness of a strong solution to (1.2.1). In particular, local existence follows from the fact that the map $y^{(k)} \mapsto y^{(k+1)}$ defined by

$$y_t^{(k+1)} = y_0 - \int_0^{t \wedge \tau} \nabla V(y_s^{(k)})\,\mathrm{d}t + \sqrt{2\varepsilon}\,W_t$$

is a contraction, via Banach's fixed point theorem.

To extend this result to global existence of a strong solution, we have to rule out finite-time explosion of solutions. This is, however, not hard to achieve, since the quartic growth of the potential prevents sample paths from straying very far from the origin (coercivity). Here are two possible ways of obtaining global existence:

1. One checks that the drift and diffusion coefficients in (1.2.1) satisfy the one-sided growth condition

$$y \cdot (-\nabla V(y)) + \frac{1}{2}(\sqrt{2\varepsilon})^2 \leqslant K(1 + \|y\|^2) \qquad \forall y \in \mathbb{R}^\Lambda \tag{1.2.2}$$

   for a constant $K > 0$. See for instance [Mao97, Theorem 3.5].

2. Let

$$\mathscr{L} := \varepsilon \Delta - \nabla V \cdot \nabla \tag{1.2.3}$$

   denote the infinitesimal generator of the diffusion (1.2.1). Then we can take advantage of the fact that V is itself a Lyapunov function for the system, in the sense that there exists a constant $c \geqslant 0$ such that

$$(\mathscr{L}V)(y) \leqslant cV(y) \qquad \forall y \in \mathbb{R}^\Lambda\,. \tag{1.2.4}$$

   This implies that the process is non-explosive according to [MT93c, Theorem 2.1].

Exercise 1.2.1. Check that the SDE (1.0.1) satisfies the one-sided growth condition (1.2.2) as well as the Lyapunov condition (1.2.4). Show that, in addition, there exists $R < \infty$ such that $\mathscr{L}V(y)$ is negative outside the ball $\{\|y\| \leqslant R\}$. Can we do better than that?   ♣

The main result of this section is thus as follows.

Proposition 1.2.2. *For any initial condition $y_0 \in \mathbb{R}^\Lambda$, the SDE (1.0.1) admits an almost surely continuous, pathwise unique strong solution $(y_t)_{t \geqslant 0}$ which is global in time.*

1.3   Invariant measure and reversibility

In order to determine an invariant probability measure for the diffusion (1.2.1), it is useful to rewrite the infinitesimal generator (1.2.3) in the form

$$\mathscr{L} = \varepsilon\, \mathrm{e}^{V/\varepsilon} \nabla \cdot \mathrm{e}^{-V/\varepsilon} \nabla\,, \tag{1.3.1}$$

with the usual convention that any differential operator acts on everything on its right. This is indeed equivalent to (1.2.3) by Leibniz' rule.

Proposition 1.3.1. *Assume $\varepsilon > 0$ and the potential V is such that*

$$\mathscr{Z} := \int_{\mathbb{R}^\Lambda} e^{-V(y)/\varepsilon}\, dy < \infty . \tag{1.3.2}$$

Then the diffusion (1.2.1) *admits a unique invariant probability measure, with density*

$$\pi(y) := \frac{1}{\mathscr{Z}} e^{-V(y)/\varepsilon} \tag{1.3.3}$$

with respect to Lebesgue measure dy, *known as a* Gibbs *measure.*

PROOF: Let f be a bounded measurable test function in the domain of $\mathscr{L}$. Writing $\mathscr{L}^\dagger$ for the adjoint of $\mathscr{L}$, we have

$$\begin{aligned}
\langle \mathscr{L}^\dagger \pi, f \rangle_{L^2} &= \langle \pi, \mathscr{L} f \rangle_{L^2} \\
&= \int_{\mathbb{R}^\Lambda} \pi(y)\varepsilon\, e^{V(y)/\varepsilon} \nabla \cdot (e^{-V(y)/\varepsilon} \nabla f(y))\, dy \\
&= \frac{\varepsilon}{\mathscr{Z}} \int_{\mathbb{R}^\Lambda} \nabla \cdot (e^{-V(y)/\varepsilon} \nabla f(y))\, dy \\
&= 0
\end{aligned}$$

by the divergence theorem. Thus $\mathscr{L}^\dagger \pi = 0$, and Kolmogorov's backward equation (or Fokker–Planck equation) shows that π is invariant.

Uniqueness of π follows from the fact that the diffusion is irreducible (with respect to Lebesgue measure) by uniform ellipticity of the noise. One way of seeing this is to apply Theorem 3.3 in [MT93c] to show that the process is Harris recurrent (it almost surely hits any set of positive measure), owing to the improved estimate on $\mathscr{L}V$ obtained in Exercise 1.2.1. □

Exercise 1.3.2. Use integration by parts to compute an explicit expression for $\mathscr{L}^\dagger$, and use it to verify that $\mathscr{L}^\dagger \pi = 0$.   ♣

Another consequence of the specific form (1.3.1) of the generator is that the diffusion is *reversible*.

Proposition 1.3.3. *The infinitesimal generator $\mathscr{L}$ is self-adjoint with respect to the weighted inner product*

$$\langle f, g \rangle_\pi := \int_{\mathbb{R}^\Lambda} e^{-V(y)/\varepsilon}\, \overline{f(y)} g(y)\, dy . \tag{1.3.4}$$

As a consequence, the transition probability density p_t of the diffusion satisfies the detailed balance condition

$$\pi(y) p_t(y,z) = \pi(z) p_t(z,y) \tag{1.3.5}$$

for all $y, z \in \mathbb{R}^\Lambda$ and all $t > 0$.

PROOF: The divergence theorem yields

$$\langle f, \mathscr{L} g \rangle_\pi = \varepsilon \int_{\mathbb{R}^\Lambda} \overline{f(y)} \nabla \cdot [e^{-V(y)/\varepsilon} \nabla g(y)]\, dy = -\varepsilon \int \overline{\nabla f(y)} \cdot \nabla g(y) e^{-V(y)/\varepsilon}\, dy .$$

Since $\overline{f}$ and g play a symmetric role in this expression, it is equal to $\langle \mathscr{L} f, g \rangle_\pi$. It follows that the Markov semigroup $P_t = e^{t\mathscr{L}}$ is also self-adjoint, that is,

$$\langle f, P_t g \rangle_\pi = \langle P_t f, g \rangle_\pi .$$

Taking for f and g functions converging to Dirac distributions δ_y and δ_z yields (1.3.5). □

Remark 1.3.4. The reason we have omitted the normalisation $\mathscr{Z}^{-1}$ in the weighted inner product (1.3.4) is purely for later notational convenience. Of course, it does not change anything in the presented results. ◇

1.4 Large deviations

We now focus on the weak-noise regime $0 < \varepsilon \ll 1$. The theory of large deviations provides a useful tool to estimate the probability of rare events as ε decreases to 0.

Definition 1.4.1 (Large-deviation principle). *Fix a time horizon $T \in (0, \infty)$. We say that a function $\mathscr{I} = \mathscr{I}_{[0,T]}$ from the space $\mathscr{C}_0$ of continuous paths $\gamma : [0, T] \to \mathbb{R}^\Lambda$ to $\mathbb{R}_+$ is a good rate function if it is lower semi-continuous (its sublevel sets are closed) and has compact level sets. The process $(y_t)_{t \in [0,T]}$ satisfies a sample-paths large-deviation principle (LDP) with good rate function $\mathscr{I}$ if*

$$\liminf_{\varepsilon \to 0} 2\varepsilon \log \mathbb{P}\{(y_t)_{t \in [0,T]} \in O\} \geqslant -\inf_{\gamma \in O} \mathscr{I}_{[0,T]}(\gamma)$$

$$\limsup_{\varepsilon \to 0} 2\varepsilon \log \mathbb{P}\{(y_t)_{t \in [0,T]} \in C\} \leqslant -\inf_{\gamma \in C} \mathscr{I}_{[0,T]}(\gamma)$$

holds for all open $O \subset \mathscr{C}_0$ and all closed $C \subset \mathscr{C}_0$.

Roughly speaking, the large-deviation principle says that for a sufficiently nice set Γ of paths $\gamma : [0, T] \to \mathbb{R}^\Lambda$, we have

$$\mathbb{P}\{(y_t)_{t \in [0,T]} \in \Gamma\} \simeq e^{-\inf_\Gamma \mathscr{I}_{[0,T]}/(2\varepsilon)}$$

in the sense of logarithmic equivalence. Here are two illustrative examples:

1. Let $\gamma_0 : [0, T] \to \mathbb{R}^\Lambda$ be a given continuous path, and let Γ be the set of paths such that $\|\gamma(t) - \gamma_0(t)\| < \delta$ for all $t \in [0, T]$. As δ decreases to 0, the infimum of $\mathscr{I}_{[0,T]}$ over Γ will converge to $\mathscr{I}(\gamma_0)$. Therefore, the probability of sample paths tracking γ_0 up to a small error δ will be close to $e^{-\mathscr{I}_{[0,T]}(\gamma_0)/(2\varepsilon)}$ for small δ.

2. Let $D \subset \mathbb{R}^\Lambda$ be a bounded, connected subset of $\mathbb{R}^\Lambda$, fix a point $y_0 \in D$, and let Γ be the set of continuous paths starting in y_0 and leaving D at least once during the time interval $(0, T)$. Then we have

$$\mathbb{P}^{y_0}\{\tau_D < T\} \simeq e^{-\inf_\Gamma \mathscr{I}_{[0,T]}/(2\varepsilon)},$$

where $\tau_D = \inf\{t > 0 : y_t \notin D\}$ is the first-exit time from D.

In the case of scaled Brownian motion $\sqrt{2\varepsilon} W_t$, Schilder [Sch66] obtained a large-deviation principle with rate function

$$\mathscr{I}_{[0,T]}(\gamma) := \begin{cases} \dfrac{1}{2} \displaystyle\int_0^T \|\dot{\gamma}(t)\|^2 \, \mathrm{d}t & \text{if } \gamma \in H^1 , \\ +\infty & \text{otherwise} . \end{cases}$$

Freidlin and Wenzell [FW98] extended this result to general diffusions. Their proof uses the Cameron–Martin–Girsanov formula to reduce the general problem to the special case of scaled Brownian motion. In the case of the SDE (1.2.1), the LDP takes the following form.

Theorem 1.4.2 (LDP for gradient diffusions). *The diffusion (1.2.1) satisfies a large-deviation principle with good rate function*

$$\mathscr{I}_{[0,T]}(\gamma) := \begin{cases} \dfrac{1}{2} \displaystyle\int_0^T \|\dot{\gamma}(t) + \nabla V(\gamma(t))\|^2 \, \mathrm{d}t & \text{if } \gamma \in H^1 , \\ +\infty & \text{otherwise} . \end{cases}$$

For SDEs with more general drift coefficients $f(y)$, the term $\|\dot{\gamma}(t) + \nabla V(\gamma(t))\|^2$ in the rate function has to be replaced by $\|\dot{\gamma}(t) - f(\gamma(t))\|^2$, and there also exists a version for SDEs with general diffusion coefficients. However, the gradient form of the system (1.2.1) entails a substantial simplification, since we have the identity

$$\mathcal{I}_{[0,T]}(\gamma) = \frac{1}{2} \int_0^T \|\dot{\gamma}(t) - \nabla V(\gamma(t))\|^2 \, dt + 2 \int_0^T \dot{\gamma}(t) \cdot \nabla V(\gamma(t)) \, dt$$

$$= \frac{1}{2} \int_0^T \|\dot{\gamma}(t) - \nabla V(\gamma(t))\|^2 \, dt + 2[V(\gamma(T)) - V(\gamma(0))] . \tag{1.4.1}$$

If for example we want to analyse properties of the first-exit time from a potential well D, starting from its bottom y^*, the integral in (1.4.1) can be made arbitrarily small by taking T large and letting $\gamma(t)$ be the solution of the equation with reversed drift $\dot{\gamma} = +\nabla V(\gamma)$. The rate function will thus be dominated by the minimum of the potential difference $2[V(y) - V(y^*)]$ over all y on the boundary of D.

Exercise 1.4.3. Compute the rate function $\mathcal{I}_{[0,T]}$ in the case of the Ornstein–Uhlenbeck process

$$dy_t = -y_t \, dt + \sqrt{2\varepsilon} \, dW_t ,$$

and use it to analyse the cumulative distribution function of $\tau = \inf\{t > 0 \colon |y_t| > L\}$ as $\varepsilon \to 0$. ♣

1.5 Metastability

The existence of a unique invariant probability measure having been settled, the next natural question to ask is whether the system will converge to this measure. In fact, as hinted at in Exercise 1.2.1, in the case of the diffusion (1.0.1) we have the even stronger Lyapunov property

$$(\mathcal{L}V)(y) \leqslant -cV(y) + d$$

for two constants $c > 0$ and $d \geqslant 0$. According to [MT93c, Theorem 6.1], we have an exponential ergodicity result of the form

$$\sup_{f \colon |f| \leqslant V+1} \left| \mathbb{E}^y[f(y_t)] - \langle \pi, f \rangle \right| \leqslant C[V(y) + 1] e^{-\beta t} \qquad \forall y \in \mathbb{R}^\Lambda$$

for some constants $\beta, C > 0$. However, this result does not yield a good control on the constants β and C, and in particular β can behave quite badly in terms of ε. This is a manifestation of the phenomenon of *metastability*, which is related to the fact that for small ε, the system may take extremely long to move between potential wells. A classical approach to quantifying this phenomenon is thus to investigate the law of this interwell transition time.

1.5.1 Arrhenius law

For simplicity, we will only consider cases where V is a double-well potential, as is the case for the system (1.0.1) for $\gamma > \gamma_1(N)$ according to Proposition 1.1.1. We can always assume that the saddle is located at the origin, and denote the two local minima by $y_\pm^*$. Assume that the diffusion starts in y_-^*, and given a small constant $\delta > 0$, denote by

$$\tau_+ := \inf\{t > 0 \colon \|y_t - y_+^*\| < \delta\}$$

the first-hitting time of a small neighbourhood of y_+^*.

A first type of results on the law of τ_+ can be obtained by applying the theory of large deviations outlined in Section 1.4.

Theorem 1.5.1 (Large-deviation results for interwell transitions). *Let $H = V(0) - V(y_-^*)$ be the potential difference between the starting minimum and the saddle. Then*

$$\lim_{\varepsilon \to 0} \varepsilon \log \mathbb{E}^{y_-^*}[\tau_+] = H \,. \tag{1.5.1}$$

Furthermore, for any $\eta > 0$,

$$\lim_{\varepsilon \to 0} \mathbb{P}^{y_-^*}\Big\{ e^{(H-\eta)/\varepsilon} \leqslant \tau_+ \leqslant e^{(H+\eta)/\varepsilon} \Big\} = 1 \,. \tag{1.5.2}$$

Finally, let D_- denote the basin of attraction of y_-^* under the deterministic dynamics. Then the location $y_{\tau_{D_-}}$ of first exit from D_- satisfies

$$\lim_{\varepsilon \to 0} \mathbb{P}^{y_-^*}\Big\{ y_{\tau_{D_-}} \in \mathcal{N} \Big\} = 0 \tag{1.5.3}$$

for any closed $\mathcal{N} \subset \partial D_-$ that does not contain the saddle at 0.

Relation (1.5.1) is called *Arrhenius' law*. It states that the expected transition time behaves like $e^{H/\varepsilon}$ in the sense of logarithmic equivalence. The exponential dependence in H/ε goes back to works by van t'Hoff and Arrhenius in the late 19th century [Arr89]. Relation (1.5.2) shows that the law of τ_+ concentrates around its expectation, albeit in a rather weak sense. The last result (1.5.3) states that the exit location from the starting potential well concentrates near the saddle in the vanishing noise limit.

Theorem 1.5.1 can be obtained in two main steps. Firstly, analogous results for the first exit from a bounded subset of D_- follow from Theorems 2.1, 4.1 and 4.2 in [FW98, Chapter 4]. The main idea of the proof is that successive attempts to exit this subset are almost independent, and have an exponentially small probability of success as discussed at the end of Section 1.4. Secondly, these estimates can be extended to the distribution of τ_+ by showing that the process behaves like a Markov chain jumping between neighbourhoods of critical points of V, see Theorems 5.1 and 5.3 in [FW98, Chapter 6].

Particularising to the diffusion (1.0.1) for $\gamma > \gamma_1(N)$, we obtain that Arrhenius' law (1.5.1) holds with

$$H = V(0) - V(-1, \ldots, -1) = \frac{N}{4} \,. \tag{1.5.4}$$

We point out that at this stage, we do not claim any control on the speed of convergence in (1.5.1), (1.5.2) and (1.5.3) as a function of N. This is a more subtle point that we will come back to later on.

Remark 1.5.2. The case where γ lies below $\gamma_1(N)$ but is still of order N^2 can be analysed in a similar way. The main difference is that instead of a single saddle located at the origin, the number of relevant saddles is proportional to N. Thus H is no longer given by (1.5.4), and the exit location from the starting well is concentrated in the union of the relevant saddles. See [BFG07a, Theorem 2.10] and [BFG07b, Theorem 2.4].   ◇

1.5.2 Potential theory

One of the approaches allowing to obtain sharper asymptotics on metastable transition times is based on potential theory. The starting point is the observation that owing to Dynkin's formula, several probabilistic quantities of interest solve boundary value problems involving the infinitesimal generator $\mathscr{L}$ of the process.

Let $A \subset \mathbb{R}^\Lambda$ be a closed set with smooth boundary, and let $\tau_A = \inf\{t > 0 \colon y_t \in A\}$ be the first-hitting time of A. Then the map $y \mapsto w_A(y) = \mathbb{E}^y[\tau_A]$ satisfies the Poisson problem

$$\begin{cases} (\mathscr{L} w_A)(y) = -1 & y \in A^c \,, \\ w_A(y) = 0 & y \in A \,. \end{cases} \tag{1.5.5}$$

This is a particular case of [Øks85, Corollary 9.1.2].
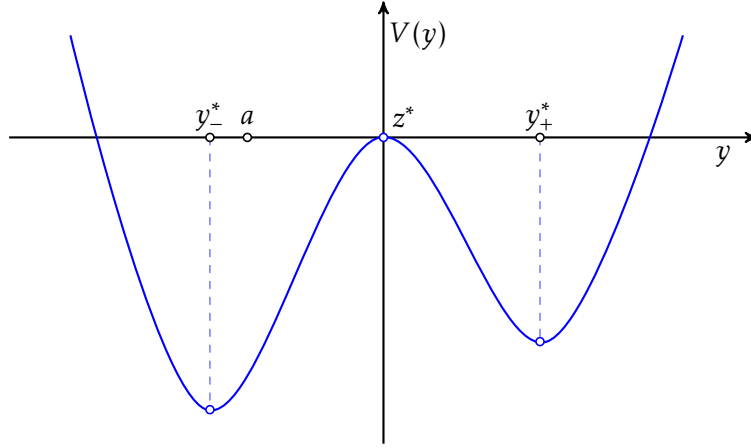
Figure 1.2 – A one-dimensional double-well potential.

Exercise 1.5.3. Consider the SDE (1.2.1) in the one-dimensional case $y \in \mathbb{R}$, for a potential V growing sufficiently fast at infinity. Show that for $A = (-\infty, a]$, the solution of (1.5.5) is given by

$$w_A(y) = \frac{1}{\varepsilon} \int_a^y \int_{y_2}^\infty \mathrm{e}^{[V(y_2) - V(y_1)]/\varepsilon} \,\mathrm{d}y_1 \,\mathrm{d}y_2 \qquad \forall y > a \,.$$

Assume now that V is a double-well potential, with two local minima located at $y_-^* < a$ and $y_+^* > a$, and a saddle (local maximum) at $z^* \in (a, y_+^*)$ (Figure 1.2). Use Laplace asymptotics to show that

$$w_A(y) = \frac{2\pi}{\sqrt{|V''(z^*)| V''(y_+^*)}} \mathrm{e}^{[V(z^*) - V(y_+^*)]/\varepsilon} [1 + \mathcal{O}(\sqrt{\varepsilon})] \,. \tag{1.5.6}$$

Relation (1.5.6) is known as *Kramers' law*.   ♣

In general, there is no explicit solution to the Poisson equation (1.5.5). However, the solution can be represented as

$$w_A(y) = -\int_{A^c} G_{A^c}(y, z) \,\mathrm{d}z \,, \tag{1.5.7}$$

where G_{A^c} is the *Green function* associated with A^c, which solves

$$\begin{cases} (\mathscr{L} G_{A^c})(y, z) = \delta(y - z) & y \in A^c \,, \\ G_{A^c}(y, z) = 0 & y \in A \,. \end{cases}$$

Reversibility of the SDE (1.2.1) implies that G_{A^c} satisfies the detailed-balance relation

$$\mathrm{e}^{-V(y)/\varepsilon} G_{A^c}(y, z) = \mathrm{e}^{-V(z)/\varepsilon} G_{A^c}(z, y) \,. \tag{1.5.8}$$

Let now A and B be two disjoint closed sets with smooth boundary. A second quantity of interest is the *committor function* $h_{AB}(y) = \mathbb{P}^y\{\tau_A < \tau_B\}$, also called *equilibrium potential*. It satisfies the Dirichlet problem

$$\begin{cases} (\mathscr{L} h_{AB})(y) = 0 & y \in (A \cup B)^c \,, \\ h_{AB}(y) = 1 & y \in A \,, \\ h_{AB}(y) = 0 & y \in B \,. \end{cases} \tag{1.5.9}$$

Exercise 1.5.4. Consider again the SDE (1.2.1) in the one-dimensional case $y \in \mathbb{R}$. Given $a < b$, let $A = (-\infty, a]$ and $B = [b, \infty)$. Show that

$$h_{AB}(y) = \frac{\displaystyle\int_y^b e^{V(z)/\varepsilon}\,\mathrm{d}z}{\displaystyle\int_a^b e^{V(z)/\varepsilon}\,\mathrm{d}z} \qquad \forall y \in [a, b] \,. \tag{1.5.10}$$

Use Laplace asymptotics to determine the behaviour of $h_{AB}(y)$ as $\varepsilon \to 0$.   ♣

There exists again an integral representation of the solution to (1.5.9) in terms of the Green function, namely

$$h_{AB}(y) = -\int_{\partial A} G_{B^c}(y, z) e_{AB}(\mathrm{d}z) \,. \tag{1.5.11}$$

Here e_{AB} is a measure concentrated on ∂A called the *equilibrium measure*. It is defined by

$$e_{AB}(\mathrm{d}y) := (-\mathscr{L} h_{AB})(\mathrm{d}y)$$

interpreted in the weak sense (i.e., both sides have to be integrated against test functions). The *capacity* is then defined by

$$\mathrm{cap}(A, B) := \int_{\partial A} e^{-V(y)/\varepsilon} e_{AB}(\mathrm{d}y) \,.$$

Note that the capacity is the normalisation required to make

$$\nu_{AB}(\mathrm{d}y) := \frac{1}{\mathrm{cap}(A, B)} e^{-V(y)/\varepsilon} e_{AB}(\mathrm{d}y)$$

a probability measure on ∂A.

Remark 1.5.5. The name *capacity* is due to an analogy with electrostatics. Indeed, in the case without potential $V = 0$, (1.5.9) is the equation for the electric potential of a capacitor, consisting of a conductor A at potential 1 and a grounded conductor B. The integral relation (1.5.11) expresses the fact that h_{AB} is the electric potential created by a charge density e_{AB} on ∂A with zero boundary conditions on B. Since the potential difference between the conductors is equal to 1, the capacity is equal to the total charge on the conductor A.   ◇

The main result that makes the potential-theoretic approach so successful is the following relation between expected first-hitting time and capacity.

Theorem 1.5.6. *For any disjoint sets A and B with smooth boundary,*

$$\mathbb{E}^{\nu_{AB}}[\tau_B] := \int_{\partial A} w_B(y) \nu_{AB}(\mathrm{d}y) = \frac{1}{\mathrm{cap}(A, B)} \int_{B^c} e^{-V(y)/\varepsilon} h_{AB}(y)\,\mathrm{d}y \,. \tag{1.5.12}$$

Proof: It follows from the integral representation (1.5.7) of w_B, the detailed balance (1.5.8) of the Green function, and the integral representation (1.5.11) of h_{AB}, that

$$
\begin{aligned}
\operatorname{cap}(A,B)\int_{\partial A} w_B(y)\nu_{AB}(\mathrm{d}y) &= -\int_{\partial A}\int_{B^c} G_{B^c}(y,z)\,\mathrm{d}z\,\mathrm{e}^{-V(y)/\varepsilon}\,e_{AB}(\mathrm{d}y) \\
&= -\int_{B^c}\int_{\partial A} G_{B^c}(z,y)e_{AB}(\mathrm{d}y)\mathrm{e}^{-V(z)/\varepsilon}\,\mathrm{d}z \\
&= \int_{B^c} h_{AB}(z)\mathrm{e}^{-V(z)/\varepsilon}\,\mathrm{d}z\,.
\end{aligned}
$$

Dividing by the capacity yields (1.5.12). □

The exact relation (1.5.12) is useful because there exist variational principles that often allow to obtain good upper and lower bounds on the capacity. The first of these principles states that the capacity is a minimiser of the *Dirichlet form*.

Definition 1.5.7 (Dirichlet form). *The* Dirichlet form *associated with the diffusion* (1.2.1) *is the quadratic form acting on test functions* $f:\mathbb{R}^\Lambda\to\mathbb{R}$ *in the domain of* $\mathscr{L}$, *defined by*

$$
\mathscr{E}(f) := \langle f,-\mathscr{L}f\rangle_\pi = \varepsilon\int_{\mathbb{R}^\Lambda}\mathrm{e}^{-V(y)/\varepsilon}\|\nabla f(y)\|^2\,\mathrm{d}y\,. \tag{1.5.13}
$$

The integral expression for the Dirichlet form in (1.5.13) is a consequence of the first Green identity, which is essentially the divergence theorem (using the fact that $\mathrm{e}^{-V/\varepsilon}[\varepsilon\nabla f\cdot\nabla g+g\mathscr{L}f]=\varepsilon\nabla\cdot(g\,\mathrm{e}^{-V/\varepsilon}\nabla f))$ and says that for a set D with smooth boundary,

$$
\int_D \mathrm{e}^{-V(y)/\varepsilon}\big[\varepsilon\nabla f(y)\cdot\nabla g(y)+g(y)(\mathscr{L}f)(y)\big]\,\mathrm{d}y = \varepsilon\int_{\partial D}\mathrm{e}^{-V(y)/\varepsilon}g(y)\partial_{n(y)}f(y)\sigma(\mathrm{d}y)\,, \tag{1.5.14}
$$

where $\partial_{n(y)}f(y)=\nabla f(y)\cdot n(y)$ denotes the derivative in the direction of the unit outer normal $n(y)$ at a point $y\in\partial D$, and $\sigma(\mathrm{d}y)$ is the Lebesgue measure on ∂D. In the particular case $D=\mathbb{R}^\Lambda$, the right-hand side of (1.5.14) vanishes and one obtains the integral expression in (1.5.13).

The quadratic form $\mathscr{E}$ can be extended by polarisation to a bilinear form given by

$$
\mathscr{E}(f,g) = \langle f,-\mathscr{L}g\rangle_\pi = \varepsilon\int_{\mathbb{R}^\Lambda}\mathrm{e}^{-V(y)/\varepsilon}\,\overline{\nabla f(y)}\cdot\nabla g(y)\,\mathrm{d}y\,,
$$

which satisfies the Cauchy–Schwarz inequality $\mathscr{E}(f,g)^2\leqslant\mathscr{E}(f)\mathscr{E}(g)$.

Theorem 1.5.8 (Dirichlet principle). *Let* $\mathscr{H}_{AB}$ *be the set of functions* $h:\mathbb{R}^\Lambda\to[0,1]$ *that are in the domain of* $\mathscr{E}$ *and such that* $h|_A=1$ *and* $h|_B=0$. *Then*

$$
\operatorname{cap}(A,B) = \mathscr{E}(h_{AB}) = \inf_{h\in\mathscr{H}_{AB}}\mathscr{E}(h)\,.
$$

Proof: Pick any $h\in\mathscr{H}_{AB}$. Since h_{AB} satisfies the Dirichlet problem (1.5.9), we have

$$
\begin{aligned}
\mathscr{E}(h,h_{AB}) &= \varepsilon\int_{\mathbb{R}^\Lambda}\mathrm{e}^{-V(y)/\varepsilon}h(y)(-\mathscr{L}h_{AB})(y)\,\mathrm{d}y \\
&= \varepsilon\int_{\partial A}\mathrm{e}^{-V(y)/\varepsilon}e_{AB}(\mathrm{d}y) \\
&= \operatorname{cap}(A,B)\,.
\end{aligned}
$$

Since $h_{AB}\in\mathscr{H}_{AB}$, we obtain $\mathscr{E}(h_{AB})=\operatorname{cap}(A,B)$. The Cauchy–Schwarz inequality then yields $\mathscr{E}(h_{AB})^2=\mathscr{E}(h,h_{AB})^2\leqslant\mathscr{E}(h)\mathscr{E}(h_{AB})$, i.e. $\mathscr{E}(h)\geqslant\mathscr{E}(h_{AB})$. A sketch is given in Figure 1.3. □
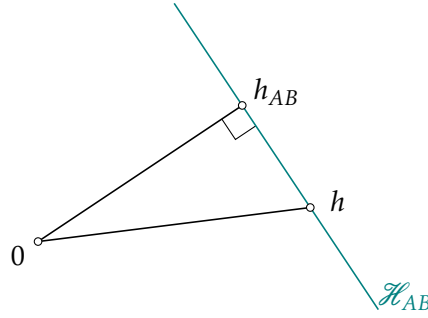
Figure 1.3 – Dirichlet principle. The capacity minimises the distance to the origin given by the Dirichlet form over all functions h satisfying the boundary conditions.

There also exists a variational principle allowing to obtain lower bounds on the capacity, the *Thomson principle*, which involves divergence-free flows.

Definition 1.5.9 (Divergence-free unit flow). *A divergence-free unit AB-flow is a vector field φ of class $\mathscr{C}^1$ on $(A \cup B)^c$ such that $\nabla \cdot \varphi = 0$ in $(A \cup B)^c$ and*

$$\int_{\partial A} \varphi(y) \cdot n_A(y)\sigma(\mathrm{d}y) = 1 = -\int_{\partial B} \varphi(y) \cdot n_B(y)\sigma(\mathrm{d}y), \qquad (1.5.15)$$

where $n_A(y)$, $n_B(y)$ denote the unit outward normal vectors at $y \in \partial A$ and $y \in \partial B$ respectively. We denote by $\mathfrak{U}^1_{AB}$ the set of divergence-free unit AB-flows. The harmonic unit flow *is defined as*

$$\varphi_{AB}(y) := -\frac{\varepsilon}{\mathrm{cap}(A, B)}\, \mathrm{e}^{-V(y)/\varepsilon}\, \nabla h_{AB}(y)$$

and it satisfies $\varphi_{AB} \in \mathfrak{U}^1_{AB}$.

The fact that φ_{AB} is divergence-free follows from the form (1.3.1) of the generator and the fact that h_{AB} is harmonic. Its satisfies (1.5.15) as a consequence of the divergence theorem.

We define a bilinear form on $\mathfrak{U}^1_{AB}$ by

$$\mathscr{D}(\varphi, \psi) := \frac{1}{\varepsilon}\int_{(A \cup B)^c} \mathrm{e}^{V(y)/\varepsilon}\, \varphi(y) \cdot \psi(y)\,\mathrm{d}y, \qquad (1.5.16)$$

and set $\mathscr{D}(\varphi, \varphi) := \mathscr{D}(\varphi)$.

Proposition 1.5.10 (Thomson principle). *The capacity satisfies*

$$\mathrm{cap}(A, B) = \frac{1}{\mathscr{D}(\varphi_{AB})} = \sup_{\varphi \in \mathfrak{U}^1_{AB}} \frac{1}{\mathscr{D}(\varphi)}.$$

PROOF: Set $\Psi_{h_{AB}} = \mathrm{cap}(A, B)\varphi_{AB} = -\varepsilon\, \mathrm{e}^{-V/\varepsilon}\, \nabla h_{AB}$. Then

$$\mathscr{D}(\Psi_{h_{AB}}) = \frac{\varepsilon^2}{\varepsilon}\int_{(A \cup B)^c} \mathrm{e}^{V(y)/\varepsilon}|\mathrm{e}^{-V(y)/\varepsilon}\, \nabla h_{AB}(y)|^2\,\mathrm{d}y = \mathrm{cap}(A, B).$$

This implies $\mathscr{D}(\varphi_{AB}) = (\mathrm{cap}(A, B))^{-1}$ by bilinearity. For any unit flow $\varphi \in \mathfrak{U}^1_{AB}$, we have

$$
\begin{aligned}
\mathscr{D}(\Psi_{h_{AB}}, \varphi) &= \frac{1}{\varepsilon} \int_{(A \cup B)^c} e^{V(y)/\varepsilon} \Psi_{h_{AB}}(y) \cdot \varphi(y) \, \mathrm{d}y \\
&= - \int_{(A \cup B)^c} \nabla h_{AB}(y) \cdot \varphi(y) \, \mathrm{d}y \\
&= - \int_{(A \cup B)^c} \nabla \cdot (h_{AB}(y) \varphi(y)) \, \mathrm{d}y + \int_{(A \cup B)^c} h_{AB}(y) \nabla \cdot \varphi(y) \, \mathrm{d}y \, .
\end{aligned}
$$

The second integral vanishes because φ is divergence-free. By the divergence theorem, the first term is equal to

$$
- \int_{\partial A \cup \partial B} h_{AB}(y) \varphi(y) \cdot n(y) \sigma(\mathrm{d}y) \, ,
$$

where $n(y)$ is the unit normal vector pointing *inside* $A \cup B$. The integral on ∂B vanishes, while the integral on ∂A is equal to 1 because $n(y) = -n_A(y)$ and φ satisfies (1.5.15). This implies $\mathscr{D}(\Psi_{h_{AB}}, \varphi) = 1$, and thus, by the Cauchy–Schwarz inequality,

$$
1 = \mathscr{D}(\Psi_{h_{AB}}, \varphi)^2 \leqslant \mathscr{D}(\Psi_{h_{AB}}) \mathscr{D}(\varphi) = \mathrm{cap}(A, B) \mathscr{D}(\varphi) \, ,
$$

showing that indeed $\mathrm{cap}(A, B) \geqslant 1/\mathscr{D}(\varphi)$. $\qquad\square$

Remark 1.5.11. The potential-theoretic approach described in this section admits a simple analogue in the setting of discrete Markov chains. This will not play any role in what follows, but for completeness and since it may help intuition, we summarise this theory in Appendix A. $\quad\diamond$

1.5.3 Eyring–Kramers law

We now apply the potential-theoretic approach to the system of interacting diffusions, in order to obtain sharper asymptotics on the transition time between the two states $y^*_\pm = \pm(1, \ldots, 1)$. For simplicity, we only consider the case $\gamma > \gamma_1(N)$, when the potential is a double-well potential with a saddle located at the origin. The theory works, however, in much greater generality.

One simplification due to the condition $\gamma > \gamma_1(N)$ is that we will be able to use the following symmetry argument.

Lemma 1.5.12. *Let A and B satisfy $B = -A$. Then*

$$
\int_{(A \cup B)^c} e^{-V(y)/\varepsilon} h_{AB}(y) \, \mathrm{d}y = \frac{1}{2} \mathscr{Z}
$$

where $\mathscr{Z}$ is the partition function *of the system defined in (1.3.2).*

Exercise 1.5.13. Prove Lemma 1.5.12, using the fact that $V(-y) = V(y)$ and the relations $h_{AB}(y) = h_{BA}(-y)$ and $h_{AB}(y) = 1 - h_{BA}(y)$. $\quad\clubsuit$

To apply Theorem 1.5.6, it is thus sufficient to estimate the partition function $\mathscr{Z}$ and the capacity $\mathrm{cap}(A, B)$. In order to do so, it turns out to be useful to make a change of variables. Let $(e_0, \ldots, e_{N-1})$ be an orthonormal basis of $\mathbb{R}^\Lambda$, where

$$
e_0 := \frac{1}{\sqrt{N}} (1, \ldots, 1)^{\mathrm{T}} \, .
$$

The precise form of the other basis vectors will not matter – one possibility is to use those appearing in the discrete Fourier transform. We denote by $\mathbb{R}_\perp^\Lambda$ the span of $e_1, \ldots, e_{N-1}$. The change of variables is given by

$$y = y_0 e_0 + \sqrt{\varepsilon} y_\perp$$

where

$$y_0 = y \cdot e_0 = \frac{1}{\sqrt{N}} \sum_{i=1}^N y^i, \qquad y_\perp = \frac{1}{\sqrt{\varepsilon}}(y - y_0 e_0) \in \mathbb{R}_\perp^\Lambda.$$

The role of the factor $\sqrt{\varepsilon}$ is to highlight the scaling properties of some quantities with ε. Due to this factor, the Jacobian of the transformation $y \mapsto (y_0, y_\perp)$ is equal to $\varepsilon^{(N-1)/2}$. Performing the change of variables in the potential V, and using the fact that the sum of the coordinates of $y_\perp$ vanishes, we obtain

$$\frac{1}{\varepsilon} V(y) = \frac{1}{\varepsilon} V_0(y_0) + \frac{1}{2} \langle y_\perp, Q_\perp(y_0) y_\perp \rangle + R_\varepsilon(y_0, y_\perp). \tag{1.5.17}$$

Here

$$V_0(y_0) := N U\left(\frac{y_0}{\sqrt{N}}\right) = \frac{1}{4N} y_0^4 - \frac{1}{2} y_0^2 + \frac{N}{4},$$

while $\langle y_\perp, Q_\perp(y_0) y_\perp \rangle$ is the quadratic form defined by

$$Q_\perp(y_0) := \left(\frac{3}{N} y_0^2 - 1\right)\mathbb{1} - \gamma \Delta_{\perp,N}$$

where $\Delta_{\perp,N}$ is the discrete Laplacian acting on $\mathbb{R}_\perp^\Lambda$, and R_ε is a remainder given by

$$R_\varepsilon(y_0, y_\perp) = \sqrt{\frac{\varepsilon}{N}} y_0 \sum_{i \in \Lambda} (y_\perp^i)^3 + \frac{\varepsilon}{4} \sum_{i \in \Lambda} (y_\perp^i)^4.$$

Proposition 1.5.14. *The partition function has the asymptotic form*

$$\mathscr{Z} = 2\sqrt{\frac{(2\pi\varepsilon)^N}{2 \det Q_\perp(-\sqrt{N})}}[1 + \mathcal{O}(\varepsilon)] = 2\sqrt{\frac{(2\pi\varepsilon)^N}{\prod_{k=0}^{N-1} v_k^N}}[1 + \mathcal{O}(\varepsilon)], \tag{1.5.18}$$

where the v_k^N are the eigenvalues (1.1.4) *of the Hessian of $V(y_-^*)$.*

PROOF: This result follows rather directly from standard Laplace asymptotics, but we will give some details of the proof as they will be useful later on. Using (1.5.17), we get

$$\mathscr{Z} = \varepsilon^{(N-1)/2} \int_{-\infty}^\infty e^{-V_0(y_0)/\varepsilon} \int_{\mathbb{R}_\perp^\Lambda} e^{-\langle y_\perp, Q_\perp(y_0) y_\perp \rangle/2} e^{-R_\varepsilon(y_0, y_\perp)} \, dy_\perp \, dy_0. \tag{1.5.19}$$

The idea is to view the integral over $\mathbb{R}_\perp^\Lambda$ as an expectation under the Gaussian measure $g_N(y_0)$ with covariance $Q_\perp(y_0)^{-1}$. Taking the normalisation into account, we obtain

$$\mathscr{Z} = \varepsilon^{(N-1)/2} \int_{-\infty}^\infty e^{-V_0(y_0)/\varepsilon} \sqrt{\frac{(2\pi)^{N-1}}{\det Q_\perp(y_0)}} \mathbb{E}^{g_N(y_0)}[e^{-R_\varepsilon(y_0, y_\perp)}] \, dy_0.$$

The expectation is bounded uniformly in ε and N, because the potential satisfies the quadratic lower bound (1.1.5) derived in Exercise 1.1.2. Therefore it converges to 1 as $\varepsilon \to 0$ by the dominated convergence theorem. The result then follows by one-dimensional Laplace asymptotics for the integral over y_0, since V_0 has quadratic minima in $\pm\sqrt{N}$. The fact that the error has order ε instead of $\sqrt{\varepsilon}$ is due to the fact that the potential is even in y_0.   □

Remark 1.5.15. We have not claimed that the error term $\mathcal{O}(\varepsilon)$ is (1.5.18) is uniform in N. In fact, this is indeed the case, but proving it needs a little bit more work. We will come back to this point in Section 2.7. ◇

To simplify the computation of the capacity, we will assume that the sets A and B are of the form

$$A = -B = \{y \colon |y_0 + \sqrt{N}| \leqslant \delta, y_\perp \in D_\perp\}, \tag{1.5.20}$$

where $\delta \ll \sqrt{N}$ and $D_\perp$ is a ball sufficiently large for $A \cup B$ to contain most of the mass of the invariant measure π, in the sense that $\pi((A \cup B)^c) = \mathcal{O}(\varepsilon)$. This holds for $D_\perp$ of radius of order $\sqrt{\log(\varepsilon^{-1})}$, see for instance [BDGW17, Lemma 5.9].

Proposition 1.5.16. *For A and B satisfying (1.5.20), one has*

$$\mathrm{cap}(A, B) = \frac{1}{2\pi} \sqrt{\frac{(2\pi\varepsilon)^N |\mu_0^N|}{2 \det Q_\perp(0)}}\, e^{-N/(4\varepsilon)}[1 + \mathcal{O}(\varepsilon)] = \frac{|\mu_0^N|}{2\pi} \sqrt{\frac{(2\pi\varepsilon)^N}{\prod_{k=0}^{N-1} \mu_k^N}}\, e^{-N/(4\varepsilon)}[1 + \mathcal{O}(\varepsilon)], \tag{1.5.21}$$

where the μ_k^N are the eigenvalues (1.1.3) of the Hessian of $V(0)$.

PROOF: We will apply the Dirichlet and Thomson principles with appropriate choices of h and φ. For the upper bound, we use a function depending only on the coordinate y_0, and which is simply given by the one-dimensional committor (1.5.10) derived in Exercise 1.5.4:

$$h(y) = h_0(y_0) = \frac{1}{c_0} \int_y^a e^{V_0(\xi)/\varepsilon}\, \mathrm{d}\xi\,, \qquad c_0 := \int_{-a}^a e^{V_0(\xi)/\varepsilon}\, \mathrm{d}\xi\,,$$

where $a = \sqrt{N} - \delta$ is the left boundary of B. For $|y_0| > a$, $h_0(y_0)$ is continuously extended by constant values 1 or 0. Inserting this in the Dirichlet form, we obtain

$$\mathscr{E}(h) = \frac{\varepsilon}{c_0^2} \varepsilon^{(N-1)/2} \int_{-a}^a e^{V_0(y_0)/\varepsilon} \int_{\mathbb{R}_\perp^\Lambda} e^{-\langle y_\perp, Q_\perp(y_0)y_\perp\rangle/2}\, e^{-R_\varepsilon(y_0, y_\perp)}\, \mathrm{d}y_\perp\, \mathrm{d}y_0\,.$$

Observe that the sign of V_0 has changed in the exponent with respect to (1.5.19), which means that y_0 close to 0 will now dominate the integral. Writing again the integral over $\mathbb{R}_\perp^\Lambda$ as an expectation under the Gaussian measure g_N yields an upper bound of the desired form, noting that $c_0 = e^{V_0(0)/\varepsilon}[2\pi\varepsilon/|\mu_0^N|]^{1/2}[1 + \mathcal{O}(\varepsilon)]$ and $V_0(0) = N/4$.

For the lower bound, we apply the Thomson principle with the unit flow[1]

$$\varphi(y) = \frac{1}{K} \mathbb{1}_{D_\perp}(y_\perp) e^{-\langle y_\perp, Q_\perp(0)y_\perp\rangle/2}\, e_0\,, \qquad K := \varepsilon^{(N-1)/2} \int_{D_\perp} e^{-\langle y_\perp, Q_\perp(0)y_\perp\rangle/2}\, \mathrm{d}y_\perp\,.$$

Since φ depends only on $y_\perp$ and is directed along e_0, it is indeed divergence-free. In addition, it has intensity 1 by definition of K. We obtain

$$\mathscr{D}(\varphi) = \frac{1}{\varepsilon K} \int_{-a}^a e^{V_0(y_0)/\varepsilon}\, \mathbb{E}^{g_N(0)}\big[\mathbb{1}_{D_\perp} e^{\langle \cdot, [Q_N(\varphi_0) - Q_\perp(0)]\cdot\rangle/2} e^{R_\varepsilon(y_0, y_\perp)}\big]\, \mathrm{d}y_0\,.$$

The expectation is bounded because $D_\perp$ is bounded, and equal to $1 + \mathcal{O}(\varepsilon)$ thanks to the assumption on $D_\perp$. The result then follows by computing the Gaussian integral K and performing Laplace asymptotics on the integral over y_0. □

---

[1]We could have included the quartic part of the potential in the exponent as well, yielding a better control on the decay for large $y_\perp$. This is, however, not needed if we do not care about uniformity in N.

Combining the above estimates, we obtain the following sharp asymptotics on the transition time, which is the main result of this section.

Theorem 1.5.17 (Eyring–Kramers law for the double-well situation). *Assume $\gamma > \gamma_1(N)$ and B is as in* (1.5.20). *Then*

$$\mathbb{E}^{y^*}[\tau_B] = \frac{2\pi}{|\mu_0^N|}\sqrt{\frac{|\det \operatorname{Hess} V(0)|}{\det \operatorname{Hess} V(y_-^*)}}\, \mathrm{e}^{[V(0)-V(y_-^*)]/\varepsilon}[1 + \mathcal{O}(\varepsilon)]\,, \tag{1.5.22}$$

where $\operatorname{Hess} V(y)$ *denotes the Hessian matrix of V at y.*

PROOF: The result when starting with the distribution ν_{AB} follows directly by inserting (1.5.18) and (1.5.21) in the exact relation obtained in Theorem 1.5.6. To extend this to solutions starting in y_-^*, there are two possibilities. One of them is to use Harnack inequalities, which bound the oscillation of harmonic functions, to show that $w_B(y)$ does not depend too badly on y, cf. [BEGK04, Lemma 4.6]. An alternative is to use a coupling argument as in [MOS89]. □

Note that (1.5.22) is indeed a generalisation to higher dimension of the one-dimensional expression (1.5.6) obtained in Exercise 1.5.3.

Remark 1.5.18. Similar results as (1.5.22) hold in much more general finite-dimensional situations, with a less sharp control of the error term, including situations with more than 2 wells [BEGK04]. Furthermore, the spectral gap of the generator $\mathscr{L}$ can be shown to be exponentially close to the inverse of the expected transition time (1.5.22) [BGK05].   ◇

1.6   Bibliographical notes

The system (1.0.1) of coupled diffusions was introduced in [BFG07a, BFG07b] to understand the general theory of metastability in a specific example. These works provide a number of results on the potential landscape, both for small and large coupling γ, and asymptotic results for large N. In particular, Proposition 1.1.1 is [BFG07a, Proposition 2.2].

General results on solutions of SDEs can be found in the monographs [McK69, Øks85, KS91, Mao97]. The use of Lyapunov functions to prove non-explosion, Harris recurrence, and various ergodicity results has been developed by Meyn and Tweedie in a series of works [MT92, MT93b, MT93c], as well as the monograph [MT93a].

The theory of large deviations for SDEs is developed by Freidlin and Wenzell in the monograph [FW98]. Other general monographs on large deviations include [DZ98, DS89].

The potential-theoretic approach to metastability was mainly developed in [BEGK02] for Markov chains, and in [BEGK04, BGK05, Eck05] for reversible diffusions. A comprehensive account of the potential-theoretic approach can be found in the monograph [BdH15]. Short overviews are also found in [Slo12b, Ber13]. The Thomson principle is proved (in a more general, non-reversible setting) in [LMS19].

Another successful approach to sharp asymptotics for metastable transition times is based on semiclassical analysis of the Witten Laplacian, and was initiated in [HKN04, HN05]. Extensions can be found, e.g., in [LP10, LP11, LPNV13].

There exist several extensions of the results on the Eyring–Kramers formula presented here. The case of saddles with vanishing Hessian determinant has been considered in [BG10]. Situations with many degeneracies due to symmetries have been considered in [BD16a] for markovian jump processes, and [Dut15, BD16b] for diffusions. Uniformity in N of the error terms in the Eyring–Kramers formula for the system (1.0.1) was obtained in [BBM10]. Results on the spectral gap that are uniform in N have been obtained in [DGLP17].

Chapter 2

Allen–Cahn SPDE in one space dimension

Consider the formal limit of the system (1.0.1) of coupled diffusions as $N \to \infty$ with $\gamma \sim N^2$. Given a parameter $L > 0$, that we will choose below as a function of the limit of γ/N^2, we define a function $\phi(t, x)$ by setting

$$y_t^i = \phi\left(t, \frac{i}{N}L\right), \qquad i \in \Lambda$$

and interpolating linearly (or with some higher-order polynomials) between lattice points. The discrete Laplacian then formally satisfies, for $x = (i/N)L \in (0, L]$,

$$y_t^{i+1} - 2y_t^i + y_t^{i-1} = \phi\left(t, x + \frac{L}{N}\right) - 2\phi(t, x) + \phi\left(t, x - \frac{L}{N}\right)$$
$$= \frac{L^2}{N^2}\partial_{xx}\phi(t, x) + \mathcal{O}\left(\frac{L^4}{N^4}\right).$$

Taking L such that

$$L^2 = \lim_{N \to \infty} \frac{2N^2}{\gamma},$$

we see that the SDE (1.0.1) converges formally, as $N \to \infty$, to the equation

$$\partial_t \phi(t, x) = \partial_{xx}\phi(t, x) + \phi(t, x) - \phi(t, x)^3 + \sqrt{2\varepsilon}\xi(t, x), \qquad (2.0.1)$$

where $\xi(t, x)$ is a stochastic process called *space-time white noise*, which we will have to properly define. In what follows, we will write $\Delta\phi$ instead of $\partial_{xx}\phi$, even though x is one-dimensional, because the same notation will apply for higher-dimensional x. Note in particular that for the critical value $\gamma_1(N)$ of γ obtained in Proposition 1.1.1, we have

$$\lim_{N \to \infty} \frac{2N^2}{\gamma_1(N)} = (2\pi)^2,$$

indicating that the value $L = 2\pi$ will play a special role.

Remark 2.0.1. The equation with a negative coefficient in front of ϕ

$$\partial_t \phi(t, x) = \Delta\phi(t, x) - m^2\phi(t, x) - \phi(t, x)^3 + \sqrt{2\varepsilon}\xi(t, x) \qquad (2.0.2)$$

(where $m^2 \geqslant 0$) is called the Φ^4 *model* [GJ68, GJS75, GJ81] (massive Φ^4 model if $m > 0$) or *stochastic quantisation equation* [PW81], and plays an important role in Quantum Field Theory. The solution theory is the same for (2.0.1) and (2.0.2), but their long-time behaviour is very different, since (2.0.1) displays metastability while (2.0.2) does not. $\diamond$
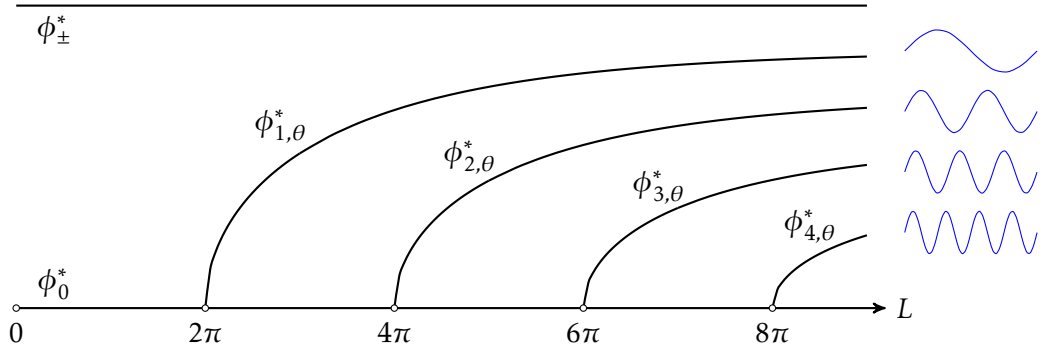
Figure 2.1 – Schematic bifurcation diagram of the deterministic Allen–Cahn equation (2.1.1). The vertical coordinate represents roughly the amplitude of the stationary solution.

2.1   Deterministic dynamics

We start by briefly analysing (2.0.1) in the deterministic case $\varepsilon = 0$, where it takes the form of the PDE

$$\partial_t \phi(t,x) = \Delta\phi(t,x) + \phi(t,x) - \phi(t,x)^3 \,. \tag{2.1.1}$$

This equation is commonly known as *Allen–Cahn equation*, though one finds other names in the literature, including *Chaffee–Infante equation*, or *real Ginzburg–Landau equation*.

In the present setting, x belongs to the scaled circle $\Lambda = \mathbb{R}/(L\mathbb{Z})$, which implicitly implies that we consider (2.1.1) with periodic boundary conditions.

A first useful observation is that the right-hand side of (2.1.1) derives again from a potential, obtained as the continuum limit of the potential (1.1.2), which is given by

$$V(\phi) := \int_0^L \left[\frac{1}{2}\|\nabla\phi(x)\|^2 + \frac{1}{4}\left(\phi(x)^2 - 1\right)^2 \right]. \tag{2.1.2}$$

Here we have written $\nabla\phi(x)$ instead of $\partial_x\phi(x)$, to have a notation also valid in higher dimensions. Indeed, the Gâteaux derivative of V at ϕ in the direction ψ is given by

$$\begin{aligned}
\nabla_\psi V(\phi) &:= \frac{\partial}{\partial\lambda} V(\phi + \lambda\psi)\Big|_{\lambda=0} \\
&= \int_0^L \left[\nabla\phi(x)\cdot\nabla\psi(x) - \phi(x)\psi(x) + \phi(x)^3\psi(x) \right] \mathrm{d}x \\
&= -\langle \Delta\phi + \phi - \phi^3, \psi\rangle_{L^2} \,,
\end{aligned}$$

where we have used integration by parts and the periodic boundary conditions in the last step. In particular, this shows that stationary solutions of (2.1.1) are critical points of V.

The role of the Hessian of V at ϕ is played by the bilinear form mapping two periodic functions ψ_1, ψ_2 to

$$\begin{aligned}
\nabla^2_{\psi_1,\psi_2} V(\phi) &:= \frac{\partial^2}{\partial\lambda_1\partial\lambda_2} V(\phi + \lambda_1\psi_1 + \lambda_2\psi_2)\Big|_{\lambda_1=\lambda_2=0} \\
&= \int_0^L \left[\nabla\psi_1(x)\cdot\nabla\psi_2(x) - \psi_1(x)\psi_2(x) + 3\phi(x)^2\psi_1(x)\psi_2(x) \right] \mathrm{d}x \\
&= \langle \psi_1, [-\Delta - 1 + 3\phi(\cdot)^2]\psi_2\rangle_{L^2} \,. \tag{2.1.3}
\end{aligned}$$

This reflects the fact that the linearisation of (2.1.1) around a stationary solution ϕ^* is given by the variational equation

$$\partial_t\psi(t,x) = \Delta\psi(t,x) + [1 - 3\phi^*(x)^2]\psi(t,x) \,. \tag{2.1.4}$$
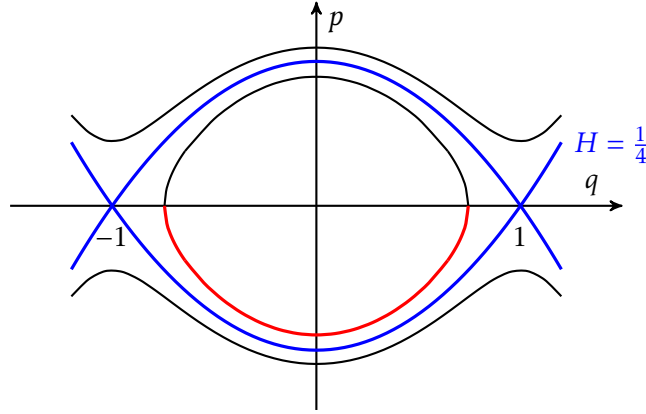
Figure 2.2 – Solutions of the Hamilton equations (2.1.5). Closed orbits correspond to periodic boundary conditions, while the orbit shown in red corresponds to Neumann boundary conditions. The curves $\{H = \frac{1}{4}\}$ are separatrices reached in the limit of the period going to infinity.

Determining the stability of ϕ^* is equivalent to solving the Sturm–Liouville problem consisting in computing the spectrum of (2.1.4) with periodic boundary conditions.

The PDE (2.1.1) has three obvious stationary solutions, which are constant in space, and natural analogues of the constant stationary solutions of the discrete system. We will denote by $\phi_\pm^*$ the constant solutions equal to ± 1, and by ϕ_0^* the constant solution equal to 0. The following result describes the set of all stationary solutions and their stability. The bifurcation diagram is sketched in Figure 2.1.

Proposition 2.1.1. *The stationary solutions $\phi_\pm^*$ are always stable. The solution ϕ_0^* has one unstable direction if $0 < L < 2\pi$, and more generally $2k + 1$ unstable directions if $2k\pi < L < 2(k+1)\pi$ for any $k \in \mathbb{N}_0$. In addition, at every multiple of 2π, a one-parameter family of non-constant solutions bifurcates from the origin. These solutions are unstable, and of the form $\phi_{k,\theta}^*(x) = \phi_{k,0}^*(x + \theta)$.*

PROOF: Setting $\phi = \pm 1$ in (2.1.3), we find that the Hessian of V at $\phi_\pm^*$ is given by $-\Delta + 2$. This has eigenvalues

$$\nu_k = \left(\frac{2k\pi}{L}\right)^2 + 2, \qquad k \in \mathbb{Z},$$

which are all positive. Therefore, the Sturm–Liouville equation (2.1.4) has negative spectrum, and $\phi_\pm^*$ are stable. In the case of ϕ_0^*, we obtain a Hessian equal to $-\Delta - 1$, which has eigenvalues

$$\mu_k = \left(\frac{2k\pi}{L}\right)^2 - 1, \qquad k \in \mathbb{Z}.$$

Therefore, μ_0 is always negative, while two additional μ_k become negative whenever L exceeds $2k\pi$. The other stationary solutions are periodic solutions of

$$\phi''(x) = \phi(x)^3 - \phi(x)$$

of period L. This second-order ODE is equivalent to the Hamiltonian system

$$q'(x) = p(x)$$
$$p'(x) = q(x)^3 - q(x), \tag{2.1.5}$$

which derives from the Hamiltonian

$$H(p,q) = \frac{1}{2}p^2 + \frac{1}{2}q^2 - \frac{1}{4}q^4.$$

The energy H is constant along solutions of (2.1.5), and therefore orbits are contained in level curves of H. Some level curves are shown in Figure 2.2. Non-constant stationary solutions of (2.1.1) correspond to closed level curves whose period is of the form L/k for some $k \in \mathbb{N}$ (as they can be tracked multiple times). Expressing $p = q'$ in terms of H and q and integrating, one obtains that the period of the closed level curve $\{H(q, p) = E\}$ is given by

$$T(E) = 2 \int_{q_-(E)}^{q_+(E)} \frac{\mathrm{d}q}{\sqrt{2E - q^2 + \frac{1}{2}q^4}}, \qquad q_\pm(E) = \pm\sqrt{1 - \sqrt{1 - 4E}},$$

which is a Jacobi elliptic integral that is known to be an increasing function of E, starting at 2π for $E = 0$.

It follows that whenever L is a multiple of 2π, a new family of non-constant stationary solutions bifurcates from ϕ_0^*. This is a family of solutions because one can choose the origin $x = 0$ anywhere on the closed level curve. The fact that these solutions are unstable has topological reasons (conservation of the degree of stationary points of functions depending on a parameter), and can be checked locally by a bifurcation analysis. $\qquad\square$

In what follows, we will again mostly concentrate for simplicity on the case $0 < L < 2\pi$, when V is a double-well potential with two local minima located at $\phi_\pm^*$ and a saddle at the origin. However, the results hold in more generality.

Exercise 2.1.2. Discuss the case of zero-flux Neumann boundary conditions

$$\frac{\partial \phi}{\partial x}(0) = \frac{\partial \phi}{\partial x}(L) = 0.$$

This setting corresponds to the orbit shown in red in Figure 2.2. Derive the bifurcation diagram representing the stationary solutions and their stability as a function of L. $\qquad\clubsuit$

2.2   Space-time white noise

We now turn to the precise definition of the space-time white noise process ξ formally introduced in (2.0.1). It should have the following properties:

- each $\xi(t, x)$ should be a Gaussian random variable;
- each $\xi(t, x)$ should be centred;
- the values of ξ at different space-time points should be independent, and thus uncorrelated, which is sometimes written informally as

$$\mathbb{E}[\xi(t, x)\xi(s, y)] = \delta(t - s)\delta(x - y). \tag{2.2.1}$$

It turns out that there is no random function having the above properties, but that there exists a random Schwartz distribution doing the job, i.e., one can consider ξ as a random linear functional acting on test functions. We will denote by $\mathscr{H}$ be the Hilbert space $L^2(\mathbb{R} \times \Lambda)$, where we recall that $\Lambda = \mathbb{R}/(L\mathbb{Z})$. Let $\mathscr{S}'(\mathscr{H})$ be the space of Schwartz distributions, and denote by $\langle \zeta, \varphi \rangle$ the duality pairing between a distribution $\zeta \in \mathscr{S}'(\mathscr{H})$ and a test function $\varphi \in \mathscr{H}$.

Definition 2.2.1 (Space-time white noise). *Space-time white noise on $\mathbb{R} \times \Lambda$ is a random distribution ξ on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ such that for any smooth test function $\varphi \in \mathscr{H}$, $\langle \xi, \varphi \rangle$ is a centred Gaussian random variable of variance $\|\varphi\|_{\mathscr{H}}^2$, and the covariances are given by*

$$\mathbb{E}[\langle \xi, \varphi_1 \rangle \langle \xi, \varphi_2 \rangle] = \langle \varphi_1, \varphi_2 \rangle_{\mathscr{H}} \tag{2.2.2}$$

for any two smooth test functions $\varphi_1, \varphi_2 \in \mathscr{H}$.

The fact that such a process ξ exists is a consequence of general results on Gaussian measure theory. Details can be found, for instance, in Chapter 3 of Martin Hairer's lecture notes [Hai09].

Note that we recover the expression (2.2.1) by formally replacing the test functions φ_1 and φ_2 by Dirac distributions. Let us consider a couple more examples.

Example 2.2.2. Let $A_1, A_2 \subset \Lambda$ be open sets, fix $T_1, T_2 > 0$ and take

$$\varphi_i(t,x) = \mathbb{1}_{\{0 \leqslant t \leqslant T_i\}} \mathbb{1}_{\{x \in A_i\}}, \qquad i = 1, 2.$$

These test functions are not smooth, but can be obtained as limits of smooth functions. Define

$$W_{T_i}^{A_i} = \langle \xi, \varphi_i \rangle = \int_0^{T_i} \int_{A_i} \xi(t,x) \, dx \, dt \qquad i = 1, 2$$

(where the last integral is of course formal). Then (2.2.2) takes the form

$$\mathbb{E}[W_{T_1}^{A_1} W_{T_2}^{A_2}] = (T_1 \wedge T_2)|A_1 \cap A_2|,$$

where $|\cdot|$ denotes the Lebesgue measure of a set. Thus the $W_{T_i}^{A_i}$ behave like Brownian motions, which are independent if the sets A_1 and A_2 are disjoint. ◇

Example 2.2.3. Let $\{e_k\}_{k \in \mathbb{Z}}$ be an orthonormal basis of $L^2(\Lambda)$, for instance a Fourier basis, fix $T_1, T_2 > 0$ and take

$$\varphi_i(t,x) = \mathbb{1}_{\{0 \leqslant t \leqslant T_i\}} e_{k_i}(x), \qquad i = 1, 2.$$

Let

$$W_{T_i}^{(k_i)} = \langle \xi, \varphi_i \rangle = \int_0^{T_i} \langle \xi(t,\cdot), e_{k_i} \rangle \, dt \qquad i = 1, 2.$$

Then (2.2.2) yields

$$\mathbb{E}[W_{T_1}^{(k_1)} W_{T_2}^{(k_2)}] = (T_1 \wedge T_2)\delta_{k_1 k_2},$$

showing that the $W_T^{(k_i)}$ are independent Brownian motions. ◇

An important property of space-time white noise is related to scaling, in a way that should be reminiscent of Brownian motion. Given $\tau, \lambda > 0$, define a scaling operator acting on test functions by

$$(\mathscr{S}^{\tau,\lambda} \varphi)(t,x) := \frac{1}{\tau \lambda} \varphi\left(\frac{t}{\tau}, \frac{x}{\lambda}\right). \tag{2.2.3}$$

Then we define a scaled version $\xi_{\tau,\lambda}$ of space-time white noise by

$$\langle \xi_{\tau,\lambda}, \varphi \rangle = \langle \xi, \mathscr{S}^{\tau,\lambda} \varphi \rangle \tag{2.2.4}$$

for any test function φ.

Exercise 2.2.4. Show that if ξ is replaced by a smooth function f in (2.2.4), then one simply has $f_{\tau,\lambda}(t,x) = f(\tau t, \lambda x)$. ♣

Proposition 2.2.5 (Scaling property of space-time white noise). *We have*

$$\xi_{\tau,\lambda} \stackrel{\text{law}}{=} \frac{1}{\sqrt{\tau \lambda}} \xi,$$

where $\stackrel{\text{law}}{=}$ *denotes equality in distribution.*

Proof: Both processes are Gaussian and centred. Therefore, it suffices to show that they have the same covariance. Given two compactly supported test functions φ_1, φ_2, we have by (2.2.2)

$$
\begin{aligned}
\mathbb{E}[\langle \xi_{\tau,\lambda}, \varphi_1 \rangle \langle \xi_{\tau,\lambda}, \varphi_2 \rangle] &= \mathbb{E}[\langle \xi, \mathscr{S}^{\tau,\lambda} \varphi_1 \rangle \langle \xi, \mathscr{S}^{\tau,\lambda} \varphi_2 \rangle] \\
&= \int_{-\infty}^{\infty} \int_{\Lambda} (\mathscr{S}^{\tau,\lambda} \varphi_1)(t,x)(\mathscr{S}^{\tau,\lambda} \varphi_2)(t,x) \, dx \, dt \\
&= \frac{1}{\tau \lambda} \langle \varphi_1, \varphi_2 \rangle_{\mathscr{H}} .
\end{aligned}
\tag{2.2.5}
$$

This is indeed the covariance of $\xi / \sqrt{\lambda \tau}$. $\qquad\square$

In what follows, it will be important to find some functions spaces to which the relevant objects belong. A first important family of such spaces are scaled versions of the classical Hölder spaces. The scaling comes from the fact that it turns out to be useful to work with the *parabolic distance*

$$
\|(t,x) - (s,y)\|_{\mathfrak{s}} := |t-s|^{1/2} + |x-y|
$$

between space-time points, owing to the parabolic nature of the SPDE (2.0.1), which is first-order in time and second-order in space. This is not a genuine distance because it fails to satisfy the triangle inequality, but this will not be important.

Definition 2.2.6 (Parabolic Hölder space). *For $0 < \alpha < 1$, the parabolic Hölder space $\mathscr{C}_{\mathfrak{s}}^{\alpha}(\mathbb{R} \times \Lambda)$ consists of all continuous functions $f : \mathbb{R} \times \Lambda \to \mathbb{R}$ such that for every compact $\mathfrak{K} \subset \mathbb{R} \times \Lambda$*

$$
\|f\|_{\mathscr{C}_{\mathfrak{s}}^{\alpha}(\mathfrak{K})} := \sup_{z \in \mathfrak{K}} |f(z)| + \sup_{\substack{z_1, z_2 \in \mathfrak{K} \\ z_1 \neq z_2}} \frac{|f(z_1) - f(z_2)|}{\|z_1 - z_2\|_{\mathfrak{s}}^{\alpha}} < \infty .
$$

There exists an extension of these spaces to negative index α, which are a scaled version of a particular class of Besov spaces, commonly denoted $\mathscr{B}_{\infty,\infty}^{\alpha}$. Since these are closely related to the above Hölder spaces, as we will see in the next section, we keep the notation $\mathscr{C}_{\mathfrak{s}}^{\alpha}$ as in [Hai14]. To define these spaces, we slightly adapt the notation (2.2.3) to allow for zooming in at any space-time point, setting

$$
(\mathscr{S}_{t,x}^{\lambda} \varphi)(s,y) := \frac{1}{\lambda^3} \varphi \left(\frac{s-t}{\lambda^2}, \frac{y-x}{\lambda} \right).
\tag{2.2.6}
$$

Given a positive integer r, we further let B_r be the set of smooth test functions $\varphi : \mathbb{R} \times \Lambda \to \mathbb{R}$ supported in the unit $\|\cdot\|_{\mathfrak{s}}$-ball, such that

$$
\|\varphi\|_{\mathscr{C}^r} := \sup_{k : |k|_{\mathfrak{s}} \leqslant r} \sup_{z \in \mathbb{R} \times \Lambda} |D^k \varphi(z)| \leqslant 1 .
\tag{2.2.7}
$$

Here $k = (k_0, k_1) \in \mathbb{N}_0^2$ is a multiindex, $|k|_{\mathfrak{s}} := 2k_0 + k_1$ is its scaled norm, and $D^k := \partial_t^{k_0} \partial_x^{k_1}$.

Definition 2.2.7 (Parabolic negative index Besov–Hölder space). *For $\alpha < 0$, the space $\mathscr{C}_{\mathfrak{s}}^{\alpha}(\mathbb{R} \times \Lambda)$ consists of all Schwartz distributions $\zeta \in \mathscr{S}'(\mathbb{R} \times \Lambda)$ such that for every compact $\mathfrak{K} \subset \mathbb{R} \times \Lambda$*

$$
\|\zeta\|_{\mathscr{C}_{\mathfrak{s}}^{\alpha}(\mathfrak{K})} := \sup_{z \in \mathfrak{K}} \sup_{\varphi \in B_r} \sup_{\lambda \in (0,1]} \left| \frac{\langle \zeta, \mathscr{S}_z^{\lambda} \varphi \rangle}{\lambda^{\alpha}} \right| < \infty ,
$$

where $r = \lceil -\alpha \rceil$.

In the case where $\zeta = \xi$ is space-time white noise, it follows from (2.2.5) that

$$\mathbb{E}[\langle \xi, \mathscr{S}_z^\lambda \varphi \rangle^2] = \frac{1}{\lambda^3} \|\varphi\|_{\mathscr{H}}^2 . \tag{2.2.8}$$

Since for Gaussian random variables, second moments determine all other moments, this suggests that ξ might belong to $\mathscr{C}_{\mathfrak{s}}^{-3/2}$. The following important result shows that this is almost the case.

Theorem 2.2.8 (Regularity of space-time white noise). *We have $\xi \in \mathscr{C}_{\mathfrak{s}}^{-3/2-\kappa}$ for any $\kappa > 0$.*

SKETCH OF PROOF: We follow the argument outlined in [CW17, Theorem 2.7]. First recall the fact that if X is a Gaussian random variable, then there exists, for every integer $p > 0$, a finite constant C_p such that

$$\mathbb{E}[|X|^p] \leqslant C_p \big(\mathbb{E}[X^2]\big)^{p/2} \tag{2.2.9}$$

(see also Exercise 3.2.4). It thus follows from (2.2.8) that for every integer $p > 0$,

$$\mathbb{E}\Big[\big|\langle \xi, \mathscr{S}_z^\lambda \varphi \rangle\big|^p\Big] \leqslant C_p \lambda^{-3p/2} \|\varphi\|_{\mathscr{H}}^p . \tag{2.2.10}$$

Fix a compact $\mathfrak{K} \subset \mathbb{R} \times \Lambda$, and let $\bar{\mathfrak{K}}$ be the set of points in $\mathbb{R} \times \Lambda$ at parabolic distance at most 1 from $\mathfrak{K}$ (called the 1-*fattening* of $\mathfrak{K}$). Given $k \in \mathbb{N}_0$, we introduce a dyadic lattice discretisation of $\bar{\mathfrak{K}}$ on scale 2^{-k}, given by

$$\mathfrak{K}_k = (2^{-2k}\mathbb{Z} \times 2^{-k}\mathbb{Z}) \cap \bar{\mathfrak{K}} .$$

The crucial observation is that one can show that for any $\alpha \in \mathbb{R}$, there exists a test function φ and a constant C such that

$$\|\xi\|_{\mathscr{C}_{\mathfrak{s}}^\alpha(\mathfrak{K})} \leqslant C \sup_{k \geqslant 0} \sup_{z \in \mathfrak{K}_k} 2^{k\alpha} \big|\langle \xi, \mathscr{S}_z^{2^{-k}} \varphi \rangle\big| .$$

The fact that φ can be chosen independent of k is not trivial, but there are various ways of constructing such a φ (using for instance wavelets, or Paley–Littlewood blocks). Bounding the suprema by sums and taking the pth power, we arrive at

$$\Big(\|\xi\|_{\mathscr{C}_{\mathfrak{s}}^\alpha(\mathfrak{K})}\Big)^p \leqslant C^p \sum_{k \geqslant 0} \sum_{z \in \mathfrak{K}_k} 2^{k\alpha p} \big|\langle \xi, \mathscr{S}_z^{2^{-k}} \varphi \rangle\big|^p .$$

Note that so far, the argument is purely deterministic. At this point, we take expectations and use (2.2.10) to arrive at

$$\mathbb{E}\Big[\big(\|\xi\|_{\mathscr{C}_{\mathfrak{s}}^\alpha(\mathfrak{K})}\big)^p\Big] \leqslant \overline{C}_p \sum_{k \geqslant 0} 2^{3k} 2^{k\alpha p} 2^{3kp/2}$$

for some constant $\overline{C}_p$, where we have used the fact that $\mathfrak{K}_k$ contains of the order of 2^{3k} points. If $\alpha < -\frac{3}{2} - \frac{3}{p}$, one can sum the geometric series, yielding the bound

$$\mathbb{E}\Big[\big(\|\xi\|_{\mathscr{C}_{\mathfrak{s}}^\alpha(\mathfrak{K})}\big)^p\Big] < \infty \qquad \forall \alpha < -\frac{3}{2} - \frac{3}{p} .$$

It then follows from a version of Kolmogorov's continuity theorem that there exists a modification of ξ with bounded $\mathscr{C}_{\mathfrak{s}}^\alpha(\mathfrak{K})$-norm. Taking p large enough yields the result. $\qquad\square$

2.3   The stochastic heat equation

Before turning to the Allen–Cahn SPDE (2.0.1), we consider the linear *stochastic heat equation* on $\mathbb{R}_+ \times \Lambda$, given by

$$\partial_t \phi(t,x) = \Delta \phi(t,x) + \xi(t,x) \,. \tag{2.3.1}$$

Here we have taken $2\varepsilon = 1$, since the case of general ε can easily be recovered by scaling.

First recall that the *heat equation*

$$\partial_t \phi(t,x) = \Delta \phi(t,x) \,, \qquad \phi(0,x) = \phi_0(x)$$

admits the solution $\phi(t,x) = (e^{t\Delta} \phi_0)(x)$, where $e^{t\Delta}$ is the *heat semigroup* defined by

$$(e^{t\Delta} \phi_0)(x) := \int_\Lambda P(t,x-y)\phi_0(y)\,\mathrm{d}y \,.$$

Here $P(t,x)$ is the *heat kernel*. If Λ were equal to $\mathbb{R}$, it would be given by

$$P_\mathbb{R}(t,x) := \frac{1}{\sqrt{4\pi t}}\, e^{-x^2/(4t)} \, \mathbb{1}_{\{t>0\}} \,.$$

Since we are working in $\Lambda = \mathbb{R}/(L\mathbb{Z})$, however, we have to use a periodicised version of the heat kernel, defined by

$$P(t,x) = P_\Lambda(t,x) := \sum_{k\in\mathbb{Z}} P_\mathbb{R}(t,x-kL) \,. \tag{2.3.2}$$

Next recall that the *forced heat equation*

$$\partial_t \phi(t,x) = \Delta \phi(t,x) + f(t,x) \,, \qquad \phi(0,x) = \phi_0(x) \,,$$

where f is some smooth forcing, can be solved by the method of variation of constant, also known as *Duhamel principle* in the theory of PDEs. The result is

$$\phi(t,x) = (e^{t\Delta} \phi_0)(x) + \int_0^t (e^{(t-s)\Delta} f)(s,x)\,\mathrm{d}s$$

$$= (e^{t\Delta} \phi_0)(x) + \int_0^t \int_\Lambda P(t-s,x-y)f(s,y)\,\mathrm{d}y\,\mathrm{d}s \,. \tag{2.3.3}$$

Note that the second term on the right-hand side is the space-time convolution $(P * f)(t,x)$. Therefore, we will sometimes write (2.3.3) in the short form

$$\phi = P\phi_0 + P * f \,,$$

keeping in mind that the star denotes space-time convolution, while no star stands for convolution in space only.

It is thus natural to define the solution of the stochastic heat equation (2.3.1) by

$$\phi = P\phi_0 + P * \xi \,, \tag{2.3.4}$$

where $(P * \xi)(t,x)$ is called the *stochastic convolution*. It seems reasonable to expect that this definition makes sense, since the stochastic convolution looks similar to pairing ξ with a test function. One should be careful, however, with the fact that P has a singularity at the origin, and is thus not strictly speaking a test function.

We will discuss two equivalent ways of analysing the stochastic convolution. The first one is based on Fourier series and fractional Sobolev spaces, while the second one uses Besov– Hölder spaces and the so-called Schauder estimate.

2.3.1 Fourier series and fractional Sobolev spaces

Define an orthonormal basis $(e_k)_{k \in \mathbb{Z}}$ of $L^2(\Lambda)$ by

$$e_k(x) := \sqrt{\frac{2}{L}} \cos\left(\frac{2k\pi x}{L}\right) \quad \text{if } k > 0 \,, \qquad e_0(x) := \frac{1}{\sqrt{L}} \,, \qquad e_k(x) := \sqrt{\frac{2}{L}} \sin\left(\frac{2k\pi x}{L}\right) \quad \text{if } k < 0 \,.$$

(2.3.5)

These basis functions satisfy the eigenvalue problem

$$\Delta e_k = -\lambda_k e_k \,, \qquad \lambda_k := \left(\frac{2k\pi}{L}\right)^2 \,.$$

(2.3.6)

We will write the expansion of $\phi \in L^2(\Lambda)$ in the Fourier basis (2.3.5)

$$\phi(x) = \sum_{k \in \mathbb{Z}} \hat{\phi}_k e_k(x) \,.$$

(2.3.7)

Remark 2.3.1. It might seem more convenient to use a complex Fourier basis of the form $e_k(x) = e^{2ik\pi x/L}/\sqrt{L}$. While this simplifies certain computations involving nonlinear terms, it has the drawback that the modes k and $-k$ become correlated. This is not really a problem, but makes it simpler to work with real Fourier series, at least in the case of linear equations.   ◇

Fourier series are intimately related to the scale of *fractional Sobolev spaces* (also called *Bessel potential spaces*).

Definition 2.3.2 (Fractional Sobolev spaces). *For $s \geqslant 0$, the fractional Sobolev space $H^s(\Lambda)$ is given by the subspace of functions $\phi \in L^2(\Lambda)$ such that*

$$\|\phi\|_{H^s}^2 := \sum_{k \in \mathbb{Z}} (1 + k^2)^s \hat{\phi}_k^2 < \infty \,.$$

(2.3.8)

In particular, $H^0(\Lambda) = L^2(\Lambda)$. For $s < 0$, $H^s(\Lambda)$ is the closure of $L^2(\Lambda)$ under the norm (2.3.8).

A first use of fractional Sobolev spaces is that they allow to quantify the regularising effect of the heat semigroup.

Lemma 2.3.3. *For any $s \geqslant 0$, there exists a constant $C(s) < \infty$ such that*

$$\|e^{t\Delta} \phi_0\|_{H^s} \leqslant (1 + C(s) t^{-s/2}) \|\phi_0\|_{L^2}$$

holds for all $\phi_0 \in L^2(\Lambda)$.

Exercise 2.3.4. Prove Lemma 2.3.3, using the Fourier representation of $e^{t\Delta} \phi_0$ and the fact that the map $x \mapsto x^s e^{-x}$ is bounded uniformly in $x > 0$ by a constant depending on s.   ♣

Projecting the stochastic heat equation (2.3.1) on the basis function e_k, and using the observation on $\langle \xi, \mathbb{1}_{\{0 \leqslant t \leqslant T\}} e_k \rangle$ made in Example 2.2.3, we obtain that each Fourier mode evolves according to the SDE

$$d\hat{\phi}_k = -\lambda_k \hat{\phi}_k \, dt + dW_t^{(k)} \,,$$

where the $(W_t^{(k)})_{t \geqslant 0}$ are independent Wiener processes. The solution is given by

$$\hat{\phi}_k(t) = e^{-\lambda_k t} \hat{\phi}_k(0) + \int_0^t e^{-\lambda_k(t-s)} dW_s^{(k)} \,.$$

Hence $\hat{\phi}_0(t)$ is a Brownian motion, while all other $\hat{\phi}_k(t)$ are Orstein–Uhlenbeck processes. The stochastic convolution can thus be defined as

$$(P * \xi)(t,x) = \sum_{k\in\mathbb{Z}} \int_0^t e^{-\lambda_k(t-s)} \, \mathrm{d}W_s^{(k)} e_k(x), \qquad (2.3.9)$$

which together with the initial-condition term $e^{t\Delta}\phi_0$ defines the so-called *mild solution* of the stochastic heat equation. Using Itô's isometry, we obtain

$$\mathbb{E}\Big[\|(P * \xi)(t,\cdot)\|_{H^s}^2\Big] = t + \sum_{k\in\mathbb{Z}^*} (1+k^2)^s \frac{1 - e^{-2\lambda_k t}}{2\lambda_k},$$

which is finite for finite t for all $s < 1/2$. In fact, we have the following result, which is a special case of [Hai09, Theorem 5.13].

Theorem 2.3.5 (Sobolev regularity of the stochastic convolution). *The stochastic convolution* $(P * \xi)(t,\cdot)$ *belongs to $H^s(\Lambda)$ for every $s < \frac{1}{2}$ and $t \in \mathbb{R}_+$.*

Exercise 2.3.6. Compute the covariance $\mathbb{E}\big[(P * \xi)(t,x)(P * \xi)(s,y)\big]$ of the stochastic convolution (2.3.9). **Hint:** It may help to distinguish the cases $t > s$, $t = s$ and $t < s$. ♣

2.3.2 Besov–Hölder spaces and Schauder estimate

It is also possible to quantify the Hölder regularity of the stochastic convolution. One way of doing this is to use Theorem 2.2.8 on the regularity of space-time white noise in conjunction with a regularising estimate for the heat semigroup, called *Schauder estimate*. To state this result in full generality, we extend the definition of parabolic Hölder spaces given in Definition 2.2.6 to exponents $\alpha > 1$ in the following way.

Definition 2.3.7 (Parabolic Hölder spaces of positive index). *Let $\alpha \geqslant 0$. The space $\mathscr{C}_{\mathfrak{s}}^\alpha(\mathbb{R}\times\Lambda)$ consists of all $f : \mathbb{R}\times\Lambda \to \mathbb{R}$ such that there exist polynomials $\{P_z\}_{z\in\mathbb{R}\times\Lambda}$ of parabolic degree less than α, such that for every compact $\mathfrak{K} \subset \mathbb{R}\times\Lambda$*

$$\|f\|_{\mathscr{C}_{\mathfrak{s}}^\alpha(\mathfrak{K})} := \sup_{z\in\mathfrak{K}} \sup_{\varphi\in B_0} \sup_{\lambda\in(0,1]} \left|\frac{\langle f - P_z, \mathscr{S}_z^\lambda\varphi\rangle}{\lambda^\alpha}\right| < \infty, \qquad (2.3.10)$$

where B_0 is defined right before (2.2.7).

At first sight, this definition may seem redundant with Definition 2.2.6 when $0 < \alpha < 1$, but it is in fact equivalent. To see this, consider the case $z = (0,0)$, and take $P_0(t,x) = f(0,0)$. Then

$$\langle f - P_0, \mathscr{S}_z^\lambda\varphi\rangle = \int_{\mathbb{R}} \int_\Lambda [f(t,x) - f(0,0)] \frac{1}{\lambda^3} \varphi\Big(\frac{t}{\lambda^2}, \frac{x}{\lambda}\Big) \mathrm{d}x\,\mathrm{d}t$$

$$= \int_{\mathbb{R}} \int_\Lambda [f(\lambda^2 t, \lambda x) - f(0,0)] \varphi(t,x)\,\mathrm{d}x\,\mathrm{d}t,$$

which has indeed order λ^α if f is α-Hölder in the parabolic norm. For $\alpha > 1$, one can easily check that P_z must be the Taylor expansion of f at z to parabolic order $\lfloor\alpha\rfloor$.

Exercise 2.3.8. Prove that if $f \in \mathscr{C}_{\mathfrak{s}}^\alpha$ and $k \in \mathbb{N}_0^2$ is a multiindex, then $\mathrm{D}^k f \in \mathscr{C}_{\mathfrak{s}}^{\alpha - |k|_{\mathfrak{s}}}$, where derivatives are understood in the sense of distributions if $\alpha < |k|_{\mathfrak{s}}$. **Hint:** Show that

$$\langle \mathrm{D}^k(f - P_z), \mathscr{S}_z^\lambda\varphi\rangle = -\frac{1}{\lambda^{|k|_{\mathfrak{s}}}} \langle f - P_z, \mathscr{S}_z^\lambda \mathrm{D}^k\varphi\rangle$$

using integration by parts. ♣

The main result quantifying the regularising properties of the heat semigroup is the following Schauder estimate.

Theorem 2.3.9 (Schauder estimate)**.** *For every $\alpha \in \mathbb{R} \setminus \mathbb{Z}$ and $t > 0$, there exists a constant C such that for every $\zeta \in \mathscr{C}_{\mathfrak{s}}^{\alpha}(\mathbb{R}_+ \times \Lambda)$, one has*

$$\|P * \zeta\|_{\mathscr{C}_{\mathfrak{s}}^{\alpha+2}([0,t] \times \Lambda)} \leqslant C \|\zeta\|_{\mathscr{C}_{\mathfrak{s}}^{\alpha}([0,t] \times \Lambda)} .$$

Therefore $P * \zeta \in \mathscr{C}_{\mathfrak{s}}^{\alpha+2}(\mathbb{R}_+ \times \Lambda)$. The heat semigroup is said to be regularising of parabolic order 2.

Remark 2.3.10. This result is in general *not* true if α is an integer. In fact, our definition of the spaces $\mathscr{C}_{\mathfrak{s}}^{\alpha}$ implies a kind of discontinuity at integer values of α. For instance, α-Hölder functions approach continuously differentiable functions as α decreases to 1 from above, and Lipschitz functions as α increase to 1 from below. This will, however, not cause any problems, since we will always be able to slightly decrease α. ◇

SKETCH OF PROOF: The proof we sketch here is not the only possible one, and perhaps not the shortest, but it will prepare us for the theory of regularity structures. The main idea is to decompose P as a sum

$$P(z) = \sum_{n \geqslant 0} P_n(z) + R(z) , \qquad (2.3.11)$$

where R is smooth and bounded, and each P_n is supported in a ball $\{z : \|z\|_{\mathfrak{s}} \leqslant 2^{-n}\}$, and satisfies the bound

$$\left| D^k P_n(z) \right| \leqslant C 2^{(1+|k|_{\mathfrak{s}})n} \qquad (2.3.12)$$

for every multiindex k. Such a decomposition is possible as shown in [Hai14, Lemma 5.5]. We then have

$$(P * \zeta)(z) = \sum_{n \geqslant 0} \langle \zeta, P_n(z - \cdot) \rangle + \langle \zeta, R(z - \cdot) \rangle .$$

The term involving R is smooth, so we only need to concern ourselves with the sum over n.

Consider first the case $\alpha < -2$. The aim is to show that for a test function φ,

$$\left| \sum_{n \geqslant 0} \int_{[0,t] \times \Lambda} \langle \zeta, P_n(\bar{z} - \cdot) \rangle (\mathscr{S}_z^{\lambda} \varphi)(\bar{z}) \, d\bar{z} \right| \lesssim \lambda^{\alpha+2} . \qquad (2.3.13)$$

The proof proceeds differently depending on the relative value of 2^{-n} and λ. If $2^{-n} > \lambda$, it follows from the fact that $\zeta \in \mathscr{C}_{\mathfrak{s}}^{\alpha}$ and the properties of P_n that

$$\left| \langle \zeta, P_n(\bar{z} - \cdot) \rangle \right| \lesssim 2^{-(\alpha+2)n} . \qquad (2.3.14)$$

Indeed, P_n can be bounded above by a scaled test function multiplied by 2^{-2n}, cf. [Hai14, Remark 2.21]. Since $\alpha + 2 < 0$, the sum over all n such that $2^{-n} > \lambda$ is dominated by the largest n, and satisfies (2.3.13) as required.

If $2^{-n} \leqslant \lambda$, we write the quantity to be bounded as $\langle \zeta, Y_n^{\lambda} \rangle$, where, for fixed z,

$$Y_n^{\lambda}(z') := \int_{[0,t] \times \Lambda} P_n(\bar{z} - z') (\mathscr{S}_z^{\lambda} \varphi)(\bar{z}) \, d\bar{z} . \qquad (2.3.15)$$

Here we note that Y_n^{λ} is supported in a ball of parabolic radius 2λ around z (see Figure 2.3). Bounding $\mathscr{S}_z^{\lambda} \varphi$ by λ^{-3} and using the properties of P_n, we obtain

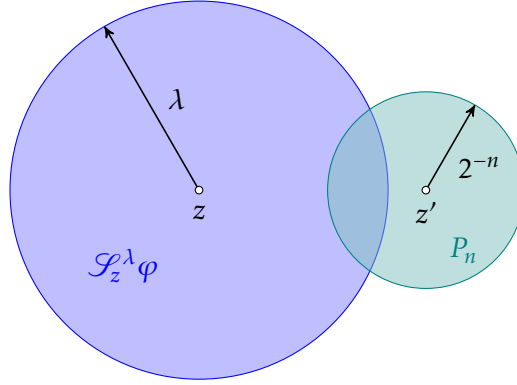$$\left| Y_n^{\lambda}(z') \right| \lesssim 2^{-2n} \lambda^{-3} .$$

Figure 2.3 – Domain of the integral (2.3.15) defining $Y_n^\lambda(z')$. The functions $\mathscr{S}_z^\lambda$ and $P_n(\bar z - \cdot)$ are supported, respectively, in a ball of (parabolic) radius λ centred in z, and in a ball of radius 2^{-n} centred in z'. Therefore, $Y_n^\lambda(z')$ vanishes if $\|z' - z\|_{\mathfrak{s}} > \lambda + 2^{-n}$. The support of the integrand lies in the intersection of the balls, and is thus contained in a ball of radius 2^{-n}.

A similar bound can be obtained for derivatives of Y_n^λ (cf. [Hai14, Lemma 5.19]), and using once again the fact that $\zeta \in \mathscr{C}_{\mathfrak{s}}^\alpha$, we arrive at

$$\left| \langle \zeta, Y_n^\lambda \rangle \right| \lesssim \lambda^\alpha 2^{-2n} .$$

Summing over all n such that $2^{-n} \leqslant \lambda$ yields again a bound of the form (2.3.13).

Consider next the case $-2 < \alpha < -1$. Then, instead of the bound (2.3.13), we have to show that

$$\left| \sum_{n \geqslant 0} \int_{[0,t] \times \Lambda} \langle \zeta, P_n(\bar z - \cdot) - P_n(z - \cdot) \rangle (\mathscr{S}_z^\lambda \varphi)(\bar z) \, d\bar z \right| \lesssim \lambda^{\alpha+2} . \tag{2.3.16}$$

The difference is that we now subtract the term $P_n(z - \cdot)$, which accounts for the polynomial P_z in (2.3.10). If $2^{-n} > \lambda$, using a Taylor expansion, one obtains instead of (2.3.14)

$$\left| \langle \zeta, P_n(\bar z - \cdot) - P_n(z - \cdot) \rangle \right| \lesssim \sum_{k \in \{(1,0),(0,1)\}} \|z - \bar z\|_{\mathfrak{s}}^{|k|_{\mathfrak{s}}} 2^{(|k|_{\mathfrak{s}} - \alpha - 2)n} .$$

Combining this with the bound

$$\int_{[0,t] \times \Lambda} \|z - \bar z\|_{\mathfrak{s}}^{|k|_{\mathfrak{s}}} (\mathscr{S}_z^\lambda \varphi)(\bar z) \, d\bar z \lesssim \lambda^{|k|_{\mathfrak{s}}}$$

and summing over n yields the result. For $2^{-n} \leqslant \lambda$, the term $Y_n^\lambda(z')$ has to be replaced by

$$Y_n^\lambda(z') - P_n(z - z')$$

since the scaled test function integrates to 1. When tested against ζ, the new term $P_n(z - \cdot)$ produces an extra term $2^{-(\alpha+2)n}$, whose sum over n again satisfies the required bound.

For large values of α, the proof is similar, except that one has to extract more terms in the Taylor expansion of the P_n. $\qquad\square$

We now directly obtain the nontrivial fact that the stochastic convolution is Hölder continuous of exponent almost $1/2$. This will become very useful when solving nonlinear equations with a fixed-point argument.

Corollary 2.3.11 (Hölder regularity of the stochastic convolution). *The stochastic convolution $P * \xi$ belongs to $\mathscr{C}_{\mathfrak{s}}^\alpha(\mathbb{R}_+ \times \Lambda)$ for all $\alpha < 1/2$.*

Proof: This follows immediately by combining Theorem 2.2.8 and the Schauder estimate of Theorem 2.3.9. $\qquad\square$

2.4   Existence and uniqueness of solutions

We turn now to semilinear SPDEs of the form

$$\partial_t \phi(t,x) = \Delta\phi(t,x) + F(\phi(t,x)) + \sqrt{2\varepsilon}\,\xi(t,x) \tag{2.4.1}$$

with initial condition $\phi_0(x)$ and locally Lipschitz nonlinearity F. For instance, in the case of the Allen–Cahn SPDE (2.0.1), $F(\phi) = \phi - \phi^3$. By analogy with the Duhamel principle yielding (2.3.4) in the case of the stochastic heat equation, one way of defining a notion of solution to (2.4.1) is as a fixed point of

$$\phi = P\phi_0 + P * F(\phi) + \sqrt{2\varepsilon}\,P * \xi\,. \tag{2.4.2}$$

A more explicit notation of this relation, analogous to (2.3.3), is

$$\phi(t) = e^{t\Delta}\phi_0 + \int_0^t e^{(t-s)\Delta} F(\phi(s))\,\mathrm{d}s + \sqrt{2\varepsilon}\int_0^t e^{(t-s)\Delta}\,\mathrm{d}\xi(s)\,. \tag{2.4.3}$$

Definition 2.4.1 (Mild solution).
- *A stochastic process $t \mapsto \phi(t)$ with values in a Banach space $\mathscr{B}$ satisfying (2.4.3) almost surely for every $t > 0$ is called a* mild solution *of the SPDE (2.4.1).*
- *A* local mild solution *of (2.4.1) is a pair (ϕ, τ), where $t \mapsto \phi(t)$ is a stochastic process and τ is a stopping time such that (2.4.3) is satisfied almost surely for all $t \leqslant \tau$.*
- *A local mild solution (ϕ, τ) is* maximal *if for any local mild solution $(\tilde{\phi}, \tilde{\tau})$, one has $\tilde{\tau} \leqslant \tau$ almost surely.*

Existence and uniqueness of a maximal local mild solution can be proved in a way which is analogous to what is done for ODEs and SDEs.

Theorem 2.4.2 (Existence of a unique maximal local mild solution). *Let $\mathscr{C}^0$ be the Banach space of continuous functions $f : \Lambda \to \mathbb{R}$, equipped with the supremum norm. For every $\phi_0 \in \mathscr{C}^0$, the semilinear equation (2.4.1) admits a unique maximal local mild solution ϕ taking values in $\mathscr{C}^0$. Moreover, this solution has continuous sample paths, and satisfies $\lim_{t \nearrow \tau} \|\phi(t)\|_{\mathscr{C}^0} = \infty$ almost surely on the set $\{\tau < \infty\}$.*

Proof: Given $T > 0$ and a continuous function $g : \mathbb{R}_+ \to \mathscr{C}^0$, we define a map $\mathscr{T}$ from the space of continuous functions from $[0, T]$ to $\mathscr{C}^0$ to itself by

$$(\mathscr{T}\phi)(t) := \int_0^t e^{(t-s)\Delta} F(\phi(s))\,\mathrm{d}s + g(t)\,, \qquad t \in [0, T]\,.$$

Hölder's inequality shows that $\|e^{t\Delta} f\|_{\mathscr{C}^0} \leqslant \|f\|_{\mathscr{C}^0}$ for any $t \geqslant 0$. Therefore, we get

$$\sup_{t \in [0,T]} \|(\mathscr{T}\phi)(t) - (\mathscr{T}\psi)(t)\|_{\mathscr{C}^0} \leqslant T \sup_{t \in [0,T]} \|F(\phi(t)) - F(\psi(t))\|_{\mathscr{C}^0}$$

$$\sup_{t \in [0,T]} \|(\mathscr{T}\phi)(t) - g(t)\|_{\mathscr{C}^0} \leqslant T \sup_{t \in [0,T]} \|F(\phi(t))\|_{\mathscr{C}^0}\,.$$

Now fix a constant $R > 0$. Since F is locally Lipschitz, we obtain the existence of a constant $M = M(R) > 0$ such that for any ϕ, ψ in a ball of radius R around g, one has

$$\sup_{t \in [0,T]} \|(\mathscr{T}\phi)(t) - (\mathscr{T}\psi)(t)\|_{\mathscr{C}^0} \leqslant TM \sup_{t \in [0,T]} \|\phi(t) - \psi(t)\|_{\mathscr{C}^0}$$

$$\sup_{t \in [0,T]} \|(\mathscr{T}\phi)(t) - g(t)\|_{\mathscr{C}^0} \leqslant TM\,.$$

Choosing T small enough, one can guarantee that $\mathcal{T}$ maps the ball of radius R into itself, and is a contraction of constant $\lambda < 1/2$ there. Therefore, Banach's fixed point theorem shows that $\mathcal{T}$ admits a unique fixed point in that ball. Choosing

$$g(t) = e^{t\Delta}\phi_0 + \sqrt{2\varepsilon}\int_0^t e^{(t-s)\Delta}\,d\xi_s\,,$$

which is in $\mathscr{C}^0$ for all t according to Corollary 2.3.11, we obtain the existence of a unique local mild solution on $[0, T]$. Note that this is a pathwise solution, that is, for each realisation of ξ, we find such a solution, where T may depend on the particular realisation of the noise.

To show maximality, we argue exactly as in the deterministic case. Since R is finite, we may restart the process to extend it to a slightly longer time interval. This procedure can be iterated as long as the solution remains bounded, proving that either $\tau = \infty$, or $\phi(t)$ diverges as t approaches τ. Finally, uniqueness and continuity can be shown in the same way as for finite-dimensional SDEs, see for instance [Øks85, Theorem 5.2.1]. $\square$

To obtain global existence of solutions, we have again to exploit the sign of the nonlinearity. There exist various one-sided conditions on the nonlinearity allowing to do that. One of them is that F satisfy the following property: for any $R > 0$, there exists a $C > 0$ such that

$$uF(u + v) \leqslant Cu^2 \qquad \text{when } |v| \leqslant R\,. \tag{2.4.4}$$

One easily checks that this holds in the case $F(u) = u - u^3$ of the Allen–Cahn equation.

Theorem 2.4.3 (Global existence of the solution). *If F is locally Lipschitz and satisfies (2.4.4), then for any $\phi_0 \in \mathscr{C}^0$, the solution of the semilinear stochastic PDE (2.4.1) exists globally in time.*

Proof: We follow the proof given in [Hai09, Proposition 6.23], using a Lyapunov function to show that $\phi(t)$ cannot blow up. Let $W_\Delta(t)$ denote the stochastic convolution in (2.4.3). Then $\psi(t) = \phi(t) - \sqrt{2\varepsilon}W_\Delta(t)$ satisfies

$$\psi(t) = e^{t\Delta}\phi_0 + \int_0^t e^{(t-s)\Delta}G(s)\,ds\,, \qquad G(s) = F(\psi(s) + \sqrt{2\varepsilon}W_\Delta(s))\,.$$

Continuity of sample paths and the Markov property imply

$$\psi(t + h) = e^{h\Delta}[\psi(t) + hG(t)] + \mathcal{O}(h)\,.$$

Consider now the Lyapunov function

$$\mathscr{V}(\psi) := \sup_{x \in \Lambda}\psi(x)^2\,.$$

Note that since the heat semigroup is contracting,[1] $\mathscr{V}(e^{h\Delta}\psi) \leqslant \mathscr{V}(\psi)$. Thus we have

$$\limsup_{h \to 0}\frac{1}{h}[\mathscr{V}(\psi(t+h)) - \mathscr{V}(\psi(t))] \leqslant \limsup_{h \to 0}\frac{1}{h}[\mathscr{V}(\psi(t) + hG(t)) - \mathscr{V}(\psi(t))]\,.$$

Now it follows from the one-sided condition (2.4.4) that as long as $\sqrt{2\varepsilon}|W_\Delta(t)| \leqslant R$, one has

$$\mathscr{V}(\psi(t) + hG(t)) = \sup_{x \in \Lambda}\left(\psi(t,x)^2 + 2h\psi(t,x)G(t,x) + h^2 G(t,x)^2\right)$$
$$\leqslant \mathscr{V}(\psi(t)) + 2hC\mathscr{V}(\psi(t)) + h^2\mathscr{V}(G(t))\,.$$

It follows that

$$\limsup_{h \to 0}\frac{1}{h}[\mathscr{V}(\psi(t+h)) - \mathscr{V}(\psi(t))] \leqslant 2C\mathscr{V}(\psi(t))\,.$$

By a Gronwall argument, $\psi(t)$ cannot blow up unless $W_\Delta(t)$ does, which is not the case. $\square$

---

[1] This holds actually for any convex function $\mathscr{V}$, cf. [Hai09, Lemma 6.22].

We note that for SPDEs with confining nonlinearities, one can obtain much better bounds on solutions. For instance, [Cer99, Proposition 3.4] applied to $F(u) = u - u^3$ provides the bound

$$\sup_{\phi_0 \in \mathscr{C}^0} \|\phi(t)\|_{\mathscr{C}^0} \leqslant \frac{C}{\sqrt{t}}\Big(1 + \sqrt{2\varepsilon} \sup_{0 \leqslant s \leqslant t} \|W_\Delta(s)\|_{\mathscr{C}^0}\Big) + \sqrt{2\varepsilon}\|W_\Delta(t)\|_{\mathscr{C}^0}\,.$$

The remarkable point is that $\phi(t)$ decreases like $1/\sqrt{t}$ for small times, whatever the size of the initial condition. This property is known as *coming down from infinity*, and is very useful to obtain continuation results for solutions with very little prior knowledge. This phenomenon actually already exists for very simple ODEs, as shows the following exercise.

Exercise 2.4.4. Show that the solution of the ODE

$$\dot{x} = -x^3\,, \qquad x(0) = x_0$$

can be bounded by a function independent of the value of x_0. ♣

2.5   Invariant measure

2.5.1   The Gaussian free field

Consider first the case of linear SPDEs. The stochastic heat equation (2.3.1) does not admit an invariant measure, since we have seen that its zeroth Fourier mode performs a Brownian motion. This problem, however, is easily cured by considering the equation

$$\partial_t \phi(t,x) = \Delta \phi(t,x) - \phi(t,x) + \sqrt{2\varepsilon}\xi(t,x)\,, \tag{2.5.1}$$

where we have reintroduced the parameter ε to keep track of its effect. The modes of the Fourier series (2.3.7) then obey the SDEs

$$\mathrm{d}\hat{\phi}_k(t) = -(\lambda_k + 1)\hat{\phi}_k(t)\,\mathrm{d}t + \sqrt{2\varepsilon}\,\mathrm{d}W_t^{(k)}\,,$$

where $-\lambda_k$ are the eigenvalues of the Laplacian, cf. (2.3.6). Hence each $\hat{\phi}_k$ is an Ornstein–Uhlenbeck process given by

$$\hat{\phi}_k(t) = \mathrm{e}^{-(\lambda_k+1)t}\,\hat{\phi}_k(0) + \sqrt{2\varepsilon}\int_0^t \mathrm{e}^{-(\lambda_k+1)(t-s)}\,\mathrm{d}W_s^{(k)}\,.$$

Asymptotically as $t \to \infty$, each $\hat{\phi}_k$ follows a centred normal law of variance $\varepsilon/(\lambda_k + 1)$. This motivates the following definition.

Definition 2.5.1 (Gaussian free field). *The* massive Gaussian free field (GFF) *with covariance* $(-\Delta + 1)^{-1}$ *on Λ is the random function $\phi_{\mathrm{GFF}} \in L^2(\Lambda)$ with Fourier series*

$$\phi_{\mathrm{GFF}}(x) := \sum_{k \in \mathbb{Z}} \frac{Z_k}{\sqrt{\lambda_k + 1}} e_k(x)\,,$$

where the $Z_k \sim \mathcal{N}(0,1)$ are i.i.d. standard normal random variables.

This definition is justified by the fact that ϕ_{GFF} is centred, and thus has covariance

$$\mathbb{E}[\phi_{\mathrm{GFF}}(x)\phi_{\mathrm{GFF}}(y)] = \sum_{k \in \mathbb{Z}} \frac{1}{\lambda_k + 1} e_k(x)e_k(y) = \sum_{k \in \mathbb{Z}} e_k(x)(-\Delta + 1)^{-1}e_k(y)\,. \tag{2.5.2}$$

Note in particular that the trace of the covariance is given by

$$\int_\Lambda \mathbb{E}[\phi_{\mathrm{GFF}}(x)^2]\,\mathrm{d}x = \mathrm{Tr}\big[(-\Delta + 1)^{-1}\big] = \sum_{k\in\mathbb{Z}} \frac{1}{\lambda_k + 1} < \infty\,.$$

The covariance operator is said to be *trace class*. This is indeed a necessary and sufficient condition for a Gaussian invariant measure to exist, cf. [DPZ96, Theorem 6.2.1]; see also [Hai09, Proposition 3.15].

Exercise 2.5.2. Show that for any test functions $\varphi_1, \varphi_2 : L^2(\Lambda) \to \mathbb{R}$, one has

$$\mathbb{E}[\langle\phi_{\mathrm{GFF}}, \varphi_1\rangle\langle\phi_{\mathrm{GFF}}, \varphi_2\rangle] = \langle\varphi_1, (-\Delta + 1)^{-1}\varphi_2\rangle_{L^2}\,,$$

where we have identified elements of $L^2(\Lambda)$ and its dual in the canonical way.   ♣

More generally, $\sqrt{\varepsilon}\phi_{\mathrm{GFF}}$ is a Gaussian free field with covariance $\varepsilon(-\Delta + 1)^{-1}$, which defines a Gaussian probability measure $\mu_{\mathrm{GFF}}^{(\varepsilon)}$ on $L^2(\Lambda)$. This measure is invariant under the dynamics of (2.5.1), meaning that

$$\mathbb{E}^{\mu_{\mathrm{GFF}}^{(\varepsilon)}}\big[f(\phi_t)\big] = \mathbb{E}\big[f(\sqrt{\varepsilon}\phi_{\mathrm{GFF}})\big] \qquad \forall t \geqslant 0 \tag{2.5.3}$$

for any integrable functional $f : L^2(\Lambda) \to \mathbb{R}$. Note that in particular, any linear functional $\ell : L^2(\Lambda) \to \mathbb{R}$ follows a centred normal distribution, which is in fact a general way to characterise Gaussian measures on infinite-dimensional spaces. See for instance [Hai09, Definition 3.2].

Exercise 2.5.3. Show that for a linear functional $\ell : L^2(\Omega) \to \mathbb{R}$, one has

$$\mathbb{E}[e^{i\langle\ell, \phi_{\mathrm{GFF}}\rangle}] = \exp\Big\{-\frac{1}{2}\langle\ell, (-\Delta + 1)^{-1}\ell\rangle_{L^2}\Big\}\,.$$

What is the variance of $\langle\ell, \phi_{\mathrm{GFF}}\rangle$?   ♣

2.5.2   Invariant Gibbs measures

We return now to semilinear SPDEs of the form

$$\partial_t\phi(t,x) = \Delta\phi(t,x) + F(\phi(t,x)) + \sqrt{2\varepsilon}\xi(t,x)\,, \tag{2.5.4}$$

where the nonlinearity F is such that Theorem 2.4.3 applies. We denote its Markov semigroup

$$P_t(\phi_0, A) := \mathbb{P}^{\phi_0}\{\phi_t \in A\}$$

(where P_t should not be confused with the heat kernel). More generally, for a test function $f : \mathcal{H} = L^2(\Lambda) \to \mathbb{R}$, we write

$$(P_t f)(\phi_0) := \mathbb{E}^{\phi_0}[f(\phi_t)]\,.$$

One can show that P_t has the *Feller property*, meaning that it maps continuous functions into continuous functions. A probability measure π on $\mathcal{H}$ is *invariant* if

$$(\pi P_t)(A) := \int_{\mathcal{H}} \pi(\mathrm{d}\phi)P_t(\phi, A) = \pi(A)$$

for every $t \geqslant 0$ and every measurable $A \subset \mathcal{H}$.

A classical way of establishing the existence of an invariant probability measure, due to Krylov and Bogoliubov, relies on a tightness argument.

Definition 2.5.4 (Tightness). *A family $\{\mu_t\}$ of probability measures on $\mathcal{H}$ is* tight *if for any $\delta > 0$, there exists a compact set $K \subset \mathcal{H}$ such that $\mu_t(K) \geqslant 1 - \delta$ for all t.*

Indeed one can show (cf. [DPZ96, Corollary 3.1.2]) that if the family of ergodic averages

$$\left\{ \frac{1}{T} \int_0^T P_t(\phi_0, \cdot) \, dt \colon T \geqslant 1 \right\}$$

is tight, then there exists an invariant probability measure. In the present situation, tightness follows from stochastic stability of ϕ_t, meaning that for any $\phi_0 \in \mathcal{H}$ and any $\delta > 0$, there exists $R > 0$ such that for all $T \geqslant 1$, one has

$$\frac{1}{T} \int_0^T \mathbb{P}^{\phi_0}\{\|\phi_t\| > R\} \, dt \leqslant \delta \, .$$

See [DPZ96, Theorem 6.1.2] for details.

In the case of the semilinear SPDE (2.5.4), we can apply [DPZ96, Theorem 6.3.3] to obtain the following exponential ergodicity result. The main ingredient of the proof is a suitable growth bound on $\mathbb{E}^{\phi_0}[\phi_t]$.

Theorem 2.5.5 (Existence of a unique invariant measure and exponential ergodicity). *When $\varepsilon > 0$, there exists a unique invariant probability measure π for (2.5.4), which is strongly mixing[a] and such that νP_t converges weakly to π as $t \to \infty$ for any probability measure ν on $\mathcal{H}$. Furthermore, there exist constants $C, \beta > 0$ such that*

$$\left| (P_t f)(\phi_0) - \langle \pi, f \rangle \right| \leqslant C(1 + \|\phi_0\|_{\mathcal{H}}) \, e^{-\beta t} \|f\|_{\mathrm{Lip}} \tag{2.5.5}$$

for all $t \geqslant 0$, all $\phi_0 \in \mathcal{H}$, and all bounded Lipschitz continuous test functions $f : \mathcal{H} \to \mathbb{R}$ with Lipschitz constant $\|f\|_{\mathrm{Lip}}$.

---

[a]Strong mixing means that $\langle P_t f, g \rangle$ converges, as $t \to \infty$, to $\langle f, 1 \rangle \langle 1, g \rangle$, where the inner products are in $L^2(\mathcal{H}, \pi)$.

In the particular case $F(\phi) = -\phi$ in (2.5.4), the invariant measure is given by $\mu_{\mathrm{GFF}}^{(\varepsilon)}(d\phi)$, the distribution of the scaled GFF defined in (2.5.3).

In the general case, the semilinear equation (2.5.4) can be written in gradient form as

$$\partial_t \phi(t, x) = [\Delta - \mathbb{1}]\phi(t, x) - \nabla \widetilde{V}(\phi(t, x)) + \sqrt{2\varepsilon}\xi(t, x) \, ,$$

where

$$\widetilde{V}(\phi) = \int_\Lambda \widetilde{U}(\phi(x)) \, dx \, , \qquad \widetilde{U}'(\phi) = \phi - F(\phi) \, .$$

Then it turns out that one has the following explicit expression for the invariant measure, which is in fact a natural extension of the Gibbs measure (1.3.3) we encountered in the finite-dimensional case. Indeed, the following result is a particularisation of [DPZ96, Theorem 8.6.3].

Theorem 2.5.6 (Invariance and reversibility of the Gibbs measure). *For $\varepsilon > 0$, the unique invariant measure of (2.5.4) is absolutely continuous with respect to the scaled Gaussian free field, with density $e^{\widetilde{V}/\varepsilon}$:*

$$\pi(d\phi) = \frac{1}{\mathscr{Z}_0} e^{-\widetilde{V}(\phi)/\varepsilon} \mu_{\mathrm{GFF}}^{(\varepsilon)}(d\phi) \, , \tag{2.5.6}$$

where $\mathscr{Z}_0$ is the normalisation. In addition, π satisfies the detailed-balance property

$$\int_{\mathcal{H}} f(\phi)(P_t g)(\phi)\pi(d\phi) = \int_{\mathcal{H}} g(\phi)(P_t f)(\phi)\pi(d\phi) \tag{2.5.7}$$

for all bounded test functions $f, g : \mathcal{H} \to \mathbb{R}$.

Remark 2.5.7. As in the case of Markov chains, reversibility of a measure implies its invariance, as can be seen by choosing $f(\phi) = 1$ and $g(\phi) = \mathbb{1}_A(\phi)$ in (2.5.7). Thus it is in fact sufficient to prove that the measure defined in (2.5.6) satisfies detailed balance to prove the theorem. The proof proceeds by using a finite-dimensional spectral Galerkin approximation of (2.5.4) and passing to the limit.   ◇

To see why (in the case $F(\phi) = \phi - \phi^3$) this is indeed the Gibbs measure associated with the potential V defined in (2.1.2), it suffices to notice that

$$V(\phi) - \widetilde{V}(\phi) = \frac{1}{2}\int_\Lambda [\|\nabla\phi(x)\|^2 + \phi(x)^2]\,\mathrm{d}x\,,$$

which is exactly the quadratic form associated with the linear case $F(\phi) = -\phi$. This suggests using the sloppy notation

$$\pi(\mathrm{d}\phi) = \frac{1}{\mathcal{Z}}\,\mathrm{e}^{-V(\phi)/\varepsilon}\,\mathrm{d}\phi\,, \qquad \mathcal{Z} = \int_{\mathcal{H}}\mathrm{e}^{-V(\phi)/\varepsilon}\,\mathrm{d}\phi\,,$$

but one should not forget that there is no Lebesgue measure on $\mathcal{H} = L^2(\Lambda)$!

2.6   Large deviations

Large deviation estimates for the SPDE

$$\partial_t\phi(t,x) = \Delta\phi(t,x) + F(\phi(t,x)) + \sqrt{2\varepsilon}\xi(t,x) \tag{2.6.1}$$

as $0 < \varepsilon \ll 1$ have been obtained in a way very similar to the finite-dimensional situation discussed in Section 1.4 [FJL82, Fre88, CM97]. One starts by showing that scaled space-time white noise $\sqrt{2\varepsilon}\xi(t,x)$ satisfies an LDP on $[0,T]$ with good rate function

$$\mathcal{I}_{[0,T]}(\gamma) = \frac{1}{2}\int_0^T\int_\Lambda\left[\frac{\partial}{\partial t}\gamma(t,x)\right]^2\mathrm{d}x\,\mathrm{d}t\,,$$

with the understanding that $\mathcal{I}_{[0,T]}(\gamma) = +\infty$ if the integral on the right-hand side does not converge. Then one uses a contraction principle to extend the LDP to the general case.

Theorem 2.6.1 (Large-deviation principle for SPDEs on the one-dimensional torus). *The semilinear SPDE (2.6.1) satisfies an LDP on $[0,T]$ with good rate function*

$$\mathcal{I}_{[0,T]}(\gamma) := \begin{cases} \dfrac{1}{2}\displaystyle\int_0^T\int_\Lambda\left[\frac{\partial}{\partial t}\gamma(t,x) - \Delta\gamma(t,x) - F(\gamma(t,x))\right]^2\mathrm{d}x\,\mathrm{d}t & \text{if the integral is finite}\,, \\ +\infty & \text{otherwise}\,. \end{cases}$$

From this result, one can for instance deduce information on the exit time and location from a domain in $\mathcal{H}$ in exactly the same way as in the finite-dimensional case.

Exercise 2.6.2. Use a similar argument as in (1.4.1) to show that the lower bound

$$\mathcal{I}_{[0,T]}(\phi) \geqslant 2[V(\phi_T) - V(\phi_0)]$$

holds for any given T, where V is the potential given by (2.1.2) in the case $F(\phi) = \phi - \phi^3$.   ♣

2.7 Metastability

We examine now the dynamics of the stochastic Allen–Cahn equation

$$\partial_t \phi(t,x) = \Delta\phi(t,x) + \phi(t,x) - \phi(t,x)^3 + \sqrt{2\varepsilon}\,\xi(t,x)$$

for small positive ε. For simplicity, we concentrate on the case $L < 2\pi$, when the associated potential V is a double-well potential, cf. Proposition 2.1.1. As in the finite-dimensional case, the estimate (2.5.5) shows that any initial measure will converge exponentially fast to the invariant measure π, but it provides no control on the dependence of the constants C and β on ε.

The large-deviation approach provides similar results as those in Theorem 1.5.1, including an Arrhenius law for the first-hitting time τ_+ of a neighbourhood of ϕ_+^* of the form

$$\lim_{\varepsilon\to 0} \varepsilon \log \mathbb{E}^{\phi_-^*}[\tau_+] = H \,,$$

where for $L < 2\pi$, one has $H = V(\phi_0^*) - V(\phi_-^*) = L/4$.

Remark 2.7.1. For $L > 2\pi$, ϕ_0^* is no longer a saddle with one unstable direction, and the role of the transition state is played by the family $\{\phi_{1,\theta}^*\}$ of non-constant stationary solutions that bifurcate from ϕ_0^* at $L = 2\pi$. The value of H can be expressed in terms of elliptic integrals, cf. [MS01, MS03].   ◇

The next question is thus whether an analogue of the Eyring–Kramers formula (1.5.22) holds. We first have to check whether the ratio of determinants of Hessians in that formula makes any sense in the infinite-dimensional case. Recall that we have shown in Proposition 2.1.1 that the Hessians are given by $-\Delta - 1$ and $-\Delta + 2$. Though each of them has an infinite determinant, it turns out that their ratio converges. To see this, let us write this ratio as

$$\det\!\big((-\Delta-1)(-\Delta+2)^{-1}\big) = \det\!\big(\mathbb{1} - 3(-\Delta+2)^{-1}\big). \tag{2.7.1}$$

This is a so-called *Fredholm determinant*. To see that it indeed converges, we note that it is equal to $-\frac{1}{2}\det(\mathbb{1} - 3(-\Delta_\perp+2)^{-1})$ where $\Delta_\perp$ acts on mean-zero functions, and write

$$\log\det\!\big(\mathbb{1} - 3(-\Delta_\perp+2)^{-1}\big) = \operatorname{Tr}\log\!\big(\mathbb{1} - 3(-\Delta_\perp+2)^{-1}\big) = -\sum_{n\geqslant 1} \frac{3^n}{n}\operatorname{Tr}\!\big((-\Delta_\perp+2)^{-n}\big).$$

The series converges because $(-\Delta_\perp+2)^{-1}$ is trace class (and therefore all its powers as well), and its eigenvalues decrease like $((2k\pi/L)^2+2)^{-1}$.

Remark 2.7.2. The ratio (2.7.1) can also be computed explicitly, writing it as

$$\prod_{k\in\mathbb{Z}} \frac{\nu_k}{\mu_k} = \prod_{k\in\mathbb{Z}} \frac{\left(\frac{2k\pi}{L}\right)^2 + 2}{\left(\frac{2k\pi}{L}\right)^2 - 1} = -2\left(\prod_{k\in\mathbb{N}} \frac{1 + 2\left(\frac{L}{2k\pi}\right)^2}{1 - \left(\frac{L}{2k\pi}\right)^2}\right)^2 = -\frac{\sinh^2(L/\sqrt{2})}{\sin^2(L/2)} \,,$$

where we used two of Euler's infinite product identities in the last step. However, the interpretation in terms of Fredholm determinant will be more useful in higher dimensions.   ◇

Exercise 2.7.3. By working in Fourier space, show that

$$\det\!\big(\mathbb{1} + c(-\Delta+a)^{-1}\big) = \frac{1}{\det\!\big(\mathbb{1} - c(-\Delta+b)^{-1}\big)}$$

for every $a, b \in \mathbb{R}$ such that $a + c = b$ and the Fredholm determinants make sense.   ♣

In order to prove an Eyring–Kramers law, we will approximate the infinite-dimensional system by a finite-dimensional one and pass to the limit. One way of doing this is to take the limit $N \to \infty$ the lattice system (1.0.1). While this is possible, we prefer the slightly easier approach of using a spectral Galerkin approximation, which has the advantage that the relevant eigenvalues are independent of N. For $\phi \in L^2(\Lambda)$, we define its Galerkin projection as

$$(\Pi_N \phi)(x) = \sum_{\substack{k \in \mathbb{Z} \\ |k| \leqslant N}} \langle \phi, e_k \rangle e_k(x), \qquad (2.7.2)$$

where the e_k denote the Fourier basis functions (2.3.5). We then consider the projected equation

$$\partial_t \phi_N(t, x) = \Delta \phi_N(t, x) + \phi_N(t, x) - \Pi_N \phi_N(t, x)^3 + \sqrt{2\varepsilon}(\Pi_N \xi)(t, x)$$

on the finite-dimensional subspace E_N of $L^2(\Lambda)$ spanned by the e_k with $|k| \leqslant N$ (we have used the fact that Π_N and Δ commute). This is equivalent to the finite-dimensional SDE for Fourier coefficients

$$d\hat{\phi}_t = -\nabla \widehat{V}(\hat{\phi}_t) dt + \sqrt{2\varepsilon}\, dW_t,$$

where $\widehat{V}$ is simply the potential (2.1.2) expressed in Fourier coordinates. This form is suitable for the potential-theoretic approach described in Section 1.5.2

One easily checks that for symmetric sets $B = -A \subset E_N$ Lemma 1.5.12 still applies here, so that it is essentially sufficient to estimate the partition function $\mathscr{Z}$ and the capacity $\operatorname{cap}(A, B)$ uniformly in the cut-off N.

As in the finite-dimensional case, it turns out to be useful to perform the change of variables

$$\phi_N(x) = \frac{1}{\sqrt{L}}\phi_0 + \sqrt{\varepsilon}\phi_\perp(x), \qquad \phi_0 = \langle e_0, \phi_N \rangle = \frac{1}{\sqrt{L}}\int_\Lambda \phi_N(x)\, dx, \qquad (2.7.3)$$

where $\phi_\perp$ has zero mean. Inserting this in the potential (2.1.2) yields

$$\frac{1}{\varepsilon}V\left(\frac{1}{\sqrt{L}}\phi_0 + \sqrt{\varepsilon}\phi_\perp(x)\right) = \frac{1}{\varepsilon}V_0(\phi_0) + \frac{1}{2}\langle \phi_\perp, Q_\perp(\phi_0)\phi_\perp \rangle + R_\varepsilon(\phi_0, \phi_\perp), \qquad (2.7.4)$$

where

$$V_0(\phi_0) = \frac{L}{4}\left(\frac{\phi_0^2}{L} - 1\right)^2,$$

$$Q_\perp(\phi_0) = -\Delta_\perp - 1 + \frac{3\phi_0^2}{L},$$

$$R_\varepsilon(\phi_0, \phi_\perp) = \sqrt{\frac{\varepsilon}{L}}\phi_0 \int_\Lambda \phi_\perp(x)^3\, dx + \frac{\varepsilon}{4}\int_\Lambda \phi_\perp(x)^4\, dx.$$

Here the Laplacian $\Delta_\perp = \Delta_{\perp,N}$ acts on $E_{\perp,N}$, the subspace of zero-mean functions in E_N. Therefore, the eigenvalues of $Q_\perp(\phi_0)$ are given by $\mu_k + 3\phi_0^2/L \geqslant 2\pi/L - 1 > 0$, where the $\mu_k = \lambda_k - 1$ are the eigenvalues of $-\Delta_\perp - 1$. We will denote by $g_N(\phi_0)$ the Gaussian measure with covariance $Q_\perp(\phi_0)^{-1}$ on $E_{\perp,N}$, called a (truncated) Gaussian free field with mass $m = (3\phi_0^2/L)^{1/2}$. A crucial estimate is the following one.

Lemma 2.7.4. *There exist constants C_n, uniform in N and ϕ_0, such that the bound*

$$\mathbb{E}^{g_N(\phi_0)}\left[\int_\Lambda \phi_\perp(x)^{2n}\, dx\right] \leqslant C_n$$

holds for all even moments of $\phi_\perp$, while the odd moments vanish.

PROOF: Let $\mathscr{K}_N = \{k \in \mathbb{Z} : k \neq 0, |k| \leqslant N\}$. Then we can represent $\phi_\perp$ as

$$\phi_\perp(x) = \sum_{k \in \mathscr{K}_N} \frac{Z_k}{\sqrt{\mu_k + m^2}} e_k(x) \, ,$$

where the Z_k are i.i.d. standard normal random variables. Therefore

$$\mathbb{E}^{g_N(\phi_0)}\left[\int_\Lambda \phi_\perp(x)^{2n}\,\mathrm{d}x\right] = \sum_{k_1,\ldots,k_{2n} \in \mathscr{K}_N} \frac{\mathbb{E}[Z_{k_1}\ldots Z_{k_{2n}}]}{\sqrt{\mu_{k_1}+m^2}\ldots\sqrt{\mu_{k_{2n}}+m^2}} \int_\Lambda e_{k_1}(x)^2 \ldots e_{k_{2n}}(x)^2\,\mathrm{d}x \, .$$

The expectation vanishes unless the k_i are pairwise equal. Therefore, we obtain

$$\mathbb{E}^{g_N(\phi_0)}\left[\int_\Lambda \phi_\perp(x)^{2n}\,\mathrm{d}x\right] \lesssim \sum_{k_1,\ldots,k_n \in \mathscr{K}_N} \frac{1}{(\mu_{k_1}+m^2)\ldots(\mu_{k_n}+m^2)} = \left(\sum_{k \in \mathscr{K}_N} \frac{1}{\mu_k + m^2}\right)^n \, ,$$

which is bounded uniformly in N and m. Odd moments vanish by symmetry. □

We will see in the next chapter that this bound can be substantially improved using Wick calculus. For now, it will be sufficient to obtain the following key estimate.

Proposition 2.7.5. *We have*

$$\mathrm{e}^{-c\varepsilon} \leqslant \mathbb{E}^{g_N(\phi_0)}\left[\mathrm{e}^{-R_\varepsilon(\phi_0,\cdot)}\right] \leqslant 1 + \mathrm{e}^{C(1+\phi_0^2)}\varepsilon \, ,$$

where c and C are uniform in N.

PROOF: By Lemma 2.7.4, $\mathbb{E}^{g_N(\phi_0)}[R_\varepsilon] = c_N\varepsilon$, where $c_N \leqslant c$ is bounded uniformly in N and does not depend on ϕ_0. Therefore the lower bound follows directly from Jensen's inequality, which yields

$$\mathbb{E}^{g_N(\phi_0)}\left[\mathrm{e}^{-R_\varepsilon(\phi_0,\cdot)}\right] \geqslant \exp\left\{-\mathbb{E}^{g_N(\phi_0)}[R_\varepsilon(\phi_0,\cdot)]\right\} \geqslant \exp\{-c\varepsilon\} \, .$$

For the other direction, we first derive a rough bound on a slightly higher exponential moment, using a change of mass in the Gaussian measure. For any $p \geqslant 1$, we have

$$\mathbb{E}^{g_N(\phi_0)}\left[\mathrm{e}^{-pR_\varepsilon(\phi_0,\cdot)}\right] = \sqrt{\frac{\det Q_\perp(\phi_0)}{\det Q_\perp(0)}}\,\mathbb{E}^{g_N(0)}\left[\exp\left\{-\frac{3\phi_0^2}{2L}\int_\Lambda \phi_\perp(x)^2\,\mathrm{d}x - pR_\varepsilon(\phi_0,\cdot)\right\}\right] \, .$$

For the ratio of determinants, we note that

$$\log\left(\frac{\det Q_\perp(\phi_0)}{\det Q_\perp(0)}\right) = \log \prod_{k \in \mathscr{K}_N}\left(1 + \frac{3\phi_0^2}{L\mu_k}\right) = \sum_{k \in \mathscr{K}_N}\log\left(1 + \frac{3\phi_0^2}{L\mu_k}\right) \leqslant \frac{3\phi_0^2}{L}\sum_{k \in \mathscr{K}_N}\frac{1}{\mu_k} =: 2C_N\phi_0^2 \, ,$$

where C_N is bounded uniformly in N. Furthermore, using completion of squares, we get

$$\frac{3\phi_0^2}{2L}\int_\Lambda \phi_\perp(x)^2\,\mathrm{d}x + pR_\varepsilon(\phi_0,\phi_\perp) = \int_\Lambda \phi_\perp(x)^2\left[\frac{3\phi_0^2}{2L} + p\sqrt{\frac{\varepsilon}{L}}\phi_0\phi_\perp(x) + p\frac{\varepsilon}{4}\phi_\perp(x)^2\right]\mathrm{d}x$$

$$= \frac{1}{4}\int_\Lambda \phi_\perp(x)^2\left[\left(\frac{2\sqrt{p}\phi_0}{\sqrt{L}} + \sqrt{\varepsilon p}\phi_\perp(x)\right)^2 + \frac{2(3-2p)\phi_0^2}{L}\right]\mathrm{d}x \, ,$$

which is positive for $p \leqslant 3/2$. Therefore $\mathbb{E}^{g_N(\phi_0)}[e^{-(3/2)R_\varepsilon(\phi_0,\cdot)}] \leqslant e^{C_N \phi_0^2}$. We now decompose

$$
\begin{aligned}
\mathbb{E}^{g_N(\phi_0)}[e^{-R_\varepsilon}] &= \mathbb{E}^{g_N(\phi_0)}[e^{-R_\varepsilon}\mathbb{1}_{\{R_\varepsilon \geqslant 0\}}] + \mathbb{E}^{g_N(\phi_0)}[e^{-R_\varepsilon}\mathbb{1}_{\{R_\varepsilon < 0\}}] \\
&\leqslant \mathbb{P}^{g_N(\phi_0)}\{R_\varepsilon \geqslant 0\} + \mathbb{E}^{g_N(\phi_0)}[(1 - R_\varepsilon + e^{-R_\varepsilon} - 1 + R_\varepsilon)\mathbb{1}_{\{R_\varepsilon < 0\}}] \\
&\leqslant \mathbb{P}^{g_N(\phi_0)}\{R_\varepsilon \geqslant 0\} + \mathbb{P}^{g_N(\phi_0)}\{R_\varepsilon < 0\} - \mathbb{E}^{g_N(\phi_0)}[R_\varepsilon] + \frac{1}{2}\mathbb{E}^{g_N(\phi_0)}[R_\varepsilon^2 e^{-R_\varepsilon}] \\
&\leqslant 1 - \mathbb{E}^{g_N(\phi_0)}[R_\varepsilon] + \frac{1}{2}\mathbb{E}^{g_N(\phi_0)}[R_\varepsilon^6]^{1/3}\mathbb{E}^{g_N(\phi_0)}[e^{-(3/2)R_\varepsilon}]^{2/3},
\end{aligned}
$$

where we have used the fact that $e^{-x} - 1 + x \leqslant \frac{1}{2}x^2 e^{-x}$ for $x < 0$ and Hölder's inequality. Since by Lemma 2.7.4, $\mathbb{E}^{g_N(\phi_0)}[R_\varepsilon^6] = \mathcal{O}(\varepsilon^3 \phi_0^6)$, the result follows from the bound on $\mathbb{E}^{g_N(\phi_0)}[e^{-(3/2)R_\varepsilon(\phi_0,\cdot)}]$ for an appropriate choice of C. $\square$

We can now estimate the partition function and capacity, in a similar way as we did in the finite-dimensional case. We define the sets A and B as

$$
A = -B = \{y \colon |\phi_0 + \sqrt{L}| \leqslant \delta, \phi_\perp \in D_\perp\}, \tag{2.7.5}
$$

with $D_\perp$ containing a ball of radius $\sqrt{\log(\varepsilon^{-1})}$ in $H^s(\Lambda)$ for some $s < \frac{1}{2}$. Then we have the following result.

Proposition 2.7.6. *The partition function and the capacity satisfy*

$$
\mathscr{Z}_N := \int_{E_N} e^{-V(\phi)/\varepsilon}\,d\phi = 2\frac{(2\pi\varepsilon)^{N+1/2}}{\sqrt{\nu_0 \det(-\Delta_\perp + 2)}}[1 + \mathcal{O}(\varepsilon)],
$$

$$
\mathrm{cap}(A,B) = \frac{(2\pi\varepsilon)^{N+1/2}\sqrt{|\mu_0|}\,e^{-V_0(0)/\varepsilon}}{2\pi\sqrt{\det(-\Delta_\perp - 1)}}[1 + \mathcal{O}(\varepsilon)],
$$

where $\Delta_\perp = \Delta_{\perp,N}$, $\nu_0 = 2$, $\mu_0 = -1$, and the error terms are uniform in the cut-off N.

PROOF: The partition function can be written as

$$
\mathscr{Z}_N = \varepsilon^N \int_{-\infty}^{\infty} e^{-V_0(\phi_0)/\varepsilon}\frac{(2\pi)^N}{\sqrt{\det Q_\perp(\phi_0)}}\mathbb{E}^{g_N(\phi_0)}[e^{-R_\varepsilon(\phi_0,\cdot)}]\,d\phi_0.
$$

The result then follows from Proposition 2.7.5, using Laplace asymptotics for the integral over ϕ_0, which is dominated by values of ϕ_0 near $\pm\sqrt{L}$. The error term $\varepsilon e^{C(1+\phi_0^2)}$ only yields an error of order ε, because it is dominated by the factor $e^{-V_0(\phi_0)/\varepsilon}$.

For the upper bound on the capacity, we use as before the Dirichlet principle, with a test function

$$
h(\phi) = h_0(\phi_0) = \frac{1}{c_0}\int_{\phi_0}^a e^{V_0(\xi)/\varepsilon}\,d\xi, \qquad c_0 = \int_{-a}^a e^{V_0(\xi)/\varepsilon}\,d\xi. \tag{2.7.6}
$$

Writing $\mathscr{E}(h)$ in terms of $\mathbb{E}^{g_N(\phi_0)}[e^{-R_\varepsilon(\phi_0,\cdot)}]$ yields the upper bound in a similar way as for the partition function. For the lower bound on the capacity, we use the Thomson principle. Here it is useful to decompose the potential as

$$
\frac{1}{\varepsilon}V\left(\frac{1}{\sqrt{L}}\phi_0 + \sqrt{\varepsilon}\phi_\perp(x)\right) = \frac{1}{\varepsilon}V_0(\phi_0) + V_\perp(\phi_\perp) + V_+(\phi_0,\phi_\perp),
$$

where

$$V_\perp(\phi_\perp) = \frac{1}{2}\langle \phi_\perp, Q_\perp(0)\phi_\perp\rangle + \frac{\varepsilon}{4}\int_\Lambda \phi_\perp(x)^4\, dx\,,$$

$$V_+(\phi_0, \phi_\perp) = \frac{3\phi_0^2}{2L}\int_\Lambda \phi_\perp(x)^2\, dx + \sqrt{\varepsilon}\phi_0\int_\Lambda \phi_\perp(x)^3\, dx\,.$$

We define the divergence-free unit flow

$$\varphi(\phi) = \frac{1}{K}\mathbb{1}_{D_\perp}(\phi_\perp)\, e^{-V_\perp(\phi_\perp)}\, e_0\,, \qquad K = \varepsilon^N\int_{D_\perp} e^{-V_\perp(\phi_\perp)}\, d\phi_\perp\,.$$

Substituting in the quadratic form (1.5.16) yields

$$\mathscr{D}(\varphi) = \frac{1}{\varepsilon K^2}\int_{-\infty}^{\infty} e^{V_0(\phi_0)/\varepsilon}\frac{(2\pi\varepsilon)^N}{\sqrt{\det Q_\perp(0)}}\mathbb{E}^{\mu_N(0)}\left[\mathbb{1}_{D_\perp}\exp\left\{-\frac{\varepsilon}{4}\int_\Lambda \phi_\perp(x)^4\, dx + V_+(\phi_0, \cdot)\right\}\right]d\phi_0\,.$$

Proceeding as in Proposition 2.7.5, one checks that the gaussian expectation is again equal to $1 + \mathcal{O}(\varepsilon)$, despite the different sign of the cubic term and the extra quadratic term. A similar analysis can be performed for K. The lower bound then follows from Thomson's principle. $\square$

Proposition 2.7.6 is one of the essential building blocks of the following theorem, which is the main result of this section.

Theorem 2.7.7 (Eyring-Kramers law for the one-dimensional Allen–Cahn SPDE)**.** *Assume that $L < 2\pi$ and A and B are as in (2.7.5). Then*

$$\mathbb{E}^{\phi_-^*}[\tau_B] = \frac{2\pi}{|\mu_0|}\frac{e^{[V(\phi_0^*) - V(\phi_-^*)]/\varepsilon}}{\sqrt{\left|\det\left(\mathbb{1} + 3(-\Delta - 1)^{-1}\right)\right|}}[1 + \mathcal{O}(\varepsilon)]\,,$$

where $V(\phi_0^*) - V(\phi_-^*) = L/4$ and the determinant is to be interpreted as the inverse of the Fredholm determinant (2.7.1) (cf. Exercise 2.7.3).

SKETCH OF PROOF: For finite cut-off N, and when starting in the equilibrium distribution on ∂A, the result follows from Proposition 2.7.6 and Theorem 1.5.6. To extend this to the infinite-dimensional equation, we have to check two more things. The first one is that first-hitting times of B in the finite-dimensional system converge to those of the infinite-dimensional one. This can be done by combining an a priori estimate on $\mathbb{E}[\tau_B^2]$ with a sample-path approximation argument given in [BJ13, Theorem 3.1], cf. [BG13, Theorem 3.1 and Proposition 3.4]. The second thing to be done is to prove that one can replace the initial distribution on ∂A by a start in ϕ_-^*. This can be achieved by using the coupling argument given in [MOS89, Corollary 3.1], cf. [BG13, Theorem 3.5 and Proposition 3.6]. $\square$

2.8 Bibliographical notes

The deterministic PDE commonly called Allen–Cahn equation was introduced in [CI75] and [AC79]. Detailed descriptions of its dynamics can be found for instance in [CP89, Che04].

The definitions of parabolic Hölder spaces and related Schauder estimates have been taken from [Hai14] and [CW17], while properties of Gaussian fields, in particular in relation with fractional Sobolev spaces, can be found in [Hai09].

Existence and uniqueness of solutions of the Allen–Cahn SPDE have been first established in [FJL82]. Regarding the general existence and uniqueness theory, as well as properties of

invariant measures, we have mainly followed [DPZ96] with some input from [Hai09]. The sharper bounds on the decay of supremum norms are from [Cer99].

The large-deviation principle for the stochastic Allen–Cahn equation has been first obtained in [FJL82], and later extended in [Fre88, Sow92, CM97].

Eyring–Kramers laws for the one-dimensional Allen–Cahn equation have been obtained in [BG13] (using spectral Galerkin approximations) and in [Bar15] (using lattice discretisations). Results in these papers are much more general than those presented here, as they apply to equations with a general confining nonlinearity, with some smoothness assumptions. Results in [BG13] include the bifurcation regime and the dynamics for $L > 2\pi$ for periodic boundary conditions, in which the transition states are degenerate. The proof presented here is somewhat different from the proofs given in those works, which did not use Gaussian measure theory and the Thomson principle, but were instead based on Hausdorff–Young inequalities and direct approximations of Dirichlet forms, in the spirit of [BEGK04].

Chapter 3

Allen–Cahn SPDE in two space dimensions

The particle system considered in Chapter 1 makes also sense in two dimensions, that is, for a two-dimensional lattice of N^2 particles, each one interacting with its 4 nearest neighbours. Proceeding as in Chapter 2, we obtain formally a continuum limit given by the SPDE

$$\partial_t \phi(t,x) = \Delta \phi(t,x) + \phi(t,x) - \phi(t,x)^3 + \sqrt{2\varepsilon}\xi(t,x)\,, \tag{3.0.1}$$

where x now belongs to the two-dimensional torus $\Lambda = (\mathbb{R}/(L\mathbb{Z}))^2$, and the Laplacian $\Delta = \partial_{x_1 x_1} + \partial_{x_2 x_2}$ acts on both components of x. The PDE obtained in the deterministic case $\varepsilon = 0$ is perfectly well-defined.

Exercise 3.0.1. Compute the eigenvalues of the Laplacian Δ for periodic boundary conditions on Λ, and use it to determine the stability of the constant stationary solutions of the deterministic PDE. Argue that for $L < 2\pi$, we are again in the double-well situation, as there can be no non-constant stationary solutions. Prove this fact, using

$$W(\phi) = \frac{1}{2}\int_\Lambda \|\nabla\phi(x)\|^2 \,\mathrm{d}x$$

as a Lyapunov function. ♣

It turns out, however, that the SPDE (3.0.1) is not well-posed. This is related to the fact that space-time white noise is more irregular in dimension 2 than in dimension 1. Before discussing a cure for this problem, we first have to understand in more detail where the problem comes from.

3.1   The need for renormalisation

Space-time white noise on $\mathbb{R} \times \Lambda$ can be defined exactly as in Definition 2.2.1, except that Λ now denotes the two-dimensional torus. The scaling operator in (2.2.3) now takes the form

$$(\mathscr{S}^{\tau,\lambda}\varphi)(t,x) := \frac{1}{\tau\lambda^2}\varphi\Big(\frac{t}{\tau}, \frac{x}{\lambda}\Big)\,,$$

and the scaling property of Proposition 2.2.5 becomes

$$\xi_{\tau,\lambda} \stackrel{\text{law}}{=} \frac{1}{\sqrt{\tau}\lambda}\xi\,.$$

41

The parabolic Hölder spaces are defined as before, with the parabolic norm defined by

$$\|(t,x) - (s,y)\|_{\mathfrak{s}} = |t-s|^{1/2} + \sum_{i=1}^{2} |x_i - y_i| \, ,$$

and the scaling operator in (2.2.6) replaced by

$$(\mathscr{S}_{t,x}^{\lambda} \varphi)(s,y) := \frac{1}{\lambda^4} \varphi\left(\frac{s-t}{\lambda^2}, \frac{y-x}{\lambda}\right).$$

Making the necessary changes in the proof of Theorem 2.2.8, one now finds that

$$\xi \in \mathscr{C}_{\mathfrak{s}}^{-2-\kappa} \qquad \text{for any } \kappa > 0 \, . \tag{3.1.1}$$

Exercise 3.1.1. Prove (3.1.1) by adapting the argument used in Theorem 2.2.8.   ♣

The Schauder estimate stated in Theorem 2.3.9 remains true in the present setting. The problem is, however, that the stochastic convolution $P * \xi$ belongs to $\mathscr{C}_{\mathfrak{s}}^{\alpha}$ for any $\alpha < 0$ only, instead any $\alpha < \frac{1}{2}$ in the one-dimensional case. Therefore, the stochastic convolution is only a distribution, albeit rather close to being a function. Since there is no canonical way of multiplying two distributions, the fixed-point equation (2.4.2) becomes ill-defined, as it requires the nonlinear function F to be applied to the stochastic convolution.

As the stochastic convolution is a Gaussian process, its singularity can be traced to the behaviour of the covariance, which diverges near the diagonal $y = x$ (see also (2.5.2)). Therefore, a regularisation becomes possible by subtracting an "infinite constant" from the right-hand side of (3.0.1), a process known as *renormalisation*. An important tool to understand how it works is Wick calculus for Gaussian random variables.

3.2   Wick calculus

3.2.1   Isserlis' theorem

Consider a Gaussian random vector X on $\mathbb{R}^N$ with covariance matrix C. Recall that this means that for any test function F, one has

$$\mathbb{E}[F(X)] = \frac{1}{\mathscr{Z}} \int_{\mathbb{R}^N} F(x) e^{-\langle x, C^{-1} x \rangle / 2} \, dx \, ,$$

where C has entries $C_{ij} = \mathbb{E}[X_i X_j]$, and $\mathscr{Z}$ is such that $\mathbb{E}[1] = 1$. Moments of X can easily be computed by the following integration-by-parts formula, which is a finite-dimensional version of the Schwinger–Dyson equations from Quantum Field Theory.

Lemma 3.2.1. *For any $i \in \{1, \dots, N\}$ we have*

$$\mathbb{E}[X_i F(X)] = \sum_{j=1}^{N} C_{ij} \mathbb{E}[\partial_j F(X)] \tag{3.2.1}$$

for all real-valued differentiable F such that both sides of the equality are well-defined.

Proof: Leibniz' rule yields

$$\frac{\partial}{\partial x_j}\left(F(x) e^{-\langle x, C^{-1} x \rangle / 2}\right) = \left[\frac{\partial F}{\partial x_j}(x) - \frac{1}{2} F(x) \frac{\partial}{\partial x_j} \langle x, C^{-1} x \rangle\right] e^{-\langle x, C^{-1} x \rangle / 2} \, .$$

Since we have

$$\frac{1}{2}\sum_{j=1}^{N} C_{ij}\frac{\partial}{\partial x_j}\langle x, C^{-1}x\rangle = \sum_{j,k=1}^{N} C_{ij}C_{jk}^{-1}x_k = \sum_{k=1}^{N}\delta_{ik}x_k = x_i\,,$$

it follows that

$$x_i F(x)\,e^{-\langle x, C^{-1}x\rangle/2} = \sum_{j=1}^{N} C_{ij}\left[-\frac{\partial}{\partial x_j}\Big(F(x)\,e^{-\langle x, C^{-1}x\rangle/2}\Big) + \frac{\partial F}{\partial x_j}\,e^{-\langle x, C^{-1}x\rangle/2}\right].$$

Integrating over the whole space, we see that the boundary terms vanish, and the result follows at once. □

An immediate consequence of this result is the following theorem, due to Leon Isserlis.

Theorem 3.2.2 (Isserlis). *For any $1 \leqslant n \leqslant N/2$, we have*

$$\mathbb{E}[X_1\ldots X_{2n}] = \sum_{\text{pairings }\mathscr{P}\text{ of }\{1,\ldots,2n\}}\ \prod_{\{i,j\}\in\mathscr{P}}\mathbb{E}[X_iX_j]\,,$$

$$\mathbb{E}[X_1\ldots X_{2n-1}] = 0\,,$$

where a pairing *of $\{1,\ldots,2n\}$ is a partition $\mathscr{P} = \{\{i_1,j_1\},\ldots,\{i_n,j_n\}\}$ of this set into disjoint subsets of two elements.*
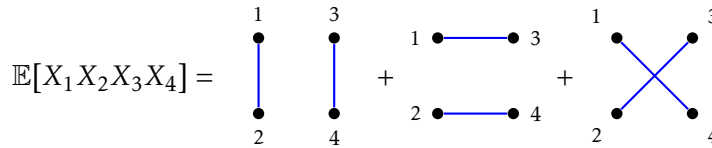
PROOF: By induction on n, applying (3.2.1) with $i = 1$ and $F(X)$ of the form $X_2\ldots X_m$. □

Remark 3.2.3. An alternative proof relies on the identity $\mathbb{E}[e^{i\langle\ell,X\rangle}] = e^{-\langle\ell,C\ell\rangle/2}$, expanding both sides in powers of $\ell = (\ell_1,\ldots,\ell_n)^{\mathrm{T}}$ and identifying coefficients. ◇

For instance, in the case $2n = 4$, we obtain

$$\mathbb{E}[X_1X_2X_3X_4] = \mathbb{E}[X_1X_2]\mathbb{E}[X_3X_4] + \mathbb{E}[X_1X_3]\mathbb{E}[X_2X_4] + \mathbb{E}[X_1X_4]\mathbb{E}[X_2X_3]\,.$$

A convenient graphical way of representing this relation is the following:

Exercise 3.2.4. Determine the number of pairings of $\{1,\ldots,2n\}$. Use this information to compute $\mathbb{E}[X^{2n}]$ when X follows a standard normal distribution, and prove (2.2.9). ♣

3.2.2 Hermite polynomials

Consider the one-dimensional Ornstein–Uhlenbeck process, solving the SDE

$$\mathrm{d}X_t = -X_t\,\mathrm{d}t + \sqrt{2\varepsilon}\,\mathrm{d}W_t\,.$$

We know from Section 1.3 that its invariant distribution π has density $e^{-x^2/(2\varepsilon)}/\sqrt{2\pi\varepsilon}$ with respect to Lebesgue measure, while its generator

$$\mathscr{L} = \varepsilon\,e^{x^2/(2\varepsilon)}\frac{\mathrm{d}}{\mathrm{d}x}\,e^{-x^2/(2\varepsilon)}\frac{\mathrm{d}}{\mathrm{d}x}$$

is self-adjoint in $L^2(\mathbb{R}, \pi(\mathrm{d}x))$. Furthermore, one easily checks that the operators

$$a = \mathrm{e}^{x^2/(2\varepsilon)}\left(x + \varepsilon\frac{\mathrm{d}}{\mathrm{d}x}\right)\mathrm{e}^{-x^2/(2\varepsilon)} = \varepsilon\frac{\mathrm{d}}{\mathrm{d}x}\,,$$

$$a^\dagger = -\mathrm{e}^{x^2/(2\varepsilon)}\varepsilon\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^{-x^2/(2\varepsilon)} = x - \varepsilon\frac{\mathrm{d}}{\mathrm{d}x}$$

are mutually adjoint in $L^2(\mathbb{R}, \pi(\mathrm{d}x))$, and satisfy the relations

$$a^\dagger a = -\varepsilon\mathscr{L}\,,\qquad aa^\dagger - a^\dagger a = \varepsilon\,.$$

These relations are useful to find eigenfunctions of $\mathscr{L}$. Indeed, we obviously have $\mathscr{L}H_0 = 0$, where $H_0(x) = 1$ is constant. Now if H is an eigenfunction of $\mathscr{L}$ with eigenvalue $-\lambda$, we have

$$-\mathscr{L}(a^\dagger H) = \frac{1}{\varepsilon}a^\dagger aa^\dagger H = \frac{1}{\varepsilon}a^\dagger(a^\dagger a + \varepsilon)H = (\lambda + 1)a^\dagger H\,,$$

showing that $a^\dagger H$ is an eigenfunction of $\mathscr{L}$ with eigenvalue $-(\lambda + 1)$. We conclude that the functions in the family $(H_n)_{n\in\mathbb{N}_0}$ defined recursively by

$$H_{n+1}(x;\varepsilon) = (a^\dagger H_n)(x;\varepsilon) = xH_n(x;\varepsilon) - \varepsilon\frac{\mathrm{d}}{\mathrm{d}x}H_n(x;\varepsilon) \qquad (3.2.2)$$

are orthogonal, and satisfy

$$\mathscr{L}H_n(x;\varepsilon) = -nH_n(x;\varepsilon)\,,\qquad n\in\mathbb{N}_0\,.$$

The corresponding eigenfunctions of $\mathscr{L}^\dagger$ are simply given by $\mathrm{e}^{-x^2/(2\varepsilon)}H_n(x;\varepsilon)$.

Remark 3.2.5. The operators $a^\dagger$ and a behave in the same way as creation and annihilation operators for the quantum harmonic oscillator. In fact, the generator $\mathscr{L}$ is conjugated to the Hamiltonian of the harmonic oscillator via the transformation $\mathscr{L} \mapsto \mathrm{e}^{-x^2/(4\varepsilon)}\mathscr{L}\,\mathrm{e}^{x^2/(4\varepsilon)}$, and the operators $a^\dagger$ and a are conjugated in the same way to those of the harmonic oscillator. ◇

In fact, one can show that the family $(H_n)_{n\in\mathbb{N}_0}$ forms a complete orthonormal basis of $L^2(\mathbb{R}, \pi(\mathrm{d}x))$. See for instance [Nua06, Lemma 1.1.2].

Definition 3.2.6 (Hermite polynomials). *The polynomial*

$$H_n(x;\varepsilon) = (a^\dagger)^n H_0(x;\varepsilon) = (-\varepsilon)^n \mathrm{e}^{x^2/(2\varepsilon)}\frac{\mathrm{d}^n}{\mathrm{d}x^n}\mathrm{e}^{-x^2/(2\varepsilon)}$$

is called the nth Hermite polynomial with variance ε.

Hermite polynomials can easily be computed by the recurrence relation (3.2.2). The first few of them are given by

$$H_0(x;\varepsilon) = 1$$
$$H_1(x;\varepsilon) = x$$
$$H_2(x;\varepsilon) = x^2 - \varepsilon$$
$$H_3(x;\varepsilon) = x^3 - 3\varepsilon x$$
$$H_4(x;\varepsilon) = x^4 - 6\varepsilon x^2 + 3\varepsilon^2\,.$$

A useful alternative expression for Hermite polynomials relies on their generating function.

Lemma 3.2.7. *For every $t, x \in \mathbb{R}$ and $\varepsilon \geqslant 0$, one has*

$$G(t, x; \varepsilon) := e^{tx - \varepsilon t^2/2} = \sum_{n=0}^{\infty} \frac{t^n}{n!} H_n(x; \varepsilon). \tag{3.2.3}$$

Exercise 3.2.8. Prove Lemma 3.2.7. **Hint:** Take the derivative with respect to t on both sides, and use the recurrence relation (3.2.2). Another proof (with a slightly different normalisation) can be found in [Nua06, Lemma 1.1.1]. ♣

A very important consequence of Lemma 3.2.7 is the following identity, which can be seen as an analogue of Isserlis' theorem for Hermite polynomials of Gaussian random variables.

Lemma 3.2.9. *Let X and Y be jointly Gaussian centred random variables, of respective variance ε_1 and ε_2. Then for any $n, m \geqslant 0$, one has*

$$\mathbb{E}[H_n(X; \varepsilon_1) H_m(Y; \varepsilon_2)] = \begin{cases} n! \mathbb{E}[XY]^n & \text{if } n = m, \\ 0 & \text{otherwise.} \end{cases} \tag{3.2.4}$$

PROOF: Let C denote the covariance matrix of $Z = (X, Y)^{\mathrm{T}}$. Then the expression

$$\mathbb{E}\left[e^{\langle \ell, Z \rangle}\right] = \exp\left\{\frac{1}{2}\langle \ell, C\ell \rangle\right\}$$

for the Laplace transform of Z, applied with $\ell = (t, s)^{\mathrm{T}}$, implies that

$$\mathbb{E}\left[e^{tX - \varepsilon_1 t^2/2} e^{sY - \varepsilon_2 s^2/2}\right] = e^{ts\mathbb{E}[XY]}.$$

Using (3.2.3) and expanding both sides yields

$$\sum_{n,m=0}^{\infty} \frac{t^n s^m}{n! m!} \mathbb{E}[H_n(X; \varepsilon_1) H_m(Y; \varepsilon_2)] = \sum_{k=0}^{\infty} \frac{(ts)^k}{k!} \mathbb{E}[XY]^k.$$

Identifying coefficients of the power series gives the result. □

A convenient graphical way of representing (3.2.4), for instance in the case $n = 4$, is

$$\mathbb{E}\left[\;\times\!\!\!\!\times\;\right] = 4! \;\ominus\;. \tag{3.2.5}$$

Here each diagram with four legs represents a term H_4, and each edge on the right-hand side represents a covariance $\mathbb{E}[XY]$. The combinatorial factor 4! counts the number of ways of pairing the four legs, with the rule that only legs from different diagrams can be paired. This is the main difference with Isserlis' theorem, which allowed for all possible pairings. Similar identities as (3.2.4) can be obtained for products of more than two Hermite polynomials, by summing over all pairings of legs belonging to different terms.

Note that if we put $m = 0$ in (3.2.4), we obtain

$$\mathbb{E}[H_n(X; \varepsilon)] = \delta_{n0}, \tag{3.2.6}$$

which shows that Hermite polynomials of Gaussian random variables are centred for $n \geqslant 1$. Another useful consequence of Lemma 3.2.7 is the following binomial formula for Hermite polynomials.

Lemma 3.2.10. *For any* $x, y \in \mathbb{R}$, $\varepsilon_1, \varepsilon_2 \geqslant 0$ *and* $n \in \mathbb{N}_0$,

$$H_n(x + y, \varepsilon_1 + \varepsilon_2) = \sum_{m=0}^{n} \binom{n}{m} H_m(x, \varepsilon_1) H_{n-m}(y, \varepsilon_2).$$

PROOF: Expanding the identity

$$e^{t(x+y)-(\varepsilon_1+\varepsilon_2)t^2/2} = e^{tx-\varepsilon_1 t^2/2} e^{ty-\varepsilon_2 t^2/2}$$

yields

$$\sum_{n=0}^{\infty} \frac{t^n}{n!} H_n(x + y; \varepsilon_1 + \varepsilon_2) = \sum_{m=0}^{\infty} \frac{t^m}{m!} H_m(x; \varepsilon_1) \sum_{k=0}^{\infty} \frac{t^k}{k!} H_k(y; \varepsilon_2).$$

Comparing coefficients of t^n gives the result.   □

3.2.3   The two-dimensional Gaussian free field

We define a Fourier basis of $\Lambda = (\mathbb{R}/(L\mathbb{Z}))^2$ by

$$e_k(x) = e_{k_1}(x_1)e_{k_2}(x_2), \qquad k = (k_1, k_2) \in \mathbb{Z}^2,$$

where the e_{k_i} are the one-dimensional basis functions defined in (2.3.5). The slight abuse of notation made by using the same symbol for one- and two-dimensional basis functions will not matter in what follows. The e_k are eigenfunctions of the Laplacian on Λ with periodic boundary conditions, satisfying

$$\Delta e_k = -\lambda_k e_k, \qquad \lambda_k = \left(\frac{2\|k\|\pi}{L}\right)^2.$$

For $N \in \mathbb{N}$, define sets of indices $\mathcal{K}_N = \{k \in \mathbb{Z}^2 : |k| \leqslant N\}$ and $\mathcal{K}_N^+ = \{k \in \mathcal{K}_N : k_1, k_2 > 0\}$, where $|k| = |k_1| + |k_2|$. Let E_N be the space spanned by $\{e_k : k \in \mathcal{K}_N\}$. Note the identities

$$e_{(k_1,k_2)}(x)^2 + e_{(k_1,-k_2)}(x)^2 + e_{(-k_1,k_2)}(x)^2 + e_{(-k_1,-k_2)}(x)^2 = \frac{4}{L^2} \qquad \forall(k_1, k_2) \in \mathcal{K}_N^+,$$

$$e_{(k_1,0)}(x)^2 + e_{(-k_1,0)}(x)^2 = e_{(0,k_2)}(x)^2 + e_{(0,-k_2)}(x)^2 = \frac{2}{L^2} \qquad \forall k_1, k_2 \neq 0. \qquad (3.2.7)$$

Definition 3.2.11 (Two-dimensional Gaussian free field)**.** *The truncated two-dimensional Gaussian free field (GFF) with covariance* $(-\Delta_N + 1)^{-1}$ *on* Λ *is defined as*

$$\phi_{\mathrm{GFF},N}(x) := \sum_{k \in \mathcal{K}_N} \frac{Z_k}{\sqrt{\lambda_k + 1}} e_k(x),$$

where Δ_N *is the restriction of* Δ *to* E_N *and the* Z_k *are i.i.d. standard normal random variables.*

Thanks to the symmetry relations (3.2.7), $\mathbb{E}[\phi_{\mathrm{GFF},N}(x)^2]$ does not depend on x, and is thus equal to

$$C_N := \frac{1}{L^2} \int_{\Lambda} \mathbb{E}[\phi_{\mathrm{GFF},N}(x)^2] \, dx = \frac{1}{L^2} \int_{\Lambda} \sum_{k \in \mathcal{K}_N} \frac{1}{\lambda_k + 1} e_k(x)^2 \, dx$$

$$= \frac{1}{L^2} \sum_{k \in \mathcal{K}_N} \frac{1}{\lambda_k + 1} = \frac{1}{L^2} \mathrm{Tr}[(-\Delta_N + 1)^{-1}]. \qquad (3.2.8)$$

Since k is two-dimensional and λ_k grows like $\|k\|^2$, this constant diverges like $\log N$.

$$C_N = \frac{\log N}{2\pi} + \mathcal{O}(1) \tag{3.2.9}$$

as $N \to \infty$. **Hint:** View (3.2.8) as a Riemann sum and integrate using polar coordinates. ♣

Thus we see that the non-truncated two-dimensional GFF has infinite variance at every point! Fortunately, since this variance is constant, we can subtract it in order to get a meaningful object in the limit, a procedure which is akin to an Itô correction. This leads to the following definition.

Definition 3.2.13 (Wick powers of the GFF). *For any integer $n \in \mathbb{N}$, the nth Wick power of the truncated GFF is defined as*

$$:\phi_{\mathrm{GFF},N}^n(x): \, = \, :\phi_{\mathrm{GFF},N}^n(x):_{C_N} \, := H_n(\phi_{\mathrm{GFF},N}(x); C_N), \qquad x \in \Lambda.$$

Note in particular that owing to (3.2.6), all Wick powers are centred random variables at any $x \in \Lambda$. Furthermore, Lemma 3.2.9 implies the following important result, which is a generalisation of Lemma 2.7.4 to the two-dimensional case.

Proposition 3.2.14. *For every $n \in \mathbb{N}$, we have*

$$\sup_{N \geqslant 1} \mathbb{E}\left[\left(\frac{1}{L^2} \int_\Lambda :\phi_{\mathrm{GFF},N}^n(x): \mathrm{d}x\right)^2\right] < \infty.$$

PROOF: By (3.2.4), we have

$$\mathbb{E}\left[\left(\int_\Lambda :\phi_{\mathrm{GFF},N}^n(x): \mathrm{d}x\right)^2\right] = \int_\Lambda \int_\Lambda \mathbb{E}[:\phi_{\mathrm{GFF},N}^n(x):: \phi_{\mathrm{GFF},N}^n(y):] \, \mathrm{d}x \, \mathrm{d}y$$

$$= n! \int_\Lambda \int_\Lambda \mathbb{E}[\phi_{\mathrm{GFF},N}(x)\phi_{\mathrm{GFF},N}(y)]^n \, \mathrm{d}x \, \mathrm{d}y$$

$$= n! \int_\Lambda \int_\Lambda \left(\sum_{k,\ell \in \mathscr{K}_N} \frac{\mathbb{E}[Z_k Z_\ell]}{\sqrt{(\lambda_k + 1)(\lambda_\ell + 1)}} e_k(x)e_\ell(y)\right)^n \, \mathrm{d}x \, \mathrm{d}y$$

$$= n! \int_\Lambda \int_\Lambda \left(\sum_{k \in \mathscr{K}_N} \frac{1}{\lambda_k + 1} e_k(x)e_k(y)\right)^n \, \mathrm{d}x \, \mathrm{d}y$$

$$= n! \sum_{k_1,\ldots,k_n \in \mathscr{K}_N} \frac{1}{(\lambda_{k_1} + 1)\ldots(\lambda_{k_n} + 1)}\left(\int_\Lambda e_{k_1}(x)\ldots e_{k_n}(x) \, \mathrm{d}x\right)^2.$$

By orthogonality of the eigenfunctions, the integral in the last expression vanishes unless some linear combination $k_1 \pm k_2 \pm \cdots \pm k_n$ is equal to zero. Therefore, we have the bound

$$\mathbb{E}\left[\left(\int_\Lambda :\phi_{\mathrm{GFF},N}^n(x): \mathrm{d}x\right)^2\right] \lesssim \sum_{\substack{k_1,\ldots,k_n \in \mathscr{K}_N \\ k_1 + \cdots + k_n = 0}} \frac{1}{(\lambda_{k_1} + 1)\ldots(\lambda_{k_n} + 1)}.$$

Thanks to the condition that the sum of the k_i vanishes, this sum is bounded uniformly in N, by a Young-type inequality, see for instance [ZZ15a, Lemma 3.10]. □

Remark 3.2.15. In the graphical notation introduced in (3.2.5), we have $C_N =$ ⬤ , since C_N involves the covariance of any $\phi_{\mathrm{GFF},N}(x)$ with itself. ◇

3.2.4   Wick powers of the stochastic convolution

We consider now the stochastic heat equation

$$\partial_t \phi(t,x) = \Delta \phi(t,x) + \xi_N(t,x)\,, \tag{3.2.10}$$

where $\xi_N = \Pi_N \xi$ is a spectral Galerkin approximation of space-time white noise (cf. (2.7.2)). Similarly to (2.3.3), the solution of (3.2.10) can be written as

$$\phi = P\phi_0 + P * \xi_N\,,$$

where P denotes the analogue of the heat kernel (2.3.2) in two dimensions, namely

$$P(t,x) := \sum_{k \in \mathbb{Z}^2} P_{\mathbb{R}^2}(t, x - kL)\,, \qquad P_{\mathbb{R}^2}(t,x) := \frac{1}{4\pi t} e^{-\|x\|^2/(4t)} \mathbb{1}_{\{t>0\}}\,.$$

The stochastic convolution $P * \xi_N$ can be analysed as in Section 2.3. In particular, subtracting the Brownian motion $W_t^{(0)}$ of the zeroth Fourier mode, we find that $(P * \xi_N)(t,x) - W_t^{(0)}$ converges as $t \to \infty$ to a Gaussian free field with covariance $(-\Delta_{\perp,N})^{-1}$, where $\Delta_{\perp,N} = \Pi_N \Delta_\perp$ acts on zero-mean functions. By Proposition 3.2.14, defining C_N as in (3.2.8), all Wick powers $:(P * \xi_N)^n:_{C_N}$ will have a variance uniformly bounded in N, and should have well-defined limits (one easily checks that adding a constant to the λ_k in (3.2.8) does not change its divergent part).

There is a slightly different way to regularise the stochastic heat equation, which turns out to be equivalent to (3.2.10). Let $\rho : \mathbb{R} \times \Lambda \to \mathbb{R}$ be smooth, compactly supported and of integral 1. Then for any $\delta \in (0,1]$, define the *mollifier*

$$\rho_\delta(t,x) = \mathscr{S}_0^\delta \rho(t,x) := \frac{1}{\delta^4} \rho\!\left(\frac{t}{\delta^2}, \frac{x}{\delta}\right).$$

We can then set $\xi^\delta(t,x) = (\rho_\delta * \xi)(t,x)$ (interpreted as testing ξ against $\rho((t,x) - \cdot)$), and consider the regularised stochastic heat equation

$$\partial_t \phi(t,x) = \Delta \phi(t,x) + \xi^\delta(t,x)\,,$$

which admits solutions in the classical sense since ξ^δ is smooth, namely

$$\phi = P\phi_0 + P * \xi^\delta\,,$$

with a stochastic convolution $P * \xi^\delta$ given by

$$(P * \xi^\delta)(t,x) = \int_{-\infty}^{\infty} \int_\Lambda P(t-s, x-y) \xi^\delta(s,y)\, \mathrm{d}y\, \mathrm{d}s = \langle \xi, P * \rho_\delta \rangle(t,x)\,. \tag{3.2.11}$$

Since this object and its powers will play a central role in the following, we introduce the graphical notation (borrowed from [Hai14, CW17])

$$(P * \xi^\delta)(t,x) = \mathord{\uparrow}_\delta(t,x)\,, \qquad (P * \xi^\delta)(t,x)^2 = \mathord{\vee}_\delta(t,x)\,, \qquad (P * \xi^\delta)(t,x)^3 = \mathord{\psi}_\delta(t,x)\,.$$

Writing $P_\delta = P * \rho_\delta$ and using the defining property (2.2.2) of space-time white noise, we obtain

$$\mathbb{E}[\mathord{\vee}_\delta(t,x)] = \mathbb{E}[\langle \xi, P_\delta \rangle(t,x)^2] = \langle P_\delta, P_\delta \rangle_{L^2(\mathbb{R} \times \Lambda)} = \int_{-\infty}^{\infty} \int_\Lambda P_\delta(t,x)^2\, \mathrm{d}x\, \mathrm{d}t =: C_\delta\,, \tag{3.2.12}$$

which is independent of (t,x).

Proposition 3.2.16. *Assume that $\rho(t,x) = \rho^{(0)}(t)\rho^{(1)}(x)$ where $\rho^{(1)}$ is even. Then the constant C_δ has the asymptotic behaviour*

$$C_\delta = \frac{\log(\delta^{-1})}{4\pi} + \mathcal{O}(1) \qquad as \ \delta \to 0 \,. \tag{3.2.13}$$

PROOF: We present a proof which is perhaps not the shortest possible, but will highlight the link with other object such as the Green function. First note that the Markov property implies that for any $t,s > 0$ and $x \in \Lambda$,

$$\int_\Lambda P(t, x-y) P(s, y) \, \mathrm{d}y = P(t+s, x) \,.$$

Furthermore, for any $x \in \Lambda$,

$$\int_{-\infty}^{\infty} P(t,x) \, \mathrm{d}t = \int_0^\infty e^{t\Delta}(x) \, \mathrm{d}t = -G(x) \,,$$

where $G = \Delta^{-1}$ is the Green function of the Laplacian acting on zero-mean functions, that is, the solution of $(\Delta G)(x) = \delta(x)$. Taking into account the properties of ρ, one finds that for $\delta > 0$, these relations become

$$\int_\Lambda P_\delta(t, x-y) P_\delta(s, y) \, \mathrm{d}y = \tilde{P}_\delta(t+s, x) \,,$$
$$\int_{-\infty}^{\infty} P_\delta(t,x) \, \mathrm{d}t = -G_\delta(x) \,, \tag{3.2.14}$$

where $G_\delta = G * \rho_\delta^{(1)}$ and $\tilde{P}_\delta = P * (\rho * \rho)_\delta$ is a regularised version of P with a different mollifier. It follows, using that P is even in x, that

$$\int_{-\infty}^{\infty} \int_\Lambda P_\delta(t,x)^2 \, \mathrm{d}x \, \mathrm{d}t = \int_{-\infty}^{\infty} \tilde{P}_\delta(2t, 0) \, \mathrm{d}t = -\frac{1}{2}\tilde{G}_\delta(0) \,,$$

where again $\tilde{G}_\delta$ is a regularisation of G with a different mollifier. Using the fact that the Green function of the two-dimensional Laplacian behaves like $\log\|x\|/(2\pi)$ near the origin, one easily obtains $\tilde{G}_\delta(0) = \log(\delta)/(2\pi) + \mathcal{O}(1)$, which proves (3.2.13). $\qquad\square$

Remark 3.2.17. Note that the expressions (3.2.13) and (3.2.9) have the same divergent behaviour. Furthermore, the divergent part of C_δ is independent of the mollifier ρ. The term of order 1, however, will depend on ρ. In other words, there is no canonical choice for the bounded part of the renormalisation constants. $\qquad\diamond$

Proposition 3.2.14 shows that the Wick powers

$$\begin{aligned}
{:}\mathord{\uparrow}_\delta{:} &:= H_1(\mathord{\uparrow}_\delta; C_\delta) = \mathord{\uparrow}_\delta \,, \\
{:}\mathord{\vee}_\delta{:} &:= H_2(\mathord{\uparrow}_\delta; C_\delta) = \mathord{\vee}_\delta - C_\delta \,, \\
{:}\mathord{\psi}_\delta{:} &:= H_3(\mathord{\uparrow}_\delta; C_\delta) = \mathord{\psi}_\delta - 3C_\delta \mathord{\uparrow}_\delta
\end{aligned} \tag{3.2.15}$$

have a variance uniformly bounded in δ. In fact, one can show that they admit well-defined limits $\mathord{\uparrow}$, $\mathord{\vee}$ and $\mathord{\psi}$ as $\delta \searrow 0$ in terms of iterated stochastic integrals, which belong to $\mathscr{C}_{\mathfrak{s}}^{-\kappa}$ for any $\kappa > 0$. See [CW17, Appendix A], [Hai14, Section 10.1] and [DPD03, Lemma 3.2] for details. The same works of course for higher powers of the stochastic convolution.

Remark 3.2.18. A formal way of writing these limits is to introduce

$$\xi(\mathrm{d}z_1) \diamond \xi(\mathrm{d}z_2) = \xi(\mathrm{d}z_1)\xi(\mathrm{d}z_2) - \delta(z_1 - z_2),$$

$$\xi(\mathrm{d}z_1) \diamond \xi(\mathrm{d}z_2) \diamond \xi(\mathrm{d}z_3) = \xi(\mathrm{d}z_1)\xi(\mathrm{d}z_2)\xi(\mathrm{d}z_3)$$
$$- \xi(\mathrm{d}z_1)\delta(z_2 - z_3) - \xi(\mathrm{d}z_2)\delta(z_3 - z_1) - \xi(\mathrm{d}z_3)\delta(z_1 - z_2).$$

The quantities on the left-hand side, which define the limiting objects $\vee$ and Ψ when integrated against a product of P_δ, can be given a rigorous meaning as elements of the Wiener chaos decomposition, see [Nua06, Section 1.1], and [CW17, Appendix A].   ◇

3.3   Existence and uniqueness of solutions

The discussion so far suggests that the renormalised version of the two-dimensional stochastic Allen–Cahn equation we should consider is

$$\partial_t \phi_\delta(t,x) = \Delta\phi_\delta(t,x) + \phi_\delta(t,x) - :\phi_\delta(t,x)^3:_{C_\delta} + \xi^\delta(t,x)$$
$$= \Delta\phi_\delta(t,x) + \phi_\delta(t,x) - \left(\phi_\delta(t,x)^3 - 3C_\delta\phi_\delta(t,x)\right) + \xi^\delta(t,x). \qquad (3.3.1)$$

For $\delta > 0$, this is a smooth PDE (for any realisation of the noise), whose solution with initial condition $\phi_{\delta,0}$ satisfies the fixed-point equation

$$\phi_\delta = P\phi_{\delta,0} + \mathord{\uparrow}_\delta + P * [\phi_\delta - \phi_\delta^3 + 3C_\delta\phi_\delta]. \qquad (3.3.2)$$

The idea exploited by Da Prato and Debussche in [DPD03] is that since the most irregular term on the right-hand side should be the stochastic convolution $\mathord{\uparrow}_\delta$, one can write

$$\phi_\delta = \mathord{\uparrow}_\delta + \psi_\delta,$$

where ψ_δ is expected to be more regular than ϕ_δ. Substituting in (3.3.2), we obtain

$$\psi_\delta = P\psi_{\delta,0} + P * [\mathord{\uparrow}_\delta + \psi_\delta - (\mathord{\uparrow}_\delta + \psi_\delta)^3 + 3C_\delta(\mathord{\uparrow}_\delta + \psi_\delta)]$$
$$= P\psi_{\delta,0} + P * [\mathord{\uparrow}_\delta + \psi_\delta - \psi_\delta^3 - 3\mathord{\uparrow}_\delta\psi_\delta^2 - 3(\vee_\delta - C_\delta)\psi_\delta - (\Psi_\delta - 3C_\delta\mathord{\uparrow})].$$

Taking formally the limit $\delta \searrow 0$ and using (3.2.15) yields the fixed-point equation

$$\psi = P\psi_0 + P * [\mathord{\uparrow} + \psi - \psi^3 - 3\mathord{\uparrow}\psi^2 - 3\vee\psi - \Psi]. \qquad (3.3.3)$$

The key result allowing to analyse this fixed-point equation is the following, see [Hai14, Proposition 4.14] as well as [BCD11, Section 2.6].

Theorem 3.3.1 (Product in Besov spaces). *Let $\alpha, \beta \in \mathbb{R}$ satisfy $\alpha + \beta > 0$. Then there exists a bilinear map $B : \mathscr{C}_{\mathfrak{s}}^\alpha \times \mathscr{C}_{\mathfrak{s}}^\beta \to \mathscr{C}_{\mathfrak{s}}^{\alpha \wedge \beta}$ with the following properties:*

1. If $f \in \mathscr{C}_{\mathfrak{s}}^\alpha$ and $g \in \mathscr{C}_{\mathfrak{s}}^\beta$ are continuous functions, then $B(f,g)(z) = f(z)g(z)$.

2. For arbitrary $f \in \mathscr{C}_{\mathfrak{s}}^\alpha$ and $g \in \mathscr{C}_{\mathfrak{s}}^\beta$, one has the bound

$$\|B(f,g)\|_{\mathscr{C}_{\mathfrak{s}}^{\alpha \wedge \beta}} \lesssim \|f\|_{\mathscr{C}_{\mathfrak{s}}^\alpha} \|g\|_{\mathscr{C}_{\mathfrak{s}}^\beta}.$$

If $\alpha + \beta \leqslant 0$, then no bilinear map satisfying these two properties exists.

We then have the following existence and uniqueness result.

Theorem 3.3.2 (Global existence and uniqueness for the two-dimensional Allen–Cahn SPDE).
Assume $\psi_0 \in \mathscr{C}_{\mathfrak{s}}^{-\kappa}(\Lambda)$. For sufficiently small $\kappa > 0$, equation (3.3.3) with initial condition ψ_0 admits an almost surely unique global solution in $\mathscr{C}_{\mathfrak{s}}^{\alpha}(\mathbb{R}_+ \times \Lambda)$ for any $\alpha < 2 - \kappa$. Therefore, the solution of (3.3.1) with initial condition $\rho_\delta * \psi_0$ converges almost surely, as $\delta \searrow 0$, to a stochastic process $(\psi_t)_{t \geqslant 0}$ such that $\psi - \uparrow \in \mathscr{C}_{\mathfrak{s}}^{\alpha}(\mathbb{R}_+ \times \Lambda)$.

SKETCH OF PROOF: We first prove a local existence result, in the spirit of Theorem 2.4.2. We know that $\uparrow, \vee, \mathbb{V}$ belong to $\mathscr{C}_{\mathfrak{s}}^{-\kappa}$ for any $\kappa > 0$, while an analogue of the Schauder estimate in Theorem 2.3.9 shows that $P\psi_0 \in \mathscr{C}_{\mathfrak{s}}^{2-\kappa}([0,T] \times \Lambda)$. Using Theorem 3.3.1, we obtain for any $\alpha, T > 0$ the estimate

$$\|\uparrow + \psi - \psi^3 - 3\uparrow\psi^2 - 3\vee\psi - \mathbb{V}\|_{\mathscr{C}_{\mathfrak{s}}^{-\kappa}([0,T] \times \Lambda)} \lesssim (\|\psi\|_{\mathscr{C}_{\mathfrak{s}}^{\alpha}([0,T] \times \Lambda)} + M)^3$$

for some constant M. The Schauder estimate shows that the convolution of this quantity with P belongs to $\mathscr{C}_{\mathfrak{s}}^{2-\kappa}([0,T] \times \Lambda)$. This is not sufficient to obtain a contraction, since the norm does not become small for small T. Such a T-dependence can, however, be obtained by sacrificing a bit of regularity. One can indeed show that, for any $\sigma > 0$,

$$\|P * [\uparrow + \psi - \psi^3 - 3\uparrow\psi^2 - 3\vee\psi - \mathbb{V}]\|_{\mathscr{C}_{\mathfrak{s}}^{2-\kappa-\sigma}([0,T] \times \Lambda)} \lesssim T^{\sigma}(\|\psi\|_{\mathscr{C}_{\mathfrak{s}}^{\alpha}([0,T] \times \Lambda)} + M)^3.$$

Choosing $\alpha = 2 - \kappa - \sigma$, we obtain that the right-hand side of (3.3.3) maps a ball into itself for sufficiently small T. A similar argument shows that this map is a contraction in that ball for small T, so that the existence of a local solution follows from Banach's fixed-point theorem. Since σ was arbitrary, the fixed point has the claimed regularity.

Global existence then follows from showing non-explosion, by using an appropriate Lyapunov function and a Gronwall-type argument. See [DPD03, p. 1914] and [TW18b, Section 3.4], as well as [MW17b, Section 9] for an argument working when Λ is replaced by $\mathbb{R}^2$. □

Remark 3.3.3. This result appears with different variants of function spaces in the literature. The proof given in [DPD03] uses spaces mixing Lebesgue spaces L^p in the time variable and Besov spaces $\mathscr{B}_{q,r}^{\alpha}$ in the space variable. In [TW18b, Section 3], one finds a formulation with Hölder regularity in space and continuity in time. The formulation here follows [Hai14]. The result remains true for any polynomial nonlinearity of odd degree and negative leading-order coefficient. ⋄

3.4 Invariant measure

We now consider the stochastic Allen–Cahn equation written in the form

$$\partial_t \phi(t,x) = \Delta\phi(t,x) + \phi(t,x) - {:}\phi(t,x)^3{:} + \sqrt{2\varepsilon}\,\xi(t,x)\,, \tag{3.4.1}$$

where we have reintroduced the small parameter in front of the noise for later use. Here the Wick power is with respect to the constant $\varepsilon\overline{C}_\delta = 2\varepsilon C_\delta$ since the variance of the noise has been multiplied by 2ε.

By analogy with (2.5.6), a natural candidate for the invariant measure of (3.4.1) is

$$\pi(\mathrm{d}\phi) := \frac{1}{\mathscr{Z}_0}\,\mathrm{e}^{-\widetilde{V}(\phi)/\varepsilon}\,\mu_{\mathrm{GFF}}^{(\varepsilon)}(\mathrm{d}\phi)\,, \qquad \widetilde{V}(\phi) := \int_\Lambda \left[\frac{1}{4}{:}\phi^4(x){:} - {:}\phi^2(x){:} + \frac{1}{4}\right]\mathrm{d}x\,,$$

where $\mu_{\mathrm{GFF}}^{(\varepsilon)}$ is the measure of the Gaussian free field with covariance $\varepsilon(-\Delta + 1)^{-1}$ (and the constant term $1/4$ plays no role but will be convenient in computations). Note that the integral over Λ has finite expectation and variance by Proposition 3.2.14. However, this does not automatically imply that the partition function $\mathscr{Z}_0$ is finite.

In the case of the Φ^4 model (without the linear term $\phi(t,x)$ in (3.4.1), cf. Remark 2.0.1), this problem is in fact an old problem in Quantum Field Theory, which has been solved by various methods. A particularly elegant approach has recently been developed by Barashkov and Gubinelli [BG18]. To adapt it to the nonconvex potential of the Allen–Cahn equation, we start by performing, in a similar way as in Section 2.7, cf. (2.7.3), the change of variables

$$\phi(x) = \frac{1}{L}\phi_0 + \sqrt{\varepsilon}\phi_\perp(x)\,, \tag{3.4.2}$$

where $\phi_\perp$ has zero mean. Using Lemma 3.2.10 to transform Wick powers, we obtain

$$\frac{1}{\varepsilon}\widetilde{V}(\phi) = \frac{1}{\varepsilon}\left[V_0(\phi_0) - \frac{1}{2}\phi_0^2\right] + \frac{1}{2}\mathcal{Q}_\perp(\phi_0,\phi_\perp) + R_\varepsilon(\phi_0,\phi_\perp)\,,$$

where

$$V_0(\phi_0) = \frac{L^2}{4}\left(\frac{\phi_0^2}{L^2} - 1\right)^2,$$

$$\mathcal{Q}_\perp(\phi_0,\phi_\perp) = \left(\frac{3\phi_0^2}{L^2} - 2\right)\int_\Lambda :\phi_\perp(x)^2:\,\mathrm{d}x\,,$$

$$R_\varepsilon(\phi_0,\phi_\perp) = \frac{\sqrt{\varepsilon}}{L}\phi_0\int_\Lambda :\phi_\perp(x)^3:\,\mathrm{d}x + \frac{\varepsilon}{4}\int_\Lambda :\phi_\perp(x)^4:\,\mathrm{d}x\,. \tag{3.4.3}$$

This yields the expression

$$\mathcal{Z}_0 = \mathbb{E}^{\mu_{\mathrm{GFF}}^{(\varepsilon)}}[\mathrm{e}^{-\widetilde{V}/\varepsilon}] = \frac{1}{\sqrt{2\pi\varepsilon}}\int_{-\infty}^{\infty} \mathrm{e}^{-V_0(\phi_0)/\varepsilon}\,\mathbb{E}^{\mu_{\mathrm{GFF},\perp}}\left[\mathrm{e}^{-\mathcal{Q}_\perp(\phi_0,\cdot)/2 - R_\varepsilon(\phi_0,\cdot)}\right]\mathrm{d}\phi_0 \tag{3.4.4}$$

for the partition function, where $\mu_{\mathrm{GFF},\perp}$ has covariance $(-\Delta_\perp + 1)^{-1}$.

Our aim is now to bound the expectation in (3.4.4). One way of doing this, which was used in [BDGW17], is based on the so-called Nelson estimate, a refinement of Proposition 3.2.14. We outline here instead the approach used in [BG18], which relies on the Boué–Dupuis formula. The idea is that if $\mu_\perp(a)$ is a Gaussian measure with covariance $(-\Delta_\perp + a)^{-1}$, it can be regularised in the following way. Let $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ be a compactly supported smooth function such that $\rho(0) = 0$, and set, for each $k \in \mathbb{Z}^2$,

$$v_k(t) := \frac{\sigma_k(t)}{\sqrt{\lambda_k^2 + a}}\,, \qquad \sigma_k(t)^2 := \frac{\mathrm{d}}{\mathrm{d}t}\rho\left(\frac{\|k\|}{t}\right).$$

Then introduce the stochastic process

$$Y_t(x) := \sum_{k\in\mathbb{Z}^2\setminus\{0\}}\int_0^t v_k(s)\,\mathrm{d}W_s^{(k)}e_k(x)\,,$$

where the $W_t^{(k)}$ are independent Wiener processes. Using Itô's isometry, we obtain

$$\mathbb{E}[\langle Y_t,\varphi\rangle\langle Y_s,\psi\rangle] = \sum_{k\in\mathbb{Z}^2\setminus\{0\}}\int_0^{t\wedge s} v_k(u)^2\,\mathrm{d}u\,\overline{\hat{\varphi}_k}\hat{\psi}_k = \sum_{k\in\mathbb{Z}^2\setminus\{0\}}\rho\left(\frac{\|k\|}{t\wedge s}\right)\overline{\hat{\varphi}_k}\hat{\psi}_k$$

for any test functions φ and ψ with Fourier coefficients $\hat{\varphi}_k$ and $\hat{\psi}_k$. In particular,

$$\lim_{t\to\infty}\mathbb{E}[\langle Y_t,\varphi\rangle\langle Y_t,\psi\rangle] = \sum_{k\in\mathbb{Z}^2\setminus\{0\}}\frac{1}{\lambda_k^2 + a}\overline{\hat{\varphi}_k}\hat{\psi}_k = \langle\varphi,(-\Delta_\perp + a)^{-1}\psi\rangle\,.$$

Then the Boué–Dupuis variational formula, whose proof relies on the Girsanov transformation [BG18, Section 2], reads as follows.

Theorem 3.4.1 (Boué–Dupuis formula, [BD98])**.** *For any $T > 0$ and functional $\mathscr{V}$, we have*

$$-\log \mathbb{E}[\mathrm{e}^{-\mathscr{V}(Y_T)}] = \inf_{u \in \mathbb{H}} \mathbb{E}\left[\mathscr{V}(Y_T + Z_T(u)) + \frac{|\Lambda|}{2}\int_0^T \|u_t\|_{L^2}^2 \, \mathrm{d}t\right], \tag{3.4.5}$$

where

$$Z_T(u, x) := \sum_{k \in \mathbb{Z}^2 \setminus \{0\}} \int_0^T u_k(t) \, \mathrm{d}t \, e_k(x)$$

and $\mathbb{H}$ is the space of progressively measurable processes which are almost surely in $L^2(\mathbb{R}_+ \times \Lambda)$.

Note that here the parameter T does not represent time, but a continuous regularisation parameter such that in the limit $T \to \infty$, one recovers the expectation under the Gaussian measure $\mu_{\mathrm{GFF},\perp}$. This leads to the following estimate, which extends Proposition 2.7.5 to the two-dimensional case (the bounds' ε-dependence can be improved, cf. Proposition 3.6.3).

Proposition 3.4.2. *Assume that $L < 2\pi$. Then there exists a constant c such that*

$$1 \leqslant \mathbb{E}^{\mu_{\mathrm{GFF},\perp}}\left[\mathrm{e}^{-\mathscr{Q}_\perp(\phi_0,\cdot)/2 - R_\varepsilon(\phi_0,\cdot)}\right] \leqslant \mathrm{e}^{c\phi_0^4}.$$

As a consequence, the partition function $\mathscr{Z}_0$ is bounded.

SKETCH OF PROOF: The lower bound follows from Jensen's inequality and the fact that Wick powers are centred. For the upper bound, we first perform a shift to the Gaussian measure $\mu_\perp$ of covariance $(-\Delta_\perp - 1)^{-1}$ (which is allowed since the eigenvalues of $-\Delta_\perp$ are greater than 1 for $L < 2\pi$). One obtains

$$\mathbb{E}^{\mu_{\mathrm{GFF},\perp}}\left[\mathrm{e}^{-\mathscr{Q}_\perp(\phi_0,\cdot)/2 - R_\varepsilon(\phi_0,\cdot)}\right] = K\mathbb{E}^{\mu_\perp}\left[\mathrm{e}^{-\mathscr{V}(\phi_0,\cdot)}\right],$$

where

$$\mathscr{V}(\phi_0, \phi_\perp) := R_\varepsilon(\phi_0, \phi_\perp) + \frac{3\phi_0^2}{2L^2}\int_\Lambda :\phi_\perp(x)^2: \mathrm{d}x,$$

and K is a constant. This is actually a nontrivial point, to which we will come back in Section 3.6, see Lemma 3.6.1. Expanding the Wick powers of $Y_T + Z_T(u)$ with the help of Proposition 3.2.14, and inserting this in (3.4.5), we get

$$-\log \mathbb{E}[\mathrm{e}^{-\mathscr{V}(\phi_0,\cdot)}] = \inf_{u \in \mathbb{H}} \mathbb{E}\left[\Psi_T(u) + \Phi_T(u)\right],$$

where, since all Wick powers of Y_T have zero expectation,

$$\Psi_T(u) = \frac{\varepsilon}{4}\int_\Lambda Z_T^4(u)\,\mathrm{d}x + \frac{3\phi_0^2}{2L^2}\int_\Lambda Z_T^2(u)\,\mathrm{d}x + \frac{|\Lambda|}{2}\int_0^T \|u_t\|_{L^2}^2\,\mathrm{d}t,$$

$$\Phi_T(u) = \varepsilon\int_\Lambda :Y_T^3:Z_T(u)\,\mathrm{d}x + \frac{3}{2}\varepsilon\int_\Lambda :Y_T^2:Z_T^2(u)\,\mathrm{d}x + \varepsilon\int_\Lambda Y_T Z_T^3(u)\,\mathrm{d}x$$

$$+ \frac{\sqrt{\varepsilon}}{L}\phi_0\left[3\int_\Lambda :Y_T^2:Z_T(u)\,\mathrm{d}x + 3\int_\Lambda Y_T Z_T^2(u)\,\mathrm{d}x + \int_\Lambda Z_T^3(u)\,\mathrm{d}x\right] + \frac{3\phi_0^2}{2L^2}\int_\Lambda Y_T Z_T(u)\,\mathrm{d}x.$$

Here one has to show that for each choice of $u \in \mathbb{H}$, the positive term $\Psi_T(u)$ dominates $\Phi_T(u)$ uniformly in T up to a remainder of order ϕ_0^4. This is a purely deterministic argument, which is based on functional inequalities, cf. [BG18, Section 3]. The claim on $\mathscr{Z}_0$ then follows from Laplace asymptotics on (3.4.4). $\qquad\square$

The existence of the Gibbs measure π being established, it is natural to ask whether π is indeed invariant under the dynamics of the Allen–Cahn equation (3.4.1), whether it is the unique invariant measure, and what are its ergodic properties. This task is not straightforward at all, since even the fact that solutions of (3.4.1) enjoy the Markov property is not automatic, but requires a proof. All these properties have in fact been established recently. We summarise them in the following theorem, which incorporates results from [RZZ17b, TW18b, RZZ17a, HM18].

Theorem 3.4.3 (Ergodic properties of the two-dimensional Allen–Cahn equation). *The solution of* (3.4.1) *defines a Markov process* $(\phi_t)_{t \geqslant 0}$ *with transition semi-group* $(P_t)_{t \geqslant 0}$ *with respect to the filtration* $(\mathscr{F}_t)_{t \geqslant 0}$ *generated by all random variables* $\langle \xi, \varphi \rangle$ *obtained by testing space-time white noise against functions* φ *supported on* $[0, t] \times \Lambda$. *The semigroup* $(P_t)_{t \geqslant 0}$ *satisfies the* strong Feller *property, that is, it maps bounded measurable functions to continuous functions for every* $t > 0$. *The Gibbs measure* $\pi(\mathrm{d}\phi)$ *is the unique invariant measure of the dynamics, and is reversible. Finally, the dynamics is exponentially mixing, in the sense that there exist constants* $C, \beta > 0$ *such that*

$$\|P_t(\phi_0) - \pi\|_{\mathrm{TV}} \leqslant C\, e^{-\beta t} \|\delta_{\phi_0} - \pi\|_{\mathrm{TV}} \tag{3.4.6}$$

holds for all $\phi_0 \in \mathscr{C}_{\mathfrak{s}}^{-\kappa}(\Lambda)$ *and all* $t > 0$, *where* $\|\cdot\|_{\mathrm{TV}}$ *denotes the total variation distance.*

3.5   Large deviations

Consider the Allen–Cahn equation on the torus Λ with weak mollified space-time white noise

$$\partial_t \phi(t, x) = \Delta\phi(t, x) + \phi(t, x) - \left(\phi(t, x)^3 - 3\varepsilon \overline{C}_\delta \phi(t, x)\right) + \sqrt{2\varepsilon}\xi^\delta(t, x), \tag{3.5.1}$$

where $\overline{C}_\delta = 2C_\delta$. Note that we have now two small parameter δ and ε, and that the counterterm $3\varepsilon \overline{C}_\delta \phi$ has been scaled accordingly (recall from (3.2.13) that $\overline{C}_\delta$ diverges logarithmically). For fixed $\varepsilon > 0$, we know that this equation admits a well-defined limit as $\delta \searrow 0$. When $\delta, \varepsilon > 0$, (3.5.1) is a well-posed PDE for any realisation of the noise. In [HW15], Hairer and Weber have established the following large-deviation principle.

Theorem 3.5.1 (Large-deviation principle for the renormalised equation). *Let* $\phi_{\delta,\varepsilon}$ *denote the solution of* (3.5.1), *and let* $\varepsilon \mapsto \delta(\varepsilon) \geqslant 0$ *be a function such that*

$$\lim_{\varepsilon \to 0} \delta(\varepsilon) = 0. \tag{3.5.2}$$

Then the family $(\phi_{\delta(\varepsilon),\varepsilon})_{\varepsilon > 0}$ *with fixed initial condition* ϕ_0 *satisfies an LDP on* $[0, T]$ *with good rate function*

$$\mathscr{I}_{[0,T]}(\gamma) := \begin{cases} \dfrac{1}{2}\displaystyle\int_0^T \int_\Lambda \left[\dfrac{\partial}{\partial t}\gamma(t, x) - \Delta\gamma(t, x) - \gamma(t, x) + \gamma(t, x)^3\right]^2 \mathrm{d}x\,\mathrm{d}t & \textit{if the integral is finite}, \\[2ex] +\infty & \textit{otherwise}. \end{cases} \tag{3.5.3}$$

The remarkable thing about this result is that the counterterm $3\varepsilon \overline{C}_\delta \phi$ does not appear in the rate function, which suggests that the Wick power consisting of cubic term and counterterm is the object that should really be considered physically relevant.

On the other hand, for $\delta > 0$, we may also consider the variant of (3.5.1) without counterterm

$$\partial_t \phi(t, x) = \Delta\phi(t, x) + C\phi(t, x) - \phi(t, x)^3 + \sqrt{2\varepsilon}\xi^\delta(t, x), \tag{3.5.4}$$

which does not admit a limit as $\delta \searrow 0$. Here, one can also obtain a large-deviation result, which reads as follows.

Theorem 3.5.2 (Large-deviation principle for equation without renormalisation)**.** *Let $\tilde{\phi}_{\delta,\varepsilon}$ denote the solution of* (3.5.4), *and let $\varepsilon \mapsto \delta(\varepsilon) \geqslant 0$ be a function satisfying* (3.5.2) *as well as*

$$\lim_{\varepsilon \to 0} \varepsilon \log\big(\delta(\varepsilon)^{-1}\big) = \lambda \in [0, \infty) \, .$$

Then the family $(\tilde{\phi}_{\delta(\varepsilon),\varepsilon})_{\varepsilon>0}$ with fixed initial condition ϕ_0 satisfies an LDP on $[0,T]$ with good rate function

$$\mathscr{I}_{[0,T]}(\gamma) := \begin{cases} \dfrac{1}{2} \displaystyle\int_0^T \int_\Lambda \Big[\dfrac{\partial}{\partial t} \gamma(t,x) - \Delta\gamma(t,x) + C^{(\lambda)}\gamma(t,x) + \gamma(t,x)^3 \Big]^2 \, \mathrm{d}x\,\mathrm{d}t & \text{if the integral is finite}, \\ +\infty & \text{otherwise}, \end{cases}$$

where $C^{(\lambda)} = C - 3\lambda^2/(4\pi)$.

Note that by the asymptotics (3.2.13) of C_δ, the results of both theorems indeed overlap if C is such that $C_\lambda = 0$.

3.6 Metastability

We consider now the long-time dynamics of the renormalised equation (3.5.1) for small positive ε, again for simplicity in the case $L < 2\pi$, when the only stationary solutions of the deterministic equation have constant value 0 or ± 1. As before, we denote these stationary solutions ϕ_0^* and $\phi_\pm^*$. Note that at first glance, one might think that the stable stationary solutions of (3.5.1) are located in $\pm[1 + 3\varepsilon\overline{C}_\delta]^{1/2}$ instead of ± 1. However, the large-deviation estimate (3.5.3) suggests otherwise, and we shall see that indeed the system behaves in some sense as if the counterterm were absent.

As in the one-dimensional case, we have no good control on the ε-dependence of the constant β controlling the speed of convergence in the exponential mixing result (3.4.6). A natural approach to characterise this speed is thus to establish asymptotic results on the expected transition time τ_+ from ϕ_-^* to ϕ_+^*. The LDP given in Theorem 3.5.1 can be used to show that, just as in the one-dimensional case, this time satisfies the Arrhenius law

$$\lim_{\varepsilon \to 0} \varepsilon \log \mathbb{E}^{\phi_-^*}[\tau_+] = H \, ,$$

where for $L < 2\pi$, one has $H = L^2/4$. Note that the fact that the rate function (3.5.3) does not depend on C_δ is crucial here.

The next step is to derive Eyring–Kramers asymptotics for $\mathbb{E}^{\phi_-^*}[\tau_+]$. One may indeed wonder whether there exists an analogue to Theorem 2.7.7 on the one-dimensional Allen–Cahn equation. Here, however, we note that the Fredholm determinant

$$\det\big((-\Delta - 1)(-\Delta + 2)^{-1}\big) = \det\big(\mathbb{1} - 3(-\Delta + 2)^{-1}\big).$$

that would appear in the prefactor is *not* well-defined, since $(-\Delta + 2)^{-1}$ is not trace class in two dimensions. But one should not forget that the Eyring–Kramers formula also involves a potential difference. If the potential V is defined by

$$V(\phi) = \int_\Lambda \Big[\frac{1}{2}\|\nabla\phi(x)\|^2 + \frac{1}{4}{:}\phi(x)^4{:} - \frac{1}{2}{:}\phi(x)^2{:} + \frac{1}{4} \Big] \mathrm{d}x \, , \tag{3.6.1}$$

one can check that

$$V(\phi_0^*) - V(\phi_-^*) = \frac{L^2}{4} + \frac{3}{2}L^2\varepsilon\overline{C}_\delta \,.$$

It turns out that the extra term in $\overline{C}_\delta$ exactly compensates the divergence of the Fredholm determinant, replacing it by a so-called *Carleman–Fredholm* (renormalised) determinant.

To see this, we will work with the spectral Galerkin approximation of the SPDE given by

$$\partial_t\phi_N(t,x) = \Delta\phi_N(t,x) + \phi_N(t,x) - \Pi_N\phi_N(t,x)^3 + 3\varepsilon C_N\phi_N(t,x) + \sqrt{2\varepsilon}\Pi_N\xi(t,x)\,, \qquad (3.6.2)$$

where Π_N is the projection on Fourier modes with wave vector k satisfying $|k| = |k_1| + |k_2| \leqslant N$. As before, we denote by $\mathscr{K}_N$ the set of these k. In order to keep track of the choice of the order-one part of the renormalisation constant, which is in principle arbitrary, it will be convenient to choose C_N such that

$$L^2 C_N = \mathrm{Tr}\big(|-\Delta_N - 1|^{-1}\big) + \theta = \sum_{k\in\mathscr{K}_N}\frac{1}{|\lambda_k - 1|} + \theta = L^2\frac{\log N}{2\pi} + \mathcal{O}(1)\,, \qquad (3.6.3)$$

where $\theta \in \mathbb{R}$ is a free parameter. The renormalised potential is then given by (3.6.1) where Wick powers are with respect to εC_N. Using the decomposition (3.4.2) of ϕ into its mean and oscillatory parts, we obtain

$$\frac{1}{\varepsilon}V(\phi) = \frac{1}{\varepsilon}V_0(\phi_0) + \frac{1}{2}\langle\phi_\perp, Q_\perp(\phi_0)\phi_\perp\rangle + \frac{1}{2}\big(1 - m^2(\phi_0)\big)L^2 C_N + R_\varepsilon(\phi_0, \phi_\perp)\,,$$

where V_0 and R_ε are as in (3.4.3), $Q_\perp(\phi_0) = -\Delta_\perp - 1 + m^2(\phi_0)$ and $m^2(\phi_0) = 3\phi_0^2/L^2$. Note that compared to the one-dimensional case, cf. (2.7.4), there is an additional term proportional to C_N, which is due to our using Wick powers in the potential.

The key observation here is the following explicit expression for a change of mass in a Gaussian measure.

Lemma 3.6.1. *For $a > 0$, let $\mu(a)$ be the Gaussian measure with covariance $(-\Delta_N + a)^{-1}$. Then for any $\mu(a)$-integrable random variable $F(\phi_N)$ and any $b > 0$, one has*

$$\mathbb{E}^{\mu(a)}[F(\phi_N)] = \sqrt{\mathfrak{D}_N(a - b; b)}\,\mathbb{E}^{\mu(b)}\bigg[\exp\Big\{-\frac{a-b}{2}\int_\Lambda :\phi_N(x)^2: \mathrm{d}x\Big\}F(\phi_N)\bigg]\,,$$

where the Wick power is with respect to $\mathrm{Tr}(-\Delta + b)^{-1}/L^2$, and for any $b \notin \mathrm{spec}(\Delta_N)$, we define

$$\mathfrak{D}_N(c; b) = \det{}_2\big(\mathbb{1} + c(-\Delta_N + b)^{-1}\big) := \det\big(\mathbb{1} + c(-\Delta_N + b)^{-1}\big)e^{-c\,\mathrm{Tr}(-\Delta_N + b)^{-1}}\,. \qquad (3.6.4)$$

The modified determinant $\det_2(\mathbb{1} + \mathscr{L}) = \det(\mathbb{1} + \mathscr{L})e^{-\mathrm{Tr}\,\mathscr{L}}$ of a linear operator $\mathscr{L}$, as appearing in (3.6.4), is called its *Carleman–Fredholm determinant*, and is convergent whenever $\mathscr{L}$ is *Hilbert–Schmidt*, meaning that only $\mathscr{L}^\dagger\mathscr{L}$ needs to be trace class, see [Sim05, Chapter 5].

Exercise 3.6.2. Use a Fourier basis $\{e_k\}_{k\in\mathscr{K}_N}$ in order to

- show that there is a constant M uniform in N, b and c such that $e^{-Mc^2/b^2} \leqslant \mathfrak{D}_N(c; b) \leqslant e^{Mc^2/b^2}$;
- show that $\mathfrak{D}_N(a - b; b)^{-1} = \mathfrak{D}_N(b - a; a)\exp\big\{(a - b)^2\,\mathrm{Tr}[(-\Delta + a)^{-1}(-\Delta + b)^{-1}]\big\}$;
- prove Lemma 3.6.1, using Parseval's identity.   ♣

As in the one-dimensional case, the Galerkin approximation (3.6.2) is equivalent to a finite-dimensional Itô SDE

$$\mathrm{d}\hat{\phi}_t = -\nabla\widehat{V}_N(\hat{\phi}_t)\,\mathrm{d}t + \sqrt{2\varepsilon}\,\mathrm{d}W_t\,,$$

with a potential $\widehat{V}_N$ depending now explicitly on C_N. We can thus apply as before the potential-theoretic approach. To be able to apply the symmetry argument of Lemma 1.5.12, we choose again symmetric sets

$$A = -B = \{y \colon |\phi_0 + L| \leqslant \delta, \phi_\perp \in D_\perp\}, \tag{3.6.5}$$

with $D_\perp$ containing this time a ball of radius $\sqrt{\log(\varepsilon^{-1})}$ in $H^s(\Lambda)$ for some $s < 0$. Then we have the following result [BDGW17].

Proposition 3.6.3. *The partition function and capacity satisfy*

$$\frac{\mathscr{Z}_N}{2\operatorname{cap}(A,B)} = \frac{2\pi}{|\mu_0|} \frac{e^{-3\theta/2}}{\sqrt{|\mathcal{D}_N(3;-1))|}} e^{V_0(0)/\varepsilon} [1 + \mathcal{O}(\varepsilon)],$$

where $\nu_0 = 2$, $\mu_0 = -1$, and the error term is uniform in the cut-off N.

SKETCH OF PROOF: We write $\mathscr{K}_N^* = \mathscr{K}_N \backslash \{0\}$, and define $\mathcal{D}_N^\perp(c;b)$ as in (3.6.4), with Δ_N replaced by $\Delta_{\perp,N}$. Interpreting as before the integral over $\phi_\perp$ as an expectation under the Gaussian measure $\mu_\perp(-1+m^2(\phi_0))$, which is well-defined since the smallest eigenvalue of $\Delta_{\perp,N}$ is $(2\pi/L)^2$, we find for the partition function

$$\mathscr{Z}_N = \int_{E_N} e^{-V(\phi)/\varepsilon} \, d\phi$$

$$= \int_{-\infty}^{\infty} e^{-V_0(\phi_0)/\varepsilon} \prod_{k \in \mathscr{K}_N^*} \sqrt{\frac{2\pi\varepsilon}{\lambda_k - 1 + m^2(\phi_0)}} \, e^{(1-m^2(\phi_0))L^2 C_N/2} \, \mathbb{E}^{\mu_\perp(-1+m^2(\phi_0))}[e^{-R_\varepsilon}] \, d\phi_0$$

$$= \prod_{k \in \mathscr{K}_N^*} \sqrt{\frac{2\pi\varepsilon}{\lambda_k - 1}} \int_{-\infty}^{\infty} e^{-V_0(\phi_0)/\varepsilon} \sqrt{\frac{e^{L^2 C_N} e^{-m^2(\phi_0)(|\mu_0|^{-1}+\theta)}}{\mathcal{D}_N^\perp(m^2(\phi_0);-1)}} \, \mathbb{E}^{\mu_\perp(-1+m^2(\phi_0))}[e^{-R_\varepsilon}] \, d\phi_0 \, .$$

Proceeding as in Proposition 2.7.5, one can show that $\mathbb{E}^{\mu_\perp(-1+m^2(\phi_0))}[e^{-R_\varepsilon}] = 1 + \mathcal{O}(\varepsilon e^{\phi_0^4})$. As in that proposition, the argument actually requires Gaussian changes of mass, using here Lemma 3.6.1, to obtain the required estimate, cf. [BDGW17, Proposition 5.7]. When performing the integral over ϕ_0, we note that $-V_0(\phi_0)$ is maximal at $\phi_0 = \pm L^2$, where $m^2(\phi_0) = 3$, so that by Laplace asymptotics we obtain

$$\mathscr{Z}_N = 2\sqrt{\frac{2\pi\varepsilon \, e^{L^2 C_N - 3(|\mu_0|^{-1}+\theta)}}{\nu_0 \mathcal{D}_N^\perp(3;-1)}} \prod_{k \in \mathscr{K}_N^*} \sqrt{\frac{2\pi\varepsilon}{\lambda_k - 1}} [1 + \mathcal{O}(\varepsilon)] \, .$$

For the capacity, using the Dirichlet and Thomson principles, we find in a similar way

$$\operatorname{cap}(A,B) = \frac{\varepsilon}{c_0^2} \prod_{k \in \mathscr{K}_N^*} \sqrt{\frac{2\pi\varepsilon}{\lambda_k - 1}} \int_{-a}^{a} e^{V_0(\phi_0)/\varepsilon} \sqrt{\frac{e^{L^2 C_N} e^{-m^2(\phi_0)(|\mu_0|^{-1}+\theta)}}{\mathcal{D}_N^\perp(m^2(\phi_0);-1)}} \, \mathbb{E}^{\mu_\perp(-1+m^2(\phi_0))}[e^{-R_\varepsilon}] \, d\phi_0 \, ,$$

where c_0 is as in (2.7.6). This time, the integral over ϕ_0 is dominated by ϕ_0 near 0, where $m^2 = 0$, and the result is

$$\operatorname{cap}(A,B) = \frac{\varepsilon}{c_0} \sqrt{\frac{e^{L^2 C_N}}{\mathcal{D}_N^\perp(0;-1)}} \prod_{k \in \mathscr{K}_N^*} \sqrt{\frac{2\pi\varepsilon}{\lambda_k - 1}} [1 + \mathcal{O}(\varepsilon)] \, .$$

The result follows, using the facts that $c_0 = \sqrt{2\pi\varepsilon |\mu_0|^{-1}} \, e^{V_0(0)/\varepsilon} [1 + \mathcal{O}(\varepsilon)]$, that $\mathcal{D}_N^\perp(0;-1) = 1$, and that $\nu_0 |\mu_0|^{-1} \mathcal{D}_N^\perp(3;-1) = e^{-3|\mu_0|^{-1}} |\mathcal{D}_N(3;-1)|$. $\qquad \square$

To turn Proposition 3.6.3 into a limiting result, one needs to perform some post-processing, as discussed in the proof of Theorem 2.7.7. Namely, one has to show that the expected transition times of the spectral Galerkin approximation converge to the one of the limiting system, and derive a coupling result to get rid of the equilibrium distribution on ∂A. This has been achieved in [TW18a], with the following result.

Theorem 3.6.4 (Eyring-Kramers law for the two-dimensional Allen–Cahn SPDE). *Assume that $L < 2\pi$ and A and B are as in (3.6.5). Then, for the choice (3.6.3) of regularisation C_N, one has*

$$\lim_{N \to \infty} \mathbb{E}^{\phi_-^*}[\tau_B] = \frac{2\pi}{|\mu_0|} \frac{e^{-3\theta/2} e^{L^2/(4\varepsilon)}}{\sqrt{\left|\det_2[1\!\!1 + 3(-\Delta - 1)^{-1}]\right|}} [1 + \mathcal{O}(\varepsilon)] \,,$$

where the determinant is to be interpreted as the Carleman–Fredholm determinant (3.6.4).

3.7   Bibliographical notes

Besides the system of coupled diffusions, the Allen–Cahn SPDE also describes particular scaling limits of Ising–Kac models, in which each spin interacts with a large number of other spins [MW17a].

Background on Wick calculus can be found in [Nua06, Chapter 1]. The graphical notation for Wick powers has been introduced in [Hai14], and we adopted the version used in [CW17] for their regularisation.

Construction of solutions to stochastic quantisation equations was investigated in [AR91], using the theory of Dirichlet forms. Existence and uniqueness of strong solutions to stochastic quantisation equations on the two-dimensional torus, as well as the absence of explosion, was first proved in [DPD03]. The result was extended to the whole plane in [MW17b].

Connections between the Gibbs measure and the stochastic quantisation equation were investigated in [JLM85]. The idea to use the Boué–Dupuis formula to bound the partition function was introduced in [BG18]. The fact that solutions to stochastic quantisation equations satisfy the Markov property and are reversible with respect to the Gibbs measure was proved in [RZZ17b] using Dirichlet forms, while uniqueness of the Gibbs measure and convergence to it were obtained in [RZZ17a]. The strong Feller property, as well as the Markov property and exponential mixing were proved in [TW18b] using a dissipative bound, while the strong Feller property was also proved (for more general equations) in [HM18], using the theory of regularity structures.

The large-deviation results are from [HW15]. The uniform-in-N Eyring–Kramers asymptotics on expected transition times were obtained in [BDGW17], and extended to a full proof of the Eyring–Kramers formula in [TW18a].

Chapter 4

Allen–Cahn SPDE in three space dimensions

We finally turn to the analysis of the three-dimensional Allen–Cahn equation, whose non-renormalised version is formally written as

$$\partial_t \phi(t,x) = \Delta \phi(t,x) + \phi(t,x) - \phi(t,x)^3 + \sqrt{2\varepsilon}\xi(t,x)\,, \tag{4.0.1}$$

where x now belongs to the torus $\Lambda = (\mathbb{R}/(L\mathbb{Z}))^3$. The deterministic PDE obtained for $\varepsilon = 0$ is again well-defined. However, we will see that the stochastic PDE remains ill-posed even when the nonlinear term $\phi(t,x)^3$ is replaced by its Wick power $:\phi(t,x)^3:$. This is a consequence of space-time white noise being even more singular in three than in two space dimensions.

There exist by now several methods to make sense of the singular equation (4.0.1) after proper renormalisation, namely *regularity structures* [Hai14], *paracontrolled calculus* [GIP15, CC18] and a Wilsonian renormalisation group approach [Kup16]. In this chapter, we will describe the approach based on regularity structures. First, however, we need to understand the reason for the ill-posedness in more detail.

4.1 Failure of the previous approaches

The definition of space-time white noise ξ in one time and three space dimensions proceeds as before, with some straightforward adjustments. In particular, ξ has now the scaling behaviour

$$\xi_{\tau,\lambda} \overset{\text{law}}{=} \frac{1}{\sqrt{\tau\lambda^3}}\xi\,.$$

With the parabolic norm

$$\|(t,x) - (s,y)\|_{\mathfrak{s}} = |t - s|^{1/2} + \sum_{i=1}^{3}|x_i - y_i|\,,$$

and the scaling operator

$$(\mathscr{S}_{t,x}^{\lambda}\varphi)(s,y) := \frac{1}{\lambda^5}\varphi\Big(\frac{s-t}{\lambda^2}, \frac{y-x}{\lambda}\Big)\,,$$

Theorem 2.2.8 becomes

$$\xi \in \mathscr{C}_{\mathfrak{s}}^{-5/2-\kappa} \qquad \text{for any } \kappa > 0\,. \tag{4.1.1}$$

The Schauder estimate stated in Theorem 2.3.9 remains true, showing that the stochastic convolution $P * \xi$ now belongs to $\mathscr{C}_{\mathfrak{s}}^{\alpha}$ for $\alpha < -\frac{1}{2}$ only. It is thus much more singular than in dimension two, where it was very close to being a function.

In analogy with what we did in dimension two, it seems natural to consider the regularised equation (for $2\varepsilon = 1$)

$$
\begin{aligned}
\partial_t \phi_\delta(t,x) &= \Delta \phi_\delta(t,x) + \phi_\delta(t,x) - \,:\!\phi_\delta(t,x)^3\!:_{C_\delta} + \xi^\delta(t,x) \\
&= \Delta \phi_\delta(t,x) + \phi_\delta(t,x) - \big(\phi_\delta(t,x)^3 - 3C_\delta \phi_\delta(t,x)\big) + \xi^\delta(t,x)\,,
\end{aligned}
\tag{4.1.2}
$$

where ξ^δ denotes a mollified noise defined by

$$
\xi^\delta(t,x) = (\rho_\delta * \xi)(t,x)\,, \qquad \rho_\delta(t,x) = \frac{1}{\delta^5}\rho\left(\frac{t}{\delta^2}, \frac{x}{\delta}\right)
$$

for a compactly supported smooth function $\rho : \mathbb{R} \times \Lambda \to \mathbb{R}$ of integral 1. The renormalisation constant C_δ defining the Wick power in (4.1.2) is given as in two dimensions by

$$
C_\delta := \int_{-\infty}^{\infty} P_\delta(t,x)^2 \,\mathrm{d}x\,\mathrm{d}t\,,
$$

where $P_\delta = P * \rho_\delta$, except that P is now the three-dimensional heat kernel

$$
P(t,x) := \sum_{k \in \mathbb{Z}^3} P_{\mathbb{R}^3}(t, x - kL)\,, \qquad P_{\mathbb{R}^3}(t,x) := \frac{1}{(4\pi t)^{3/2}}\,\mathrm{e}^{-\|x\|^2/(4t)}\,\mathbb{1}_{\{t>0\}}\,.
$$

Exercise 4.1.1. Use the fact that the Green function of the three-dimensional Laplacian is the periodicised version of

$$
G_{\mathbb{R}^3}(x) = -\frac{1}{4\pi\|x\|}
$$

to show that C_δ diverges like δ^{-1} as $\delta \to 0$. **Hint:** Recall Proposition 3.2.16. ♣

Making the ansatz $\phi_\delta = \mathord{\uparrow}_\delta + \psi_\delta$, where $\mathord{\uparrow}_\delta = P * \xi^\delta$, and taking the limit $\delta \to 0$, we obtain as in Section 3.3 the fixed-point equation

$$
\psi = P\psi_0 + P * \big[\,\mathord{\uparrow} + \psi - \psi^3 - 3\,\mathord{\uparrow}\psi^2 - 3\,\mathord{\vee}\psi - \mathord{\vee\!\!\!\vee}\,\big]\,.
\tag{4.1.3}
$$

The term with the worst regularity on the right-hand side should be

$$
P * \mathord{\vee\!\!\!\vee} =: \mathord{\Psi}\,.
$$

Here $\mathord{\vee\!\!\!\vee}$ belongs to $\mathscr{C}_{\mathfrak{s}}^{\alpha}$ for any $\alpha < -\frac{3}{2}$, so that the Schauder estimate shows that $\mathord{\Psi}$ has Hölder regularity almost $\frac{1}{2}$, and the same is expected to hold for ψ. This is a problem, however, for the definition of the product $\mathord{\vee}\psi$. Indeed, since $\mathord{\vee}$ has regularity almost -1, Theorem 3.3.1 implies that this product is ill-defined!

One may think of circumventing this problem by making a second change of variables

$$
\psi = -\,\mathord{\Psi} + \eta\,,
$$

in the hope that η will be more regular than ψ. This turns the fixed-point equation into

$$
\eta = P\eta_0 + P * \big[\,\mathord{\uparrow} - \mathord{\Psi} + \eta - \eta^3 + 3\eta^2(\mathord{\Psi} - \mathord{\uparrow}) - 3\eta(\mathord{\vee} - 2\,\mathord{\uparrow}\mathord{\Psi} + \mathord{\Psi}^2) + \mathord{\Psi}^3 - 3\,\mathord{\uparrow}\mathord{\Psi}^2 + 3\,\mathord{\vee}\mathord{\Psi}\,\big]\,.
\tag{4.1.4}
$$

The problem now is that the product $\mathord{\vee}\,\mathord{\Psi}$ is not well-defined, since the regularity of the two factors is slightly less than -1 and $\frac{1}{2}$ respectively. This is a first indication that a second renormalisation constant will be needed. It further suggests that one has to rethink the way solutions are represented as linear combinations of iterated stochastic integrals.

4.2 Perturbation theory for the Gibbs measure

One way to understand the need for a further renormalisation constant is to consider, as in Section 3.6, the Wick-renormalised potential

$$V(\phi) := \int_\Lambda \left[\frac{1}{2} \|\nabla\phi(x)\|^2 + \frac{1}{4} :\phi(x)^4: - \frac{1}{2} :\phi(x)^2: + \frac{1}{4} \right] dx, \qquad (4.2.1)$$

where we use a spectral Galerkin approximation of order N for ϕ, and Wick powers are with respect to εC_N. Here we choose

$$L^2 C_N = \mathrm{Tr}\left((-\Delta_{\perp,N} - 1)^{-1}\right) = \sum_{k \in \mathscr{K}_N^*} \frac{1}{\lambda_k - 1} = \mathcal{O}(N),$$

where, like before, $\lambda_k = (2\|k\|\pi/L)^2$ and $\mathscr{K}_N^* = \{k \in \mathbb{Z}^3 : 0 < |k| \leqslant N\}$ with $|k| = |k_1| + |k_2| + |k_3|$.

To lighten the notation, we assume from now on that $L = 1$. As we have seen, e.g., in the proof of Proposition 3.6.3, showing that the potential (4.2.1) is well-defined involves computing the expectation of e^{-R_ε} under the Gaussian measure $\mu_\perp = \mu_\perp(\phi_0)$ having covariance $(-\Delta_{\perp,N} - 1 + 3\phi_0^2)^{-1}$, where

$$R_\varepsilon(\phi_0, \phi_\perp) := \sqrt{\varepsilon}\phi_0 \int_\Lambda :\phi_\perp(x)^3: dx + \frac{\varepsilon}{4} \int_\Lambda :\phi_\perp(x)^4: dx.$$

We will represent this graphically as

$$R_\varepsilon(\phi_0, \phi_\perp) = \sqrt{\varepsilon}\phi_0 \; \rangle\!\!-\!\!\bullet\!\!-\!\! + \frac{\varepsilon}{4} \; \rangle\!\!\!\bullet\!\!\!\langle \; .$$

We can compute $\mathbb{E}^{\mu_\perp}[e^{-R_\varepsilon}]$ perturbatively in ε by using the *cumulant expansion*

$$\begin{aligned}
-\log \mathbb{E}^{\mu_\perp}[e^{-R_\varepsilon}] = \; & \mathbb{E}^{\mu_\perp}[R_\varepsilon] \\
& + \frac{1}{2!}\left(\mathbb{E}^{\mu_\perp}[R_\varepsilon]^2 - \mathbb{E}^{\mu_\perp}[R_\varepsilon^2]\right) \\
& + \frac{1}{3!}\left(2\mathbb{E}^{\mu_\perp}[R_\varepsilon]^3 - 3\mathbb{E}^{\mu_\perp}[R_\varepsilon]\mathbb{E}^{\mu_\perp}[R_\varepsilon^2] + \mathbb{E}^{\mu_\perp}[R_\varepsilon^3]\right) \\
& + \dots
\end{aligned} \qquad (4.2.2)$$

In order to compute these expectations, we recall (see (2.5.2)) that[1]

$$\mathbb{E}^{\mu_\perp}[\phi_\perp(x)\phi_\perp(y)] = G_N(x - y),$$

where G_N is the Green function of $\Delta_{\perp,N} - 1$. In a Fourier basis $\{e_k\}_{k \in \mathscr{K}_N^*}$, it is given by

$$G_N(x) := \sum_{k \in \mathscr{K}_N^*} \frac{1}{\lambda_k - 1} e_k(x)$$

(note that we recover the fact that $G_N(0) = C_N$).

Consider first the case $\phi_0 = 0$, when only the quartic term is present in R_ε. We know that the fourth Wick power has zero expectation, while according to Lemma 3.2.9, its variance is given by

$$\int_\Lambda \int_\Lambda \mathbb{E}^{\mu_\perp}[:\phi_\perp(x)^4::\phi_\perp(y)^4:]\,dx\,dy = 4!\int_\Lambda G_N(x)^4\,dx$$

---

[1] We are going to be slightly sloppy here, by ignoring the fact that the covariance of the Gaussian measure $\mu_\perp$ depends on ϕ_0, which affects G_N and C_N. However, this results only in negligible extra terms, essentially thanks to Lemma 3.6.1.

(we have used translation invariance of the Green function and the fact that $L = 1$). We denote this graphically (cf. (3.2.5)) by

$$\mathbb{E}^{\mu_\perp}\left[\times\!\!\!\times \quad \times\!\!\!\times \right] = 4! \; \bigcirc\!\!\!\!\!\ominus\!\!\!\!\!\bigcirc \; . \tag{4.2.3}$$

In a similar way, we find

$$\mathbb{E}^{\mu_\perp}\left[\times\!\!\!\times \quad \times\!\!\!\times \quad \times\!\!\!\times \right] = \binom{4}{2}^3 2^3 \; \bigtriangleup\!\!\!\!\!\bigcirc \; , \tag{4.2.4}$$

where the combinatorial factor counts the number of pairings of the legs, and

$$\bigtriangleup\!\!\!\!\!\bigcirc := \int_\Lambda \int_\Lambda G_N(x)^2 G_N(y)^2 G_N(x-y)^2 \, dx \, dy \, . \tag{4.2.5}$$

The graphs occurring in (4.2.3) and (4.2.5) are special cases of Feynman diagrams, called *vacuum diagrams*.

Inserting (4.2.3) and (4.2.4) in the cumulant expansion (4.2.2), we obtain the asymptotic expansion

$$-\log \mathbb{E}^{\mu_\perp(0)}\left[e^{-R_\varepsilon(0,\cdot)}\right] = -\frac{4!}{2!4^2}\varepsilon^2 \; \bigcirc\!\!\!\!\!\ominus\!\!\!\!\!\bigcirc \; + \frac{2^3}{3!4^3}\binom{4}{2}^3 \varepsilon^3 \; \bigtriangleup\!\!\!\!\!\bigcirc \; + \mathcal{O}(\varepsilon^4) \, , \tag{4.2.6}$$

where the two vacuum diagrams diverge, respectively, like N and like $\log N$. This suggests that for $\phi_0 = 0$, the potential (4.2.1) can be renormalised by subtracting two terms of respective order $\varepsilon^3 N$ and $\varepsilon^4 \log(N)$, called *energy renormalisation* terms. Of course, at this point we can tell nothing about the N-dependence of the higher-order terms in ε.

If $\phi_0 \neq 0$, the variance of the third Wick power in R_ε will add a new divergent term of order $\varepsilon \phi_0^3 \log N$ to the cumulant expansion. This is a symptom of the fact that the Gibbs measures obtained for different values of ϕ_0 are not absolutely continuous with respect to one another. We cannot merely subtract this ϕ_0-dependent term from the potential, as this would not result

from a decomposition into mean and oscillating part. However, we can modify the potential (4.2.1) by adding a renormalisation term to the coefficient of the second Wick power, that is,

$$V(\phi) = \int_\Lambda \left[\frac{1}{2}\|\nabla\phi(x)\|^2 + \frac{1}{4}{:}\phi(x)^4{:} - \frac{1}{2}\left(1 - \varepsilon^2 C_N^{(2)}\right){:}\phi(x)^2{:} + \frac{1}{4} \right]\mathrm{d}x\,.$$

This procedure, called *mass renormalisation*, results in the expression

$$R_\varepsilon(\phi_0, \phi_\perp) = \frac{1}{2}\varepsilon C_N^{(2)}\phi_0^2 + \frac{1}{2}\varepsilon^2 C_N^{(2)}\; \vcenter{\hbox{⬤—●}} \; + \sqrt{\varepsilon}\phi_0 \; \vcenter{\hbox{⧼—}} \; + \frac{\varepsilon}{4} \; \vcenter{\hbox{✕}} \;.$$

In particular, we now find for the first two moments

$$\mathbb{E}^{\mu_\perp}[R_\varepsilon] = \frac{1}{2}\varepsilon C_N^{(2)}\phi_0^2\,,$$

$$\mathbb{E}^{\mu_\perp}[R_\varepsilon^2] = \left(\frac{1}{2}\varepsilon C_N^{(2)}\phi_0^2\right)^2 + \varepsilon\phi_0^2 3! \; \vcenter{\hbox{⬭}} \; + \frac{\varepsilon^2}{4^2}4! \; \vcenter{\hbox{⬯}} \; + \mathcal{O}(\varepsilon^4)\,,$$

where we have used the fact that terms with an odd number of legs have zero expectation (cf. Lemma 3.2.9). Substituting in the cumulant expansion (4.2.2), we obtain

$$-\log\mathbb{E}^{\mu_\perp}[\mathrm{e}^{-R_\varepsilon}] = \frac{1}{2}\varepsilon C_N^{(2)}\phi_0^2 + \frac{1}{2!}\left(-\varepsilon\phi_0^2 3! \; \vcenter{\hbox{⬭}} \; - \frac{\varepsilon^2}{4^2}4! \; \vcenter{\hbox{⬯}} \; \right) + \mathcal{O}(\varepsilon^3)\,.$$

We thus see that choosing

$$C_N^{(2)} = 3! \; \vcenter{\hbox{⬭}}$$

allows to kill the divergent term of order ε. Pushing the expansion one step further, one can show that the divergent part of the term of order ε^3 is the same as in (4.2.6).

What this purely formal computation suggests, is that by choosing the renormalised potential of the form

$$V(\phi) = \int_\Lambda \left[\frac{1}{2}\|\nabla\phi(x)\|^2 + \frac{1}{4}{:}\phi(x)^4{:}_{\varepsilon C_N^{(1)}} - \frac{1}{2}\left(1 - \varepsilon^2 C_N^{(2)}\right){:}\phi(x)^2{:}_{\varepsilon C_N^{(1)}} + \frac{1}{4} + \varepsilon^3 C_N^{(3)} - \varepsilon^4 C_N^{(4)} \right]\mathrm{d}x\,, \quad (4.2.7)$$

the limit $N \to \infty$ may be well-defined, provided we choose the counterterms

$$C_N^{(1)} := \vcenter{\hbox{◯}} = \mathcal{O}(N)\,,$$

$$C_N^{(2)} := 3! \; \vcenter{\hbox{⬭}} = \mathcal{O}(\log N)\,,$$

$$C_N^{(3)} := \frac{4!}{2!4^2} \; \vcenter{\hbox{⬯}} = \mathcal{O}(N)\,,$$

$$C_N^{(4)} := \frac{2^3}{3!4^3}\binom{4}{2}^3 \; \vcenter{\hbox{△}} = \mathcal{O}(\log N)\,. \quad (4.2.8)$$

Of course, this perturbative computation up to order ε^3 does not imply that no other counterterms of higher order in ε are needed. The fact that this is not the case is a highly nontrivial fact, which makes the static Φ_3^4 model one of the important solved cases of (Euclidean) Quantum Field Theory. See Section 4.8 for some historical notes. A recent proof, which is relatively compact, of the well-posedness of the Gibbs measure is provided in [BG18].

4.3   Regularity structures: introduction

4.3.1   Hölder spaces revisited

One of the key ideas of the theory of regularity structures is to replace expansions as appearing in the fixed-point equation (4.1.3) by more general expansions, which separate the roles of the coefficients, such as $\psi(t,x)$, and of the singular distributions, such as $\vee$. These objects can be viewed as a kind of generalised Taylor expansion.

Consider for simplicity the case of a function $f_0 \in \mathscr{C}^{2+\alpha}(I)$, where $I \subset \mathbb{R}$ is an interval, and $\alpha \in (0,1)$. One way of defining $\mathscr{C}^{2+\alpha}(I)$ is to require that f_0 is twice continuously differentiable, with a second derivative in $\mathscr{C}^{\alpha}(I)$, cf. Exercise 2.3.8. An alternative definition, which is more in line with Definition 2.3.7, is to require the existence of a triple (f_0, f_1, f_2) of functions from I to $\mathbb{R}$ satisfying for all $x, y \in I$ the bounds on Taylor expansions

$$\left| f_0(y) - f_0(x) - (y-x)f_1(x) - \tfrac{1}{2}(y-x)^2 f_2(x) \right| \leqslant C|x-y|^{2+\alpha} \,,$$
$$\left| f_1(y) - f_1(x) - (y-x)f_2(x) \right| \leqslant C|x-y|^{1+\alpha} \,,$$
$$\left| f_2(y) - f_2(x) \right| \leqslant C|x-y|^{\alpha} \,. \tag{4.3.1}$$

Note that this definition, while it seems redundant, completely avoids the use of derivatives, though of course we necessarily have $f_1(x) = f_0'(x)$ and $f_2(x) = f_0''(x)$.

In the theory of regularity structures, one uses the abstract notation

$$f(x) = f_0(x)\mathbf{1} + f_1(x)\mathbf{X} + \tfrac{1}{2}f_2(x)\mathbf{X}^2 \,, \tag{4.3.2}$$

where the symbols $\mathbf{1}$, $\mathbf{X}$ and $\mathbf{X}^2$ are basis vectors of a vector space T of dimension 3, called a *model space*, and $f : I \to T$ is considered as a map from I to T. We will denote the collection of basis vectors by $\mathscr{F} = \{\mathbf{1}, \mathbf{X}, \mathbf{X}^2\}$.

In order to be able to evaluate the abstract object (4.3.2) at a particular point, we introduce the notion of *model*. This is given by a collection of maps $\{\Pi_x \tau : x \in I, \tau \in \mathscr{F}\}$ defined for all $z \in I$ by

$$(\Pi_x \mathbf{1})(z) := 1 \,, \qquad (\Pi_x \mathbf{X})(z) := z - x \,, \qquad (\Pi_x \mathbf{X}^2)(z) := (z-x)^2 \,. \tag{4.3.3}$$

In this way, we have

$$(\Pi_x f(x))(z) := f_0(x)(\Pi_x \mathbf{1})(z) + f_1(x)(\Pi_x \mathbf{X})(z) + \tfrac{1}{2}f_2(x)(\Pi_x \mathbf{X}^2)(z)$$
$$= f_0(x) + f_1(x)(z-x) + \tfrac{1}{2}f_2(x)(z-x)^2 \,,$$

which is the second-order Taylor expansion of f at x. In particular, we can recover f_0 thanks to the *diagonal identity*

$$(\Pi_x f(x))(x) = f_0(x) \,. \tag{4.3.4}$$

However, f contains more information than just the value of f_0 at any point x, since the higher-order terms of the Taylor expansion provide some description of how f behaves in the vicinity of x.

Remark 4.3.1. An aspect of the theory that may be confusing at first is that it is possible to define the quantity $(\Pi_x f(y))(z)$ which depends on three different arguments. This is not really of interest in the case of polynomial models considered here, but will play a role later on. $\diamond$

It will also be important to encode how the model Π_x depends on the base point x. For this purpose, note that if we define a collection $\{\Gamma_{xy} : x, y \in I\}$ of linear maps from T to T by

$$\Gamma_{xy}\mathbf{1} := \mathbf{1} \,,$$
$$\Gamma_{xy}\mathbf{X} := \mathbf{X} + (x-y)\mathbf{1} \,,$$
$$\Gamma_{xy}\mathbf{X}^2 := \left[\mathbf{X} + (x-y)\mathbf{1}\right]^2 := \mathbf{X}^2 + 2(x-y)\mathbf{X} + (x-y)^2\mathbf{1} \,, \tag{4.3.5}$$

then we have the relations

$$\Pi_y \tau = \Pi_x \Gamma_{xy} \tau \qquad\qquad \forall \tau \in \mathscr{F}, \forall x, y \in I,$$
$$\Gamma_{xy} \Gamma_{yz} \tau = \Gamma_{xz} \tau \qquad\qquad \forall \tau \in \mathscr{F}, \forall x, y, z \in I. \qquad (4.3.6)$$

Exercise 4.3.2. Check that the relations (4.3.6) hold. What is the natural generalisation of (4.3.3) and (4.3.5) to general monomials $\mathbf{X}^k$, $k \in \mathbb{N}_0$? Check the analogue of (4.3.6) for these general values of $k \in \mathbb{N}_0$ as well. ♣

The interest of this framework is that by combining (4.3.2) and (4.3.5), we obtain

$$f(y) - \Gamma_{yx} f(x) = [f_0(y) - f_0(x) - (y-x)f_1 - x) - \tfrac{1}{2}(y-x)^2 f_2(x)]\mathbf{1}$$
$$+ [f_1(y) - f_1(x) - (y-x)f_2(x)]\mathbf{X}$$
$$+ \tfrac{1}{2}[f_2(y) - f_2(x)]\mathbf{X}^2.$$

In view of (4.3.1), we see that if $\mathcal{Q}_k$ denotes the projection on the subspace of T spanned by $\mathbf{X}^k$, then we have

$$f \in \mathscr{C}^{2+\alpha} \qquad \Leftrightarrow \qquad \big| \mathcal{Q}_k[f(y) - \Gamma_{yx} f(x)] \big| \lesssim |x-y|^{2+\alpha-k} \qquad \forall k \in \{0, 1, 2\}. \qquad (4.3.7)$$

A central idea in the theory of regularity structures is that this characterisation of regularity can be extended to situations in which the monomials $\mathbf{X}^k$ are complemented by more singular objects, such as the stochastic convolution ⇂ and its powers.

4.3.2 Overview of the approach

Our primary aim will be to construct solutions for the regularised Allen–Cahn SPDE on $\mathbb{R}_+ \times \Lambda$ given by

$$\partial_t \phi_\delta(t, x) = \Delta \phi_\delta(t, x) + \phi_\delta(t, x) - \phi_\delta(t, x)^3 + [3C_\delta^{(1)} - 9C_\delta^{(2)}]\phi_\delta(t, x) + \xi^\delta(t, x), \qquad (4.3.8)$$

where the renormalisation constants $C_\delta^{(1)} = \mathcal{O}(\delta^{-1})$ and $C_\delta^{(2)} = \mathcal{O}(\log(\delta^{-1}))$ are analogues of the first two constants in (4.2.8), while $\xi^\delta = \rho_\delta * \xi$ is a mollified version of space-time white noise. Indeed, the right-hand side corresponds to the derivative of the renormalised potential (4.2.7), which does not see the constant terms.

The general idea of the method developed in [Hai14] can be described by the following commutative diagram:

$$
\begin{array}{ccc}
(\phi_0, Z^\delta) & \overset{\mathscr{S}}{\longmapsto} & \Phi \\[4pt]
\Psi \uparrow & & \downarrow \mathscr{R} \\[4pt]
(\phi_0, \xi^\delta) & \underset{\overline{\mathscr{S}}}{\longmapsto} & \phi_\delta
\end{array}
\qquad (4.3.9)
$$

The maps appearing in this diagram are as follows:

- The *classical solution map* $\overline{\mathscr{S}}$ takes an initial condition ϕ_0 and a realisation ξ^δ of mollified space-time white noise, and associates with them the solution $(\phi_\delta(t, x))_{t \geqslant 0}$ of the smooth PDE (4.3.8). This solution exists globally in time for any positive δ and almost every realisation of the noise, because the PDE is smooth with a confining nonlinearity.
- The lift Ψ maps ξ^δ to a so-called *canonical model* $Z^\delta = (\Pi^\delta, \Gamma^\delta)$, which extends the collections of maps introduced in (4.3.3) and (4.3.5) to the stochastic convolution, its powers, and similar data. We will explain the construction of Z^δ in Section 4.5.1.

- The solution map $\mathscr{S}$ associates with the initial data and the model an element Φ of an abstract space $\mathscr{D}^\gamma$ of so-called *modelled distributions* of regularity γ. This space can be thought of as an abstract analogue of the Hölder space $\mathscr{C}_{\mathfrak{s}}^\alpha$ on the level of coefficients of Taylor series. We will describe it in Section 4.5.2. The element $\Phi \in \mathscr{D}^\gamma$ is obtained as the fixed point of a contracting map on $\mathscr{D}^\gamma$.
- The *reconstruction operator* $\mathscr{R}$ takes a modelled distribution $\Phi \in \mathscr{D}^\gamma$ and maps it to a distribution in a Hölder space $\mathscr{C}_{\mathfrak{s}}^\alpha$. We will describe it in Section 4.5.3.

  The point of the whole procedure is that we have the relation

$$\overline{\mathscr{S}} = \mathscr{R} \circ \mathscr{S} \circ \Psi,$$

meaning that the fixed-point problems defining the classical solution ϕ_δ and the modelled distribution Φ are equivalent for any $\delta > 0$. In addition, both maps $\mathscr{S}$ and $\mathscr{R}$ are continuous in an appropriate topology. We explain this part of the theory in more detail in Section 4.5.4.

Renormalisation comes into play when one modifies the lift Ψ. This will be an extension of the transformation (3.2.15) from powers of stochastic convolutions to their Wick powers, which removes divergencies. One can then show that under appropriate conditions, the renormalised models $\widehat{Z}^\delta = \Psi_\delta \xi^\delta$ converge, as $\delta \to 0$, to some limiting model $\widehat{Z}$. The image of this limit under the map $\mathscr{R} \circ \mathscr{S}$ is then defined as the solution of the limiting stochastic PDE. We give more details on this procedure in Section 4.5.5.

4.4 Regularity structures: algebraic aspects

At the core of the theory lies the notion of a *regularity structure*, which is the abstract Banach space T spanned by the basis elements of generalised Taylor expansions. This space is equipped with two additional algebraic structures:

- the *structure group* $\mathscr{G}$, which allows to represent the link between expansions around different base points;
- a *renormalisation group*, which allows to encode the renormalisation procedure.

We point out that while the model space is by definition infinite-dimensional, in practice only a finite-dimensional subspace of T will matter. This is in contrast with Feynman diagram expansions, which are in principle infinite. Both the structure group and renormalisation group will be idempotent (a sufficiently high power of any group element is the identity), which implies that the algebraic structures remain reasonably simple.

4.4.1 The model space

The starting point of the construction is the following definition, cf. [Hai14, Definition 2.1].

Definition 4.4.1 (Regularity structure)**.** *A regularity structure is a triple $(A, T, \mathscr{G})$ consisting of*
1. *an index set $A \subset \mathbb{R}$, containing 0, which is bounded from below and locally finite;*
2. *a model space T, which is a graded vector space $T = \bigoplus_{\alpha \in A} T_\alpha$, where each T_α is a Banach space; the space T_0 is isomorphic to $\mathbb{R}$ and its unit is denoted $\mathbf{1}$;*
3. *a structure group $\mathscr{G}$ of linear operators acting on T, such that*

$$\Gamma \tau - \tau \in \bigoplus_{\beta < \alpha} T_\beta =: T_\alpha^- \tag{4.4.1}$$

holds for every $\Gamma \in \mathscr{G}$, every $\alpha \in A$ and every $\tau \in T_\alpha$; furthermore, $\Gamma \mathbf{1} = \mathbf{1}$ for every $\Gamma \in \mathscr{G}$.

We have already encountered an example of regularity structure in Section 4.3.1:

Example 4.4.2 (Polynomal regularity structure on $\mathbb{R}$). This regularity structure is defined as follows:

- The index set $A = \mathbb{N}_0$ is the set of all degrees of monomials on $\mathbb{R}$.
- For each $k \in \mathbb{N}_0$, T_k is the one-dimensional real vector space spanned by a basis vector denoted $\mathbf{X}^k$ if $k > 0$ and $\mathbf{1}$ if $k = 0$. This is indeed a Banach space for the norm $|\cdot|$.
- The structure group is the group $\mathscr{G} = \{\Gamma_h \colon h \in \mathbb{R}\}$ defined by

$$\Gamma_h \mathbf{X}^k := (\mathbf{X} - h)^k := \sum_{\ell=0}^{k} \binom{k}{\ell} (-h)^\ell \mathbf{X}^{k-\ell} \,. \tag{4.4.2}$$

Note that the relation (4.4.1) is indeed satisfied, and that $\mathscr{G}$ is isomorphic to the group of translations on $\mathbb{R}$ via $\Gamma_{h_1} \circ \Gamma_{h_2} = \Gamma_{h_1 + h_2}$.

The model space T is naturally equipped with the commutative product defined by $\mathbf{X}^k \mathbf{X}^\ell = \mathbf{X}^{k+\ell}$, with neutral element $\mathbf{1}$.   ◇

This construction can be easily extended to higher dimensions.

Example 4.4.3 (Polynomal regularity structure on $\mathbb{R}^{d+1}$). Fix a scaling $\mathfrak{s} = (\mathfrak{s}_0, \dots, \mathfrak{s}_d) \in \mathbb{N}^{d+1}$. Then this regularity structure is defined in the following way:

- The index set is again $A = \mathbb{N}_0$.
- The model space is now the direct sum of T_ℓ, $\ell \in \mathbb{N}_0$, where each T_ℓ is spanned by monomials $\mathbf{X}^k = \mathbf{X}_0^{k_0} \dots \mathbf{X}_d^{k_d}$ of scaled degree $|k|_\mathfrak{s} = \mathfrak{s}_0 k_0 + \dots + \mathfrak{s}_d k_d = \ell$.
- The structure group is again defined by (4.4.2), where the binomial coefficients are defined via $k! = k_0! \dots k_d!$.

Again, T is naturally equipped with a commutative product, and $\mathscr{G}$ is now isomorphic to the group of translations on $\mathbb{R}^{d+1}$.   ◇

The *degree* of a monomial $\mathbf{X}^k$ is by definition the quantity $|\mathbf{X}^k|_\mathfrak{s} = |k|_\mathfrak{s}$, which simply tells us to which T_ℓ the monomial belongs.

In the case of the three-dimensional Allen–Cahn equation, we take $d = 3$ and the parabolic scaling $\mathfrak{s} = (2, 1, 1, 1)$. The polynomial regularity structure now has to be enriched with additional elements representing the noise, the stochastic convolution, its powers and so on. We start by introducing a symbol Ξ representing space-time white noise. In view of (4.1.1), we declare its degree to be given by $|\Xi|_\mathfrak{s} := \alpha_0$, where $\alpha_0 := -\frac{5}{2} - \kappa$ for a fixed $\kappa > 0$ that we can take as small as we like.

The next step is to introduce an operator $\mathcal{I}$ creating new symbols $\mathcal{I}(\tau)$ from existing ones, which represent their convolution with the heat kernel. In order to avoid redundancies, we declare that $\mathcal{I}(\mathbf{X}^k) = 0$ for elements of the polynomial structure. All these symbols can be multiplied, in a commutative way, yielding still more symbols. Their degree is defined by the two rules

$$|\mathcal{I}(\tau)|_\mathfrak{s} := |\tau|_\mathfrak{s} + 2 \,, \qquad |\tau_1 \tau_2|_\mathfrak{s} := |\tau_1|_\mathfrak{s} + |\tau_2|_\mathfrak{s} \,.$$

It will also be convenient to use graphical notations such as $\vee := \mathcal{I}(\Xi)^2$ and $\pmb{\Psi} = \vee \, \Psi$. Unlike in (4.1.4), this last object is just a new symbol instead of an undefined product of distributions.

At this point, we realise that there may be a problem with the requirement that the index set A be bounded below, since multiplying symbols of negative degree allows to generate terms whose degree is arbitrarily negative. The solution is to set $T := \mathrm{span}(\mathscr{F}_\mathrm{AC})$, where $\mathscr{F}_\mathrm{AC}$ is the restriction of the set $\mathscr{F}$ of all symbols to the set of those which can be generated by the map

$$\tau \mapsto \mathcal{I}(\Xi + \tau^3) + \sum_k \mathbf{X}^k \,,$$

τ	Symbol	$|\tau|_{\mathfrak{s}}$
Ξ	Ξ	$-\frac{5}{2} - \kappa$
$\mathcal{I}(\Xi)^3$		$-\frac{3}{2} - 3\kappa$
$\mathcal{I}(\Xi)^2$		$-1 - 2\kappa$
$\mathcal{I}(\mathcal{I}(\Xi)^3)\mathcal{I}(\Xi)^2$		$-\frac{1}{2} - 5\kappa$
$\mathcal{I}(\Xi)$		$-\frac{1}{2} - \kappa$
$\mathcal{I}(\mathcal{I}(\Xi)^3)\mathcal{I}(\Xi)$		$0 - 4\kappa$
$\mathcal{I}(\mathcal{I}(\Xi)^2)\mathcal{I}(\Xi)^2$		$0 - 4\kappa$
$\mathcal{I}(\Xi)^2 \mathbf{X}_i$	$\mathbf{X}_i$	$0 - 2\kappa$
$\mathbf{1}$	$\mathbf{1}$	0
$\mathcal{I}(\mathcal{I}(\Xi)^3)$		$\frac{1}{2} - 3\kappa$
$\mathcal{I}(\mathcal{I}(\Xi)^2)\mathcal{I}(\Xi)$		$\frac{1}{2} - 3\kappa$
$\mathcal{I}(\mathcal{I}(\Xi))\mathcal{I}(\Xi)^2$		$\frac{1}{2} - 3\kappa$
$\mathcal{I}(\mathcal{I}(\Xi)^2)$		$1 - 2\kappa$
$\mathcal{I}(\mathcal{I}(\Xi))\mathcal{I}(\Xi)$		$1 - 2\kappa$
$\mathbf{X}_i$	$\mathbf{X}_i$	1
$\mathcal{I}(\mathcal{I}(\Xi))$		$\frac{3}{2} - \kappa$

Table 4.1 – Elements of $\mathscr{F}_{\mathrm{AC}}$ of degree up to $\frac{3}{2}$. Here $\mathbf{X}_i$ denotes any element of the form $\mathbf{X}^k$ where $k = e_i$ is a canonical basis vector.

which is an abstract representation of the fixed-point map (3.3.2).

When iterating this map, starting with the empty set, one can check that indeed all symbols in $\mathscr{F}_{\mathrm{AC}}$ have degree at least α_0. We have listed all symbols of degree up to $\frac{3}{2}$ in Table 4.1. The fact that the degree remains bounded below is a consequence of *local subcriticality*, also called *super-renormalisability* in Physics, cf. [Hai14, Assumption 8.3 and Lemma 8.10]. To define this property, consider the formal Allen–Cahn equation

$$\partial_t \phi = \Delta \phi + \phi - \phi^3 + \xi$$

on the d-dimensional torus, and perform a scaling $\tilde{\phi}(t, x) = \lambda^\alpha \phi(\lambda^\beta t, \lambda^\gamma x)$, ignoring boundary conditions. This yields the rescaled equation

$$\partial_t \tilde{\phi} = \lambda^{\beta - 2\gamma} \Delta \tilde{\phi} + \lambda^\beta \tilde{\phi} - \lambda^{\beta - 2\alpha} \tilde{\phi}^3 + \lambda^{\alpha + \beta} \xi_{\lambda^\beta, \lambda^\gamma} \, .$$

The scaling property of space-time white noise shows that $\tilde{\xi} = \lambda^{(\beta + \gamma d)/2} \xi_{\lambda^\beta, \lambda^\gamma}$ has the same law as ξ (cf. Proposition 2.2.5). Choosing $\gamma = 1$, $\beta = 2$ and $\alpha = d/2 - 1$ thus yields

$$\partial_t \tilde{\phi} = \Delta \tilde{\phi} + \lambda^2 \tilde{\phi} - \lambda^{4-d} \tilde{\phi}^3 + \tilde{\xi} \, .$$

The equation is called *locally subcritical* if the coefficient of the nonlinear term vanishes in the limit $\lambda \to 0$, which corresponds to zooming in on small scales. This is the case if and only if $d < 4$, which is why the Allen–Cahn equation can be renormalised in dimensions two and three, but not in higher dimension.

Exercise 4.4.4. Determine the degree of the symbols listed in Table 4.1 in the case of the two-dimensional Allen–Cahn equation, when Ξ has degree $\alpha_0 = -\frac{3}{2} - \kappa$. What happens in the four-dimensional case?   ♣

4.4.2 The structure group

We have already defined the structure group for the polynomial part $\overline{T} = \mathrm{span}\{\mathbf{X}^k : k \in \mathbb{N}_0^4\}$ of the model space as being given by all maps Γ_h acting as

$$\Gamma_h \mathbf{X}^k = (\mathbf{X} - h)^k, \qquad h \in \mathbb{R}^4.$$

Extending the structure group to the non-polynomial elements is a somewhat tricky point, since it has to reflect the way Taylor expansions change under translations of the base point. Since space-time white noise is translation-invariant, it seems reasonable to set

$$\Gamma \Xi = \Xi \qquad \forall \Gamma \in \mathscr{G}. \tag{4.4.3}$$

It turns out that the same holds for the element $\curlyvee$. This is related to the fact that the stochastic convolution has negative regularity, so that there is no need to subtract any polynomial terms when testing it against a scaled test function. We have encountered this situation in (2.3.13) in the proof of Schauder's Theorem 2.3.9. Therefore, we set

$$\Gamma \curlyvee = \curlyvee, \qquad \Gamma \vee = \vee, \qquad \Gamma \curlyvee\!\!\!\vee = \curlyvee\!\!\!\vee \qquad \forall \Gamma \in \mathscr{G}. \tag{4.4.4}$$

The situation is different, however, for elements such as $\curlyvee\!\!\!\curlyvee$. Indeed, $\curlyvee\!\!\!\curlyvee$ represents a function with positive regularity, from which one has to subtract a term before testing it, as in (2.3.16), a procedure known as *recentering*. Therefore, the structure group cannot act trivially on such elements.

To extend the structure group to all elements of T, it is helpful to take a look at the structure group of $\overline{T}$ in a more algebraic way. The dual $\overline{T}^*$ can be identified with the space $\{\mathrm{D}^k : k \in \mathbb{N}_0^4\}$ of differential operators with constant coefficients via

$$\langle \mathrm{D}^k, \mathbf{X}^\ell \rangle := \left.\frac{\partial \mathbf{X}^\ell}{\partial \mathbf{X}^k}\right|_{\mathbf{X}=0} = k! \delta_{k\ell}. \tag{4.4.5}$$

We can define an action $(\mathrm{D}^\ell, \tau) \mapsto \Gamma_{\mathrm{D}^\ell} \tau$ of $\overline{T}^*$ onto $\overline{T}$ by

$$\langle \mathrm{D}^k, \Gamma_{\mathrm{D}^\ell} \mathbf{X}^m \rangle := \langle \mathrm{D}^k \mathrm{D}^\ell, \mathbf{X}^m \rangle = m! \delta_{k+\ell, m} \qquad \forall \mathrm{D}^k \in \overline{T}^*, \forall \mathbf{X}^m \in \overline{T}, \tag{4.4.6}$$

which by (4.4.5) is equivalent to

$$\Gamma_{\mathrm{D}^\ell} \mathbf{X}^m = \frac{m!}{(m-\ell)!} \mathbf{X}^{m-\ell},$$

with the convention that the right-hand side vanishes unless $\ell_i \leqslant m_i$ for each $i \in \{0, 1, 2, 3\}$, which we write for short $\ell \leqslant m$. Note that we have the semigroup property $\Gamma_{\mathrm{D}^\ell \mathrm{D}^m} = \Gamma_{\mathrm{D}^\ell} \Gamma_{\mathrm{D}^m}$, since for every $\mathrm{D}^k \in \overline{T}^*$ and $\mathbf{X}^n \in \overline{T}$,

$$\langle \mathrm{D}^k, \Gamma_{\mathrm{D}^\ell \mathrm{D}^m} \mathbf{X}^n \rangle = \langle \mathrm{D}^k \mathrm{D}^\ell \mathrm{D}^m, \mathbf{X}^n \rangle = \langle \mathrm{D}^k \mathrm{D}^\ell, \Gamma_{\mathrm{D}^m} \mathbf{X}^n \rangle = \langle \mathrm{D}^k, \Gamma_{\mathrm{D}^\ell} \Gamma_{\mathrm{D}^m} \mathbf{X}^n \rangle.$$

However, not all Γ_{D^k} are invertible, so that the set $\{\Gamma_g : g \in \overline{T}^*\}$ does not form a group. We thus define $\mathscr{G}$ as the set of $g \in \overline{T}^*$ which are *group-like*, meaning that they satisfy

$$\langle g, \mathbf{X}^\ell \mathbf{X}^m \rangle = \langle g, \mathbf{X}^\ell \rangle \langle g, \mathbf{X}^m \rangle \qquad \forall \mathbf{X}^\ell, \mathbf{X}^m \in \overline{T}. \tag{4.4.7}$$

One can show (cf. [Hai14, Section 4.3]) that $\mathscr{G}$ is exactly the Lie group generated by the Lie algebra of first-order differential operators. Indeed, if for $h \in \mathbb{N}_0^4$ we set

$$\langle h, \mathrm{D} \rangle := \sum_{i=0}^{3} h_i \mathrm{D}^{e_i} \in \overline{T}^*, \qquad g = \mathrm{e}^{-\langle h, \mathrm{D} \rangle} := \sum_{k \in \mathbb{N}_0^4} \frac{(-\langle h, \mathrm{D} \rangle)^k}{k!}, \tag{4.4.8}$$

then we have

$$\Gamma_g(\mathbf{X}^\ell) = \sum_{k \in \mathbb{N}_0^4} \frac{(-h)^k}{k!} \Gamma_{\mathrm{D}^k}(\mathbf{X}^\ell) = \sum_{k \leqslant \ell} \frac{(-h)^k}{k!} \frac{\ell!}{(\ell-k)!} \mathbf{X}^{\ell-k} = (\mathbf{X}-h)^\ell\,,$$

showing that the group action Γ of $\mathscr{G}$ onto $\overline{T}$ is indeed equivalent to the action of the structure group of the polynomial regularity structure. This is related to the fact that first-order derivatives generate the group of translations.

> **Exercise 4.4.5.** Prove that the element g defined by (4.4.8) is group-like in the sense of (4.4.7) in the case where the indices k belong to $\mathbb{N}_0$ instead of $\mathbb{N}_0^4$. ♣

A useful way of encoding the Leibniz rule is to define a bilinear operator $\Delta^+ : \overline{T} \to \overline{T} \otimes \overline{T}$ by

$$\Delta^+(\mathbf{1}) := \mathbf{1} \otimes \mathbf{1}\,, \qquad \Delta^+(\mathbf{X}_i) := \mathbf{X}_i \otimes \mathbf{1} + \mathbf{1} \otimes \mathbf{X}_i\,, \tag{4.4.9}$$

which is then extended to all of $\overline{T}$ by requiring that

$$\Delta^+(\mathbf{X}^k \mathbf{X}^\ell) := \Delta^+(\mathbf{X}^k)\Delta^+(\mathbf{X}^\ell) \qquad \forall \mathbf{X}^k, \mathbf{X}^\ell \in \overline{T}\,. \tag{4.4.10}$$

The operator Δ^+ is called a *coproduct*. One can indeed check that Δ^+ is *coassociative*, meaning that

$$(\mathrm{Id} \otimes \Delta^+)\Delta^+ \mathbf{X}^k = (\Delta^+ \otimes \mathrm{Id})\Delta^+ \mathbf{X}^k \qquad \forall \mathbf{X}^k \in \overline{T}\,,$$

which provides $\overline{T}$ with a *Hopf algebra* structure[2]. Furthermore, we have the duality property

$$\langle \mathrm{D}^k \mathrm{D}^\ell, \mathbf{X}^m \rangle = \langle \mathrm{D}^k \otimes \mathrm{D}^\ell, \Delta^+ \mathbf{X}^m \rangle\,, \tag{4.4.11}$$

where by definition

$$\langle \mathrm{D}^k \otimes \mathrm{D}^\ell, Y^{(1)} \otimes Y^{(2)} \rangle := \langle \mathrm{D}^k, Y^{(1)} \rangle \langle \mathrm{D}^\ell, Y^{(2)} \rangle\,.$$

Note that in general, $\Delta^+ \mathbf{X}^m$ is a sum of several terms $Y_i^{(1)} \otimes Y_i^{(2)}$, but it is convenient to use *Sweedler's notation* suppressing the sum. The point is that by combining (4.4.11) with (4.4.6), one obtains the relation

$$\Gamma_{\mathrm{D}^\ell} \mathbf{X}^m = (\mathrm{Id} \otimes \mathrm{D}^\ell)\Delta^+ \mathbf{X}^m\,,$$

where by definition, $(\mathrm{Id} \otimes \mathrm{D}^\ell)Y^{(1)} \otimes Y^{(2)} := Y^{(1)} \langle \mathrm{D}^\ell, Y^{(2)} \rangle$.

The advantage of this rather formal approach, which may seem like overkill in the polynomial case, is that it allows to define the structure group in the general case. Define a set of formal expressions

$$\mathscr{F}_{\mathrm{AC}}^+ = \left\{ \mathbf{X}^k \prod_j \mathcal{J}_{k_i} \tau_j : k \in \mathbb{N}_0^4,\ \tau_j \in \mathscr{F}_{\mathrm{AC}},\ |\tau_j|_{\mathfrak{s}} + 2 - |k_j|_{\mathfrak{s}} > 0 \right\}. \tag{4.4.12}$$

Here the $\mathcal{J}_k \tau$ are new symbols of degree $|\tau|_{\mathfrak{s}} + 2 - |k|_{\mathfrak{s}}$, which is strictly positive by assumption. Their meaning will become clearer later on. By convention, $\mathscr{F}_{\mathrm{AC}}^+$ also contains the unit $\mathbf{1}$ and all monomials $\mathbf{X}^k$, corresponding to an empty product.

---

[2]In addition to a product and coproduct and associated neutral elements, a Hopf algebra is characterised by a linear map $\mathscr{A} : \overline{T} \to \overline{T}$ called the *antipode*. It should satisfy the identity $M(\mathscr{A} \otimes \mathrm{Id})\Delta^+ = M(\mathrm{Id} \otimes \mathscr{A})\Delta^+$, where M is the multiplication map defined by $M(\tau_1 \otimes \tau_2) = \tau_1 \tau_2$. In this setting, it is given by $\mathscr{A}(\mathbf{X}^k) = (-1)^k \mathbf{X}^k$.

τ	$\Delta^+(\tau)$	$\Gamma_g\tau$
τ_1	$\tau_1\otimes\mathbf{1}+\vee\otimes\mathcal{J}_0\Psi$	$\tau_1+\vee\langle g,\mathcal{J}_0\Psi\rangle$
τ_2	$\tau_2\otimes\mathbf{1}+\uparrow\otimes\mathcal{J}_0\Psi$	$\tau_2+\uparrow\langle g,\mathcal{J}_0\Psi\rangle$
τ_3	$\tau_3\otimes\mathbf{1}+\vee\otimes\mathcal{J}_0\vee$	$\tau_3+\vee\langle g,\mathcal{J}_0\vee\rangle$
$\vee\mathbf{X}_i$	$\vee\mathbf{X}_i\otimes\mathbf{1}+\vee\otimes\mathbf{X}_i$	$\vee\mathbf{X}_i+\vee\langle g,\mathbf{X}_i\rangle$
τ_4	$\tau_4\otimes\mathbf{1}+\mathbf{1}\otimes\mathcal{J}_0\Psi$	$\tau_4+\mathbf{1}\langle g,\mathcal{J}_0\Psi\rangle$
τ_5	$\tau_5\otimes\mathbf{1}+\mathbf{1}\otimes\mathcal{J}_0\vee$	$\tau_5+\uparrow\langle g,\mathcal{J}_0\vee\rangle$
τ_6	$\tau_6\otimes\mathbf{1}+\vee\otimes\mathcal{J}_0\uparrow+\vee\mathbf{X}_i\otimes\mathcal{J}_i\uparrow$ $+\,\vee\otimes\mathbf{X}_i\mathcal{J}_i\uparrow$	$\tau_6+\vee\langle g,\mathcal{J}_0\uparrow\rangle+\vee\mathbf{X}_i\langle g,\mathcal{J}_i\uparrow\rangle$ $+\,\vee\langle g,\mathbf{X}_i\mathcal{J}_i\uparrow\rangle$
τ_7	$\tau_7\otimes\mathbf{1}+\mathbf{1}\otimes\mathcal{J}_0\vee$	$\tau_7+\mathbf{1}\langle g,\mathcal{J}_0\vee\rangle$
τ_8	$\tau_8\otimes\mathbf{1}+\uparrow\otimes\mathcal{J}_0\uparrow+\uparrow\mathbf{X}_i\otimes\mathcal{J}_i\uparrow+\uparrow\otimes\mathbf{X}_i\mathcal{J}_i\uparrow$	$\tau_8+\uparrow\langle g,\mathcal{J}_0\uparrow\rangle+\uparrow\mathbf{X}_i\langle g,\mathcal{J}_i\uparrow\rangle+\uparrow\langle g,\mathbf{X}_i\mathcal{J}_i\uparrow\rangle$
$\mathbf{X}_i$	$\mathbf{X}_i\otimes\mathbf{1}+\mathbf{1}\otimes\mathbf{X}_i$	$\mathbf{X}_i+\mathbf{1}\langle g,\mathbf{X}_i\rangle$
τ_9	$\tau_9\otimes\mathbf{1}+\mathbf{1}\otimes\mathcal{J}_0\uparrow+\mathbf{X}_i\otimes\mathcal{J}_i\uparrow+\mathbf{1}\otimes\mathbf{X}_i\mathcal{J}_i\uparrow$	$\tau_9+\mathbf{1}\langle g,\mathcal{J}_0\uparrow\rangle+\mathbf{X}_i\langle g,\mathcal{J}_i\uparrow\rangle+\mathbf{1}\langle g,\mathbf{X}_i\mathcal{J}_i\uparrow\rangle$

Table 4.2 – Elements of $\mathscr{F}_{\mathrm{AC}}$ of degree up to $\frac{3}{2}$ on which the structure group has a non-trivial effect and their coproducts. Repeated indices i imply summation over all values of i, and we have written $\mathcal{J}_i$ instead of $\mathcal{J}_{e_i}$.

The coproduct Δ^+ is extended from $\overline{T}$ to $T=\mathrm{span}(\mathscr{F}_{\mathrm{AC}})$ by requiring that, in addition to (4.4.9) and multiplicativity as in (4.4.10), it satisfies

$$\Delta^+(\Xi):=\Xi\otimes\mathbf{1}\,,\tag{4.4.13}$$

$$\Delta^+(\mathcal{I}\tau):=(\mathcal{I}\otimes\mathrm{Id})\Delta^+(\tau)+\sum_{\ell,m}\frac{\mathbf{X}^\ell}{\ell!}\otimes\frac{\mathbf{X}^m}{m!}\mathcal{J}_{\ell+m}\tau\,.\tag{4.4.14}$$

The sum in the last equation is necessarily finite, because of the limitation on the degree in (4.4.12). Denoting by T_+^* the set of linear maps $\tau\mapsto\langle h,\tau\rangle$ from $T_+=\mathrm{span}(\mathscr{F}_{\mathrm{AC}}^+)$ to $\mathbb{R}$, the structure group $\mathscr{G}$ is given by the set of group-like elements in T^*, acting on T as

$$(g,\tau)\mapsto\Gamma_g\tau:=(\mathrm{Id}\otimes g)\Delta^+\tau\,.\tag{4.4.15}$$

Note that since $\langle g,\mathbf{1}\rangle=1$ for any group-like g, (4.4.13) yields $\Gamma_g(\Xi)=\Xi$, so that this definition is compatible with the invariance of Ξ required in (4.4.3). Furthermore, the invariance of $\uparrow$, $\vee$ and Ψ required in (4.4.4) is satisfied, because

$$\Delta^+(\uparrow)=\uparrow\otimes\mathbf{1}\,,\qquad\Delta^+(\vee)=\vee\otimes\mathbf{1}\,,\qquad\Delta^+(\Psi)=\Psi\otimes\mathbf{1}\,.$$

Indeed, since $\uparrow$ has negative degree, the sum in (4.4.14) is empty, and the other two identities follow from the product rule (4.4.10). A first non-trivial case is

$$\Delta^+(\tau_4)=(\mathcal{I}\otimes\mathrm{Id})(\Delta^+(\Psi))+\mathbf{1}\otimes\mathcal{J}_0\Psi=\tau_4\otimes\mathbf{1}+\mathbf{1}\otimes\mathcal{J}_0\Psi\,.$$

Indeed, since $|\Psi|_{\mathfrak{s}}+2=\frac{1}{2}-3\kappa$, the term $\ell=m=0$ is permitted in (4.4.14). Substituting in (4.4.15) yields

$$\Gamma_g\tau_4=\tau_4+\mathbf{1}\langle g,\mathcal{J}_0\Psi\rangle\,.\tag{4.4.16}$$

The additional term $\mathbf{1}\langle g,\mathcal{J}_0\Psi\rangle$ will be the one responsible for recentering. More examples of nontrivial structure group actions are listed in Table 4.2. Note that the indempotency requirement (4.4.1) is indeed always satisfied.

Exercise 4.4.6. Check the expressions for $\Delta^+\tau$ and $\Gamma_g\tau$ listed in Table 4.2. What is the matrix representing Γ_g in a basis given by the elements in Table 4.1? ♣

Remark 4.4.7. If $\mathscr{F}_{AC}^+$ were equal to $\mathscr{F}_{AC}$, the coproduct Δ^+ would again endow T with a Hopf-algebra structure. Since these sets are different, T is now a comodule over $T_+ = \mathrm{span}(\mathscr{F}_{AC}^+)$. However, T_+ can also be endowed with a Hopf algebra structure — see [Hai14, Section 8.1]. Strictly speaking, since $\mathscr{F}_{AC}$ and $\mathscr{F}_{AC}^+$ are not stable under multiplication, the comodule should rather be defined on $\mathrm{span}(\mathscr{F})$ and its analogue $\mathrm{span}(\mathscr{F}^+)$, but this has no influence on the theory. ◇

4.4.3 The renormalisation group

The renormalisation group also acts on the model space T, but has a different purpose than the structure group. In the two-dimensional case, we have seen that Wick renormalisation, as appearing in (3.2.15), allows to convert divergent stochastic integrals into convergent ones by subtracting appropriate terms. Here we will perform the same subtractions, but as we have seen in Section 4.2 and in (4.3.8), it is necessary to introduce a second renormalisation constant.

The second renormalisation constant $C_\delta^{(2)}$ is in fact related to the symbol $\overset{\bullet}{\vee}$. This is because similarly to (3.2.11), this symbol represents the stochastic integral

$$\left(P * (P_\delta * \xi)^2 (P_\delta * \xi)^2\right)(z) := \int \cdots \int P(z - z_0) P_\delta(z_0 - z_1) \xi(dz_1) P_\delta(z_0 - z_2) \xi(dz_2)$$
$$\times P_\delta(z - z_3) \xi(dz_3) P_\delta(z - z_4) \xi(dz_4) \, dz_0 \,. \tag{4.4.17}$$

By Isserlis' theorem, we formally have

$$\mathbb{E}[\xi(dz_1)\xi(dz_2)\xi(dz_3)\xi(dz_4)] = \delta(z_1 - z_2)\delta(z_3 - z_4) \, dz_1 \, dz_3$$
$$+ \delta(z_1 - z_3)\delta(z_2 - z_4) \, dz_1 \, dz_2$$
$$+ \delta(z_1 - z_4)\delta(z_2 - z_3) \, dz_1 \, dz_2 \,.$$

The expectation of (4.4.17) is thus given by the sum of three integrals. The first one actually vanishes for symmetry reasons, while the last two integrals are equal, and given by

$$\iiint P(z - z_0) P_\delta(z_0 - z_1) P_\delta(z - z_1) P_\delta(z_0 - z_2) P_\delta(z - z_2) \, dz_0 \, dz_1 \, dz_2 \,. \tag{4.4.18}$$

We can then use the identities (3.2.14) from the proof of Proposition 3.2.16 to obtain

$$\int P_\delta(z_0 - z_1) P_\delta(z - z_1) \, dz_1 = \int_{-\infty}^{\infty} \int_\Lambda P_\delta(t_0 - t_1, x_0 - x_1) P_\delta(t - t_1, x - x_1) \, dx_1 \, dt_1$$
$$= \int_{-\infty}^{\infty} \tilde{P}_\delta(t_0 + t - 2t_1, x_0 - x) \, dt_1 = -\frac{1}{2}\tilde{G}_\delta(x_0 - x) \,.$$

Substituting in (4.4.18) and changing variables from $x - x_0$ to $-x_0$, we thus obtain

$$C_\delta^{(2)} = \mathbb{E}\left[\overset{\bullet}{\vee}_\delta\right] = 2\left(-\frac{1}{2}\right)^2 \int_{-\infty}^{\infty} \int_\Lambda P(t - t_0, -x_0) \tilde{G}_\delta(-x_0)^2 \, dx_0 \, dt_0 = \frac{1}{2} \int_\Lambda \tilde{G}_\delta(x_0)^3 \, dx_0 = \frac{1}{2} \bigoplus \,. \tag{4.4.19}$$

Note that graphically, this amounts to pairing leaves of $\overset{\bullet}{\vee}$, which represent the noise.

The renormalisation group will primarily concern the elements of T having negative degree, namely those in the set

$$\mathscr{F}_- = \{\uparrow, \vee, \overset{\bullet}{\vee}, \overset{\bullet}{\vee}, \overset{\bullet}{\vee}, \overset{\bullet}{\vee}\} \,.$$

Formally, an element M^δ of the renormalisation group is defined as

$$M^\delta := \exp\left\{-C_\delta^{(1)} L_1 - C_\delta^{(2)} L_2\right\},$$

where L_1 and L_2 represent the substitutions

$$L_1 : \mathsf{V} \mapsto \mathbf{1}, \qquad L_2 : \mathsf{W} \mapsto \mathbf{1}.$$

The understanding is that each substitution is applied as often as possible, so that $L_1(\mathsf{W}) = 3\mathsf{I}$, because there are three ways to extract V from W, each one leaving exactly I, which is exactly what happens in Wick renormalisation, cf. (3.2.15). Applying M^δ to the elements of $\mathscr{F}_-$, we get

$$M^\delta(\mathsf{I}) = \mathsf{I}$$
$$M^\delta(\mathsf{V}) = \mathsf{V} - C_\delta^{(1)} \mathbf{1},$$
$$M^\delta(\mathsf{W}) = \mathsf{W} - 3 C_\delta^{(1)} \mathsf{I},$$
$$M^\delta(\mathsf{W}_\bullet) = \mathsf{W}_\bullet - C_\delta^{(1)} \mathsf{Y} - C_\delta^{(2)} \mathbf{1},$$
$$M^\delta(\mathsf{W}_\bullet) = \mathsf{W}_\bullet - 3 C_\delta^{(1)} \mathsf{L}_\bullet - C_\delta^{(1)} \mathsf{W} + 3 \left(C_\delta^{(1)}\right)^2 \mathsf{I} - 3 C_\delta^{(2)} \mathsf{I},$$
$$M^\delta(\mathsf{V}_\bullet) = \mathsf{V}_\bullet - 3 C_\delta^{(1)} \mathsf{L}_\bullet. \tag{4.4.20}$$

Observe that now, the idempotency results from the fact that all additional terms have a *higher* degree than the original term.

> **Remark 4.4.8.** The renormalisation group can also be introduced in a more abstract way, via a coproduct Δ^- and an associated Hopf algebra. See in particular [BHZ19, CH16]. In some sense, this indicates that recentering and renormalisation are "dual" procedures.   ◇

4.5 Regularity structures: analytic aspects

Now that the abstract algebraic framework is in place, we have to complete it with analytic objects, describing the different maps in the commutative diagram (4.3.9).

4.5.1 Models

We describe in this section the map denoted Ψ in the diagram (4.3.9), which associates to a realisation of (mollified) space-time white noise a collection $Z^\delta = (\Pi^\delta, \Gamma^\delta)$ of objects called a *model*. The following definition is [Hai14, Definition 2.17] particularised to the parabolic scaling $\mathfrak{s} = (2,1,1,1)$, where we denote space-time points by $z = (t,x) \in \mathbb{R} \times \mathbb{R}^3 \simeq \mathbb{R}^4$.

> **Definition 4.5.1** (Model). *A model for a regularity structure $(A, T, \mathscr{G})$ is a pair $Z = (\Pi, \Gamma)$, defined by a collection $\{\Pi_z : T \to \mathscr{S}'(\mathbb{R}^4)\}_{z \in \mathbb{R}^4}$ of continuous linear maps and a map $\Gamma : \mathbb{R}^4 \times \mathbb{R}^4 \to \mathscr{G}$ with the following properties.*
>
> 1. *$\Gamma_{zz} = \mathrm{Id}$ is the identity of $\mathscr{G}$ and $\Gamma_{zz'}\Gamma_{z'z''} = \Gamma_{zz''}$ for all $z, z', z'' \in \mathbb{R}^4$.*
> 2. *$\Pi_{z'} = \Pi_z \Gamma_{zz'}$ for all $z, z' \in \mathbb{R}^4$.*
> 3. *Let $r = \lceil -\inf A \rceil$. For any $\gamma \in \mathbb{R}$ and any compact set $\mathfrak{K} \subset \mathbb{R}^4$, one has*
>
> $$\|\Pi\|_{\gamma;\mathfrak{K}} := \sup_{z \in \mathfrak{K}} \sup_{\alpha < \gamma} \sup_{\tau \in T_\alpha} \sup_{\varphi \in B_r} \sup_{0 < \lambda \leqslant 1} \frac{|\langle \Pi_z \tau, \mathscr{S}_z^\lambda \varphi \rangle|}{\lambda^\alpha \|\tau\|_\alpha} < \infty. \tag{4.5.1}$$
>
> 4. *For any $\gamma \in \mathbb{R}$ and any compact set $\mathfrak{K} \subset \mathbb{R}^4$, one has*
>
> $$\|\Gamma\|_{\gamma;\mathfrak{K}} := \sup_{z,z' \in \mathfrak{K}} \sup_{\alpha < \gamma} \sup_{\beta < \alpha} \sup_{\tau \in T_\alpha} \frac{\|\Gamma_{zz'}\tau\|_\beta}{\|\tau\|_\alpha \|z - z'\|_\mathfrak{s}^{\alpha - \beta}} < \infty. \tag{4.5.2}$$

We recall that the set B_r occurring in (4.5.1) is the set of smooth test functions supported in the unit $\|\cdot\|_\mathfrak{s}$-ball and of unit $\mathscr{C}^r$-norm, already encountered in Definition 2.2.7. The norm $\|\tau\|_\alpha$ in (4.5.1) and (4.5.2) is any norm on the finite-dimensional vector space T_α.

The requirement (4.5.1) essentially asks that $\Pi_z\tau$ should belong to $\mathscr{C}_\mathfrak{s}^\alpha$ whenever $\tau \in T_\alpha$. There is a slight difference, however, since the bound is only required to hold when zooming in at the particular point z where the model is evaluated.

An important feature of the theory is that the definition allows for different models for a given regularity structure. This will become important when considering renormalisation. A particular role is played by the *canonical model* $Z^\delta = (\Pi^\delta, \Gamma^\delta)$ for mollified noise $\xi^\delta = \rho^\delta * \xi$, which is defined by

$$
\begin{aligned}
(\Pi_z^\delta\Xi)(\bar{z}) &:= \xi^\delta(\bar{z}) \,, \\
(\Pi_z^\delta\mathbf{X}^k)(\bar{z}) &:= (\bar{z} - z)^k & \forall k \in \mathbb{N}_0^4 \,, \\
(\Pi_z^\delta\tau_1\tau_2)(\bar{z}) &:= (\Pi_z^\delta\tau_1)(\bar{z})(\Pi_z^\delta\tau_2)(\bar{z}) & \forall \tau_1, \tau_2 \in T \,.
\end{aligned}
\tag{4.5.3}
$$

Note that for polynomial symbols $\mathbf{X}^k$, this definition is exactly as in (4.3.3).

The tricky part is again to extend this to elements of the form $\mathcal{I}(\tau)$. This is done in [Hai14, Section 5] via a decomposition

$$
P(t,x) = K(t,x) + R(t,x) \tag{4.5.4}
$$

of the heat kernel P into a smooth part R, and a singular part K with special algebraic properties. These properties are as follows:

- K is supported in the set $\{|x|^2 + |t| \leqslant 1\}$ where $|x| = \sum_{j=1}^3 |x_j|$;
- $K(t,x) = 0$ for $t \leqslant 0$ and $K(t,-x) = K(t,x)$ for all (t,x);
- we have

$$
K(t,x) = \frac{1}{|4\pi t|^{3/2}}\, \mathrm{e}^{-|x|^2/(4t)} \qquad \text{for } |x|^2 + |t| \leqslant \frac{1}{2}
$$

  and $K(t,x)$ is smooth for $|x|^2 + |t| > \frac{1}{2}$;
- finally, the integral of K against any polynomial of parabolic degree less or equal some fixed $\zeta \geqslant 2$ vanishes; this last condition is for compatibility with the non-redundancy condition $\mathcal{I}(\mathbf{X}^k) = 0$.

See [Hai14, Lemma 5.5] for a proof that such a decomposition is indeed possible. This lemma also shows that K can be decomposed, as in (2.3.11), as a sum of kernels K_n, where each K_n is supported in a ball of radius 2^{-n}, and has kth derivatives bounded by $C2^{(1+|k|_\mathfrak{s})n}$, as in (2.3.12). Then the construction of the canonical model is completed by requiring that

$$
(\Pi_z^\delta\mathcal{I}\tau)(\bar{z}) := \langle \Pi_z^\delta\tau, K(\bar{z} - \cdot)\rangle + \sum_\ell \frac{(\bar{z}-z)^\ell}{\ell!}\langle f_z^\delta, \mathcal{J}_\ell\tau\rangle \qquad \forall \tau \in T \,, \tag{4.5.5}
$$

where the $f_z^\delta \in T_+^*$ are linear forms defined in (4.5.6) below. This definition makes sense, even if $\Pi_z^\delta\tau$ is replaced by a genuine distribution $\Pi_z\tau$, by interpreting (4.5.5) as a sum over all K_n, each of which is tested against $\Pi_z\tau$.

As suggested by the construction of the structure group in Section 4.4.2, the linear forms f_z^δ will be used to construct the elements $\Gamma_{zz'}^\delta$ of the model. They are defined by

$$
\begin{aligned}
\langle f_z^\delta, \mathbf{1}\rangle &:= 1 \,, \\
\langle f_z^\delta, \mathbf{X}_i\rangle &:= -z_i \,, \\
\langle f_z^\delta, \tau_1\tau_2\rangle &:= \langle f_z^\delta, \tau_1\rangle\langle f_z^\delta, \bar{\tau}_2\rangle & \forall \tau_1, \tau_2 \in T \,, \\
\langle f_z^\delta, \mathcal{J}_\ell\tau\rangle &:= -\langle \Pi_z^\delta\tau, \mathrm{D}^\ell K(z - \cdot)\rangle & \forall \tau \in T \,.
\end{aligned}
\tag{4.5.6}
$$

The maps $\Gamma^\delta_{zz'}$ of the canonical model are then defined, as in (4.4.15), by

$$\Gamma^\delta_{zz'} := (F^\delta_z)^{-1} F^\delta_{z'} \qquad \text{where } F^\delta_z := \Gamma_{f^\delta_z} = (\mathrm{Id} \otimes f^\delta_z)\Delta^+ . \tag{4.5.7}$$

Note that the first property of Definition 4.5.1 is automatically satisfied. To check the second property, let $\gamma_{zz'}$ denote the element of the structure group such that $\Gamma_{zz'} = \Gamma_{\gamma_{zz'}}$. Writing $\Delta^+\tau = \tau^{(1)} \otimes \tau^{(2)}$ in Sweedler's notation, we have by (4.4.15)

$$\Pi^\delta_z \Gamma^\delta_{zz'}\tau = \Pi^\delta_z(\mathrm{Id} \otimes \gamma^\delta_{zz'})\Delta^+\tau = \Pi^\delta_z \tau^{(1)}\langle \gamma^\delta_{zz'}, \tau^{(2)}\rangle .$$

The second property of Definition 4.5.1 thus amounts to the relation

$$\Pi^\delta_{z'}\tau = \Pi^\delta_z \tau^{(1)}\langle \gamma^\delta_{z\bar{z}}, \tau^{(2)}\rangle ,$$

which provides some intuition for the meaning of Δ^+. The fact that $(\Pi^\delta, \Gamma^\delta)$ is indeed a model is proved in [Hai14, Proposition 8.27].

Exercise 4.5.2.
- Check that in the case $\tau = \mathbf{X}_i$, the definition (4.5.6) is compatible with the expressions (4.3.5) of the Γ_{xy} introduced in the one-dimensional polynomial case.
- Compute $(\Pi^\delta_z \overset{\bullet\!\bullet}{\mathsf{Y}})(\bar{z})$. Compare with (2.3.16) and interpret the result.
- Compute $F^\delta_z \overset{\bullet\!\bullet}{\mathsf{Y}}$. Representing this in the basis $(\mathbf{1}, \overset{\bullet\!\bullet}{\mathsf{Y}})$, compute its inverse, and determine a function χ^δ such that

$$\Gamma^\delta_{zz'} \overset{\bullet\!\bullet}{\mathsf{Y}} = \overset{\bullet\!\bullet}{\mathsf{Y}} + [\chi^\delta(z) - \chi^\delta(z')]\mathbf{1} .$$

What is the associated expression for $\gamma^\delta_{zz'}$? Compare with (4.4.16).   ♣

Remark 4.5.3. The only rule that we are going to bend for general models is the product rule (4.5.3). For instance, in the case of Wick renormalisation, the model for V is different from the product of two models for I, owing to the subtraction a counterterm.   ◇

4.5.2 Modelled distributions

In this section, we describe the space in which the fixed-point equation defining the solution map $\mathscr{S}$ lives, which appears in the upper part of the commutative diagram (4.3.9). The following definition is a particularisation of [Hai14, Definition 3.1].

Definition 4.5.4 (Modelled distributions). *Let $\gamma \in \mathbb{R}$. The space of modelled distributions $\mathscr{D}^\gamma$ consists of all function $f : \mathbb{R}^4 \to T^-_\gamma$ such that for every compact set $\mathfrak{K} \subset \mathbb{R}^4$ one has*

$$\|f\|_{\gamma;\mathfrak{K}} := \sup_{z \in \mathfrak{K}} \sup_{\beta < \gamma} \|f(z)\|_\beta + \sup_{\substack{z,z' \in \mathfrak{K} \\ \|z-z'\|_{\mathsf{s}} \leqslant 1}} \sup_{\beta < \gamma} \frac{\|f(z) - \Gamma_{zz'}f(z')\|_\beta}{\|z-z'\|_{\mathsf{s}}^{\gamma - \beta}} < \infty . \tag{4.5.8}$$

Here $\|f(z)\|_\beta$ denotes the norm of the projection of $f(z)$ on T_β. A modelled distribution can thus be written as a Taylor-series-like expression

$$f(z) = \sum_{\alpha < \gamma} \sum_{\tau \in T_\alpha} f_\tau(z)\tau , \tag{4.5.9}$$

where the projection of $f(z) - \Gamma_{zz'}f(z')$ on T_β satisfies a Hölder condition of exponent $\gamma - \beta$. This is indeed a generalisation of (4.3.2), while the condition (4.5.8) is of the same form as the condition (4.3.7) that ensured that f belong to a given Hölder space. One important difference

between expansions of the form (4.5.9) and usual Taylor expansions is that the "basis vectors" τ in (4.5.9) can be associated with very irregular distributions, while Taylor expansions use monomials which are analytic functions. One should thus distinguish between the *regularity* $\alpha_0 = \inf A$ of the expansion (4.5.9), and the Hölder exponent γ, which in applications will always be larger than α_0 (and, in fact, positive).

It may happen that the smallest value of α occurring in the expansion (4.5.9) is some degree $\alpha_1 > \alpha_0$. In that case, we say that f has regularity α_1 (or that it is defined on a *sector* $V \subset T$ of regularity α_1). This is important, in particular, for the following result, which corresponds to [Hai14, Theorem 4.7].

Theorem 4.5.5 (Multiplication of modelled distributions). *Let $f_1 \in \mathscr{D}^{\gamma_1}$ and $f_2 \in \mathscr{D}^{\gamma_2}$ be modelled distributions of respective regularity $\alpha_1 \leqslant 0$ and $\alpha_2 \leqslant 0$, and let $\gamma = (\gamma_1 + \alpha_1) \wedge (\gamma_2 + \alpha_2)$. Then one can define a modelled distribution $f_1 f_2 \in \mathscr{D}^{\gamma}$ of regularity $\alpha_1 + \alpha_2$, which satisfies*

$$\|f_1 f_2\|_{\gamma;\bar{\mathfrak{K}}} \lesssim \|f_1\|_{\gamma_1;\bar{\mathfrak{K}}} \|f_2\|_{\gamma_2;\bar{\mathfrak{K}}} \left(1 + \|\Gamma_{\gamma_1 + \gamma_2;\bar{\mathfrak{K}}}\|\right)^2 .$$

for every compact set $\bar{\mathfrak{K}} \subset \mathbb{R}^4$. The product $f_1 f_2$ is obtained by multiplying formally the expansions of f_1 and f_2 and truncating them to order γ.

Note that in Definition 4.5.4, $\mathscr{D}^{\gamma}$ depends on the Γ of the chosen model, so that it should be written $\mathscr{D}^{\gamma}(\Gamma)$ in case one works with several models, as will be the case when dealing with renormalisation. In order to compare elements $f \in \mathscr{D}^{\gamma}(\Gamma)$ and $\bar{f} \in \mathscr{D}^{\gamma}(\bar{\Gamma})$, we introduce the quantities

$$\|f - \bar{f}\|_{\gamma;\bar{\mathfrak{K}}} := \sup_{z \in \bar{\mathfrak{K}}} \sup_{\beta < \gamma} \|f(z) - \bar{f}(z)\|_\beta ,$$

$$\|f ; \bar{f}\|_{\gamma;\bar{\mathfrak{K}}} := \|f - \bar{f}\|_{\gamma;\bar{\mathfrak{K}}} + \sup_{\substack{z, z' \in \bar{\mathfrak{K}} \\ \|z - z'\|_{\mathfrak{s}} \leqslant 1}} \sup_{\beta < \gamma} \frac{\|f(z) - \bar{f}(z) - \Gamma_{zz'} f(z') + \bar{\Gamma}_{zz'} \bar{f}(z')\|_\beta}{\|z - z'\|_{\mathfrak{s}}^{\gamma - \beta}} .$$

The reason for the semicolon in the last expression is that it is not a function of $f - \bar{f}$, and hence is not symmetric under exchange of f and $\bar{f}$. The quantity $\|\cdot ; \cdot\|_{\gamma;\bar{\mathfrak{K}}}$ is thus not strictly speaking a norm, though it has similar properties for all practical purposes.

4.5.3   The reconstruction theorem

A key result of the theory of regularity structures is the reconstruction theorem, which states the existence of a map $\mathscr{R}$ from a space $\mathscr{D}^{\gamma}$ of modelled distributions to a "classical" Hölder space. This map appears on the right-hand side of the commutative diagram (4.3.9). The following result is part of [Hai14, Theorem 3.10].

Theorem 4.5.6 (Reconstruction theorem). *Let $(A, T, \mathscr{G})$ be a regularity structure, let $\alpha_0 = \inf A$, and fix $r > |\alpha_0|$ as well as a model (Π, Γ). Then for every $\gamma \in \mathbb{R}$, there exists a continuous map $\mathscr{R} : \mathscr{D}^{\gamma} \to \mathscr{C}_{\mathfrak{s}}^{\alpha_0}$ such that for every compact set $\bar{\mathfrak{K}} \subset \mathbb{R}^4$,*

$$\left|\langle \mathscr{R} f - \Pi_z f(z), \mathscr{S}_z^\lambda \varphi \rangle\right| \lesssim \lambda^\gamma \|\Pi\|_{\gamma;\bar{\mathfrak{K}}} \|f\|_{\gamma;\bar{\mathfrak{K}}} \tag{4.5.10}$$

holds uniformly over all test functions $\varphi \in B_r$, all $\lambda \in (0,1]$, all $f \in \mathscr{D}^{\gamma}$ and all $z \in \bar{\mathfrak{K}}$. Here $\bar{\mathfrak{K}}$ is the 1-fattening of $\bar{\mathfrak{K}}$, that is, the set of points at $\|\cdot\|_{\mathfrak{s}}$-distance at most 1 from $\bar{\mathfrak{K}}$. Furthermore, if $\gamma > 0$, then $\mathscr{R} f$ is unique.

If $f(z)$ is written as in (4.5.9), then the term $\Pi_z f(z)$ in (4.5.10) has the expression

$$\Pi_z f(z) = \sum_{\alpha < \gamma} \sum_{\tau \in T_\alpha} f_\tau(z) \Pi_z \tau \, ,$$

which is in general a distribution. In the particular case where f has non-negative regularity, one has in fact

$$\mathscr{R} f(z) = (\Pi_z f(z))(z) = \langle \mathbf{1}, f(z) \rangle := f_{\mathbf{1}}(z) \, , \tag{4.5.11}$$

cf. [Hai14, Proposition 3.28], which is a generalisation of (4.3.4). More generally, if every $\Pi_z \tau$ happens to belong to a Hölder space $\mathscr{C}_{\mathfrak{s}}^\alpha$ with $\alpha > 0$, then one has

$$\mathscr{R} f(z) = (\Pi_z f(z))(z) = \sum_{\alpha < \gamma} \sum_{\tau \in T_\alpha} f_\tau(z) (\Pi_z \tau)(z) \, ,$$

see [Hai14, Remark 3.15]. This holds in particular for the canonical model for mollified noise Π^δ constructed in the previous section. In general, $\mathscr{R} f$ and $\Pi_z f(z)$ may differ by a remainder term of degree γ, which is small if γ is large.

The proof of Theorem 4.5.6 given in [Hai14] is based on wavelet analysis. There exists by now an alternative proof, presented in [OW19], which is based on smoothing properties of the heat kernel.

4.5.4 Multilevel Schauder estimates

The last missing piece needed to lift the classical solution map $\overline{\mathscr{S}}$ to the space of modelled distributions, and defining the map $\mathscr{S}$ making the diagram (4.3.9) commute, is an operator lifting the operation of convolution with the heat kernel P. Owing to the decomposition (4.5.4) of P into a singular part K and a smooth part R, this problem can be decomposed into two separate problems, which are to make the following diagrams commute:

$$
\begin{array}{ccc}
\mathscr{D}^\gamma & \xrightarrow{\mathscr{K}_\gamma} & \mathscr{D}^{\gamma+2} \\
\downarrow{\scriptstyle \mathscr{R}} & & \downarrow{\scriptstyle \mathscr{R}} \\
\mathscr{C}_{\mathfrak{s}}^{\alpha_0} & \xrightarrow{K*} & \mathscr{C}_{\mathfrak{s}}^{\alpha_0+2}
\end{array}
\qquad
\begin{array}{ccc}
\mathscr{D}^\gamma & \xrightarrow{R_\gamma \circ \mathscr{R}} & \mathscr{D}^{\gamma+2} \\
\downarrow{\scriptstyle \mathscr{R}} & \nearrow{\scriptstyle R_\gamma} & \downarrow{\scriptstyle \mathscr{R}} \\
\mathscr{C}_{\mathfrak{s}}^{\alpha_0} & \xrightarrow{R*} & \mathscr{C}_{\mathfrak{s}}^{\alpha_0+2}
\end{array}
$$

The smooth part R is actually comparatively easy to deal with. Indeed, one can define an operator R_γ, acting on distributions $\zeta \in \mathscr{C}_{\mathfrak{s}}^{\alpha_0}$ via the Taylor-series-like expression

$$(R_\gamma \zeta)(z) := \sum_{|k|_{\mathfrak{s}} < \gamma} \frac{\mathbf{X}^k}{k!} \langle \zeta, \mathrm{D}^k R(z - \cdot) \rangle \, . \tag{4.5.12}$$

The fact that the right-hand side belongs to $\mathscr{D}^{\gamma+2}$ is shown in [Hai14, Lemma 7.3]. Since $R_\gamma \zeta$ has positive regularity, it follows from (4.5.11) that $(\mathscr{R} R_\gamma \zeta)(z)$ is just the convolution of ζ with $R(z - \cdot)$, so that for $\zeta = \mathscr{R} f$ we have indeed

$$\mathscr{R} R_\gamma \mathscr{R} f = R * \mathscr{R} f \tag{4.5.13}$$

as required.

The construction of the operator $\mathscr{K}_\gamma$ lifting the convolution with the singular part K of the heat kernel is substantially more involved. It turns out that $\mathscr{K}_\gamma$ has to be defined as follows. For any $f \in \mathscr{D}^\gamma$, it is composed of three parts, that is

$$(\mathscr{K}_\gamma f)(z) := \mathcal{I} f(z) + \mathcal{J}(z) f(z) + (\mathcal{N}_\gamma f)(z) \, . \tag{4.5.14}$$

The first term $\mathcal{I}f(z)$ is simply given by applying the abstract integration operator $\mathcal{I}$ to each symbol τ in the expansion (4.5.9) of $f(z)$. The second term, which is again associated with recentering, is obtained by setting, for each $\tau \in T_\alpha$,

$$\mathcal{J}(z)\tau := \sum_{|k|_s < \alpha+2} \frac{\mathbf{X}^k}{k!} \langle \Pi_z \tau, D^k K(z-\cdot) \rangle. \tag{4.5.15}$$

Note the similarity with (4.5.6), which is no accident! The last part of $\mathcal{K}_\gamma$ is a nonlocal operator given by

$$(\mathcal{N}_\gamma f)(z) := \sum_{|k|_s < \gamma+2} \frac{\mathbf{X}^k}{k!} \langle \mathcal{R}f - \Pi_z f(z), D^k K(z-\cdot) \rangle. \tag{4.5.16}$$

As before, both operators (4.5.15) and (4.5.16) actually are to be interpreted by replacing K by a sum of K_n. Note that they both have values in $\overline{T}$, the polynomial part of the regularity structure.

With all these objects in place, we have the following far-reaching generalisation of the Schauder estimate of Theorem 2.3.9, which is an instance of [Hai14, Theorem 5.12].

Theorem 4.5.7 (Multilevel Schauder estimate)**.** *If $\gamma + 2 \notin \mathbb{N}$, then the operator $\mathcal{K}_\gamma$ defined in (4.5.14) maps $\mathcal{D}^\gamma$ into $\mathcal{D}^{\gamma+2}$, and satisfies*

$$\mathcal{R}\mathcal{K}_\gamma f = K * \mathcal{R}f \tag{4.5.17}$$

for any $f \in \mathcal{D}^\gamma$. Furthermore, if $\overline{Z} = (\overline{\Pi}, \overline{\Gamma})$ is a second model, then one has

$$\|\mathcal{K}_\gamma f ; \mathcal{K}_\gamma \bar{f}\|_{\gamma+2;\bar{\mathfrak{K}}} \lesssim \|f ; \bar{f}\|_{\gamma;\bar{\mathfrak{K}}} + \|\Pi - \overline{\Pi}\|_{\gamma;\bar{\mathfrak{K}}} + \|\Gamma - \overline{\Gamma}\|_{\gamma;\bar{\mathfrak{K}}}$$

for any $\bar{f} \in \mathcal{D}^\gamma(\overline{\Gamma})$, where $\bar{\mathfrak{K}}$ is the 1-fattening of $\mathfrak{K}$.

The proof of this result is similar in spirit to the proof we have given of Theorem 2.3.9, but is significantly more involved because of the required bookkeeping of all the polynomial terms. In fact, bounding components of $\mathcal{K}_\gamma f$ in T_α of non-integer degree α is relatively straightforward, and all the difficulty lies in integer values of α.

4.5.5   Convergence of renormalised models

The tools introduced so far suffice to set up a fixed-point equation in a space $\mathcal{D}^\gamma$, associated with the canonical model $Z^\delta = (\Pi^\delta, \Gamma^\delta)$ for mollified noise ξ^δ, built in Section 4.5.1. In order to be able to take the limit $\delta \searrow 0$, one has to incorporate the renormalisation procedure.

We already introduced a group of renormalisation transformations M^δ in Section 4.4.3, see in particular (4.4.20). The M^δ are linear maps from the model space T into itself. It is now necessary to translate this to a transformation from the model Z^δ to a renormalised model $\widehat{Z}^\delta = (\widehat{\Pi}^\delta, \widehat{\Gamma}^\delta)$.

A useful remark here is that the second property in Definition 4.5.1 of models and (4.5.7) imply that for all $z, z' \in \mathbb{R}^4$,

$$\Pi_z^\delta (F_z^\delta)^{-1} = \Pi_{z'}^\delta (F_{z'}^\delta)^{-1} =: \mathbf{\Pi}^\delta,$$

where $\mathbf{\Pi}^\delta$ is independent of z. Therefore, the model can also be specified by the pair $(\mathbf{\Pi}^\delta, f_z^\delta)$, and $(\Pi^\delta, \Gamma^\delta)$ can be recovered via

$$\Pi_z^\delta = \mathbf{\Pi}^\delta \Gamma_{f_z^\delta}, \qquad \Gamma_{zz'}^\delta = \Gamma_{f_z^\delta}^{-1} \Gamma_{f_{z'}^\delta}.$$

The renormalised model associated with the renormalisation transformation M^δ will thus be defined in such a way that

$$\widehat{\boldsymbol{\Pi}}^\delta \tau = \boldsymbol{\Pi}^\delta M^\delta \tau \qquad \forall \tau \in T .$$

While it is possible to compute the renormalised model "by hand", there exists a more efficient way of doing this algebraically, which is described in [Hai14, Section 8.3]. Here we just illustrate the result by giving the expressions of $\widehat{\Pi}_z^\delta \tau$ for the negative-degree elements of $\mathscr{F}$:

$$\widehat{\Pi}_z^\delta(\curlyvee) = \Pi_z^\delta(\curlyvee) ,$$

$$\widehat{\Pi}_z^\delta(\vee) = \Pi_z^\delta(\curlyvee)^2 - C_\delta^{(1)} ,$$

$$\widehat{\Pi}_z^\delta(\veee) = \Pi_z^\delta(\curlyvee)^3 - 3C_\delta^{(1)}\Pi_z^\delta(\curlyvee) ,$$

$$\widehat{\Pi}_z^\delta(\curlyvee\!\!\bullet) = \Pi_z^\delta(\curlyvee)\widehat{\Pi}_z^\delta(\vee) - C_\delta^{(2)} ,$$

$$\widehat{\Pi}_z^\delta(\veee\!\!\bullet) = \Pi_z^\delta(\curlyvee)\widehat{\Pi}_z^\delta(\vee) - 3C_\delta^{(2)}\Pi_z^\delta(\curlyvee) + 3C_\delta^{(1)}\Pi_z^\delta(\vee X_i)\langle f_z, \mathcal{J}_i\curlyvee\rangle ,$$

$$\widehat{\Pi}_z^\delta(\veee\!\!\bullet) = \Pi_z^\delta(\curlyvee)\Pi_z^\delta(\curlyvee) - 3C_\delta^{(1)}\Pi_z^\delta(\curlyvee\!\!\bullet) + 3C_\delta^{(1)}\Pi_z^\delta(\curlyvee X_i)\langle f_z, \mathcal{J}_i\curlyvee\rangle . \qquad (4.5.18)$$

Note that the first three relations are compatible with Wick renormalisation as introduced in (3.2.15), while the last three relations are specific to the three-dimensional case.

The question that arises now is whether this sequence of models converges to a meaningful limit as $\delta \searrow 0$. Here the key result is the following one, see [Hai14, Theorem 10.7].

Theorem 4.5.8 (Convergence criterion for renormalised models). *Assume that there exists $\kappa > 0$ such that, for every test function $\varphi \in B_r$, every $z \in \mathbb{R}^4$ and every $\tau \in T$ of negative degree, there exists a random variable $\langle \widehat{\Pi}_z \tau, \varphi \rangle$ such that*

$$\mathbb{E}\left[|\langle \widehat{\Pi}_z \tau, \mathscr{S}_z^\lambda \varphi \rangle|^2\right] \lesssim \lambda^{2|\tau|_s + \kappa} ,$$

$$\mathbb{E}\left[|\langle \widehat{\Pi}_z \tau - \widehat{\Pi}_z^\delta \tau, \mathscr{S}_z^\lambda \varphi \rangle|^2\right] \lesssim \delta^{2\theta} \lambda^{2|\tau|_s + \kappa} \qquad (4.5.19)$$

holds for some $\theta > 0$. Then there exists a unique model $\widehat{Z} = (\widehat{\Pi}, \widehat{\Gamma})$ such that

$$\mathbb{E}\left[\|\widehat{Z}\|_{\gamma;\mathfrak{K}}^p\right] \lesssim 1 , \qquad \mathbb{E}\left[\|\widehat{Z};\widehat{Z}^\delta\|_{\gamma;\mathfrak{K}}^p\right] \lesssim \delta^{\theta p}$$

holds for every compact $\mathfrak{K} \subset \mathbb{R}^4$ and every $p \geqslant 1$. Here the quantities $\|Z\|_{\gamma;\mathfrak{K}} = \|\Pi\|_{\gamma;\mathfrak{K}} + \|\Gamma\|_{\gamma;\mathfrak{K}}$ and $\|Z;\overline{Z}\|_{\gamma;\mathfrak{K}} = \|\Pi - \overline{\Pi}\|_{\gamma;\mathfrak{K}} + \|\Gamma - \overline{\Gamma}\|_{\gamma;\mathfrak{K}}$ measure the magnitude of models and their difference.

To be more precise, each random variable $\langle \widehat{\Pi}_z \tau, \varphi \rangle$ should belong to a specific Wiener chaos, namely the nth inhomogeneous Wiener chaos if τ contains n symbols Ξ. This means that it can be written as an integral involving n space-time white noise variables, just as in (4.4.17) which is an instance with $n = 4$. The conditions (4.5.19) can then be reformulated in terms of bounds on the Wiener chaos expansion coefficients, (cf. [Hai14, Proposition 10.11]), which can be represented by Feynman diagrams similar to those we encountered in Section 4.2.

4.6 Existence and uniqueness of solutions

We return now to our aim of constructing limits of solutions of regularised Allen–Cahn SPDEs

$$\partial_t \phi_\delta(t,x) = \Delta\phi_\delta(t,x) + \phi_\delta(t,x) - \phi_\delta(t,x)^3 + [3C_\delta^{(1)} - 9C_\delta^{(2)}]\phi_\delta(t,x) + \xi^\delta(t,x) , \qquad (4.6.1)$$

with renormalisation constants $C_\delta^{(1)} = \mathcal{O}(\delta^{-1})$ and $C_\delta^{(2)} = \mathcal{O}(\log(\delta^{-1}))$ defined as in (3.2.12) and (4.4.19). One of the main results of [Hai14] is the following local existence theorem (which was formulated for the Φ^4 model, but the proof extends trivially to the Allen–Cahn equation).

Theorem 4.6.1 (Local existence of solutions for the three-dimensional Allen–Cahn equation).
Assume $\phi_0 \in \mathscr{C}_s^\eta$ for some $\eta > -\frac{2}{3}$. Then the Allen–Cahn equation (4.6.1) *with initial condition ϕ_0
admits a sequence ϕ_δ of local solutions converging in probability, as $\delta \searrow 0$, to a limit ϕ. This limit
is independent of the choice of mollifier ρ.*

Sketch of proof: We will only outline the main steps of the proof, ignoring some technical
subtleties, which are discussed in detail in [Hai14, Sections 9.4 and 10.5].

In a first step, we ignore the counterterms $C_\delta^{(i)}$ in (4.6.1). Given $\gamma > \bar{\gamma} > 0$, consider the
fixed-point equation

$$\Phi = P\phi_0 + (\mathscr{K}_{\bar{\gamma}} + \mathscr{R}R_\gamma\mathscr{R})[\Xi + \Phi - \Phi^3]\,, \tag{4.6.2}$$

where we make a slight abuse of notation by writing $P\phi_0$ also for the lift of $P\phi_0$ to $\mathscr{D}^\gamma$, defined
as in (4.5.12). First we want to show that solving (4.6.2) is indeed equivalent to solving the
regularised equation (4.6.1) without the counterterms. Applying the reconstruction operator
$\mathscr{R}$ for the canonical model $(\Pi^\delta, \Gamma^\delta)$ to both sides of (4.6.2) and using (4.5.17) and (4.5.13), we
obtain

$$\phi := \mathscr{R}\Phi = P\phi_0 + P * \mathscr{R}[\Xi + \Phi - \Phi^3]\,.$$

Since we work with the canonical model, we can identify $P * \mathscr{R}\Xi$ with the stochastic convo-
lution $P * \xi^\delta$. It thus remains to show that $\mathscr{R}(\Phi^3)$ is equal to $(\mathscr{R}\Phi)^3 = \phi^3$. To this end, we
rewrite (4.6.2), by separating non-polynomial from polynomial terms, as

$$\Phi = \mathcal{I}[\Xi + \Phi - \Phi^3] + \sum_k \Phi_{\mathbf{X}^k}\mathbf{X}^k\,. \tag{4.6.3}$$

Iterating this map, it is not hard to see that any fixed point of (4.6.3) should have the expression

$$\Phi(z) = \mathord{\uparrow} + \Phi_{\mathbf{1}}(z)\mathbf{1} - \mathord{\Psi} - 3\Phi_{\mathbf{1}}(z)\mathord{\Upsilon} + \Phi_{\mathbf{X}_i}(z)\mathbf{X}_i + \dots \tag{4.6.4}$$

and thus

$$(\mathscr{R}\Phi)(z) = \Pi_z^\delta\mathord{\uparrow}(z) + \Phi_{\mathbf{1}}(z)\,, \tag{4.6.5}$$

since all other terms in (4.6.4) have strictly positive degree, cf. (4.5.11). Furthermore, we have

$$\Phi(z)^3 = \mathord{\Psi} + 3\Phi_{\mathbf{1}}(z)\mathord{\vee} + 3\Phi_{\mathbf{1}}(z)^2\mathord{\uparrow} - 3\mathord{\Psi} - 9\Phi_{\mathbf{1}}(z)\mathord{\Psi} + 3\Phi_{\mathbf{X}_i}(z)\mathord{\vee}\mathbf{X}_i + \Phi_{\mathbf{1}}(z)^3\mathbf{1} + \dots\,,$$

where the dots stand for terms of strictly positive degree. Applying the reconstruction operator
and comparing to the cube of (4.6.5), we see that $\mathscr{R}(\Phi^3) - (\mathscr{R}\Phi)^3$ is a linear combination of

$$(\Pi_z^\delta\mathord{\Psi})(z)\,, \qquad (\Pi_z^\delta\mathord{\Psi})(z)\,, \qquad \text{and} \qquad (\Pi_z^\delta\mathord{\vee}\mathbf{X}_i)(z)\,.$$

However, all these terms are equal to zero. For instance, $(\Pi_z^\delta\mathord{\Psi})(z) = (\Pi_z^\delta\mathord{\Upsilon})(z)(\Pi_z^\delta\mathord{\vee})(z)$,
where the first term vanishes as shown in Exercise 4.5.2. We conclude that indeed $\phi = \mathscr{R}\Phi$
solves the regularised equation without counterterms.

The second step is to show that (4.6.2) admits indeed a unique fixed point. First we de-
termine admissible values of γ and $\bar{\gamma}$. Assume that $\Phi \in \mathscr{D}^\gamma$ has regularity $\alpha \leqslant 0$. Then The-
orem 4.5.5 shows that $\Phi^3 \in \mathscr{D}^{\gamma+2\alpha}$, and has regularity 3α. The multilevel Schauder estimate
of Theorem 4.5.7 shows that the right-hand side of (4.6.2) is in $\mathscr{D}^{\gamma+2\alpha+2}$ and has regularity
$(-\frac{1}{2} - \kappa) \wedge (3\alpha + 2)$, provided $\bar{\gamma} \geqslant \gamma + 2\alpha$. We can thus choose $\alpha = -\frac{1}{2} - \kappa$, which is the regularity
of the stochastic convolution, $\gamma > 1 + 2\kappa$ and $\bar{\gamma} = \gamma - 1 - 2\kappa$, to ensure that both sides of (4.6.2)
belong to the same spaces.

Here we encounter a new difficulty, which is that the term $P\phi_0$ is not necessarily small in
$\mathscr{D}^\gamma$ near $t = 0$. The solution is to modify Definition 4.5.4 of the spaces of modelled distributions

$\mathscr{D}^\gamma$ in order to allow coefficients that diverge near $t = 0$ in a way controlled by a power η. This yields new spaces $\mathscr{D}^{\gamma,\eta}$ with similar results on multiplication, reconstruction and Schauder estimates as for $\mathscr{D}^\gamma$, discussed in [Hai14, Section 6]. By slightly changing the parameter η, one can also obtain bounds which are small for small times, as required to show that the fixed-point map is a contraction for small enough times.

The third step is to incorporate the counterterms $C_\delta^{(i)}$ and take the limit $\delta \searrow 0$. This implies that we replace the commutative diagram (4.3.9) by

$$
\begin{array}{ccc}
(\phi_0, \widehat{Z}^\delta) & \xmapsto{\;\mathscr{S}_{M^\delta}\;} & \Phi_{M^\delta} \\[4pt]
{\scriptstyle \Psi \circ M^\delta} \Big\uparrow & & \Big\downarrow {\scriptstyle \widehat{\mathscr{R}}} \\[4pt]
(\phi_0, \xi^\delta) & \xmapsto[\;\overline{\mathscr{S}_{M^\delta}}\;] & \phi_\delta
\end{array}
\tag{4.6.6}
$$

where $\overline{\mathscr{S}_{M^\delta}}$ is the solution map of the renormalised equation (4.6.1), $\widehat{Z}^\delta = Z^\delta M^\delta$ is the renormalised model described in (4.5.18), and $\widehat{\mathscr{R}}$ is the corresponding reconstruction operator. The map $\mathscr{S}_{M^\delta}$ is defined by solving the fixed-point equation (4.6.2), with $\mathscr{R}$ replaced by $\widehat{\mathscr{R}}$. Then a similar computation as above, taking into account the additional terms in the model (4.5.18), shows that the diagram (4.6.6) indeed commutes.

Finally, we have to check that the renormalised model satisfies the convergence conditions in Theorem 4.5.8. This is done in [Hai14, Section 10.5] via a lengthy computation involving Feynman diagrams. However, recent results in [CH16, BHZ19, BCCH17] based on so-called BPHZ renormalisation allow to make the procedure much more systematic. □

Remark 4.6.2. One of the technicalities we have glossed over is that the term in brackets on the right-hand side of (4.6.2) has to be multiplied by an indicator function $\mathbb{1}_{\{t>0\}}$. The resulting objects have then to be shown to belong to the right spaces of distributions. ◇

4.7 Large deviations

We finally consider the weak-noise variant of (4.6.1), given by

$$
\partial_t \phi(t,x) = \Delta \phi(t,x) + \phi(t,x) - \phi(t,x)^3 + \big[(2\varepsilon)3C_\delta^{(1)} - (2\varepsilon)^2 9C_\delta^{(2)}\big]\phi(t,x) + \sqrt{2\varepsilon}\,\xi^\delta(t,x)\,. \tag{4.7.1}
$$

The work [HW15] also contains the following three-dimensional analogue of Theorem 3.5.1.

Theorem 4.7.1 (Large-deviation principle for the renormalised equation)**.** *Let $\phi_{\delta,\varepsilon}$ denote the solution of (4.7.1), and let $\varepsilon \mapsto \delta(\varepsilon) \geqslant 0$ be a function such that*

$$
\lim_{\varepsilon \to 0} \delta(\varepsilon) = 0\,. \tag{4.7.2}
$$

Then the family $(\phi_{\delta(\varepsilon),\varepsilon})_{\varepsilon>0}$ with fixed initial condition ϕ_0 satisfies an LDP on $[0,T]$ with good rate function

$$
\mathscr{I}_{[0,T]}(\gamma) := \begin{cases} \dfrac{1}{2}\displaystyle\int_0^T \int_\Lambda \Big[\dfrac{\partial}{\partial t}\gamma(t,x) - \Delta\gamma(t,x) - \gamma(t,x) + \gamma(t,x)^3\Big]^2 \,\mathrm{d}x\,\mathrm{d}t & \text{if the integral is finite}\,, \\[6pt] +\infty & \text{otherwise}\,. \end{cases}
$$

Note that just as in the two-dimensional case, the counterterms are absent from the large-deviation rate function.

As in Section 3.5, for $\delta > 0$, we may also consider the variant of (4.7.1) without counterterms

$$\partial_t \phi(t,x) = \Delta \phi(t,x) + C\phi(t,x) - \phi(t,x)^3 + \sqrt{2\varepsilon}\xi^\delta(t,x)\,, \qquad (4.7.3)$$

which does not admit a limit as $\delta \searrow 0$. Here, the large-deviation result reads as follows.

Theorem 4.7.2 (Large-deviation principle for equation without renormalisation)**.** *Let $\tilde{\phi}_{\delta,\varepsilon}$ denote the solution of* (4.7.3)*, and let $\varepsilon \mapsto \delta(\varepsilon) \geqslant 0$ be a function satisfying* (4.7.2) *as well as*

$$\lim_{\varepsilon \to 0} \varepsilon \delta(\varepsilon)^{-1} = \lambda \in [0,\infty)\,.$$

Then the family $(\tilde{\phi}_{\delta(\varepsilon),\varepsilon})_{\varepsilon>0}$ with fixed initial condition ϕ_0 satisfies an LDP on $[0,T]$ with good rate function

$$\mathscr{I}_{[0,T]}(\gamma) := \begin{cases} \dfrac{1}{2}\displaystyle\int_0^T \int_\Lambda \Big[\dfrac{\partial}{\partial t}\gamma(t,x) - \Delta\gamma(t,x) + C^{(\lambda)}\gamma(t,x) + \gamma(t,x)^3\Big]^2 \,\mathrm{d}x\,\mathrm{d}t & \text{if the integral is finite}\,, \\ +\infty & \text{otherwise}\,, \end{cases}$$

where $C^{(\lambda)} = C - 3\lambda^2 C_\delta^{(1)}$.

4.8   Bibliographical notes

Perturbation theory for the Gibbs measure of the Φ^4 model in three dimensions is an important problem in Euclidean Quantum Field theory that has been studied for a long time with various methods. The earliest works by Glimm and Jaffe and by Feldman approached the problem via a detailed combinatorial analysis of Feynman diagrams [GJ68, GJ73, Fel74, GJ81]. The works [BCG$^+$78, BCG$^+$80] introduced the idea of using a renormalisation group approach, consisting in a decomposition of the covariance of the Gaussian field into scales, which then allows to integrate sucessively over one scale after the other. This method was further perfected in [BDH95], using polymers to control error terms, an approach based on ideas from Statistical Physics [GK71].

In another direction, the approach provided in [BFS83a, BFS83b] allows to bound correlation functions without having to compute the partition function explicitly, by using it as a generating function. This involves the derivation of skeleton inequalities, which were obtained up to third order in [BFS83a], and later extended to all orders in [BF84]. A relatively compact derivation of bounds on the partition function based on the Boué–Dupuis formula was recently obtained in [BG18].

The solution theory for the three-dimensional Φ^4 or Allen–Cahn equation we presented is based on the theory of regularity structures developed in [Hai14]. Alternative approaches include paracontrolled calculus [GIP15, CC18] and a Wilsonian renormalisation group [Kup16]. The large-deviation results are from [HW15].

Regularity structures have been successfully applied to several other equations, including the KPZ equation [Hai13, HS17] and its generalisations to polynomial nonlinearities [HQ18], the continuum parabolic Anderson model [HL15], the Navier–Stokes equation [ZZ15b], the motion of a random string on a curved surface [Hai16], the FitzHugh–Nagumo SPDE [BK16], the dynamical Sine–Gordon model [HS16], the heat equation driven by space-time fractional noise [Dey16], reaction-diffusion equations with a fractional Laplacian [BK17], and the multiplicative stochastic heat equation [HL18]. Recent progress on the theory includes a systematic approach to renormalisation [CH16, BHZ19, BCCH17] based on BPHZ-renormalisation, see also [Hai18, BGHZ19].

Appendix A

Potential theory for reversible Markov chains

A.1 First-hitting time and capacity

Consider a reversible Markov chain on a finite set $\mathcal{X}$, with transition probabilities $p(x,y)$ and invariant probability measure π. Reversibility means that

$$\pi(x)p(x,y) = \pi(y)p(y,x) \qquad \forall x,y \in \mathcal{X}\,,$$

which is equivalent to the transition matrix $P = (p(x,y)_{x,y\in\mathcal{X}})$ being self-adjoint with respect to the $L^2(\mathcal{X},\pi)$ inner product

$$\langle f,g \rangle_\pi := \sum_{x\in\mathcal{X}} \pi(x)f(x)g(x)\,.$$

The *generator* of the chain is given by

$$(Lf)(x) := \sum_{y\in\mathcal{X}} p(x,y)[f(y) - f(x)]\,,$$

and is also self-adjoint in $L^2(\mathcal{X},\pi)$.

Fix a nonempty set $A \subset \mathcal{X}$ and let

$$w_A(x) := \mathbb{E}^x[\tau_A]\,, \qquad \tau_A = \inf\{n \geqslant 0\colon X_n \in A\}$$

be the expected first-hitting time of A. It satisfies the Poisson boundary value problem

$$\begin{cases} (Lw_A)(x) = -1 & x \in A^c\,, \\ w_A(x) = 0 & x \in A\,. \end{cases}$$

The solution of this problem can be written as

$$w_A(x) = -\sum_{y\in A^c} G_{A^c}(x,y)\,,$$

where the Green function (or fundamental matrix) $G_{A^c} = L_{A^c}^{-1}$ is the inverse of the restriction of the generator L to A^c.

Fix now two disjoint nonempty sets $A, B \subset \mathcal{X}$ and let

$$h_{AB}(x) := \mathbb{P}^x\{\tau_A < \tau_B\}$$

be the *committor function* (or *equilibrium potential*), where τ_A is the first-hitting time of A. It satisfies the Dirichlet boundary value problem

$$\begin{cases} (Lh_{AB})(x) = 0 & x \in (A \cup B)^c\,, \\ h_{AB}(x) = 1 & x \in A\,, \\ h_{AB}(x) = 0 & x \in B\,. \end{cases}$$

The *equilibrium measure* is defined by

$$e_{AB}(x) := -(Lh_{AB})(x) \qquad \forall x \in A\,.$$

This measure also has a probabilistic interpretation, namely

$$e_{AB}(x) = \mathbb{P}^x\{\tau_A^+ < \tau_B^+\}\,,$$

where $\tau_A^+ = \inf\{n \geqslant 1 : X_n \in A\}$ denotes the first-return time to A.

Proposition A.1.1. *The committor has the representation*

$$h_{AB}(x) = -\sum_{y \in A} G_{B^c}(x,y)e_{AB}(y) \qquad \forall x \in (A \cup B)^c\,.$$

PROOF: The restriction of Lh_{AB} to B^c satisfies

$$L_{B^c}\begin{pmatrix} h_{AB} \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -e_{AB} \end{pmatrix}$$

where the upper part of the vectors represents elements of $(A \cup B)^c$, and the lower part elements of A. Multiplying on the left by $G_{B^c} = L_{B^c}^{-1}$ yields the result.   □

The *capacity* is defined by

$$\mathrm{cap}(A,B) := \sum_{x \in A} \pi(x)e_{AB}(x)\,. \tag{A.1.1}$$

As a consequence, $\nu_{AB}(x) = \pi(x)e_{AB}(x)/\mathrm{cap}(A,B)$ defines a probability measure on A.

Lemma A.1.2. *The capacity admits the two equivalent representations*

$$\mathrm{cap}(A,B) = \langle h_{AB}, -Lh_{AB}\rangle_\pi\,,$$
$$\mathrm{cap}(A,B) = \frac{1}{2}\sum_{x,y \in \mathcal{X}} \pi(x)p(x,y)[h_{AB}(x) - h_{AB}(y)]^2\,.$$

PROOF: For the first identity, we write

$$\langle h_{AB}, -Lh_{AB}\rangle_\pi = \sum_{x \in A \cup B \cup (A \cup B)^c} \pi(x)h_{AB}(x)(-Lh_{AB})(x)\,.$$

Since $h_{AB}(x) = 1$ in A, while $h_{AB}(x) = 0$ in B and $Lh_{AB}(x) = 0$ in $(A \cup B)^c$, we recover (A.1.1). The second identity is a consequence of summation by parts (or discrete Green's identity). Indeed,

$$\langle h_{AB}, -Lh_{AB}\rangle_\pi = \sum_{x,y \in \mathcal{X}} \pi(x)p(x,y)h_{AB}(x)[h_{AB}(x) - h_{AB}(y)]$$

$$= \sum_{x,y \in \mathcal{X}} \pi(x)p(x,y)[\frac{1}{2}h_{AB}(x)^2 - h_{AB}(x)h_{AB}(y) + \frac{1}{2}h_{AB}(x)^2]$$

$$= \sum_{x,y \in \mathcal{X}} \pi(x)p(x,y)[\frac{1}{2}h_{AB}(x)^2 - h_{AB}(x)h_{AB}(y) + \frac{1}{2}h_{AB}(y)^2]\,,$$

where we have used reversibility and a permutation of summation indices to replace x by y in the last line. $\square$

The main result making the potential-theoretic approach work is the following link between expected hitting times and capacity.

Theorem A.1.3. *For any disjoint sets $A, B \subset \mathcal{X}$, one has*

$$\mathbb{E}^{\nu_{AB}}[\tau_B] := \sum_{x \in A} \nu_{AB}(x) w_B(x) = \frac{1}{\text{cap}(A, B)} \sum_{x \in B^c} \pi(x) h_{AB}(x).$$

Proof: We have the identity

$$\sum_{x \in A} \pi(x) e_{AB}(x) w_B(x) = -\sum_{x \in A} \sum_{y \in B^c} \pi(x) G_{B^c}(x, y) e_{AB}(x)$$

$$= -\sum_{y \in B^c} \sum_{x \in A} \pi(y) G_{B^c}(y, x) e_{AB}(x) = \sum_{y \in B^c} \pi(y) h_{AB}(y).$$

Dividing on both sides by the capacity yields the result. $\square$

A more probabilistic proof of this result can be found in [Slo12b].

A.2   Dirichlet principle

Define the *Dirichlet form* (or energy) by

$$\mathscr{E}(f) := \langle f, -Lf \rangle_\pi = \frac{1}{2} \sum_{x, y \in \mathcal{X}} \pi(x) p(x, y) [f(x) - f(y)]^2.$$

We associate with it the bilinear form

$$\mathscr{E}(f, g) := \frac{1}{2} \sum_{x, y \in \mathcal{X}} \pi(x) p(x, y) [f(x) - f(y)][g(x) - g(y)].$$

Note that the polarisation identity and self-adjointness yield

$$\mathscr{E}(f, g) = \frac{1}{4} (\mathscr{E}(f + g) - \mathscr{E}(f - g)) = \frac{1}{2} (\langle f, -Lg \rangle_\pi + \langle g, -Lf \rangle_\pi) = \langle f, -Lg \rangle_\pi.$$

By the Cauchy–Schwarz inequality (and reversibility), we have

$$\mathscr{E}(f, g)^2 \leqslant \mathscr{E}(f) \mathscr{E}(g). \tag{A.2.1}$$

Furthermore, Lemma A.1.2 shows that

$$\text{cap}(A, B) = \mathscr{E}(h_{AB}).$$

Proposition A.2.1 (Dirichlet principle). *We have*

$$\text{cap}(A, B) = \min_{h \in \mathscr{H}_{AB}} \mathscr{E}(h),$$

where $\mathscr{H}_{AB} = \{h : \mathcal{X} \to [0, 1] : h|_A = 1, h|_B = 0\}$. The minimum is reached in $h = h_{AB}$.

PROOF: Pick any $h \in \mathscr{H}_{AB}$ and observe that

$$
\begin{aligned}
\mathscr{E}(h, h_{AB}) = \langle h, -Lh_{AB}\rangle_\pi &= \sum_{x \in \mathcal{X}} \pi(x)h(x)(-Lh_{AB})(x) \\
&= \sum_{x \in A} \pi(x)(-Lh_{AB})(x) \\
&= \sum_{x \in A} \pi(x)e_{AB}(x) \\
&= \operatorname{cap}(A, B) \, ,
\end{aligned}
$$

where we have used the boundary conditions for h and the fact that h_{AB} vanishes in $(A \cup B)^c$ to obtain the second line. The Cauchy–Schwarz inequality (A.2.1) yields

$$
\operatorname{cap}(A, B)^2 = \mathscr{E}(h, h_{AB})^2 \leqslant \mathscr{E}(h)\mathscr{E}(h_{AB}) = \mathscr{E}(h)\operatorname{cap}(A, B)
$$

and hence $\operatorname{cap}(A, B) \leqslant \mathscr{E}(h)$. We already know that equality is reached for $h = h_{AB}$. □

A.3   Thomson principle

The *Thomson principle* provides a complementary variational principle for the capacity. It comes in two versions.

Proposition A.3.1 (Thomson principle, super-harmonic version)**.** *We have*

$$
\operatorname{cap}(A, B) = \max_{h \in \mathfrak{H}_{AB}} \frac{\left(\sum_{x \in A} \pi(x)(-Lh)(x)\right)^2}{\mathscr{E}(h)} \, ,
$$

where $\mathfrak{H}_{AB} = \{h : \mathcal{X} \to [0, 1] : (Lh)(x) \leqslant 0 \ \forall x \in B^c\}$*. The maximum is reached for* $h = h_{AB}$*.*

PROOF: Pick $h \in \mathfrak{H}_{AB}$. We have

$$
\mathscr{E}(h, h_{AB}) = \langle -Lh, h_{AB}\rangle_\pi = \sum_{x \in \mathcal{X}} \pi(x)(-Lh)(x)h_{AB}(x) \, .
$$

Since π, $-Lh$ and h_{AB} are non-negative, we obtain

$$
\mathscr{E}(h, h_{AB}) \geqslant \sum_{x \in A} \pi(x)(-Lh)(x) \geqslant 0 \, .
$$

It follows from the Cauchy–Schwarz inequality that

$$
\left(\sum_{x \in A} \pi(x)(-Lh)(x)\right)^2 \leqslant \mathscr{E}(h, h_{AB})^2 \leqslant \mathscr{E}(h)\operatorname{cap}(A, B) \, .
$$

This proves the inequality, while (A.1.1) shows that equality is again reached for $h = h_{AB}$. □

The second version involves the notion of (discrete) unit flow.

Definition A.3.2 (Unit *AB*-flow). *A* unit *AB*-flow *is a map* $\varphi : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ *satisfying*

- *antisymmetry:* $\varphi(x,y) = -\varphi(y,x)$ *for all* $x,y \in \mathcal{X}$;
- *compatibility with p: if* $p(x,y) = 0$, *then* $\varphi(x,y) = 0$;
- *Kirchhoff's law (divergence-freeness):*

$$(\operatorname{div}\varphi)(x) := \sum_{y \in \mathcal{X}} \varphi(x,y) = 0 \qquad \forall x \in (A \cup B)^c \, ;$$

- *unit intensity:*

$$\sum_{x \in A}(\operatorname{div}\varphi)(x) = 1 = -\sum_{x \in B}(\operatorname{div}\varphi)(x) \, .$$

We denote by $\mathfrak{U}^1_{AB}$ *the set of unit AB-flows.*

One should think of φ as being defined on the edges of the graph associated with P (i.e., the (x,y) on which $p(x,y) > 0$).

A special role is played by the *harmonic unit flow*

$$\varphi_{AB}(x,y) := \frac{\pi(x)p(x,y)}{\operatorname{cap}(A,B)}[h_{AB}(x) - h_{AB}(y)] \, .$$

Lemma A.3.3. φ_{AB} *is a unit AB-flow.*

Proof: φ_{AB} is clearly antisymmetric and compatible with p. It is divergence-free, since for all $x \in (A \cup B)^c$,

$$(\operatorname{div}\varphi_{AB})(x) = \frac{\pi(x)}{\operatorname{cap}(A,B)}\sum_{y \in \mathcal{X}} p(x,y)[h_{AB}(x) - h_{AB}(y)] = \frac{\pi(x)}{\operatorname{cap}(A,B)}(-Lh_{AB})(x) = 0 \, .$$

Finally, it has unit intensity, because

$$\sum_{x \in A}(\operatorname{div}\varphi_{AB})(x) = \frac{1}{\operatorname{cap}(A,B)}\sum_{x \in A}\pi(x)(-Lh_{AB})(x)$$

$$= \frac{1}{\operatorname{cap}(A,B)}\sqrt{\operatorname{cap}(A,B)\mathscr{E}(h_{AB})} = 1$$

by Proposition A.3.1. The analogous statement for B holds by a discrete version of the divergence theorem. $\qquad\square$

We now define a bilinear form $\mathscr{D}$ on the set of unit *AB*-flows by

$$\mathscr{D}(\varphi,\psi) := \frac{1}{2}\sum_{x,y \in \mathcal{X}}\frac{1}{\pi(x)p(x,y)}\varphi(x,y)\psi(x,y) \, ,$$

and introduce the shortcut $\mathscr{D}(\varphi,\varphi) =: \mathscr{D}(\varphi)$. Again, the Cauchy–Schwarz inequality applies, yielding

$$\mathscr{D}(\varphi,\psi)^2 \leqslant \mathscr{D}(\varphi)\mathscr{D}(\psi) \, .$$

Proposition A.3.4 (Thomson principle, unit flow version). *We have*

$$\operatorname{cap}(A,B) = \max_{\varphi \in \mathfrak{U}^1_{AB}} \frac{1}{\mathscr{D}(\varphi)} \, .$$

The maximum is reached for $\varphi = \varphi_{AB}$.

Proof: For a function $h : \mathcal{X} \to \mathbb{R}$, define the flow

$$\Psi_h(x, y) := \pi(x) p(x, y) [h(x) - h(y)] .$$

Then $\varphi_{AB} = \Psi_{h_{AB}} / \mathrm{cap}(A, B)$. Thus $\Psi_{h_{AB}}$ is a (non-unit) AB-flow, and we find

$$\begin{aligned}
\mathscr{D}(\Psi_{h_{AB}}) &= \frac{1}{2} \sum_{x,y \in \mathcal{X}} \frac{1}{\pi(x) p(x, y)} \Psi_{h_{AB}}(x, y)^2 \\
&= \frac{1}{2} \sum_{x,y \in \mathcal{X}} \pi(x) p(x, y) [h_{AB}(x) - h_{AB}(y)]^2 \\
&= \mathrm{cap}(A, B) .
\end{aligned}$$

By bilinearity, we have

$$\mathscr{D}(\varphi_{AB}) = \frac{1}{\mathrm{cap}(A, B)} .$$

Furthermore, given any unit AB-flow φ, we obtain

$$\begin{aligned}
\mathscr{D}(\Psi_{h_{AB}}, \varphi) &= \frac{1}{2} \sum_{x,y \in \mathcal{X}} [h_{AB}(x) - h_{AB}(y)] \varphi(x, y) \\
&= \sum_{x,y \in \mathcal{X}} h_{AB}(x) \varphi(x, y) \\
&= \sum_{x \in A} \sum_{y \in \mathcal{X}} \varphi(x, y) + \sum_{x \in (A \cup B)^c} h_{AB}(x) \sum_{y \in \mathcal{X}} \varphi(x, y) \\
&= \sum_{x \in A} (\mathrm{div}\, \varphi)(x) + \sum_{x \in (A \cup B)^c} h_{AB}(x) \underbrace{(\mathrm{div}\, \varphi)(x)}_{=0} \\
&= 1 .
\end{aligned}$$

The Cauchy–Schwarz inequality thus yields

$$1 = \mathscr{D}(\Psi_{h_{AB}}, \varphi)^2 \leqslant \mathscr{D}(\Psi_{h_{AB}}) \mathscr{D}(\varphi) = \mathrm{cap}(A, B) \mathscr{D}(\varphi) ,$$

showing that $\mathrm{cap}(A, B) \geqslant 1 / \mathscr{D}(\varphi)$ as required.   □

A.4 Bibliographical notes

A comprehensive account of the potential-theoretic approach to reversible Markov chains is given in [BdH15, Sections 7.1 and 7.3]. An overview can also be found in [Slo12b].

Let us mention that there exists an improved version of the Thomson principle, called the *Berman–Konsowa principle* [BK90]. The fact that the Berman–Konsowa principle is sharper than the Thomson principle is a consequence of Jensen's inequality, as shown in [Slo12a]. An extension of the Berman–Konsowa principle to continuous-space Markov chains has been obtained in [dHJ16].

Bibliography

[AC79]   S.M. Allen and J.W. Cahn. A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening. *Acta Metallurgica*, 27(6):1085–1905, 1979.

[AR91]   Sergio Albeverio and Michael Röckner. Stochastic differential equations in infinite dimensions: solutions via Dirichlet forms. *Probab. Theory Related Fields*, 89(3):347–386, 1991.

[Arr89]   Svante Arrhenius. On the reaction velocity of the inversion of cane sugar by acids. *J. Phys. Chem.*, 4:226, 1889. In German. Translated and published in: Selected Readings in Chemical Kinetics, M.H. Back and K.J. Laider (eds.), Pergamon, Oxford, 1967.

[Bar15]   Florent Barret. Sharp asymptotics of metastable transition times for one dimensional SPDEs. *Ann. Inst. Henri Poincaré Probab. Stat.*, 51(1):129–166, 2015.

[BBM10]   Florent Barret, Anton Bovier, and Sylvie Méléard. Uniform estimates for metastable transition times in a coupled bistable system. *Electron. J. Probab.*, 15:no. 12, 323–345, 2010.

[BCCH17]   Y. Bruned, A. Chandra, I. Chevyrev, and M. Hairer. Renormalising SPDEs in regularity structures. *arXiv:1711.10239*, pages 1–85, 2017. To appear in J. Eur. Math. Soc.

[BCD11]   Hajer Bahouri, Jean-Yves Chemin, and Raphaël Danchin. *Fourier analysis and nonlinear partial differential equations*, volume 343 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, 2011.

[BCG⁺78]   G. Benfatto, M. Cassandro, G. Gallavotti, F. Nicolò, E. Olivieri, E. Presutti, and E. Scacciatelli. Some probabilistic techniques in field theory. *Comm. Math. Phys.*, 59(2):143–166, 1978.

[BCG⁺80]   G. Benfatto, M. Cassandro, G. Gallavotti, F. Nicolò, E. Olivieri, E. Presutti, and E. Scacciatelli. Ultraviolet stability in Euclidean scalar field theories. *Comm. Math. Phys.*, 71(2):95–130, 1980.

[BD98]   Michelle Boué and Paul Dupuis. A variational representation for certain functionals of Brownian motion. *Ann. Probab.*, 26(4):1641–1659, 1998.

[BD16a]   Nils Berglund and Sébastien Dutercq. The Eyring–Kramers law for Markovian jump processes with symmetries. *J. Theoret. Probab.*, 29(4):1240–1279, 2016.

[BD16b]   Nils Berglund and Sébastien Dutercq. Interface dynamics of a metastable mass-conserving spatially extended diffusion. *J. Stat. Phys.*, 162(2):334–370, 2016.

[BDGW17] Nils Berglund, Giacomo Di Gesù, and Hendrik Weber. An Eyring–Kramers law for the stochastic Allen–Cahn equation in dimension two. *Electron. J. Probab.*, 22:1–27, 2017.

[BDH95]   D. Brydges, J. Dimock, and T. R. Hurd. The short distance behavior of $(\phi^4)_3$. *Comm. Math. Phys.*, 172(1):143–186, 1995.

[BdH15]   Anton Bovier and Frank den Hollander. *Metastability. A potential-theoretic approach*, volume 351 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Cham, 2015.

[BEGK02]   Anton Bovier, Michael Eckhoff, Véronique Gayrard, and Markus Klein. Metastability and low lying spectra in reversible Markov chains. *Comm. Math. Phys.*, 228(2):219–255, 2002.

[BEGK04]   Anton Bovier, Michael Eckhoff, Véronique Gayrard, and Markus Klein. Metastability in reversible diffusion processes. I. Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc. (JEMS)*, 6(4):399–424, 2004.

[Ber13]   Nils Berglund. Kramers' law: Validity, derivations and generalisations. *Markov Process. Related Fields*, 19(3):459–490, 2013.

[BF84]   Anton Bovier and Giovanni Felder. Skeleton inequalities and the asymptotic nature of perturbation theory for Φ^4-theories in two and three dimensions. *Comm. Math. Phys.*, 93(2):259–275, 1984.

[BFG07a]   Nils Berglund, Bastien Fernandez, and Barbara Gentz. Metastability in interacting nonlinear stochastic differential equations: I. From weak coupling to synchronization. *Nonlinearity*, 20(11):2551–2581, 2007.

[BFG07b]   Nils Berglund, Bastien Fernandez, and Barbara Gentz. Metastability in interacting nonlinear stochastic differential equations II: Large-N behaviour. *Nonlinearity*, 20(11):2583–2614, 2007.

[BFS83a]   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. *Comm. Math. Phys.*, 91(2):141–186, 1983.

[BFS83b]   David C. Brydges, Jürg Fröhlich, and Alan D. Sokal. The random-walk representation of classical spin systems and correlation inequalities. II. The skeleton inequalities. *Comm. Math. Phys.*, 91(1):117–139, 1983.

[BG10]   Nils Berglund and Barbara Gentz. The Eyring-Kramers law for potentials with nonquadratic saddles. *Markov Process. Related Fields*, 16(3):549–598, 2010.

[BG13]   Nils Berglund and Barbara Gentz. Sharp estimates for metastable lifetimes in parabolic SPDEs: Kramers' law and beyond. *Electron. J. Probab.*, 18:no. 24, 58, 2013.

[BG18]   N. Barashkov and M. Gubinelli.   A variational method for ϕ_3^4.   Preprint
            arXiv:1805.1081, 2018.

[BGHZ19]   Yvain Bruned, Franck Gabriel, Martin Hairer, and Lorenzo Zambotti. Geometric
            stochastic heat equations. 2019.

[BGK05]   Anton Bovier, Véronique Gayrard, and Markus Klein. Metastability in reversible
            diffusion processes. II. Precise asymptotics for small eigenvalues. *J. Eur. Math. Soc.
            (JEMS)*, 7(1):69–99, 2005.

[BHZ19]   Y. Bruned, M. Hairer, and L. Zambotti. Algebraic renormalisation of regularity
            structures. *Invent. Math.*, 215(3):1039–1156, Mar 2019.

[BJ13]   Dirk Blömker and Arnulf Jentzen. Galerkin approximations for the stochastic
            Burgers equation. *SIAM J. Numer. Anal.*, 51(1):694–715, 2013.

[BK90]   Kenneth A. Berman and Mokhtar H. Konsowa. Random paths and cuts, electrical
            networks, and reversible Markov chains. *SIAM J. Discrete Math.*, 3(3):311–319,
            1990.

[BK16]   Nils Berglund and Christian Kuehn. Regularity structures and renormalisation of
            FitzHugh–Nagumo SPDEs in three space dimensions. *Electron. J. Probab.*, 21:Paper
            No. 18, 48, 2016.

[BK17]   Nils Berglund and Christian Kuehn. Model spaces of regularity structures for
            space-fractional SPDEs. *Journal of Statistical Physics*, pages 331–368, 2017.

[CC18]   Rémi Catellier and Khalil Chouk.   Paracontrolled distributions and the 3-
            dimensional stochastic quantization equation. *Ann. Probab.*, 46(5):2621–2679,
            2018.

[Cer99]   Sandra Cerrai. Smoothing properties of transition semigroups relative to SDEs
            with values in Banach spaces. *Probab. Theory Related Fields*, 113(1):85–114, 1999.

[CH16]   Ajay Chandra and Martin Hairer. An analytic BPHZ theorem for regularity struc-
            tures. arXiv:1612.08138, 2016.

[Che04]   Xinfu Chen. Generation, propagation, and annihilation of metastable patterns. *J.
            Differential Equations*, 206(2):399–437, 2004.

[CI75]   N. Chafee and E. F. Infante. A bifurcation problem for a nonlinear partial differ-
            ential equation of parabolic type. *Applicable Anal.*, 4:17–37, 1974/75.

[CM97]   Fabien Chenal and Annie Millet. Uniform large deviations for parabolic SPDEs
            and applications. *Stochastic Process. Appl.*, 72(2):161–186, 1997.

[CP89]   J. Carr and R. L. Pego. Metastable patterns in solutions of $u_t = \epsilon^2 u_{xx} - f(u)$. *Comm.
            Pure Appl. Math.*, 42(5):523–576, 1989.

[CW17]   Ajay Chandra and Hendrik Weber. Stochastic PDEs, regularity structures, and
            interacting particle systems. *Ann. Fac. Sci. Toulouse Math. (6)*, 26(4):847–909, 2017.

[Dey16]   Aurélien Deya. On a modelled rough heat equation. *Probab. Theory Related Fields*,
            166(1-2):1–65, 2016.

[DGLP17]   Giacomo Di Gesù and Dorian Le Peutrec. Small noise spectral gap asymptotics for
            a large system of nonlinear diffusions. *J. Spectr. Theory*, 7(4):939–984, 2017.

[dHJ16]   F. den Hollander and S. Jansen. Berman–Konsowa principle for reversible Markov jump processes. *Markov Process. Related Fields*, 22(3):409–442, 2016.

[DPD03]   Giuseppe Da Prato and Arnaud Debussche. Strong solutions to the stochastic quantization equations. *Ann. Probab.*, 31(4):1900–1916, 2003.

[DPZ96]   G. Da Prato and J. Zabczyk. *Ergodicity for infinite-dimensional systems*, volume 229 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1996.

[DS89]   Jean-Dominique Deuschel and Daniel W. Stroock. *Large deviations*. Academic Press, Boston, 1989. Reprinted by the American Mathematical Society, 2001.

[Dut15]   Sébastien Dutercq. *Métastabilité dans les systèmes avec lois de conservation*. PhD thesis, Université d'Orléans, 2015.

[DZ98]   Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38 of *Applications of Mathematics*. Springer-Verlag, New York, second edition, 1998.

[Eck05]   Michael Eckhoff. Precise asymptotics of small eigenvalues of reversible diffusions in the metastable regime. *Ann. Probab.*, 33(1):244–299, 2005.

[Fel74]   Joel Feldman. The $\lambda\varphi_3^4$ field theory in a finite volume. *Comm. Math. Phys.*, 37:93–120, 1974.

[FJL82]   William G. Faris and Giovanni Jona-Lasinio. Large fluctuations for a nonlinear heat equation with noise. *J. Phys. A*, 15(10):3025–3055, 1982.

[Fre88]   Mark I. Freidlin. Random perturbations of reaction-diffusion equations: the quasideterministic approximation. *Trans. Amer. Math. Soc.*, 305(2):665–697, 1988.

[FW98]   M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*. Springer-Verlag, New York, second edition, 1998.

[GIP15]   Massimiliano Gubinelli, Peter Imkeller, and Nicolas Perkowski. Paracontrolled distributions and singular PDEs. *Forum Math. Pi*, 3:e6, 75, 2015.

[GJ68]   James Glimm and Arthur Jaffe. A $\lambda\phi^4$ quantum field without cutoffs. I. *Phys. Rev. (2)*, 176:1945–1951, 1968.

[GJ73]   James Glimm and Arthur Jaffe. Positivity of the ϕ_3^4 Hamiltonian. *Fortschr. Physik*, 21:327–376, 1973.

[GJ81]   James Glimm and Arthur Jaffe. *Quantum physics. A functional integral point of view*. Springer-Verlag, New York-Berlin, 1981.

[GJS75]   James Glimm, Arthur Jaffe, and Thomas Spencer. Phase transitions for ϕ_2^4 quantum fields. *Communications in Mathematical Physics*, 45(3):203–216, Oct 1975.

[GK71]   C. Gruber and H. Kunz. General properties of polymer systems. *Comm. Math. Phys.*, 22:133–161, 1971.

[Hai09]   Martin Hairer. An introduction to stochastic PDEs. Lecture notes, 2009. `http://arxiv.org/abs/0907.4178`.

[Hai13]   Martin Hairer. Solving the KPZ equation. *Ann. of Math. (2)*, 178(2):559–664, 2013.

[Hai14]   Martin Hairer. A theory of regularity structures. *Invent. Math.*, 198(2):269–504, 2014.

[Hai15]   Martin Hairer. Regularity structures and the dynamical Φ_3^4 model. Review article, `arXiv:1508.05261`, 2015.

[Hai16]   Martin Hairer. The motion of a random string. *arXiv:1605.02192*, pages 1–20, 2016. To appear in Proceedings of the XVIII ICMP.

[Hai18]   Martin Hairer. An analyst's take on the BPHZ theorem. In *Comput. Combin. Dyn. Stoch. Control*, pages 429–476, Cham, 2018. Springer International Publishing.

[HKN04]   Bernard Helffer, Markus Klein, and Francis Nier. Quantitative analysis of metastability in reversible diffusion processes via a Witten complex approach. *Mat. Contemp.*, 26:41–85, 2004.

[HL15]   Martin Hairer and Cyril Labbé. A simple construction of the continuum parabolic Anderson model on $\mathbf{R}^2$. *Electron. Commun. Probab.*, 20:no. 43, 11, 2015.

[HL18]   Martin Hairer and Cyril Labbé. Multiplicative stochastic heat equations on the whole space. *J. Eur. Math. Soc. (JEMS)*, 20(4):1005–1054, 2018.

[HM18]   Martin Hairer and Jonathan Mattingly. The strong Feller property for singular stochastic PDEs. *Ann. Inst. Henri Poincaré Probab. Stat.*, 54(3):1314–1340, 2018.

[HN05]   Bernard Helffer and Francis Nier. *Hypoelliptic estimates and spectral theory for Fokker-Planck operators and Witten Laplacians*, volume 1862 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2005.

[HQ18]   Martin Hairer and Jeremy Quastel. A class of growth models rescaling to KPZ. *Forum Math. Pi*, 6:e3, 112, 2018.

[HS16]   Martin Hairer and Hao Shen. The dynamical sine-Gordon model. *Comm. Math. Phys.*, 341(3):933–989, 2016.

[HS17]   Martin Hairer and Hao Shen. A central limit theorem for the KPZ equation. *Ann. Probab.*, 45(6B):4167–4221, 2017.

[HW15]   Martin Hairer and Hendrik Weber. Large deviations for white-noise driven, nonlinear stochastic PDEs in two and three dimensions. *Ann. Fac. Sci. Toulouse Math. (6)*, 24(1):55–92, 2015.

[JLM85]   G. Jona-Lasinio and P. K. Mitter. On the stochastic quantization of field theory. *Comm. Math. Phys.*, 101(3):409–436, 1985.

[KS91]   Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[Kup16]   Antti Kupiainen. Renormalization group and stochastic PDEs. *Ann. Henri Poincaré*, 17(3):497–535, 2016.

[LMS19]   C. Landim, M. Mariani, and I. Seo. Dirichlet's and Thomson's principles for nonselfadjoint elliptic operators with application to non-reversible metastable diffusion processes. *Arch. Ration. Mech. Anal.*, 231(2):887–938, 2019.

[LP10]   Dorian Le Peutrec. Local WKB construction for Witten Laplacians on manifolds
            with boundary. *Anal. PDE*, 3(3):227–260, 2010.

[LP11]   Dorian Le Peutrec. Small eigenvalues of the Witten Laplacian acting on p-forms
            on a surface. *Asymptot. Anal.*, 73(4):187–201, 2011.

[LPNV13]   Dorian Le Peutrec, Francis Nier, and Claude Viterbo. Precise Arrhenius law for p-
            forms: the Witten Laplacian and Morse-Barannikov complex. *Ann. Henri Poincaré*,
            14(3):567–610, 2013.

[Mao97]   Xuerong Mao. *Stochastic differential equations and their applications*. Horwood
            Publishing Series in Mathematics & Applications. Horwood Publishing Limited,
            Chichester, 1997.

[McK69]   H. P. McKean, Jr. *Stochastic integrals*. Probability and Mathematical Statistics,
            No. 5. Academic Press, New York, 1969.

[MOS89]   Fabio Martinelli, Enzo Olivieri, and Elisabetta Scoppola. Small random pertur-
            bations of finite- and infinite-dimensional dynamical systems: unpredictability of
            exit times. *J. Statist. Phys.*, 55(3-4):477–504, 1989.

[MS01]   Robert S. Maier and D. L. Stein. Droplet nucleation and domain wall motion in a
            bounded interval. *Phys. Rev. Lett.*, 87:270601–1, 2001.

[MS03]   Robert S. Maier and D. L. Stein. The effects of weak spatiotemporal noise on a
            bistable one-dimensional system. In L. Schimanski-Geier, D. Abbott, A. Neimann,
            and C. Van den Broeck, editors, *Noise in complex systems and stochastic dynamics*,
            volume 5114 of *SPIE Proceedings Series*, pages 67–78. 2003.

[MT92]   Sean P. Meyn and R. L. Tweedie. Stability of Markovian processes. I. Criteria for
            discrete-time chains. *Adv. in Appl. Probab.*, 24(3):542–574, 1992.

[MT93a]   Sean P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Commu-
            nications and Control Engineering Series. Springer-Verlag London, Ltd., London,
            1993.

[MT93b]   Sean P. Meyn and R. L. Tweedie. Stability of Markovian processes. II. Continuous-
            time processes and sampled chains. *Adv. in Appl. Probab.*, 25(3):487–517, 1993.

[MT93c]   Sean P. Meyn and R. L. Tweedie. Stability of Markovian processes. III. Foster-
            Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.*, 25(3):518–
            548, 1993.

[MW17a]   Jean-Christophe Mourrat and Hendrik Weber. Convergence of the two-
            dimensional dynamic Ising-Kac model to Φ_2^4. *Comm. Pure Appl. Math.*, 70(4):717–
            812, 2017.

[MW17b]   Jean-Christophe Mourrat and Hendrik Weber. Global well-posedness of the dy-
            namic Φ^4 model in the plane. *Ann. Probab.*, 45(4):2398–2476, 2017.

[Nua06]   David Nualart. *The Malliavin calculus and related topics*, volume 1995. Springer,
            2006.

[Øks85]   Bernt Øksendal. *Stochastic differential equations*. Springer-Verlag, Berlin, 1985.

[OW19]   Felix Otto and Hendrik Weber. Quasilinear SPDEs via rough paths. *Archive for Rational Mechanics and Analysis*, 232(2):873–950, May 2019.

[PW81]   G. Parisi and Yong Shi Wu. Perturbation theory without gauge fixing. *Sci. Sinica*, 24(4):483–496, 1981.

[RZZ17a] Michael Röckner, Rongchan Zhu, and Xiangchan Zhu. Ergodicity for the stochastic quantization problems on the 2D-torus. *Comm. Math. Phys.*, 352(3):1061–1090, 2017.

[RZZ17b] Michael Röckner, Rongchan Zhu, and Xiangchan Zhu. Restricted Markov uniqueness for the stochastic quantization of $P(\Phi)_2$ and its applications. *J. Funct. Anal.*, 272(10):4263–4303, 2017.

[Sch66]   M. Schilder. Some asymptotic formulas for Wiener integrals. *Trans. Amer. Math. Soc.*, 125:63–85, 1966.

[Sim05]   Barry Simon. *Trace ideals and their applications*, volume 120 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, second edition, 2005.

[Slo12a] Martin Slowik. *Contributions to the potential theoretic approach to metastability with applications to the random Curie–Weiss–Potts model*. PhD thesis, Technische Universität Berlin, 2012.

[Slo12b] Martin Slowik. A note on variational representations of capacities for reversible and nonreversible Markov chains. Technical report, Technische Universität Berlin, 2012.

[Sow92]   Richard Sowers. Large deviations for the invariant measure of a reaction-diffusion equation with non-Gaussian perturbations. *Probab. Theory Related Fields*, 92(3):393–421, 1992.

[TW18a]   Pavlos Tsatsoulis and Hendrik Weber. Exponential loss of memory for the 2-dimensional Allen–Cahn equation with small noise. `arXiv:1808.04171`, 2018.

[TW18b]   Pavlos Tsatsoulis and Hendrik Weber. Spectral gap for the stochastic quantization equation on the 2-dimensional torus. *Ann. Inst. H. Poincaré Probab. Statist.*, 54(3):1204–1249, 08 2018. `arXiv:1609.08447`.

[ZZ15a]   Rongchan Zhu and Xiangchan Zhu. Three-dimensional Navier–Stokes equations driven by space–time white noise. *Journal of Differential Equations*, 259(9):4443–4508, 2015.

[ZZ15b]   Rongchan Zhu and Xiangchan Zhu. Three-dimensional Navier-Stokes equations driven by space-time white noise. *J. Differential Equations*, 259(9):4443–4508, 2015.

Nils Berglund
Institut Denis Poisson (IDP)
Université d'Orléans, Université de Tours, CNRS – UMR 7013
Bâtiment de Mathématiques, B.P. 6759
45067 Orléans Cedex 2, France
E-mail address: `nils.berglund@univ-orleans.fr`
`https://www.idpoisson.fr/berglund`