Simulation of Stochastic Partial Differential Equations and Stochastic Active Contours

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
der Universität Mannheim

vorgelegt von

Diplom-Mathematikerin
Annika Lang
aus Münster

Mannheim, 2007

Dekan: Prof. Dr. Matthias Krause, Universität Mannheim
Referent: Prof. Dr. Jürgen Potthoff, Universität Mannheim
Korreferent: Prof. Dr. Christoph Schnörr, Universität Mannheim

Tag der mündlichen Prüfung: 6. Dezember 2007

Abstract

This thesis discusses several aspects of the simulation of stochastic partial differential equations. First, two fast algorithms for the approximation of infinite dimensional Gaussian random fields with given covariance are introduced. Later Hilbert space-valued Wiener processes are constructed out of these random fields. A short introduction to infinite-dimensional stochastic analysis and stochastic differential equations is given. Furthermore different definitions of numerical stability for the discretization of stochastic partial differential equations are presented and the numerical stability of the heat equation with additive and multiplicative noise is explicitely computed using semigroup theory. Finally stochastic active contours are used for segmentation. This thesis generalizes work done by Juan et al. and does the simulation of different stochastic partial differential equations. The results are compared to equations without stochastics.

Zusammenfassung

In dieser Dissertation werden verschiedene Aspekte der Numerik von stochastischen partiellen Differentialgleichungen betrachtet. Dabei fasst man stochastische partielle Differentialgleichungen als hilbertraumwertige stochastische Differentialgleichungen im Sinne von Da Prato und Zabczyk auf. Das erste Kapitel beschäftigt sich zunächst mit der Approximation von hilbertraumwertigen Wiener-Prozessen. Es werden zwei Algorithmen entwickelt, die mittels FFT effizient und schnell Gaußsche Zufallsfelder mit vorgegebener Kovarianz erzeugen. Simulationen zeigen, dass die erzeugten Felder die gewünschten Eigenschaften haben. Außerdem werden zur Verdeutlichung Bilder der Kovarianzen und der erzeugten Zufallsfelder gezeigt. Im zweiten Kapitel werden anfangs die Theorie von hilbertraumwertigen stochastischen Prozessen und Integralen, sowie Differentialgleichungen mit additivem und multiplikativem Rauschen behandelt. Weiter wird beschrieben, wie man diese Gleichungen diskretisiert, und insbesondere, wie aus den im ersten Kapitel erzeugten Zufallsfeldern Wiener-Prozesse generiert werden können. Schließlich wird numerische Stabilität von Diskretisierungen von stochastischen partiellen Differentialgleichungen thematisiert. In der Literatur findet man bisher nur Untersuchungen bzgl. der Konvergenz von Diskretisierungen, die nicht alle äquivalent sind. Diese werden mit den damit erzielten Ergebnissen zusammengefasst, bevor eine Definition von numerischer Stabilität für stochastische Differentialgleichungen von Kloeden und Platen und eine andere aus dem Bereich der deterministischen partiellen Differentialgleichungen von Sewell für stochastische partielle Differentialgleichungen verallgemeinert werden. Die eine Definition erweist sich in späteren Rechnungen als ungeeignet, da bei Simulationen von errechnet stabilen Schemata Effekte auftreten, die man verhindern möchte. Die beiden Definitionen werden anhand unterschiedlicher Gleichungen überprüft, insbesondere wird eine mit einem expliziten Verfahren diskretisierte Wärmeleitungsgleichung mit additivem und multiplikativem Rauschen auf numerische Stabilität untersucht. Die numerischer Stabilität nach Sewell wird für diese Gleichungen mit Restriktionen an die Gitterfeinheit und die Kovarianz im Ort des Rauschens mit Hilfe von Halbgruppentheorie bewiesen. Im Anhang findet sich zudem eine Möglichkeit, die Gleichung mit additivem Rauschen mit Fouriermethoden und mit Hilfe von Fundamentallösungen auf numerische Stabilität zu untersuchen. Das dritte Kapitel widmet sich der Anwendung im Bereich

der Bildverarbeitung. Die Segmentierung von Bildern in zwei Teilbereiche wird durch implizit dargestellte Kurven berechnet. Dabei liegt dem Ganzen ein zu minimierendes Energiefunktional zugrunde, welches mit Hilfe von Variationsrechnung in eine Euler-Lagrange-Gleichung umgewandelt wird und durch Einführen einer künstlichen Zeit eine partielle Differentialgleichung ergibt. Beim Simulieren dieser Differentialgleichung tritt häufig das Problem auf, dass die Kurve in lokalen Minima stecken bleibt. Juan et al. haben vorgeschlagen, dies durch das Einführen von zusätzlichem endlichdimensionalem Rauschen zu überwinden. In dieser Arbeit wird der Vorschlag zu hilbertraumwertigem Rauschen verallgemeinert, welches auf die Theorie in den ersten beiden Kapiteln zurückzuführen ist. Welche Auswirkungen verschiedenes Rauschen auf Differentialgleichungen hat, wird mit Hilfe einer Wärmeleitungsgleichung mit gradientengekoppeltem Rauschen verdeutlicht. Die Arbeit schließt mit Segmentierungsergebnissen, die vor allem auch die Unterschiede von verschiedenem Rauschen verdeutlichen sollen.

Acknowledgements

First I want to thank my supervisor Prof. Dr. Jürgen Potthoff who supported me for the last years. Thanks for the many scientific discussions, hints, ideas, and especially for the motivation to learn new mathematical topics and not to be scared of those. He gave me a feeling of how to address different types of problems and how to "see" the real problems to be solved by doing small and relatively easy calculations, especially in the case of error estimations. In addition I would like to thank him for the ability to give talks at conferences and other universities. Next I thank my second supervisor Prof. Dr. Christoph Schnörr for the discussions concerning the application of SPDEs in computer vision. I also thank his PhD student Christian Gosch for the collaboration and discussions concerning level sets and segmentation. It was a great help for me to work together with him on simulations. I also thank both of them for the very helpful hints concerning literature on level sets and active contours. They helped me to overcome the problem that in the beginning it was hard for me to understand the approach to this topic from the perspective of people in computer vision. Furthermore I would like to thank all people from the same corridor who were bothered so many times with my stupid questions. It helped me a lot that there was always somebody who listened to me. Financially I have to thank Baden-Württemberg for the "Promotionsstipendium nach dem Landesgraduiertenförderungsgesetz" which supported me during the first period of my thesis. Finally I would like to thank everybody who supported me with the project of writing a PhD thesis, especially Jörg Zentgraf and Matthias Lang for proof-reading. The last thing to be mentioned here is that the work was done by Annika Niehage but she married before handing in this thesis.

Contents

List of Figures

Introduction

The simulation of stochastic partial differential equations (SPDEs) is a relatively new field which is used in different applications, e.g. in image processing and in finance. In this thesis we will focus on image processing, especially on active contours and segmentation. There are different approaches to the problem of segmentation which is still not solved satisfactorily in all cases. One idea is to use a closed curve and define inside and outside of the curve. An image energy depending on the curve is calculated. The goal is to place the curve on the image such that the energy is minimized. By calculus of variation the minimization problem can be transformed to an Euler–Lagrange equation which results in a partial differential equation (PDE) by parameterizing the descent direction with an artificial time. Thereby the problem of segmentation is transformed to a PDE that has to be solved. This PDE describes the evolution of the curve but this evolution might get stuck in a "false" local minimum. To overcome this problem, Juan, Keriven, and Postelnicu suggested in "Stochastic Motion and the Level Set Method in Computer Vision: *Stochastic Active Contours*" [47] to add noise to the PDEs resulting from variational problems and deterministic approaches made so far. The authors base their theory on recent work by Lions and Souganidis concerning viscosity solutions of SPDEs with finite-dimensional stochastic Stratonovich integrals [57, 58, 60, 61]. Juan et al. simulate the following SPDE

$$du(t,x) = F(D^2u(t,x), Du(t,x), x, t)\,dt + |Du(t,x)| \sum_{i=1}^{m} \phi_i(x) \circ dW_i(t),$$

where $W(t) = (W_1(t), \ldots, W_m(t))$ is an m-dimensional Brownian motion, F is the diffusion coefficient, and the elements $\phi_i : \mathbb{R}^N \to \mathbb{R}$ are smooth functions with compact support. The authors choose an equally distributed grid $\{x_i, i = 1, \ldots, n\}$ on the region of simulation. They set

$$\phi_i(x) = \phi(x - x_i),$$

where ϕ is some convenient regular function, and for all i, j it holds that $\phi_i(x_j) = \delta_{ij}$. In this thesis this approach is generalized to infinite-dimensional Brownian motions, also called Q-Wiener processes, where Q is the operator describing the space correlation. The theory is based on the work of Da Prato and Zabczyk [16] on infinite-dimensional stochastic equations. In order to do fast simulations of these SPDEs which can be understood as Hilbert space-valued stochastic differential equations, the question of how to model infinite-dimensional noise arises. It turns out that this process can only be approximated by a finite-dimensional random field because a computer is only able to do simulations on a finite number of points. In our case these points are arranged on a grid because the theory is supposed to be applied to images which consist of a finite grid of pixels. Therefore the infinite-dimensional processes are approximated by Gaussian random fields on a grid and its dimension is the number of grid points while Juan et al. take a smaller subgrid for the noise model which implies a smaller noise dimension. The fast algorithms that are developed in this thesis will generate Gaussian random fields with

mean zero and given covariance which is one of the input parameters of the algorithms. These random fields ensure that segmentation does not take too much longer than simulating the corresponding deterministic equations.

Once the implementation of the Gaussian random fields and of different SPDEs — especially those usable for segmentation — is done, it turns out that the discretized time step size $\triangle t$ has to be pretty small in comparison to the space grid size $\triangle x$. Otherwise the simulations will be numerically instable, i.e. approximation errors will dominate the calculations and the simulation results are useless. There are different ways to overcome this problem. One way to ensure stability is to implement a CFL condition [72, 84] in the following way: The condition requires that the curve, also called front, should not move more than one grid point per simulation step. Therefore using an explicit scheme, the time step size is chosen such that the increment of the current simulation step is not larger than one. Similarly this can be done for SPDEs but one has to keep in mind that the stochastic increment is scaled with $\sqrt{\triangle t}$ and not like the deterministic part with $\triangle t$. This approach of forcing the simulation scheme to be numerically stable will always work but the question arises of how to prove numerical stability mathematically. There exist different definitions of numerical stability that all try to describe the phenomenon that approximation errors spread and finally dominate the simulation. A summary of the work done so far for SPDEs is given in this thesis. So far there does not exist a method to check numerical stability for nonlinear SPDEs especially if the noise is coupled to the size of the gradient of the function. To give an idea of how calculations can look like, these are explicitely done for the case of the heat equation perturbed with additive and multiplicative noise. It will be proved that two different definitions of numerical stability lead to different results. So finally we will see that one of the definitions seems to be useless in the context of SPDEs. It should be possible to get a result that yields for a given linear SPDE bounds on the discretization step sizes in space and time, and accuracy conditions on the approximation of the operators such that the resulting approximation scheme is numerically stable.

This thesis is structured in the following way and covers the topics presented next. Chapter 1 introduces Gaussian random fields. We give heuristics that use Dirac distributions and give an idea of how to derive efficient methods for generating special random fields. Therefore let C be a positive continuous semi-definite function on $\mathbb{R}^d$. According to Bochner's theorem C can be written as the Fourier transform of a positive measure μ_C on $\mathbb{R}^d$. For the practical purposes addressed in this thesis, there is no loss of generality if we assume that μ_C has a Lebesgue density denoted by γ such that C can be written as

$$C(x) = \int_{\mathbb{R}^d} e^{-2\pi i(p,x)} \gamma(p)\, dp,$$

where γ is an even, positive function and $(\cdot,\cdot)$ denotes the inner product on $\mathbb{R}^d$. Then a Gaussian random field φ with covariance C can be written as

$$\varphi(x) = (\mathcal{F}^{-1}\gamma^{1/2}\mathcal{F}W)(x),$$

where W is white noise and $\mathcal{F}$ denotes the Fourier transform. The expression $\mathcal{F}W$ can be rewritten without Fourier transform such that one Fourier transform can be omitted for the calculation of φ, i.e.

$$\varphi(x) = (\mathcal{F}^{-1}\gamma^{1/2}(\pi^+ W + i\pi^- W))(x),$$

where π^+ and π^- denote the projection onto the even and odd part of W. These two formulae lead to two algorithms for fast random field generation via FFT with independent normal distributed random numbers on input. Using a mathematically correct approach to generalized random fields and white noise, the same algorithms are achieved. In that case a generalized random field φ is defined as a real- or complex-valued mapping from the Schwartz space on $\mathbb{R}^d$ to the real- or complex-valued random variables, respectively. A generalized random field W on $\mathbb{R}^d$ is then called white noise, if its characteristic function is given by

$$\Xi_{\mathrm{WN}}(f) = e^{-\frac{1}{2}(f,f)_{L^2(\mathbb{R}^d)}}.$$

Moreover operations on generalized random fields are introduced and used for the construction of stationary centered Gaussian random fields φ defined by their covariance

$$\mathrm{Cov}\left(\varphi(f), \varphi(g)\right) = (f, Cg)_{L^2(\mathbb{R}^d)},$$

where C is assumed to be given by an integral kernel K which is supposed to have a density γ. Then

$$\psi(f) = (\mathcal{F}^{-1}\gamma^{1/2}\mathcal{F}W)(f)$$

has the same properties as φ. Another realization of φ is given by

$$\chi(f) = (\mathcal{F}^{-1}\gamma^{1/2}(\pi_+ W + i\pi_- W))(f),$$

where similarly to π^+ and π^-, π_+ and π_- denote the projection onto the even and the odd part but this time in the language of generalized random fields. Next two efficient algorithms are formulated explicitly. Afterwards a class of covariance functions is presented that is symmetric under rotation and that has exponential decay, which might be desirable in order to have only local perturbations. At the end of the chapter the implementations of the algorithms are tested and the resulting random fields are visualized.

In Chapter 2 SPDEs are introduced as Hilbert space-valued SDEs and their simulations including numerical stability are discussed. A short summary about the theory of Hilbert space-valued stochastic processes and integrals is done in the sense of Da Prato and Zabczyk [16]. Furthermore existence and uniqueness results are given for linear differential equations with additive noise of the form

$$\begin{cases} dX(t) = [AX(t) + f(t)]\,dt + B\,dW(t), \\ X(0) = \xi, \end{cases}$$

where $A : D(A) \subset H \to H$ and $B : U \to H$ are linear operators, f is an H-valued stochastic process. We assume that the deterministic Cauchy problem

$$\begin{cases} u'(t) = Au(t), \\ u(0) = x \in H \end{cases}$$

is uniformly well posed and that B is bounded, i.e. A generates a strongly continuous semigroup $S(\cdot)$ in H and $B \in L(U, H)$. Moreover we require that f is a predictable process with Bochner integrable trajectories on an arbitrary finite interval $[0, T]$ and that ξ is $\mathcal{F}_0$-measurable. Then

the equation has exactly one weak solution and under further assumptions also a strong solution exists. A lemma states under which conditions A^k can be applied to the solution of the SPDE. Next a similar differential equation with multiplicative noise is examined. Let

$$\begin{cases} dX(t) = [AX(t) + f(t)]\,dt + B(X(t))\,dW(t), \\ X(0) = \xi, \end{cases}$$

be given on a finite time interval $[0, T]$, where again $A : D(A) \subset H \to H$ is the infinitesimal generator of a strongly continuous semigroup $S(\cdot)$, ξ is an H-valued $\mathcal{F}_0$-measurable random variable, f is a predictable process with locally integrable trajectories and $B : D(B) \to L_2^0$ is a linear operator. For $f \equiv 0$ this SPDE has a unique strong solution that is also a weak solution and the unique mild solution. This concludes the section about the theory of SPDEs and we continue with numerical discretizations. For the deterministic part we give a short summary on existing discretization methods and introduce the approximation of the Laplace operator via finite differences. Furthermore it is discussed how to approximate a stochastic process with the aid of the Gaussian random fields introduced in the first chapter. The main section of the second chapter is about numerical stability of SPDEs. This section starts with a summary of the literature that exists about convergence of approximations. Papers by Hausenblas [39], Hofmann et al. [44], Tocino [90] on SDEs and by Gyöngy [35], Hausenblas and Marchis [40], Lord and Rougemont [63], and Shardlow [87] on SPDEs are reviewed. Then a definition of numerical stability for SDEs by Kloeden and Platen [52] is generalized to SPDEs and checked on different SDEs as well as on the heat equation with additive and multiplicative noise. It turns out that this definition does not seem to catch all effects that should be omitted because the heat equation is always stochastically numerically stable, independent of the choice of $\triangle x$ and $\triangle t$, but the simulations show effects that one would call numerically instable. Finally another definition of numerical stability for ODEs and PDEs by Sewell [85] is generalized to SPDEs and checked on the same equations. For the heat equation with additive noise

$$du(t) = \frac{1}{2}\Delta u(t)\,dt + \sigma(t)\,dW(t),$$

the discretization of the Laplacian with finite differences is numerically stable if $\triangle t \leq (\triangle x)^2$, where the noise is approximated by the Gaussian random fields introduced in Chapter 1. For the corresponding equation with multiplicative noise

$$du(t) = \frac{1}{2}\Delta u(t)\,dt + u(t)\,dW(t)$$

stability is achieved if $\triangle t < (\triangle x)^2$. Therefore the bound is in this case a bit more restrictive than the one for the heat equation with additive noise and the one known from deterministic stability analysis.

Finally Chapter 3 is about the application of the results of the previous chapters to segmentation. First zero level sets are introduced and calculus of variations is described. Then two different energy functionals and their corresponding Euler–Lagrange equations used in deterministic active contours are derived. The first was introduced by Chan and Vese [10] and is

based on work by Mumford and Shah [68]. Their energy functional is given by

$$\begin{aligned} J(c_1, c_2, \phi) := \mu \int_\Lambda \delta(\phi(x,y)) \, |\nabla \phi(x,y)| \, dx \, dy \\ + \ \nu \int_\Lambda H(\phi(x,y)) \, dx \, dy \\ + \ \lambda_1 \int_\Lambda |u_0(x,y) - c_1|^2 \, H(\phi(x,y)) \, dx \, dy \\ + \ \lambda_2 \int_\Lambda |u_0(x,y) - c_2|^2 \, (1 - H(\phi(x,y))) \, dx \, dy. \end{aligned}$$

where Λ is the area, $\mu, \nu, \lambda_1, \lambda_2 > 0$ are weighting coefficients, H denotes the Heavyside function, c_1 is the average gray value inside the curve, c_2 is the average gray value outside the curve, and $u_0(x, y)$ is the gray value of the pixel (x, y). The Euler–Lagrange equation is given by

$$\delta_\varepsilon(\phi) \, (\mu \, \kappa(\phi) - \nu - \lambda_1 \, |u_0 - c_1|^2 + \lambda_2 \, |u_0 - c_2|^2) = 0,$$

where $\kappa(\phi)$ denotes the mean curvature of ϕ and δ_ε is a regularized delta function. The second approach by Rousson and Deriche [83] models each region by a Gaussian distribution of unknown mean and variance. The energy to be minimized is given by

$$E(\Gamma, \mu_1, Q_1, \mu_2, Q_2) = \int_{in(\Gamma)} e_1(x) \ dx + \int_{out(\Gamma)} e_2(x) \ dx + \nu L(\Gamma),$$

where Γ is the curve, integration is done inside and outside the curve, $L(\Gamma)$ is the length of the curve, and $e_i(x) = -\log p_{\mu_i Q_i}(I(x))$ with

$$p_{\mu_i Q_i}(I(x)) = ((2\pi)^m \det(Q_i))^{-1/2} e^{-\frac{1}{2}\langle I(x) - \mu_i, Q_i^{-1}(I(x) - \mu_i)\rangle}$$

being the m-dimensional Gaussian density for a given value $I(x)$ with respect to the hypothesis (μ_i, Q_i). The parameters (μ_i, Q_i), estimated from the pixel currently inside and outside Γ, are functions of Γ. The corresponding Euler–Lagrange equation is given by

$$e_2(x) - e_1(x) + \nu\kappa(u) = 0.$$

Then randomness is added to the resulting PDEs. This was suggested by Juan et al. [47] with finite-dimensional noise and will be modified to Hilbert space-valued SPDEs and infinite-dimensional noise introduced in the second chapter of this thesis and approximated with Gaussian random fields presented in Chapter 1. Equations of the form

$$du(t,x) = F(D^2u, Du, x, t) \, dt + |Du(t,x)| \, dW(t,x)$$

are simulated, where F is given by one of the two presented deterministic minimization problems. In order to get a feeling how different types of noise and covariances affect the simulations, a stochastic heat equation and stochastic motion by mean curvature are implemented. Finally different examples of image segmentation are given and the influence of the noise on the result is tested.

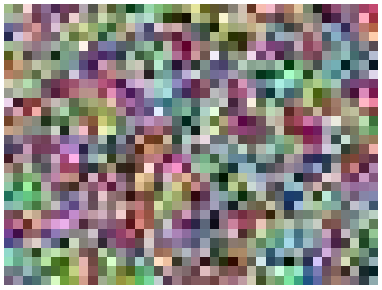
The Appendix A gives an introduction to strongly continuous semigroups. In Appendix B a stability proof of the heat equation with additive noise in the sense of Sewell using Fourier methods and fundamental solutions is given. It turns out that this approach is a lot more laborious than using semigroups as done in the main part of this thesis. Appendix C summarizes different representations of curvature, e.g. of a parameterized curve and of an embedded curve, and their connections.

1. Gaussian Random Fields

This chapter gives an introduction to Gaussian random fields (GRFs) and leads finally to fast algorithms for the generation of random field with given spacial correlation using white noise and FFT. There are two major parts in this chapter. The first part introduces GRFs in the language of applications with delta functions. In the second part another approach to the same topic is given. Both approaches lead to the same algorithms presented afterwards. Finally some possible and in the context of this thesis useful covariance functions are studied and simulations as well as statistical tests on a two-dimensional square are done.

1.1. Introduction

Assume that φ is a real-valued stationary Gaussian random field (GRF) on $\mathbb{R}^d$, i.e. let $(X_i, i \in \mathbb{R}^d)$ be a family of equally distributed random variables on a probability space $(\Omega, \mathcal{F}, P)$ that satisfies for any subset $\{i_1, \ldots, i_n \in \mathbb{R}^d\}$ and $\alpha_k \in \mathbb{R}$ that the random variable $\sum_{k=1}^{n} \alpha_k X_{i_k}$ is normal distributed. Then its only statistical parameters are its mean $m = \mathbb{E}(\varphi(x)) = \mathbb{E}(\varphi(0))$, $x \in \mathbb{R}^d$, and its covariance $C(x) = \mathbb{E}(\varphi(0)\varphi(x)) - m^2 = \mathbb{E}(\varphi(y)\varphi(x+y)) - m^2$, $x, y \in \mathbb{R}^d$. Without loss of generality we will assume that $m = 0$. For simulations on a computer we will later assume GRFs on a finite grid, i.e. discrete random fields. Figure 1.1 shows two two-dimensional examples of how discrete GRFs in applications may look like.



(a) picture with random pixels,

81	105	255	111	33	143	136	101
120	231	85	140	118	180	134	93
99	134	180	83	75	168	166	153
57	130	76	100	217	156	107	69
70	183	118	110	75	146	135	129
92	141	26	85	80	71	133	0
95	120	68	97	56	124	76	181
79	69	169	36	71	61	160	145

(b) random number array,

Figure 1.1.: Example of a discrete GRF.

1.2. Heuristics

Our goal is to construct GRFs with given Covariance C. For the one-dimensional case, the construction ideas presented in this section are given in [81]. We first make some observations and assumptions concerning C. We note that C is a positive semi-definite function and assume that it is continuous. Then Bochner's theorem [8] states that C is the Fourier transform of a

positive measure μ_C on $\mathbb{R}^d$, i.e. C can be written as

$$C(x,y) = \int e^{-2\pi i(p,x-y)}\, d\mu_C(p), \qquad x \in \mathbb{R}^d$$

where $(\cdot,\cdot)$ denotes the Euclidean inner product on $\mathbb{R}^d$. We observe that C is even because of the stationarity and therefore so is μ_C. For the practical purposes addressed in this thesis, there is no loss of generality if we assume that μ_C has a Lebesgue density denoted by γ which also has to be an even and positive function. Then C can be written in the following way:

$$C(x,y) = \int_{\mathbb{R}^d} e^{-2\pi i(p,x-y)}\gamma(p)\, dp, \qquad x \in \mathbb{R}^d.$$

Let W be a Gaussian white noise random field on $\mathbb{R}^d$, i.e. informally, W is a centered Gaussian family $\{W(x), x \in \mathbb{R}^d\}$ with covariance $\mathbb{E}(W(x)W(y)) = \delta(x-y)$, $x, y \in \mathbb{R}^d$. Set

$$\varphi(x) = (\mathcal{F}^{-1}\gamma^{1/2}\mathcal{F}W)(x), \qquad x \in \mathbb{R}^d, \tag{1.1}$$

where $\mathcal{F}$ denotes the d-dimensional Fourier transform and $\mathcal{F}^{-1}$ denotes its inverse. Then, since W is centered Gaussian, so is φ. Next we will calculate the covariance of φ.

$$\begin{aligned}
&\mathbb{E}(\varphi(x)\varphi(y)) \\
&\quad = \iiiint e^{-2\pi i((p,x)+(q,y))}\gamma(p)^{1/2}\gamma(q)^{1/2}e^{2\pi i((p,x')+(q,y'))}\mathbb{E}(W(x')W(y'))\, dx'\, dy'\, dp\, dq \\
&\quad = \iint e^{-2\pi i((p,x)+(q,y))}\gamma(p)^{1/2}\gamma(q)^{1/2}\int e^{2\pi i(p+q,x')}\, dx'\, dp\, dq \\
&\quad = \int e^{-2\pi i(p,x-y)}\gamma(p)\, dp \\
&\quad = C(x,y).
\end{aligned}$$

Thus Equation (1.1) provides our first algorithm to generate samples of φ with given covariance C. This algorithm is presented in Section 1.4.

Next we will work out a more efficient algorithm by replacing one Fourier transform with a faster operation. The goal is to construct $\mathcal{F}W$ directly, i.e. a complex GRF with the same properties as $\mathcal{F}W$ has to be constructed. We consider the complex-valued random field $\mathcal{F}W$:

$$\mathcal{F}W(p) = \int_{\mathbb{R}^d} e^{2\pi i(p,x)}W(x)\, dx$$

This is obviously centered Gaussian and the covariance is given by

$$\begin{aligned}
\mathbb{E}(\mathcal{F}W(p)\overline{\mathcal{F}W(q)}) &= \iint e^{2\pi i((p,x)-(q,y))}\mathbb{E}(W(x)W(y))\, dx\, dy \\
&= \int e^{2\pi i(p-q,x)}\, dx \\
&= \delta(p-q),
\end{aligned}$$

and similarly

$$\mathbb{E}(\mathcal{F}W(p)\mathcal{F}W(q)) = \delta(p+q).$$

The following lemma gives a direct construction of $\mathcal{F}W$.

Lemma 1.1. *Let V be a complex-valued random field defined by*

$$\begin{aligned}\operatorname{Re} V &= \pi^+ W,\\ \operatorname{Im} V &= \pi^- W,\end{aligned}$$

where

$$\begin{aligned}\pi^+ W(p) &= \tfrac{1}{2}\left(W(p) + W(-p)\right),\\ \pi^- W(p) &= \tfrac{1}{2}\left(W(p) - W(-p)\right).\end{aligned}$$

Then V and $\mathcal{F}W$ have the same law.

The following argumentation explains the claim. As both random fields are centered Gaussian, one only has to show that they have the same covariance. First we observe that $\pi^+ W$ and $\pi^- W$ are uncorrelated:

$$\begin{aligned}&\mathbb{E}(\pi^+ W(p)\pi^- W(q))\\ &\quad= \tfrac{1}{4}\,\mathbb{E}((W(p) + W(-p))(W(q) - W(-q)))\\ &\quad= \tfrac{1}{4}\left(\mathbb{E}(W(p)W(q)) + \mathbb{E}(W(-p)W(q)) - \mathbb{E}(W(p)W(-q)) - \mathbb{E}(W(-p)W(-q))\right)\\ &\quad= \tfrac{1}{4}\left(\delta(p-q) + \delta(p+q) - \delta(p+q) - \delta(p-q)\right) = 0.\end{aligned}$$

Next we use this observation to see that

$$\begin{aligned}\mathbb{E}(V(p)V(q)) &= \mathbb{E}(\pi^+ W(p)\pi^+ W(q)) - \mathbb{E}(\pi^- W(p)\pi^- W(q))\\ &= \tfrac{1}{4}(2\delta(p-q) + 2\delta(p+q) - (2\delta(p-q) - 2\delta(p+q)))\\ &= \delta(p+q),\end{aligned}$$

and similarly

$$\mathbb{E}(V(p)\overline{V(q)}) = \delta(p-q).$$

.

Therefore we can write φ in the following way:

$$\varphi(x) = (\mathcal{F}^{-1}\gamma^{1/2}(\pi^+ W + i\pi^- W))(x). \tag{1.2}$$

From the definition of V it follows that $V(-x) = \overline{V(x)}$. Since $\operatorname{Re} V(x)$ and $\operatorname{Im} V(x)$ are uncorrelated and Gaussian which implies that they are independent, as shown before, we set

$$\begin{aligned}V(x) &= U_x + iV_x,\\ V(-x) &= U_x - iV_x,\end{aligned}$$

where U_x and V_x are independent centered Gaussian random variables. We know that $W(x)$ is white noise for all x. To calculate the scaling factor of U_x and V_x, we have for $x \neq 0$

$$\mathbb{E}(U_x^2) = \mathbb{E}((\tfrac{1}{2}(W(x) + W(-x)))^2) = \tfrac{1}{4}(\mathbb{E}(W(x)^2) + 2\,\mathbb{E}(W(x)W(-x)) + \mathbb{E}(W(-x)^2)) = \tfrac{1}{2}\delta(0)$$

and for $x = 0$

$$\mathbb{E}(U_0^2) = \delta(0).$$

Similarly we calculate

$$\mathbb{E}(V_x^2) = \frac{1}{2}\delta(0), \qquad x \neq 0$$
$$\mathbb{E}(V_0^2) = \delta(0).$$

In the case of a d-dimensional cube with periodic boundary conditions, the corners also satisfy $W(x) = W(-x)$. Therefore similar to $x = 0$ for all corners c of the cube U_c and V_c have a scaling factor of 1. The corresponding algorithm implementing Equation (1.2) in an efficient way and using the observations made above can be found in Section 1.4.

1.3. Mathematical Approach

This section is devoted to give a mathematically rigorous approach to random fields and white noise. It avoids the problems with delta functions and is based on distribution theory. For further details, the reader is referred to [27], [42], [93], and [98].

1.3.1. Generalized Random Fields

Consider the real Hilbert space $L^2(\mathbb{R}^d)$ with inner product $(\cdot, \cdot)$. Let $\alpha = (\alpha_1, \ldots, \alpha_d)$ and $\beta = (\beta_1, \ldots, \beta_d)$ denote multiindices in $\mathbb{N}_0^d$, and set

$$D^\beta = \partial_1^{\beta_1} \cdots \partial_d^{\beta_d},$$

where $\partial_j, j = 1, 2, \ldots, d$ is the partial derivative on $\mathbb{R}^d$ with respect to the j-th coordinate. Similarly, if $u = (u_1, \ldots, u_d) \in \mathbb{R}^d$, then u^α is the abbreviation for

$$u^\alpha := u_1^{\alpha_1} \cdots u_d^{\alpha_d}.$$

Furthermore define the seminorm

$$\| \varphi \|_{\alpha,\beta} := \sup_{u \in \mathbb{R}^d} \left| u^\alpha D^\beta \varphi(u) \right|. \tag{1.3}$$

Definition 1.2. The *Schwartz space* $\mathcal{S}(\mathbb{R}^d)$ *of smooth functions of rapid decrease* is the set of all (real- or complex-valued) functions in $C^\infty(\mathbb{R}^d)$, having finite seminorm (1.3) for every $\alpha, \beta \in \mathbb{N}_0^d$. $\mathcal{S}(\mathbb{R}^d)$ is topologized by the system of seminorms (1.3). The *Schwartz space* $\mathcal{S}'(\mathbb{R}^d)$ *of tempered distributions* is the dual of $\mathcal{S}(\mathbb{R}^d)$.

We remark that $\mathcal{S}(\mathbb{R}^d)$ is a Fréchet space, i.e. a quasi-normed complete linear space. Moreover the chain rule and the triangle inequality imply that it is an algebra under pointwise multiplication.

The above defined topology on $\mathcal{S}(\mathbb{R}^d)$ is known to be equivalent to the one described in the following (see e.g. [78]). Let H be the Hamiltonian of the harmonic oscillator. The operator H is the unique self-adjoint operator on $L^2(\mathbb{R}^d)$, which on $\mathcal{S}(\mathbb{R}^d)$ is given by

$$H\varphi(u) = (-\Delta + (|u|^2 + 1))\varphi(u), \quad \varphi \in \mathcal{S}(\mathbb{R}^d),$$

where Δ is the Laplacian on $\mathbb{R}^d$. Define a system of seminorms $\{\| \cdot \|_{2,p}, p \in \mathbb{N}_0\}$ on $\mathcal{S}(\mathbb{R}^d)$ by setting $\| \varphi \|_{2,p} = \| H^p \varphi \|_2$, were the last norm is the one of $L^2(\mathbb{R}^d)$. Then this system is equivalent to the initial system of seminorms on $\mathcal{S}(\mathbb{R}^d)$ and therefore it defines the same topology on $\mathcal{S}(\mathbb{R}^d)$.

Given $p \in \mathbb{N}_0$, we denote by $\mathcal{S}_p(\mathbb{R}^d)$ the Hilbert space which is the completion of $\mathcal{S}(\mathbb{R}^d)$ (more precisely of the equivalence class of $\mathcal{S}(\mathbb{R}^d)$ in $L^2(\mathbb{R}^d)$) under the norm $\| \cdot \|_{2,p}$. Denote by $\check{\mathcal{S}}(\mathbb{R}^d)$ the projective limit of the chain

$$L^2(\mathbb{R}^d) \supset \mathcal{S}_1(\mathbb{R}^d) \supset \cdots \supset \mathcal{S}_p(\mathbb{R}^d) \supset \cdots,$$

i.e. we put $\check{\mathcal{S}}(\mathbb{R}^d) = \cap_p \mathcal{S}_p(\mathbb{R}^d)$ as a set, and equip this space with the topology generated by the neighborhoods of zero which are determined by the choice of $p \in \mathbb{N}_0$, $\varepsilon > 0$, and are given by $U_{p,\varepsilon} = \{\check{\varphi} \in \check{\mathcal{S}}(\mathbb{R}^d), \| \check{\varphi} \|_{2,p} < \varepsilon\}$. (Note that the system of seminorms $\{\| \cdot \|_{2,p}, p \in \mathbb{N}_0\}$ is ordered.)

Then $\check{\mathcal{S}}(\mathbb{R}^d)$ and $\mathcal{S}(\mathbb{R}^d)$ are in one-to-one correspondence: every equivalence class $\check{\varphi} \in \check{\mathcal{S}}(\mathbb{R}^d)$ contains exactly one representative $\varphi \in \mathcal{S}(\mathbb{R}^d)$, and to each $\varphi \in \mathcal{S}(\mathbb{R}^d)$ corresponds exactly one class in $\check{\mathcal{S}}(\mathbb{R}^d)$. Moreover, since on $\mathcal{S}(\mathbb{R}^d)$ the topologies defined by both systems of seminorms above are equivalent, we see that this correspondence is a topological isomorphism. Unless there is danger of confusion, we shall identify these two spaces.

It is well-known that the spectrum of H is given by the set

$$\{2(n_1 + \cdots + n_d) + d + 1,\, n_1, \ldots, n_d \in \mathbb{N}\},$$

e.g. [78] or any introductory book on quantum mechanics. Thus $H^{-(d+1)}$ is a trace class operator on $L^2(\mathbb{R}^d)$, and it follows that $\mathcal{S}(\mathbb{R}^d)$ is a nuclear countably Hilbert space.

Furthermore we may identify $L^2(\mathbb{R}^d)$ with its dual by Riesz' representation theorem [98, Thm. III.6], and this entails that we can consider $L^2(\mathbb{R}^d)$ as a subspace of regular elements in $\mathcal{S}'(\mathbb{R}^d)$, the dual of $L^2(\mathbb{R}^d)$. Thus we have

$$\mathcal{S}(\mathbb{R}^d) \subset L^2(\mathbb{R}^d) \subset \mathcal{S}'(\mathbb{R}^d).$$

We shall equip $\mathcal{S}'(\mathbb{R}^d)$ with its weak topology. The dual pairing between $\mathcal{S}'(\mathbb{R}^d)$ and $\mathcal{S}(\mathbb{R}^d)$ will be denoted by $\langle \cdot, \cdot \rangle$. For $f \in L^2(\mathbb{R}^d)$, $g \in \mathcal{S}(\mathbb{R}^d)$, we have

$$\langle f, g \rangle = (f, g)\,.$$

Consider now the complexifications of $L^2(\mathbb{R}^d)$ and $\mathcal{S}(\mathbb{R}^d)$

$$L^2_{\mathbb{C}}(\mathbb{R}^d) = L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d),$$
$$\mathcal{S}_{\mathbb{C}}(\mathbb{R}^d) = \mathcal{S}(\mathbb{R}^d) \times \mathcal{S}(\mathbb{R}^d),$$

equipped with the product topology and a pointwise multiplication of functions which follows the pattern of multiplication of complex numbers. If $f \in \mathcal{S}_{\mathbb{C}}(\mathbb{R}^d)$, we write $f_1 = \operatorname{Re} f$, $f_2 = \operatorname{Im} f$, and $f = \operatorname{Re} f + i \operatorname{Im} f$, and similarly for $f \in L^2_{\mathbb{C}}(\mathbb{R}^d)$. The inner product on $L^2_{\mathbb{C}}(\mathbb{R}^d)$ is defined such that it is conjugate linear in the first and linear in the second argument, that is for $f, g \in L^2_{\mathbb{C}}(\mathbb{R}^d)$:

$$(f, g) = (\operatorname{Re} f, \operatorname{Re} g) + i\,(\operatorname{Re} f, \operatorname{Im} g) - i\,(\operatorname{Im} f, \operatorname{Re} g) + (\operatorname{Im} f, \operatorname{Im} g)\,, \tag{1.4}$$

where the inner products on the right-hand side are those of $L^2(\mathbb{R}^d)$. $\mathcal{S}'_{\mathbb{C}}(\mathbb{R}^d)$ inherits a complex structure in a natural way, compatible with $L^2_{\mathbb{C}}(\mathbb{R}^d)$ as a subspace of complex, regular distributions and (1.4) as follows. By construction, the projection

$$\begin{aligned} \mathcal{S}_{\mathbb{C}}(\mathbb{R}^d) &\to \mathcal{S}(\mathbb{R}^d) \\ f &\mapsto \operatorname{Re} f \end{aligned}$$

is $\mathbb{R}$-linear and continuous, hence

$$\begin{aligned} T : \mathcal{S}(\mathbb{R}^d) &\to \mathbb{C} \\ f &\mapsto \langle T, f\rangle \end{aligned}$$

is well-defined, $\mathbb{R}$-linear and continuous. Therefore $\operatorname{Re}\langle T, \cdot\rangle$ and $\operatorname{Im}\langle T, \cdot\rangle$ define two elements in $\mathcal{S}'(\mathbb{R}^d)$, and we set

$$\langle \operatorname{Re} T, f\rangle := \operatorname{Re}\langle T, f\rangle, \qquad f \in \mathcal{S}(\mathbb{R}^d),$$

and

$$\langle \operatorname{Im} T, f\rangle := -\operatorname{Im}\langle T, f\rangle, \qquad f \in \mathcal{S}(\mathbb{R}^d). \tag{1.5}$$

Then for $f \in \mathcal{S}_{\mathbb{C}}(\mathbb{R}^d)$,

$$\langle T, f\rangle = \langle \operatorname{Re} T, \operatorname{Re} f\rangle + i\langle \operatorname{Re} T, \operatorname{Im} f\rangle - i\langle \operatorname{Im} T, \operatorname{Re} f\rangle + \langle \operatorname{Im} T, \operatorname{Im} f\rangle,$$

i.e. the dual pairing $\langle \cdot, \cdot\rangle$ between $\mathcal{S}'_{\mathbb{C}}(\mathbb{R}^d)$ and $\mathcal{S}_{\mathbb{C}}(\mathbb{R}^d)$ is conjugate linear in its first and linear in its second argument.

Let $(\Omega, \mathcal{F}, P)$ be a probability space, and denote by $L^0(P)$ the space of real- or complex-valued random variables. If it is necessary to distinguish these cases, we shall write $L^0_{\mathbb{R}}(P)$ and $L^0_{\mathbb{C}}(P)$.

Definition 1.3. A *generalized random field* on $\mathbb{R}^d$ is a $\mathbb{K}$-linear mapping

$$\varphi : \mathcal{S}_{\mathbb{K}}(\mathbb{R}^d) \to L^0_{\mathbb{K}}(P),$$

where $\mathbb{K}$ stands either for $\mathbb{R}$ or for $\mathbb{C}$.

In the following we consider only $\mathbb{K} = \mathbb{R}$. A concrete realization of certain generalized random fields can be constructed as follows.

Let σ_w denote the σ-algebra on $\mathcal{S}'(\mathbb{R}^d)$ generated by the weak topology. The following theorem is a generalization of Bochner's theorem [8] and states that characteristic functions can be rewritten as Fourier transforms. This will be important in the algorithms for efficient random field generation with given covariance.

Theorem 1.4 (Minlos' theorem [27, 42])**.** *Let Ξ be a characteristic function on $\mathcal{S}(\mathbb{R}^d)$, i.e.*

1. *Ξ is continuous in $\mathcal{S}(\mathbb{R}^d)$,*
2. *Ξ is positive definite,*
3. *$\Xi(0) = 1$.*

Then there exists a unique probability measure μ on $(\mathcal{S}'(\mathbb{R}^d), \sigma_w)$, such that for all $f \in \mathcal{S}(\mathbb{R}^d)$

$$\int_{\mathcal{S}'(\mathbb{R}^d)} e^{i\langle \omega, f\rangle}\, d\mu(\omega) = \Xi(f), \tag{1.6}$$

i.e. $\Xi(f)$ is the Fourier transform of a countably additive positive normalized measure.

Using this theorem, we set

$$\begin{aligned} \varphi : \mathcal{S}(\mathbb{R}^d) &\to L^0(\mathcal{S}'(\mathbb{R}^d), \sigma_w, \mu) \\ f &\mapsto \varphi(f)(\omega) := \langle \omega, f\rangle \end{aligned} \tag{1.7}$$

and obtain a generalized random field φ with characteristic function Ξ:

$$\mathbb{E}\big(e^{i\lambda\varphi(f)}\big) = \int_{\mathcal{S}'(\mathbb{R}^d)} e^{i\lambda\varphi(f)(\omega)}\, d\mu(\omega) = \int_{\mathcal{S}'(\mathbb{R}^d)} e^{i\langle \omega, \lambda f\rangle}\, d\mu(\omega) = \Xi(\lambda f), \quad \lambda \in \mathbb{R}.$$

Corollary 1.5. *Let Q be a non-degenerate continuous symmetric bilinear form on $\mathcal{S}(\mathbb{R}^d)$, and consider*

$$\Xi(f) := e^{-\frac{1}{2}Q(f,f)}, \qquad f \in \mathcal{S}(\mathbb{R}^d). \tag{1.8}$$

Then Ξ is a characteristic function on $\mathcal{S}(\mathbb{R}^d)$.

Proof. First, we observe that

$$\Xi(0) = e^{-\frac{1}{2}Q(0,0)} = e^0 = 1.$$

It also holds that Ξ is continuous because of the continuity of the exponential function and Q. Finally it has to be shown that Ξ is positive definite. Therefore choose for $n \in \mathbb{N}$ $f_1, \dots, f_n \in \mathcal{S}(\mathbb{R}^d)$. Without loss of generality we may assume that the elements $f_1, \dots, f_n$ are linearly independent. Define an $n \times n$ matrix K by

$$K_{lm} = Q(f_l, f_m).$$

K is symmetric and positive definite because of the properties of Q. Assume that there is a non-trivial eigenvector $e = (e_1, \dots, e_n)$ with eigenvalue zero. Then for all $l \in \{1, \dots, n\}$ it holds

$$0 = Ke = \sum_{m=1}^{n} K_{lm} e_m = \sum_{m=1}^{n} Q(f_l, f_m) e_m = Q(f_l, \sum_{m=1}^{n} f_m e_m) =: Q(f_l, g)$$

where $g = \sum_{m=1}^{n} f_m e_m$. Therefore

$$Q(g, g) = \sum_{l=1}^{n} Q(f_l, g) e_l = 0$$

and by the properties of Q and e this gives $g = 0$. This implies that $f_1, \dots, f_n$ must be linearly dependent and leads to a contradiction. Therefore all eigenvalues are strictly greater than zero and K is invertible. Consider now the symmetric, strictly positive definite $n \times n$ matrix K^{-1}, and let μ_n denote the measure on $\mathbb{R}^n$ with density

$$((2\pi)^n \det(K))^{-1/2} \exp(-\tfrac{1}{2}(x, K^{-1}x)), \quad x \in \mathbb{R}^n.$$

Then μ_n is a Gaussian probability measure on $\mathbb{R}^n$ with

$$\int_{\mathbb{R}^n} e^{i(\lambda,x)}\, d\mu_n(x) = \mathbb{E}(e^{i(\lambda,X)}) = e^{-\frac{1}{2}(\lambda,K\lambda)}, \qquad \lambda \in \mathbb{R}^n,$$

and for $l, m \in \{1, \dots, n\}$

$$\begin{aligned}\int_{\mathbb{R}^n} e^{i(e_l-e_m,x)}\, d\mu_n(x) &= e^{-\frac{1}{2}(e_l-e_m,K(e_l-e_m))} = e^{-\frac{1}{2}(K_{ll}-2K_{lm}+K_{mm})} \\ &= e^{-\frac{1}{2}(Q(f_l,f_l)-Q(f_l,f_m)-(Q(f_m,f_l)-Q(f_m,f_m)))} \\ &= e^{-\frac{1}{2}Q(f_l-f_m,f_l-f_m)}.\end{aligned}$$

Finally let $z_1, \dots, z_n \in \mathbb{C}$, then

$$\begin{aligned}\sum_{l,m=1}^{n} z_l \bar{z}_m e^{-\frac{1}{2}Q(f_l-f_m,f_l-f_m)} &= \sum_{l,m=1}^{n} z_l \bar{z}_m \int_{\mathbb{R}^n} e^{i(e_l-e_m,x)}\, d\mu_n(x) \\ &= \int_{\mathbb{R}^n} \sum_{l,m=1}^{n} z_l e^{i\langle e_l,x\rangle} \overline{z_m e^{i\langle e_m,x\rangle}}\, d\mu_n(x) \\ &= \int_{\mathbb{R}^n} \Big|\sum_{l=1}^{n} z_l e^{i\langle e_l,x\rangle}\Big|^2\, d\mu_n(x) \;\geq\; 0,\end{aligned}$$

and $\exp(-1/2\, K)$ is positive semi-definite. Therefore Ξ is positive definite and the assertion is true. □

If the random field φ is constructed by (1.6), (1.7), and (1.8), then it is centered Gaussian with covariance given by

$$\operatorname{Cov}(\varphi(f), \varphi(g)) = \mathbb{E}(\varphi(f)\varphi(g)) = Q(f, g).$$

By the nuclear theorem [78, Thm. V.12], Q can be represented by $\hat{Q} \in \mathcal{S}'(\mathbb{R}^d \times \mathbb{R}^d)$ as

$$Q(f, g) = \langle \hat{Q}, f \times g \rangle$$

and vice versa. Let $\hat{Q} = \operatorname{tr} \in \mathcal{S}'(\mathbb{R}^d \times \mathbb{R}^d)$ be the trace operator, then

$$\langle \operatorname{tr}, f \times g \rangle = \int_{\mathbb{R}^d} f(x)g(x)\, dx = (f, g)_{L^2(\mathbb{R}^d)}.$$

Informally, tr is given by the integral kernel $\delta(x-y)$. Thus we are now able to define a rigorous version of white noise on $\mathbb{R}^d$.

Definition 1.6. A generalized random field W on $\mathbb{R}^d$ is called *white noise*, if its characteristic function is given by

$$\Xi_{\mathrm{WN}}(f) = e^{-\frac{1}{2}(f,f)_{L^2(\mathbb{R}^d)}}.$$

Next we construct a complex-valued generalized random field ψ using real-valued generalized random fields φ_1 and φ_2, (1.5), and (1.7). Therefore for $f \in \mathcal{S}_{\mathbb{C}}(\mathbb{R}^d)$, we set

$$\psi(f) = \varphi_1(\operatorname{Re} f) + i\,\varphi_1(\operatorname{Im} f) - i\,\varphi_2(\operatorname{Re} f) + \varphi_2(\operatorname{Im} f). \tag{1.9}$$

Then $\psi \mapsto \psi(f)$ is conjugate linear: Let $\lambda = \lambda_1 + i\lambda_2 \in \mathbb{C}$, ψ and $\hat{\psi}$ as defined before,

$$\begin{aligned}
(\lambda\psi + \hat{\psi})(f) &= ((\lambda_1 + i\lambda_2)(\varphi_1 + i\varphi_2))(f) + \hat{\psi}(f) \\
&= \lambda_1(\psi(f)) + \lambda_2(-\varphi_2 + i\varphi_1)(f) + \hat{\psi}(f) \\
&= \lambda_1(\psi(f)) + \lambda_2(-\varphi_2(\operatorname{Re} f) - i\varphi_2(\operatorname{Im} f) - i\varphi_1(\operatorname{Re} f) + \varphi_1(\operatorname{Im} f)) + \hat{\psi}(f) \\
&= \lambda_1(\psi(f)) - i\lambda_2(\varphi_1(\operatorname{Re} f) + i\varphi_1(\operatorname{Im} f) - i\varphi_2(\operatorname{Re} f) + \varphi_2(\operatorname{Im} f)) + \hat{\psi}(f) \\
&= \bar{\lambda}(\psi(f)) + \hat{\psi}(f).
\end{aligned}$$

It is easy to show that $f \mapsto \psi(f)$ is linear. Note that if φ_1, φ_2 are independent, then $\operatorname{Re}\psi(f)$ and $\operatorname{Im}\psi(f)$ will not necessarily be independent. If however φ_1 and φ_2 are in addition Gaussian, and f is real, or f is purely imaginary, or φ_1 and φ_2 have the same law, then $\operatorname{Re}\psi(f)$ and $\operatorname{Im}\psi(f)$ are independent.

In what follows, we will show that the domain of W can be extended to elements of $L^2(\mathbb{R}^d)$. Therefore consider real-valued white noise on $\mathbb{R}^d$, i.e. we are given a linear mapping

$$W : \mathcal{S}(\mathbb{R}^d) \to L^0(\Omega, \mathcal{F}, P),$$

where $(\Omega, \mathcal{F}, P)$ is some probability space, so that the family $\{W(f),\, f \in \mathcal{S}(\mathbb{R}^d)\}$ is a centered Gaussian family with

$$\operatorname{Cov}(W(f), W(g)) = \mathbb{E}(W(f)W(g)) = (f,g)_{L^2(\mathbb{R}^d)}.$$

A possible construction of these random fields was shown before.

Lemma 1.7. *Let $f \in L^2(\mathbb{R}^d)$ and assume that $(f_n, n \in \mathbb{N})$ is a sequence in $\mathcal{S}(\mathbb{R}^d)$ converging in $L^2(\mathbb{R}^d)$ to f. Let $p \geq 1$, then $(W(f_n), n \in \mathbb{N})$ converges in $L^p(P)$ to some random variable $W_p(f)$. Furthermore for $p, q \geq 1$: $P(W_p(f) = W_q(f)) = 1$, i.e. $W(f_n) \to W(f)$ in all $L^p(P), p \geq 1$.*

Finally for $f, g \in L^2(\mathbb{R}^d)$

$$\|W(f) - W(g)\|_2^2 = \|f - g\|_2^2$$

and therefore W extends to an isometric embedding of $L^2(\mathbb{R}^d)$ into $L^2(P)$.

Proof. First of all $W(f_n) \in L^p(P)$ for all $n \in \mathbb{N}$ because the law of $W(f_n)$ has Gaussian density with mean zero and variance $\|f_n\|_2^2$, and therefore the inequality of arithmetic and geometric

means holds. Hölder's inequality allows to assume that $p = 2m$, $m \in \mathbb{N}$. As W is linear, we have for $n, l \in \mathbb{N}$

$$\begin{aligned} \| W(f_n) - W(f_l) \|_{2m}^{2m} &= \| W(f_n - f_l) \|_{2m}^{2m} \\ &= \big(\sqrt{2\pi \, \|f_n - f_l\|_2^2}\big)^{-1} \int_\Omega z^{2m} \, \exp\big(-\tfrac{z^2}{2\|f_n - f_l\|_2^2}\big) \, dz \\ &= (2m-1)!! \, \|f_n - f_l\|_2^m \end{aligned}$$

which tends to zero with $n, l \to +\infty$ by the assumption that $(f_n, n \in \mathbb{N})$ converges in $L^2(\mathbb{R}^d)$. This implies that $(W(f_n), n \in \mathbb{N})$ is Cauchy in $L^p(P)$. Because of the completeness of $L^p(P)$, $(W(f_n), n \in \mathbb{N})$ converges in every $L^p(P)$ to some random variable $W_p(f)$. Let p, $q \geq 1$, $q > p$, then $L^q(P) \subseteq L^p(P)$ and $P(W_p(f) = W_q(f)) = 1$.

Finally the calculation above gives for $f, g \in L^2(\mathbb{R}^d)$

$$\|W(f) - W(g)\|_2^2 = \|f - g\|_2^2 ,$$

and therefore the map W extends to an isometric embedding of $L^2(\mathbb{R}^d)$ into $L^2(P)$. □

In the following we assume that f is continuous and that its support is in the d-dimensional cube $C = ([a_1, b_1], \ldots, [a_d, b_d])$, $a_i, b_i \in \mathbb{R}$. Let $(\mathcal{Z}_n, n \in \mathbb{N})$ be a sequence of partitions of C, such that the maximum over the mesh sizes $\triangle_n^i$, $i = 1, \ldots, d$ in all directions converges to zero with $n \to +\infty$, and that the grid points are given by x_k^n with $x_0^n = (a_1, \ldots, a_d)$, $x_{N(n)+(1,\ldots,1)}^n = (b_1, \ldots, b_d)$, and $N(n) = (N_1(n), \ldots, N_d(n))$, $k = (k_1, \ldots, k_d)$, $k_i \in \{0, \ldots, N_i(n)\}$. Define

$$f_n := \sum_{k=0}^{N(n)} f(x_k^n) 1_{\triangle_k^n},$$

where the small cube $\triangle_k^n = ([x_{k_1}, x_{k_1+1}), \ldots, [x_{k_d}, x_{k_d+1}))$, except if there exists $i \in \{1, \ldots, d\}$ such that $k_i = N_i(n)$, then $\triangle_k^n = ([x_{k_1}, x_{k_1+1}), \ldots, [x_{N_i(n)}, b_i], \ldots, [x_{k_d}, x_{k_d+1}))$.

Corollary 1.8. *With the definitions made above, it holds that*

$$f_n \underset{n\to\infty}{\longrightarrow} f \qquad \text{in } L^2(\mathbb{R}^d)$$

and

$$W(f_n) \underset{n\to\infty}{\longrightarrow} W(f) \qquad \text{in } L^2(P)$$

with

$$W(f_n) = \sum_{k=0}^{N(n)} f(x_k^n) W_k^n \, |\triangle_k^n| , \qquad W_k^n := |\triangle_k^n|^{-1} \, W(1_{\triangle_k^n}),$$

where $|\triangle_k^n|$ denotes the volume of the cube $\triangle_k^n$.

Proof. Since f_n and f are zero outside of C, we can restrict the proof to C. Let $\varepsilon > 0$ be given, then there exists $\delta > 0$ such that for all $x, y \in C$, $|x - y| < \delta$, $|f(x) - f(y)| < \varepsilon$ because f is uniformly continuous on C. For n_0 large enough so that $\max_{i=1,\ldots,d} \triangle_{n_0}^i < \delta$, we have for

all $n \geq n_0$, $\sup_{x \in C} |f_n(x) - f(x)| < \varepsilon$. But then $\|f_n - f\|_2 < \varepsilon \cdot V(C)^{1/2}$ where $V(C)$ denotes the volume of the cube.

The convergence of $(W(f_n), n \in \mathbb{N})$ follows applying Lemma 1.7 and

$$W(f_n) = \sum_{k=0}^{N(n)} f(x_k^n) W(1_{\triangle_k^n}) = \sum_{k=0}^{N(n)} f(x_k^n)(|\triangle_k^n|^{-1} W(1_{\triangle_k^n}))\, |\triangle_k^n|$$

implies the last open claim of the corollary. □

Next we will give some details on W_k^n. First we observe that the family of random variables $(W_k^n, k = 0, \ldots, N(n))$ forms a centered Gaussian family with

$$\mathbb{E}(W_k^n W_l^n) = (|\triangle_k^n|\,|\triangle_l^n|)^{-1} \left(1_{\triangle_k^n}, 1_{\triangle_l^n}\right)_{L^2(\mathbb{R}^d)} = |\triangle_k^n|^{-1}\, \delta_{kl},$$

i.e. the family is independent. A special case consists of a sequence of partitions $(\mathcal{Z}_n, n \in \mathbb{N})$ with equidistant mesh sizes defined by $\prod_{i=1}^d (N_i(n) + 1)$ points and the elements $N_i(n)$ are fixed by $\triangle_n^i = (b_1 - a_1)/(N_1(n) + 1)$ for all $k \in \{1, \ldots, d\}$. Then $|\triangle_k^n| = \prod_{i=1}^d \triangle_n^i = ((b_1 - a_1)/(N_1(n) + 1))^d$. In this case, the family $W^n = (W_k^n, k = 0, \ldots, N(n))$ is i.i.d. and centered Gaussian with variance $((N_1(n) + 1)/(b_1 - a_1))^d$.

Definition 1.9. The i.i.d., centered Gaussian family W^n as defined above associated with the equidistant partition $\mathcal{Z}_n$ is called *discrete white noise* on a cube $C \subset \mathbb{R}^d$.

1.3.2. Operations on Generalized Random Fields

Let φ be a generalized random field. In analogy of the realization of white noise W as the canonical coordinate map on $\mathcal{S}'(\mathbb{R}^d)$ under the white noise measure, where

$$W(f)(\omega) = \langle \omega, f \rangle,$$

we define operations on φ via dual pairing.

Definition 1.10.

1. The *complex conjugation* $\mathcal{C}$ is defined by
$$(\mathcal{C}\varphi)(f) := \mathcal{C}(\varphi(\mathcal{C}f)).$$
2. The *reflection operator* ρ, which is defined on $\mathcal{S}(\mathbb{R}^d)$ resp. $\mathcal{S}_{\mathbb{C}}(\mathbb{R}^d)$ as $(\rho f)(x) = f(-x)$, is defined on φ as
$$(\rho\,\varphi)(f) := \varphi(\rho f).$$
3. The *Fourier transform* $\mathcal{F}$ and its inverse $\mathcal{F}^{-1}$ are defined on φ via
$$(\mathcal{F}\varphi)(f) := \varphi(\mathcal{F}^{-1} f),$$
$$(\mathcal{F}^{-1}\varphi)(f) := \varphi(\mathcal{F} f).$$

4. If $g \in C^\infty(\mathbb{R}^d)$ with at most polynomial growth, then

$$(g\varphi)(f) := \varphi(\bar{g}f).$$

Remark 1.11. The Fourier transform and its inverse on $\mathcal{S}_\mathbb{C}(\mathbb{R}^d)$ are defined by

$$\mathcal{F}f(p) = \int_{\mathbb{R}^d} e^{2\pi i(p,x)} f(x)\, dx, \qquad f \in \mathcal{S}_\mathbb{C}(\mathbb{R}^d),\ p \in \mathbb{R}^d,$$
$$\mathcal{F}^{-1}g(x) = \int_{\mathbb{R}^d} e^{-2\pi i(p,x)} g(p)\, dp, \qquad g \in \mathcal{S}_\mathbb{C}(\mathbb{R}^d),\ x \in \mathbb{R}^d.$$

Moreover we have

$$\mathcal{C}\rho = \rho\,\mathcal{C}, \qquad \mathcal{F}\rho = \rho\mathcal{F}$$

on $\mathcal{S}_\mathbb{C}(\mathbb{R}^d)$. We remark that the real and imaginary part of elements of $\mathcal{S}_\mathbb{C}(\mathbb{R}^d)$ are given by

$$\mathrm{Re} = \tfrac{1}{2}(\mathbb{1} + \mathcal{C}), \qquad \mathrm{Im} = \tfrac{1}{2i}(\mathbb{1} - \mathcal{C}).$$

1.3.3. Construction of Stationary Gaussian Random Fields

Let φ be a stationary, real-valued, centered Gaussian random field with covariance

$$\mathrm{Cov}\,(\varphi(f), \varphi(g)) = (f, Cg)_{L^2(\mathbb{R}^d)}, \qquad f, g \in \mathcal{S}(\mathbb{R}^d)$$

with $C : \mathcal{S}(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$. Assume that C is given by an integral kernel, which by stationarity can be written as

$$Cf(x) = \int_{\mathbb{R}^d} K(x-y) f(y)\, dy,$$

where K is even because of the symmetry of C, and positive definite. Hence if K is continuous, K can be written as

$$K(x) = \int e^{-2\pi i(x,p)}\, d\Gamma(p)$$

for some positive measure Γ by Bochner's theorem [98, Thm. XI.2]. Henceforth we consider only the case where Γ has density γ which is supposed to be strictly positive and C^∞. Then $\gamma^{1/2}$ is a strictly positive smooth root of γ. Assume that W is white noise as defined in Section 1.3.1 and set

$$\psi(f) := (\mathcal{F}^{-1}\gamma^{1/2}\mathcal{F}W)(f), \qquad f \in \mathcal{S}(\mathbb{R}^d). \tag{1.10}$$

Corollary 1.12. *ψ is centered Gaussian and*

$$\psi(f) = W(\mathcal{F}^{-1}\gamma^{1/2}\mathcal{F}f), \qquad f \in \mathcal{S}(\mathbb{R}^d).$$

Proof. We calculate for $f \in \mathcal{S}(\mathbb{R}^d)$

$$\begin{aligned}\psi(f) &= (\mathcal{F}^{-1}\gamma^{1/2}\mathcal{F}W)(f)\\ &= W(\mathcal{F}^{-1}\bar{\gamma}^{1/2}\mathcal{F}f)\end{aligned}$$

by Definition 1.10. The assertion follows by the observation that $\gamma^{1/2}$ is real. □

Next we set

$$\pi_+ = \tfrac{1}{2}(\mathbb{1} + \rho),$$
$$\pi_- = \tfrac{1}{2}(\mathbb{1} - \rho).$$

In order to prove that ψ is a realization of φ, we need the following corollary. For further information on Fourier transforms, we refer to [79].

Corollary 1.13. *The Fourier transform has the following properties:*

1. $\mathcal{C}\mathcal{F} = \mathcal{F}\mathcal{C}\rho$,
2. $\mathrm{Im}\,\mathcal{F} = \mathcal{F}\frac{1}{2i}(\mathbb{1} - \mathcal{C}\rho)$ *and* $\mathrm{Re}\,\mathcal{F} = \mathcal{F}\frac{1}{2}(\mathbb{1} + \mathcal{C}\rho)$,
3. $\pi_+\mathrm{Im} = \mathrm{Im}\,\pi_+$,
4. $\pi_\pm\mathcal{F} = \mathcal{F}\pi_\pm$,
5. $\pi_+\mathrm{Im}\,\mathcal{F} = \mathcal{F}\pi_+\mathrm{Im}$,
6. $\pi_-\mathrm{Re}\,\mathcal{F} = i\mathcal{F}\pi_-\mathrm{Im}$ *and* $\pi_-\mathrm{Im}\,\mathcal{F} = -i\mathcal{F}\pi_-\mathrm{Re}$,
7. *for* $f \in \mathcal{S}_{\mathbb{R}}(\mathbb{R}^d)$ *it holds* $\pi_-\mathrm{Re}\,\mathcal{F}f = 0$ *and* $\pi_+\mathrm{Im}\,\mathcal{F}f = 0$.

Proof.

1. For $f \in \mathcal{S}_{\mathbb{C}}(\mathbb{R}^d)$ we have
$$\mathcal{C}\mathcal{F}f(p) = \int_{\mathbb{R}^d} e^{-2\pi i(p,x)}(\mathcal{C}f(p))\,dp = \int_{\mathbb{R}^d} e^{2\pi i(p,x)}(\mathcal{C}\rho f(p))\,dp = \mathcal{F}\mathcal{C}\rho f(p).$$
2. Using the previous statement and the definition of the imaginary part (see Remark 1.11), we have
$$\mathrm{Im}\,\mathcal{F} = \frac{1}{2i}(\mathbb{1} - \mathcal{C})\mathcal{F} = \tfrac{1}{2i}(\mathcal{F} - \mathcal{F}\mathcal{C}\rho) = \mathcal{F}\tfrac{1}{2i}(\mathbb{1} - \mathcal{C}\rho),$$
and
$$\mathrm{Re}\,\mathcal{F} = \tfrac{1}{2}(\mathbb{1} + \mathcal{C})\mathcal{F} = \tfrac{1}{2}(\mathcal{F} + \mathcal{F}\mathcal{C}\rho) = \mathcal{F}\tfrac{1}{2}(\mathbb{1} + \mathcal{C}\rho).$$
3. Remark 1.11 implies
$$\pi_+\mathrm{Im} = (\mathbb{1} + \rho)\tfrac{1}{2i}(\mathbb{1} - \mathcal{C}) = \tfrac{1}{2i}(\mathbb{1} + \rho - \mathcal{C} - \mathcal{C}\rho) = \tfrac{1}{2i}(\mathbb{1} - \mathcal{C})(\mathbb{1} + \rho) = \mathrm{Im}\,\pi_+.$$
4. The commutativity of the Fourier transform and ρ leads to
$$\pi_\pm\mathcal{F} = (\mathbb{1} \pm \rho)\mathcal{F} = \mathcal{F} \pm \mathcal{F}\rho = \mathcal{F}\pi_\pm.$$
5. The following calculations use the second, third, and forth statement of this corollary, $\rho^2 = \mathbb{1}$, and the commutativity of $\mathcal{C}$ and ρ.
$$\pi_+\mathrm{Im}\,\mathcal{F} = \mathrm{Im}\,\mathcal{F}\pi_+ = \mathcal{F}\tfrac{1}{2i}(\mathbb{1} - \mathcal{C}\rho)(\mathbb{1} + \rho) = \mathcal{F}\tfrac{1}{2i}(\pi_+ - (\rho + \mathbb{1})\mathcal{C}) = \mathcal{F}\pi_+\mathrm{Im}$$

6. With *2.* and *4.* it follows

$$\pi_- \mathrm{Re}\, \mathcal{F} = \pi_- \mathcal{F} \tfrac{1}{2} (\mathbb{1} + \mathcal{C}\rho) = \mathcal{F} \tfrac{1}{4} (\mathbb{1} - \rho)(\mathbb{1} + \mathcal{C}\rho) = \mathcal{F} \tfrac{1}{4} (\mathbb{1} - \rho - \mathcal{C} + \rho\mathcal{C}) = i\, \mathcal{F} \pi_- \mathrm{Im}\,.$$

Similarly we have

$$\pi_- \mathrm{Im}\, \mathcal{F} = \pi_- \mathcal{F} \tfrac{1}{2i} (\mathbb{1} - \mathcal{C}\rho) = -i\, \mathcal{F} \tfrac{1}{4} (\mathbb{1} - \rho)(\mathbb{1} - \mathcal{C}\rho) = -i\, \mathcal{F} \tfrac{1}{4} (\mathbb{1} - \rho + \mathcal{C} - \rho\mathcal{C}) = -i\, \mathcal{F} \pi_- \mathrm{Re}\,.$$

7. Finally using the previous statement, for $f \in \mathcal{S}_{\mathbb{R}}(\mathbb{R}^d)$ it holds

$$\pi_- \mathrm{Re}\, \mathcal{F} f = i\, \mathcal{F} \pi_- \mathrm{Im}\, f = 0,$$

and the fifth statement of this corollary implies for $f \in \mathcal{S}_{\mathbb{R}}(\mathbb{R}^d)$

$$\pi_+ \mathrm{Im}\, \mathcal{F} f = \mathcal{F} \pi_+ \mathrm{Im}\, f = 0.$$

□

Using this corollary, we can prove some properties of ψ.

Corollary 1.14. *$\mathcal{F}^{-1} \gamma^{1/2} \mathcal{F} f$ is a real function in $\mathcal{S}(\mathbb{R}^d)$ for $f \in \mathcal{S}(\mathbb{R}^d)$ and therefore ψ is a real random field.*

Proof. First we observe that the Fourier transform is a one-to-one map of the Schwartz space onto itself [79, Thm. XI.1]. As $\gamma^{1/2}$ is positive, bounded, even, and in C^∞, its multiplication with $\mathcal{F} f$, $f \in \mathcal{S}(\mathbb{R}^d)$ is again in the Schwartz space with even real and odd imaginary part, and thus its inverse Fourier transform is in $\mathcal{S}(\mathbb{R}^d)$ and real. Therefore let f be a real function, then

$$\mathcal{F} f = \mathrm{Re}\, \mathcal{F} f + i\, \mathrm{Im}\, \mathcal{F} f = (\pi_+ + \pi_-) \mathrm{Re}\, \mathcal{F} f + i\, (\pi_+ + \pi_-) \mathrm{Im}\, \mathcal{F} f = \pi_+ \mathrm{Re}\, \mathcal{F} f + i\, \pi_- \mathrm{Im}\, \mathcal{F} f$$

where we used Corollary 1.13 in the last step. This implies that $\mathcal{F} f$ has even real and odd imaginary part. Next let $g \in \mathcal{S}_{\mathbb{C}}(\mathbb{R}^d)$ have even real and odd imaginary part, then its inverse Fourier transform is real by Corollary 1.13 and

$$\mathrm{Im}\, \mathcal{F}^{-1} g = (\pi_+ + \pi_-) \mathrm{Im}\, \mathcal{F}^{-1} g = \mathcal{F}^{-1} \pi_+ \mathrm{Im}\, g - i\, \mathcal{F}^{-1} \pi_- \mathrm{Re}\, g = 0.$$

Moreover ψ is a real-valued random field because of the random field construction in Section 1.3.1. □

Corollary 1.15. *It holds* $\mathrm{Cov}\,(\psi(f), \psi(g)) = (f, Cg)_{L^2(\mathbb{R}^d)}$ *and therefore ψ is a realization of φ.*

Proof. Using Plancherel's theorem [98, Ch. VI.2] twice and the natural embedding of $L^2_{\mathbb{R}}(\mathbb{R}^d)$ into $L^2_{\mathbb{C}}(\mathbb{R}^d)$, we calculate

$$\begin{aligned}
\mathbb{E}(\psi(f)\psi(g)) &= \mathbb{E}(W(\mathcal{F}^{-1}\gamma^{1/2}\mathcal{F}f)W(\mathcal{F}^{-1}\gamma^{1/2}\mathcal{F}g)) \\
&= \left(\mathcal{F}^{-1}\gamma^{1/2}\mathcal{F}f, \mathcal{F}^{-1}\gamma^{1/2}\mathcal{F}g\right)_{L^2_{\mathbb{R}}(\mathbb{R}^d)} \\
&= \left(\gamma^{1/2}\mathcal{F}f, \gamma^{1/2}\mathcal{F}g\right)_{L^2_{\mathbb{C}}(\mathbb{R}^d)} \\
&= (\mathcal{F}f, \gamma\mathcal{F}g)_{L^2_{\mathbb{C}}(\mathbb{R}^d)} \\
&= \left(\mathcal{F}f, \mathcal{F}(\mathcal{F}^{-1}\gamma) * g\right)_{L^2_{\mathbb{C}}(\mathbb{R}^d)} \\
&= (\mathcal{F}f, \mathcal{F}Cg)_{L^2_{\mathbb{C}}(\mathbb{R}^d)} \\
&= (f, Cg)_{L^2_{\mathbb{R}}(\mathbb{R}^d)} = \mathrm{Cov}\,(\varphi(f), \varphi(g)).
\end{aligned}$$

□

Another realization of φ will be constructed next. Let W be a white noise random field as constructed in Section 1.3.1 and set

$$V := \pi_+ W + i\pi_- W.$$

Define

$$\chi := \mathcal{F}^{-1}\gamma^{1/2}V. \tag{1.11}$$

Corollary 1.16. *χ is another realization of φ.*

Proof. The proof will be divided into several parts. First note that Equation (1.9) and Definition 1.10 imply for $f \in \mathcal{S}_{\mathbb{C}}(\mathbb{R}^d)$,

$$\begin{aligned}
\chi(f) &= (\mathcal{F}^{-1}\gamma^{1/2}V)(f) \\
&= V(\gamma^{1/2}\mathcal{F}f) \\
&= \pi_+ W(\gamma^{1/2}\mathcal{F}f) + i\,\pi_- W(\gamma^{1/2}\mathcal{F}f) \\
&= W(\pi_+\gamma^{1/2}\mathcal{F}f) + i\,W(\pi_-\gamma^{1/2}\mathcal{F}f) \\
&= W(\pi_+\mathrm{Re}\,\gamma^{1/2}\mathcal{F}f) - i\,W(\pi_-\mathrm{Re}\,\gamma^{1/2}\mathcal{F}f) + i\,W(\pi_+\mathrm{Im}\,\gamma^{1/2}\mathcal{F}f) + W(\pi_-\mathrm{Im}\,\gamma^{1/2}\mathcal{F}f).
\end{aligned}$$

Assuming now that f is real-valued it follows — using the fact that $\pi_\pm\mathrm{Re}$ and $\pi_\pm\mathrm{Im}$ commute with the multiplication by $\gamma^{1/2}$ — with Corollary 1.13 that

$$\begin{aligned}
\pi_-\mathrm{Re}\,\gamma^{1/2}\mathcal{F}f = 0, \\
\pi_+\mathrm{Im}\,\gamma^{1/2}\mathcal{F}f = 0,
\end{aligned}$$

and therefore $\chi(f)$ is real-valued. Furthermore $f \mapsto \chi(f)$ is linear and χ is a real-valued generalized random field for real-valued f.

Next we will calculate the covariance of χ. Let $f, g \in \mathcal{S}(\mathbb{R}^d)$, then

$$\begin{aligned}
&\mathbb{E}(\chi(f)\chi(g)) \\
&\quad = \mathbb{E}\Big(\big(W(\pi_+\mathrm{Re}\,\gamma^{1/2}\mathcal{F}f) + W(\pi_-\mathrm{Im}\,\gamma^{1/2}\mathcal{F}f)\big)\big(W(\pi_+\mathrm{Re}\,\gamma^{1/2}\mathcal{F}g) + W(\pi_-\mathrm{Im}\,\gamma^{1/2}\mathcal{F}g)\big)\Big).
\end{aligned}$$

Note that π_+ and π_- are orthogonal projections in $L^2_{\mathbb{C}}(\mathbb{R}^d)$, i.e. for $a, b \in L^2_{\mathbb{C}}(\mathbb{R}^d)$, $(\pi_+ a, \pi_- b) = 0$. This implies

$$\left(\pi_+ \mathrm{Re}\,\gamma^{1/2}\mathcal{F}f, \pi_- \mathrm{Im}\,\gamma^{1/2}\mathcal{F}g\right) = \left(\pi_- \mathrm{Im}\,\gamma^{1/2}\mathcal{F}f, \pi_+ \mathrm{Re}\,\gamma^{1/2}\mathcal{F}g\right) = 0.$$

Thus the calculations of the covariance can be reduced to

$$\mathbb{E}(\chi(f)\chi(g)) = \left(\pi_+ \mathrm{Re}\,\gamma^{1/2}\mathcal{F}f, \pi_+ \mathrm{Re}\,\gamma^{1/2}\mathcal{F}g\right)_{L^2_{\mathbb{R}}(\mathbb{R}^d)} + \left(\pi_- \mathrm{Im}\,\gamma^{1/2}\mathcal{F}f, \pi_- \mathrm{Im}\,\gamma^{1/2}\mathcal{F}g\right)_{L^2_{\mathbb{R}}(\mathbb{R}^d)}.$$

Since $\pi_- \mathrm{Re}\,\gamma^{1/2}\mathcal{F}f = \pi_+ \mathrm{Im}\,\gamma^{1/2}\mathcal{F}f = \pi_- \mathrm{Re}\,\gamma^{1/2}\mathcal{F}g = \pi_+ \mathrm{Im}\,\gamma^{1/2}\mathcal{F}g = 0$ as seen before, we can also rewrite this as

$$\begin{aligned}
\mathbb{E}(\chi(f)\chi(g)) &= \left(\mathrm{Re}\,\gamma^{1/2}\mathcal{F}f, \mathrm{Re}\,\gamma^{1/2}\mathcal{F}g\right)_{L^2_{\mathbb{R}}(\mathbb{R}^d)} + \left(\mathrm{Im}\,\gamma^{1/2}\mathcal{F}f, \mathrm{Im}\,\gamma^{1/2}\mathcal{F}g\right)_{L^2_{\mathbb{R}}(\mathbb{R}^d)} \\
&= \left(\gamma^{1/2}\mathcal{F}f, \gamma^{1/2}\mathcal{F}g\right)_{L^2_{\mathbb{C}}(\mathbb{R}^d)} \\
&= (\mathcal{F}f, \gamma\mathcal{F}g)_{L^2_{\mathbb{C}}(\mathbb{R}^d)} \\
&= \left(\mathcal{F}f, \mathcal{F}(\mathcal{F}^{-1}\gamma) * g\right)_{L^2_{\mathbb{C}}(\mathbb{R}^d)} \\
&= (f, Cg)_{L^2_{\mathbb{R}}(\mathbb{R}^d)},
\end{aligned}$$

where we used Plancherel's theorem [98, Ch. VI.2] in the last step. Therefore χ gives another realization of φ. □

1.4. Algorithms

The following algorithms are based on the previous sections. They allow a fast generation of d-dimensional stationary Gaussian random fields with given covariance using the advantage of fast Fourier transforms. The boundary conditions implemented here are periodic. For Neumann or Dirichlet boundary conditions, fast sine and cosine transformations have to be used.

The first algorithm is an implementation of Equation (1.1) in Section 1.2 and of Equation (1.10) in Section 1.3.

Algorithm 1.17.
Remarks:

1. The functions FFT and FFT^{-1} include all necessary rescaling depending on the used FFT algorithm and the integers N_i.
2. A is a d-dimensional complex-valued array, B is real-valued.
3. $x_{k_1 \cdots k_d}$ denotes the grid point corresponding to the integers $(k_1, \ldots, k_d)$. The grid points are distributed equidistantly in each direction.

Input:

1. d-dimensional cube D, where $l_1, \ldots, l_d$ is the length of the edges,

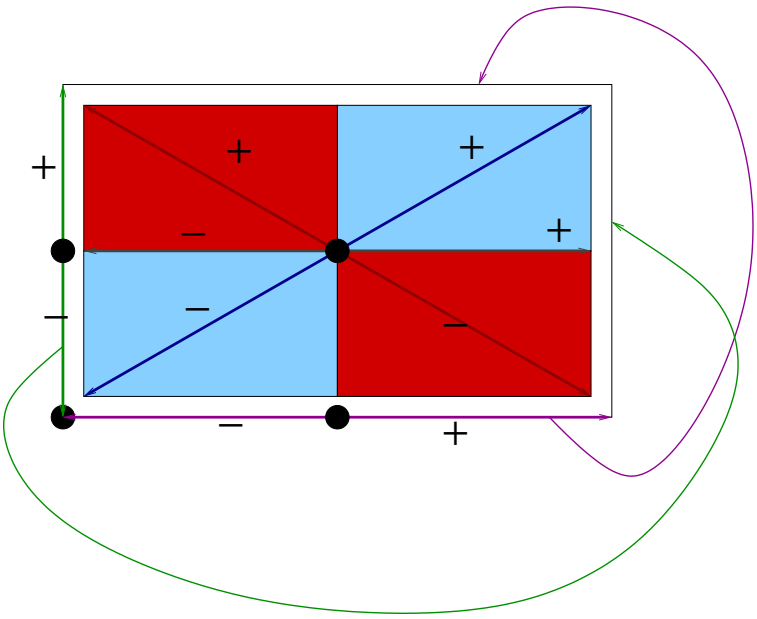


Figure 1.2.: Alignment of the random numbers in Algorithm 1.18 in the two-dimensional case.

2. $N_1, \ldots, N_d$ number of discretization points in each direction, all even,
3. $\gamma^{1/2}$ a symmetric, positive function on D,
4. $R()$ a function that generates independent $\mathcal{N}(0,1)$-distributed random numbers.

Output: GRF B with covariance depending on γ.

for $k_i = 0, \ldots, \frac{N_i}{2}$, $i = 1, \ldots, d$ **do**
  $B(k_1, \ldots, k_d) \leftarrow R()$;
end for
$A \leftarrow \mathrm{FFT}\, B$;
for $k_i = 0, \ldots, \frac{N_i}{2}$, $i = 1, \ldots, d$ **do**
  $A(k_1, \ldots, k_d) \leftarrow A(k_1, \ldots, k_d) \cdot \gamma(x_{k_1 \cdots k_d})^{1/2}$;
end for
$B \leftarrow \mathrm{FFT}^{-1} A$;

The second algorithm is based on Equation (1.2) in Section 1.2 or Equation (1.11) in Section 1.3. The important fact is that the Fourier transform of the white noise random field has even real and odd imaginary part. Figure 1.2 shows how the random numbers are organized in the Fourier transformed array in the two-dimensional case. The black dots are the real-valued grid points. Lines and areas with the same color but arrows in different directions denote the conjugate pairs which is emphasized by the + and − signs. The long arrows in green and magenta indicate the periodic boundary conditions.

If we use d-dimensional DFT on a grid with N_i discrete points in direction e_i, where the set $\{e_i,\, i = 1, \ldots, d\}$ is the standard basis in $\mathbb{R}^d$, the following algorithm generates a GRF with

approximated covariance C, where the covariance of points having a distance larger than half of the length of the edges of the cube should be small, in order to be locally in the situation of $\mathbb{R}^d$. The exact implemented covariance is given by

$$\sum_{k=0}^{N-1} \gamma(k) e_k \otimes e_k,$$

where $k, N \in \mathbb{N}_0^d$ and 1 is the d-dimensional vector with all entries equal to one and $\left\{e_k,\, k \in \mathbb{N}^d\right\}$ is a Hilbert space basis that induces periodic boundary conditions. Our choice is the basis generated by the exponential function.

Algorithm 1.18.
Remarks:

1. All calculations have to be done modulo N_i in the i-th direction.
2. The function FFT^{-1} includes all necessary rescaling depending on the used FFT algorithm and the integers N_i.
3. A is a d-dimensional complex-valued array, B is real-valued.
4. $x_{k_1 \cdots k_d}$ denotes the grid point corresponding to the integers $(k_1, \ldots, k_d)$. The grid points are distributed equidistantly in each direction.

Input:

1. d-dimensional cube D with $l_1, \ldots, l_d$ length of the edges,
2. $N_1, \ldots, N_d$ number of discretization points in each direction, all even,
3. $\gamma^{1/2}$ a symmetric, positive function on D,
4. $R()$ a function that generates independent $\mathcal{N}(0,1)$-distributed random numbers.

Output: GRF B with covariance depending on γ.

for $k_i = 0, \ldots, \frac{N_i}{2}$, $i = 1, \ldots, d$ **do**
  if ($k_i \in \{0, N_i\}$, for all $i = 1, \ldots, d$) or $k_i = \frac{N_i}{2}$, for all $i = 1, \ldots, d$) **then**
    $\mathrm{Re}\, A(k_1, \ldots, k_d) \leftarrow R() \cdot \gamma(x_{k_1 \cdots k_d})^{1/2}$;
    $\mathrm{Im}\, A(k_1, \ldots, k_d) \leftarrow 0$;
  else
    $\mathrm{Re}\, A(k_1, \ldots, k_d) \leftarrow \frac{1}{\sqrt{2}} R() \cdot \gamma(x_{k_1 \cdots k_d})^{1/2}$;
    $\mathrm{Im}\, A(k_1, \ldots, k_d) \leftarrow \frac{1}{\sqrt{2}} R() \cdot \gamma(x_{k_1 \cdots k_d})^{1/2}$;
    $\mathrm{Re}\, A(N_1 - k_1, \ldots, N_d - k_d) \leftarrow \mathrm{Re}\, A(k_1, \ldots, k_d)$;
    $\mathrm{Im}\, A(N_1 - k_1, \ldots, N_d - k_d) \leftarrow -\mathrm{Im}\, A(k_1, \ldots, k_d)$;
  end if
end for
$B \leftarrow \mathrm{FFT}^{-1} A$;

1.5. Covariance Functions

This section presents a class of covariance functions on $\mathbb{R}^2 \times \mathbb{R}^2$ that can be used for the algorithms presented in the previous section and that will be useful in Chapter 2 and 3.

Let f be a positive symmetric function and define the covariance function by

$$C(x,y) = \int_{\mathbb{R}^2} e^{-2\pi i(p,x-y)} f(p)\, dp.$$

As we chose f to be positive and symmetric, $C(x,y) = C(y,x)$. Moreover we want C to have exponential decay for $|x-y| \gg 1$ and vice versa (which follows automatically because of the symmetry property). It seems to be reasonable to have the same covariance in every direction. Therefore we will choose f to be also symmetric under rotation. The following lemma introduces a class of functions f that satisfy the required assumptions.

Lemma 1.19. *Let $C(x,y)$ be a covariance function defined by*

$$C(x,y) = \int_{\mathbb{R}^2} e^{-2\pi i(p,x-y)} f(p)\, dp.$$

This covariance function is positive, symmetric under rotation and has exponential decay in $|x-y|$ if

1. *$f(p) = (1+|p|^2)^{-1}$ with rate bounded by*
$$\text{const.}\, |x-y|\, e^{-2\pi|x-y|},$$
2. *$f(p) = (|p|^2+1)^{-n}$ for $n \geq 2$ with rate bounded by*
$$\pi\, e^{-2\pi|x-y|}(\pi^2\,|x-y|+1),$$
3. *$f(p) = (|p|^{2n}+1)^{-1}$ for $n \geq 2$ with rate bounded by*
$$5\pi/\sqrt{2}\; e^{-\sqrt{2}\pi|x-y|}.$$

Proof.

1. The proof can be found in [29]. We note that this function is not defined for $x = y$ and therefore moments higher than one do not exist.
2. It is easy to see that the function is positive and symmetric under rotation. Without loss of generality we may assume that $y = 0$ and prove the claim for $C(x,0)$. Next we reduce the statement to an easier one. Therefore notice that for $n > 2$
$$\frac{1}{(1+|p|^2)^n} \leq \frac{1}{(1+|p|^2)^2}$$
for all $p \in \mathbb{R}^2$. It follows that
$$\int_{\mathbb{R}^2} e^{-2\pi i(p,x)}(|p|^2+1)^{-n}\, dp \leq \int_{\mathbb{R}^2} e^{-2\pi i(p,x)}(|p|^2+1)^{-2}\, dp,$$

and therefore if this function has exponential decay for $n = 2$, it will have exponential decay for all $n > 2$.

$$\begin{aligned} C(x,0) &= \int_{\mathbb{R}^2} e^{-2\pi i(p,x)} (|p|^2 + 1)^{-2} \, dp \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-2\pi i p_1 |x|} (p_1^2 + p_2^2 + 1)^{-2} \, dp_1 \, dp_2, \end{aligned}$$

where p_1 is chosen to be in x direction. We define $\mu^2 := \mu(p_2)^2 := p_2^2 + 1$. Then C can be rewritten in the following way:

$$C(x,0) = \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-2\pi i p_1 |x|} (p_1 + i\mu)^{-2} (p_1 - i\mu)^{-2} \, dp_1 \, dp_2.$$

Let $g(p) = \exp(-2\pi i \, p_1 \, |x|)(p_1 + i\mu)^{-2}(p_1 - i\mu)^{-2}$. By the residue theorem we can calculate the inner integral

$$\int_{\mathbb{R}} g(p) \, dp_1 = -2\pi i \, \operatorname{res}\left(g(p), \, -i\mu\right) = \tfrac{\pi}{2} \, \mu^{-2} \left(2\pi \, |x| + \mu^{-1}\right) e^{-2\pi\mu|x|}.$$

Finally we work out some estimates using $\mu = \sqrt{p_2^2 + 1} \geq 1$ and the convergence of the integral, so we get

$$\begin{aligned} C(x,0) &= \int_{\mathbb{R}} \tfrac{\pi}{2} \, \mu^{-2} e^{-2\pi\mu|x|} (2\pi \, |x| + \mu^{-1}) \, dp_2 \\ &\leq \tfrac{\pi}{2} \, e^{-2\pi|x|} \int_{\mathbb{R}} 2\pi \, |x| \, \mu^{-2} + \mu^{-3} \, dp_2 \\ &= \pi \left(\pi^2 \, |x| + 1\right) e^{-2\pi|x|}. \end{aligned}$$

3. This part is similar to the second part of this proof. Again it is easy to see that C is symmetric under rotation and positive. Simple calculations show that

$$\frac{1}{1 + |p|^{2n}} \leq \frac{d}{1 + |p|^4}$$

for $n \geq 2$, $p \in \mathbb{R}^2$ and

$$d \geq \left(1 - (2/n)^{2/(n-2)} \, (1 - 2/n)\right)^{-1}.$$

Therefore again it is sufficient to show that C has exponential decay for $n = 2$. For this function the poles are given by

$$\begin{aligned} |p|^4 + 1 &= p_1^4 + p_2^4 + 2p_1^2 p_2^2 + 1 \\ &= \left(p_1 - \sqrt{-p_2^4 + i}\right)\left(p_1 + \sqrt{-p_2^4 + i}\right)\left(p_1 - \sqrt{-p_2^4 - i}\right)\left(p_1 + \sqrt{-p_2^4 - i}\right). \end{aligned}$$

The elements with negative imaginary part are $\sqrt{-p_2^4 - i}$ and $-\sqrt{-p_2^4 + i}$. Define $g(p) = \exp(-2\pi i p_1 |x|)\,(|p|^4+1)^{-1}$, $\rho_+ = \sqrt{\sqrt{p_2^4+1} + p_2^2}$, and $\rho_- = \sqrt{\sqrt{p_2^4+1} - p_2^2}$. Therefore we can calculate the integral by the residues and get

$$\begin{aligned}
C(x,0) &= \int_{\mathbb{R}^2} e^{-2\pi i(p,x)}\,(|p|^4+1)^{-1}\,dp \\
&= \int_{\mathbb{R}} -2\pi i \left(\operatorname{res}\big(g(p), \sqrt{-p_2^4 - i}\big) + \operatorname{res}\big(g(p), -\sqrt{-p_2^4 + i}\big)\right) dp_2 \\
&= \int_{\mathbb{R}} \pi\, e^{-\sqrt{2}\pi\rho_+|x|}\Big(\cos(\sqrt{2}\pi\,|x|\,\rho_-)\big(\sqrt{2}\sqrt{p_2^4+1}\;\rho_+\big)^{-1} \\
&\qquad + \sin(\sqrt{2}\pi\,|x|\,\rho_-)\big(\sqrt{2}\sqrt{p_2^4+1}\;\rho_-\big)^{-1}\Big)\,dp_2 \\
&\le \tfrac{\pi}{\sqrt{2}} \int_{\mathbb{R}} e^{-\sqrt{2}\pi\rho_-|x|}\big(\sqrt{p_2^4+1}\big)^{-1}(\rho_+^{-1} + \rho_-^{-1})\,dp_2.
\end{aligned}$$

Again we note that similar to the second part of this lemma $\rho_- \ge 1$ and that the integral converges. We finish the proof with

$$C(x,0) \le \pi/\sqrt{2} e^{-\sqrt{2}\pi|x|} \int_{\mathbb{R}} \big(\sqrt{p_2^4+1}\big)^{-1}(\rho_+^{-1} + \rho_-^{-1})\,dp_2 \le \tfrac{5\pi}{\sqrt{2}}\, e^{-\sqrt{2}\pi|x|}.$$

□

We can modify and generalize C and f with a parameter m. Then Lemma 1.19 extends to the following lemma.

Lemma 1.20. *Let $C(m,x,y)$ be a covariance function defined by*

$$C(m,x,y) = \int_{\mathbb{R}^2} e^{-2\pi i(p,x-y)} f(m,p)\,dp$$

with $m \in \mathbb{R}_{>0}$. This covariance function is positive, symmetric under rotation and has exponential decay in $|x-y|$ if

1. *$f(m,p) = (m^2 + |p|^2)^{-1}$ with rate bounded by*

$$\text{const.}\,|x-y|\,e^{-2\pi\, m|x-y|},$$

2. *$f(m,p) = (|p|^2 + m^2)^{-n}$ for $n \ge 2$ with rate bounded by*

$$m^{1-2n}\,\pi\,e^{-2\pi\, m|x-y|}(\pi^2\, m\,|x-y| + 1),$$

3. *$f(m,p) = (|p|^{2n} + m^{2n})^{-1}$ for $n \ge 2$ with rate bounded by*

$$5\,\pi/\sqrt{2}\, m^{1-2n} e^{-\sqrt{2}\pi\, m|x-y|}.$$

Proof. We will only show that for these choices of f we can reduce the covariance functions to those of Lemma 1.19.

1. We can rewrite $C(m,x,y)$ in the form

$$\begin{aligned} C(m,x,y) &= \int_{\mathbb{R}^2} e^{-2\pi i(p,x-y)} \left(|p|^2 + m^2\right)^{-1} dp \\ &= \int_{\mathbb{R}^2} e^{-2\pi i\,(p,x-y)} m^{-2} \left(\left(|p|/m\right)^2 + 1\right)^{-1} dp. \end{aligned}$$

   Next a variable transformation is done by $\tilde{p} = p/m$ and $dp = m d\tilde{p}$. Therefore we get

$$C(m,x,y) = m^{-2} \int_{\mathbb{R}^2} e^{-2\pi i\,(\tilde{p},m(x-y))} (|\tilde{p}|^2 + 1)^{-1} m \, d\tilde{p} = m^{-1}\, C(1, m(x-y))$$

   Applying Lemma 1.19 to $C(1, m(x-y))$ yields the claim.

2. This proof is similar to the previous one.

$$\begin{aligned} C(m,x-y) &= \int_{\mathbb{R}^2} e^{-2\pi i(p,x-y)} \left(|p|^2 + m^2\right)^{-n} dp \\ &= \int_{\mathbb{R}^2} e^{-2\pi i\,(p,x-y)} m^{-2n} \left(\left(|p|/m\right)^2 + 1\right)^{-n} dp. \end{aligned}$$

   Let $\tilde{p} = p/m$, then the equation transforms to

$$C(m,x,y) = m^{-2n} \int_{\mathbb{R}^2} e^{-2\pi i\,(\tilde{p},m(x-y))} (|\tilde{p}|^2 + 1)^{-n} m \, d\tilde{p} = m^{1-2n}\, C(1, m(x-y))$$

   and the claim is proven.

3. Similarly we prove

$$\begin{aligned} C(m,x,y) &= \int_{\mathbb{R}^2} e^{-2\pi i(p,x-y)} \left(|p|^{2n} + m^{2n}\right)^{-1} dp \\ &= \int_{\mathbb{R}^2} e^{-2\pi i\,(p,x-y)} m^{-2n} \left(\left(|p|/m\right)^{2n} + 1\right)^{-1} dp. \end{aligned}$$

   The variable transformation $\tilde{p} = p/m$ leads to

$$C(m,x,y) = m^{-2n} \int_{\mathbb{R}^2} e^{-2\pi i\,(\tilde{p},m(x-y))} (|\tilde{p}|^{2n} + 1)^{-1} m \, d\tilde{p} = m^{1-2n}\, C(1, m(x-y)),$$

   and applying Lemma 1.19 finishes the proof.

□

1.6. Simulations and Statistical Tests

This section shows the results of an implementation in C++ of the algorithms presented in Section 1.4. The simulations were done on a square in $\mathbb{R}^2$. It turns out that both algorithms lead to the same results but that Algorithm 1.18 is twice as fast as Algorithm 1.17.

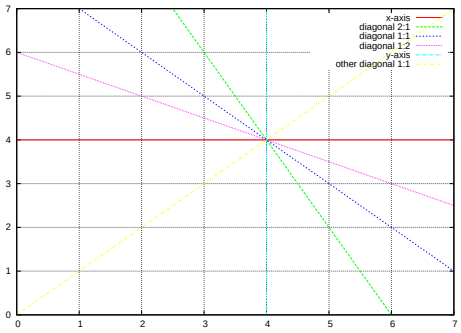
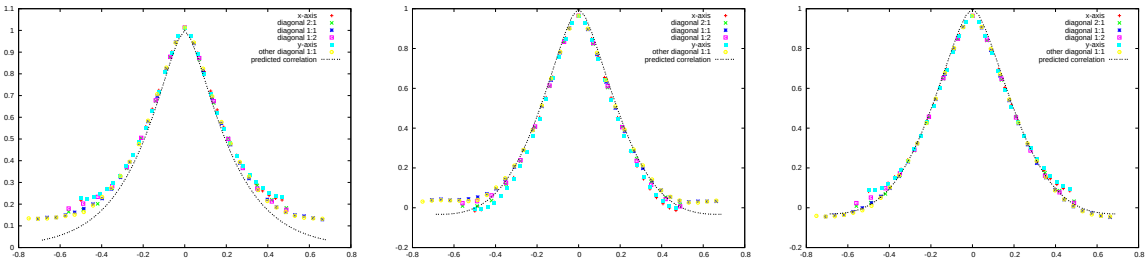


Figure 1.3.: Example of how statistical tests on lines were done using the center point.

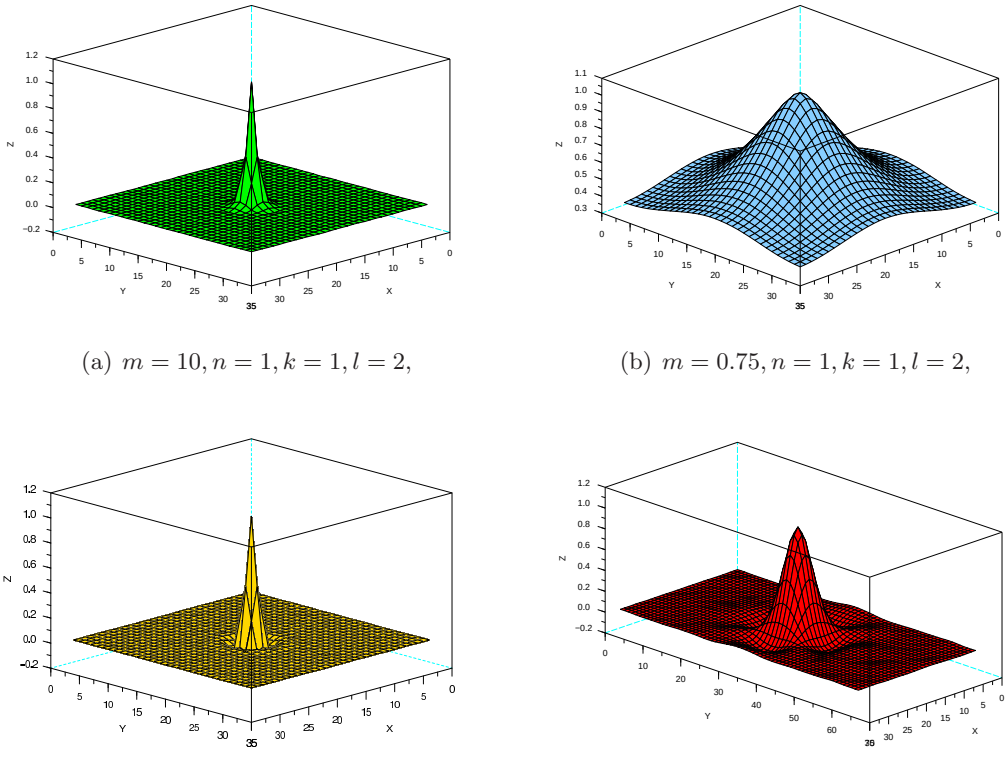
In order to test the covariance or correlation — there is no difference between the two in the simulations because the variance is scaled to 1 — statistical tests on lines were done. This means that one grid point x_0 was picked and the product was calculated with points on the horizontal, the vertical, and different diagonal lines. The lines on which statistical tests where done for the two-dimensional graphics are illustrated in Figure 1.3 in the case of a middle point. The results are shown in Figure 1.4 where the solid line shows the covariance function and the dots are the results of the statistical tests with the simulated random fields. The predictions and the simulations are pretty much the same. In the center part they are the same while at the edges there are some differences which are due to the periodic boundary conditions. We remark that the covariance was plotted as a function of $\mathbb{R}^2$. Therefore it is necessary that the values of the covariance function at the edges are small in order to have very small perturbations of the covariance because of the boundary conditions. In Figure 1.5, four different covariance functions are presented which are the result of simulations. It is an interesting observation that the yellow graph has negative covariance while f is positive. A physical interpretation would be some kind of countermove against the displacement of a point like an elastic band. The red graph shows a covariance function that is not symmetric under rotation. Functions like that might be interesting in physical and engineering applications with various correlations in different directions.

The simulated random fields are visualized in Figure 1.6, 1.7, and 1.8. Two different possibilities how to visualize the data are presented in Figure 1.6. On the left-hand side the random numbers are plotted as the graph of a function from $\mathbb{R}^2$ to $\mathbb{R}$. On the contrary, the figure on the right-hand side presents the random numbers in the form of colors where blue stands for small numbers and red for the large ones. Figure 1.7 uses this hot color map to compare correlated random fields that use the same white noise but different covariance functions. There are differences between all of the pictures but it stands out that there are hardly any differences between the three pictures where f is of degree -4, while Picture (a) where f is of degree -2 is completely different. This phenomenon can be observed for all tested other degrees as well.



(a) $f(p) = (1 + |p|^2)^{-2}$, (b) $f(p) = (1 + p_1^4 + p_2^4)^{-1}$, (c) $f(p) = (1 + |p|^4)^{-1}$,

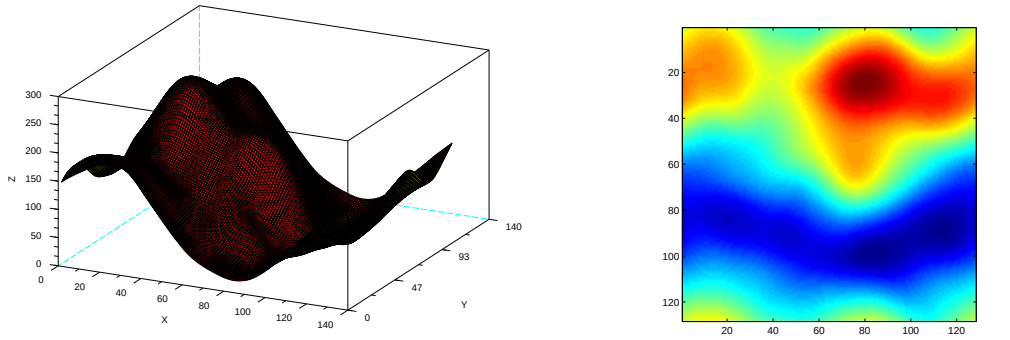
Figure 1.4.: Tested covariances on lines from the center point.



(a) $m = 10, n = 1, k = 1, l = 2$, (b) $m = 0.75, n = 1, k = 1, l = 2$,

(c) $m = 10, n = 3, k = 1, l = 1$, (d) $m = 5, n = 1, k = 3, l = 1$,

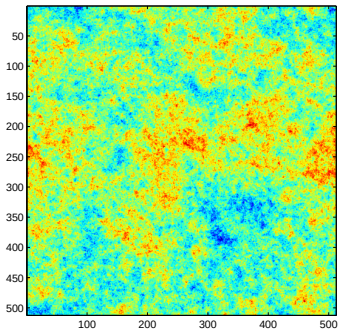
Figure 1.5.: Simulation results using different covariance functions.



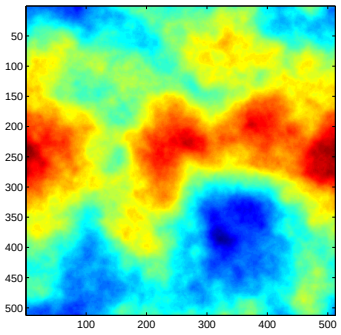
(a) 3-dimensional view, (b) 2-dimensional implicit function,

Figure 1.6.: Two different ways visualizing the same GRF.

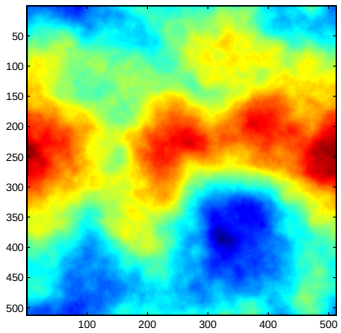
Figure 1.8 shows that using the same white noise and the same type of covariance function with different degrees of f leads to similar random fields but the higher the degree of the polynomial, the smoother the resulting noise. Setting m to something larger than 1 will result in faster decreasing correlations which can be seen in Figure 1.5.



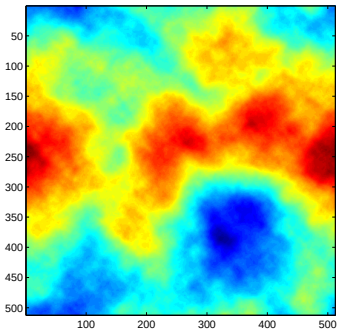
(a) $f(p) = (1 + |p|^2)^{-1}$,



(b) $f(p) = (1 + |p|^2)^{-2}$,

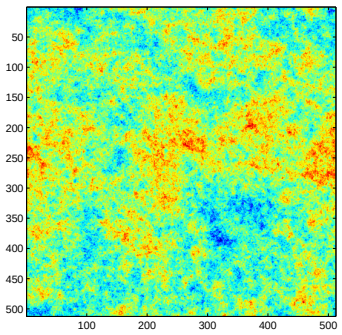


(c) $f(p) = (1 + |p|^4)^{-1}$,

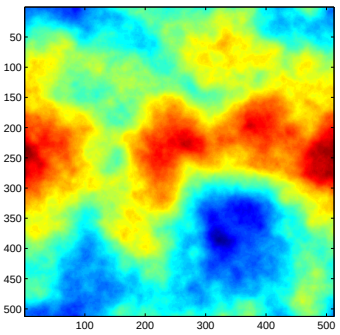


(d) $f(p) = (1 + p_1^4 + p_2^4)^{-1}$,

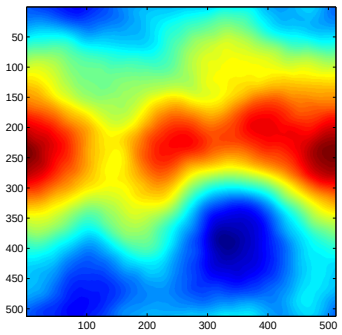
Figure 1.7.: Images of size 512×512 with the same white noise.



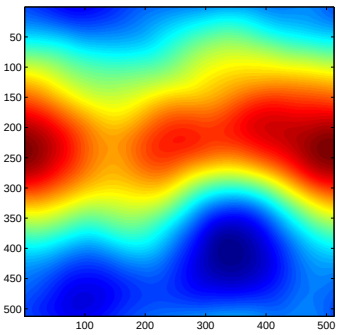
(a) $f(p) = (1 + |p|^2)^{-1}$,



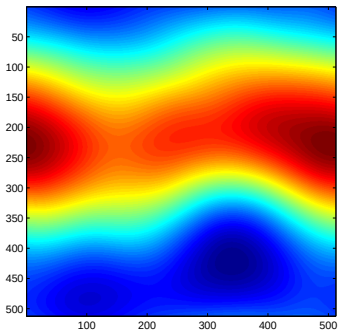
(b) $f(p) = (1 + |p|^2)^{-2}$,



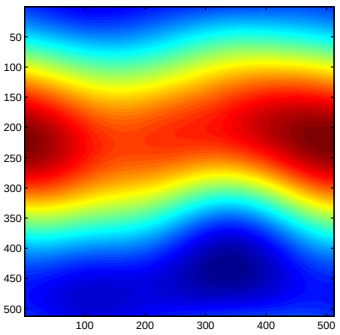
(c) $f(p) = (1 + |p|^2)^{-3}$,



(d) $f(p) = (1 + |p|^2)^{-4}$,



(e) $f(p) = (1 + |p|^2)^{-5}$,



(f) $f(p) = (1 + |p|^2)^{-6}$,

Figure 1.8.: Images of size 512×512 with the same white noise.

2. Simulation of Stochastic Partial Differential Equations

In this chapter we will present stochastic partial differential equations (SPDEs) in the terms of Da Prato and Zabczyk [16]. The discretization of differential equations of this type is discussed including stochastic processes with given covariance in space and time. One section summarizes the literature on convergence of discretizations of SDEs and SPDEs. At the end of this chapter, a definition which is well known for SDEs and a classical definition of numerical stability for PDEs are generalized to SPDEs and tested on SDEs as well as on the heat equation with additive and multiplicative noise.

2.1. SDEs in Infinite Dimensions

This section will give a short introduction to Banach space resp. Hilbert space-valued Wiener processes and integrals. Moreover, stochastic partial differential equations (SPDEs) will be presented in the sense of Hilbert space-valued SDEs and the existence of different types of solutions will be summarized. Further background material can be found in [16].

Assume that E is a separable Banach space and $\mathcal{B}(E)$ is the σ-field of its Borel subsets. Denote the corresponding norm by $\|\cdot\|$. Let $(\Omega, \mathcal{F}, P)$ be a probability space with filtration and let I be an interval of $\mathbb{R}_+$. If $(\Omega, \mathcal{F})$ and $(E, \mathcal{G})$ are two measurable spaces, then a mapping X from Ω into E, such that the set $\{\omega \in \Omega : X(\omega) \in A\} = \{X \in A\}$ belongs to $\mathcal{F}$ for arbitrary $A \in \mathcal{G}$, is called a *random variable* from $(\Omega, \mathcal{F})$ into $(E, \mathcal{G})$. A random variable is called *simple* if it takes on only a finite number of values. An important lemma which is also known from the theory of real-valued random variables is the following.

Lemma 2.1. *Let E be a separable metric space with metric ρ and let X be an E-valued random variable. Then there exists a sequence $(X_m)_{m\in\mathbb{N}}$ of simple E-valued random variables such that, for arbitrary $\omega \in \Omega$, the sequence $(\rho(X(\omega), X_m(\omega)))_{m\in\mathbb{N}}$ is monotonically decreasing to zero.*

Before starting with stochastic processes we define independence of Banach space-valued random variables and how to integrate them.

Definition 2.2. Let $\{\mathcal{F}_i\}_{i\in I}$ be a family of sub σ-fields of $\mathcal{F}$. These σ-fields are said to be *independent*, if for every finite subset $J \subset I$ and every family $\{A_i\}_{i\in J}$ such that $A_i \in \mathcal{F}_i$,

$$P\Big(\bigcap_{i\in J} A_i\Big) = \prod_{i\in J} P(A_i).$$

Random variables $\{X_i\}_{i\in I}$ are *independent* if the σ-fields $\{\sigma(X_i)\}_{i\in I}$ are independent.

Assume now that E is a separable Banach space. For a simple random variable X

$$X = \sum_{i=1}^{N} x_i \, \chi_{A_i}, \quad A_i \in \mathcal{F},\ x_i \in E,$$

set

$$\int_B X(\omega)\, dP(\omega) = \int_B X\, dP = \sum_{i=1}^{N} x_i\, P(A_i \cap B)$$

for all $B \in \mathcal{F}$. Here χ_{A_i} denotes the indicator function of A_i. The integral for general random variables uses Lemma 2.1 which states that every random variable can be approximated by a sequence of simple random variables $\{X_n\}_{n\in\mathbb{N}}$. Under the assumption that X is *Bochner integrable* or simply *integrable*, i.e.

$$\int_\Omega \| X(\omega) \|\, dP(\omega) < +\infty,$$

the integral of X is defined by

$$\int_B X(\omega)\, dP(\omega) = \lim_{n\to\infty} \int_B X_n(\omega)\, dP(\omega)$$

where $X_n \to X$ and $B \in \mathcal{F}$. This limit exists because the sequence of simple random variables (X_m) satisfies that $(\| X(\omega) - X_m(\omega) \|)$ decreases to zero for all $\omega \in \Omega$ and

$$\begin{aligned} \| \int_\Omega X_m(\omega) dP(\omega) - \int_\Omega X_n(\omega) dP(\omega) \| \\ \leq \int_\Omega \| X(\omega) - X_m(\omega) \|\, dP(\omega) + \int_\Omega \| X(\omega) - X_n(\omega) \|\, dP(\omega) \downarrow 0 \end{aligned}$$

for $m, n \to +\infty$.

Definition 2.3. A family $X = \{X(t)\}_{t\in I}$ of E-valued random variables $X(t)$, defined on Ω, is called a *stochastic process*. Set $X(t,\omega) = X(t)(\omega)$ for all $t \in I$ and $\omega \in \Omega$. The functions $X(\cdot,\omega)$ are called the *trajectories* of $X(t)$. A stochastic process Y is called a *modification* of X if

$$P(\omega \in \Omega : X(t,\omega) \neq Y(t,\omega)) = 0 \qquad \forall t \in I.$$

In the following we will summarize several definitions of regularity for Banach space-valued processes which are similar to those known from real-valued stochastic processes.

Definition 2.4.

1. X is *measurable* if the mapping $X(\cdot,\cdot) : I \times \Omega \to E$ is $\mathcal{B}(I) \otimes \mathcal{F}$-measurable.
2. X is *stochastically continuous* at $t_0 \in I$ if for all $\varepsilon > 0$ and all $\delta > 0$, there exists $\rho > 0$ such that

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

3. X is *stochastically continuous* in I if it is stochastically continuous at every point of I.
4. X is *stochastically uniformly continuous* in $I = [0,T]$ if for all $\varepsilon > 0$ and all $\delta > 0$, there exists $\rho > 0$ such that

$$P(\| X(t) - X(s) \| \geq \varepsilon) \leq \delta, \qquad \forall t, s \in [0,T],\ |t - s| < \rho.$$

5. X is *mean square continuous* at $t_0 \in I$ if
$$\lim_{t \to t_0} \mathbb{E}(\| X(t) - X(t_0) \|^2) = 0.$$
6. X is *mean square continuous* in I if it is mean square continuous at every point of I.
7. X is *continuous with probability* 1 (or *continuous*) if its trajectories $X(\cdot, \omega)$ are continuous almost surely.
8. X is *α-Hölder continuous with probability* 1 (or *α-Hölder continuous*) if its trajectories $X(\cdot, \omega)$ are α-Hölder continuous almost surely.

Theorem 2.5 (Kolmogorov-Chentsov)**.** *Let $X(t)$, $t \in [0,T]$ be a stochastic process with values in a separable Banach space E such that, for some positive constants $C > 0$, $\varepsilon > 0$, $\delta > 1$ and all $t, s \in [0,T]$*
$$\mathbb{E}(\| X(t) - X(s) \|^\delta) \leq C |t - s|^{1+\varepsilon}.$$
Then there exists a modification of X with P-almost all trajectories being Hölder continuous functions with an arbitrary exponent smaller than ε/δ. In particular X has a continuous modification.

From this point on, we will focus on Hilbert spaces. If the results also apply to more general spaces, this will be mentioned in the corresponding theorems. In order to be able to construct the Hilbert space-valued stochastic integral next, we need some more prerequisites. Therefore let H and U be two Hilbert spaces and $Q \in L(U)$ a symmetric nonnegative operator. Furthermore let $\operatorname{tr}(Q) < +\infty$. Then there exists a complete orthonormal system $\{e_k\}$ in U, and a bounded sequence of nonnegative real numbers λ_k such that
$$Qe_k = \lambda_k e_k, \qquad k = 1, 2, \ldots$$

Proposition 2.6. *Assume that E and F are separable Banach spaces and $A : D(A) \subset E \to F$ is a linear closed operator with domain $D(A)$ which is a Borel subset of E. If $X : [0,T] \times \Omega \to E$ is a stochastic process such that $X(t,\omega) \in D(A)$ almost surely, then AX is an F-valued random variable and X is a $D(A)$-valued random variable, where $D(A)$ is endowed with the graph norm of A, i.e. $\| x \|_{D(A)} := \| x \|_E + \| Ax \|_F$. If moreover*
$$\int_0^t \| AX(s,\omega) \| \, ds < +\infty, \qquad t \in [0,T],$$
then
$$A \int_0^t X(s,\omega)\, ds = \int_0^t AX(s,\omega)\, ds. \tag{2.1}$$
Furthermore for $X \in D(A^i)$, $i = 1, \ldots, k$, $k \in \mathbb{N}$, $A : D(A) \to F := E$ and
$$\int_0^t \left\| A^i X(s,\omega) \right\| ds < +\infty, \quad i = 1, \ldots, k,$$
it holds
$$A^k \int_0^t X(s,\omega)\, ds = \int_0^t A^k X(s,\omega)\, ds.$$

Proof. As $D(A)$ is a separable metric space endowed with the graph norm, by Lemma 2.1 there exists a sequence $(X_m)_{m\in\mathbb{N}}$ of simple $D(A)$-valued random variables with

$$\| X_m(s,\omega) - X(s,\omega) \|_{D(A)} = \| X_m(s,\omega) - X(s,\omega) \|_E + \| AX_m(s,\omega) - AX(s,\omega) \|_F \downarrow 0$$

decreases monotonicly to zero for $m \to +\infty$ and for all $s \in [0,T]$, $\omega \in \Omega$.

Consequently, the application of the monotone convergence theorem [5, Thm. 11.4],

$$\int_0^t \| X_m(s,\omega) - X(s,\omega) \|_E \, ds + \int_0^t \| AX_m(s,\omega) - AX(s,\omega) \|_F \, ds \downarrow 0$$

decreases monotonically to zero as well. Then

$$\int_0^t X_m(s,\omega)\, ds \to \int_0^t X(s,\omega)\, ds \quad \text{in } E,$$
$$\int_0^t AX_m(s,\omega)\, ds \to \int_0^t AX(s,\omega)\, ds \quad \text{in } F$$

because

$$\left\| \int_0^t X_m(s,\omega)\, ds - \int_0^t X(s,\omega)\, ds \right\|_E \leq \int_0^t \| X_m(s,\omega) - X(s,\omega) \|_E \, ds \downarrow 0$$

and similarly for the second integral. The definition of the integral and the linearity of A imply

$$\int_0^t AX_m(s,\omega)\, ds = A \int_0^t X_m(s,\omega)\, ds.$$

Therefore the facts that $\int_0^t X_m(s,\omega)\, ds$ is a Cauchy sequence in $D(A)$ and convergent in E, $A\int_0^t X_m(s,\omega)\, ds$ is convergent in F to $\int_0^t AX(s,\omega)\, ds$, and that A is a closed operator imply that $\int_0^t X(s,\omega)\, ds \in D(A)$ and that

$$A\int_0^t X(s,\omega)\, ds = \lim_{m\to+\infty} A\int_0^t X_m(s,\omega)\, ds = \lim_{m\to+\infty} \int_0^t AX_m(s,\omega)\, ds = \int_0^t AX(s,\omega)\, ds.$$

The last claim follows by the application of Equation (2.1) recursively. □

In the more general case of a separable Banach space E, Da Prato and Zabczyk define in [16] that a probability measure μ on $(E, \mathcal{B}(E))$ is said to be a *Gaussian measure*, if and only if the law of an arbitrary linear function $h \in E^\star$ considered as a random variable on $(E, \mathcal{B}(E), \mu)$, is a Gaussian measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Following this approach and observing that in a Hilbert space $\mathcal{H}$ with inner product $(\cdot,\cdot)$ the Hilbert space $\mathcal{H}$ and its dual $\mathcal{H}^\star$ are isomorphic, we have the following definition.

Definition 2.7. A Hilbert space-valued random variable $X : \Omega \to \mathcal{H}$ is called *Gaussian* with law $\mathcal{N}(\mu, Q)$, if and only if for all $h \in \mathcal{H}$, (h, X) has law $\mathcal{N}((h,\mu), (h, hQh))$ in $\mathbb{R}$. Furthermore a finite family of random variables $(X_1, \ldots, X_n)$ is called *Gaussian*, if for all $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$, $\sum_{i=1}^n \alpha_i X_i$ is Gaussian in the sense of the first phrase. And finally a family of random variables $(X_i)_{i\in\mathbb{N}}$ is called *Gaussian*, if any finite subset is Gaussian.

This definition of Hilbert space-valued Gaussian random variables implies for the covariance Q the following.

Proposition 2.8 ([16, Prop. 2.15])**.** *Let μ be a Gaussian probability measure on a Hilbert space with mean 0 and covariance Q. Then Q is a trace class operator, i.e.* $\operatorname{tr} Q = \mathbb{E}(\| X \|^2) < +\infty$.

For real-valued random variables it is known that a convergent sequence of Gaussian random variables $(X_i)_{i\in\mathbb{N}}$ is Gaussian in the limit, see e.g. [41, 70]. In Chapter 2 we will need a similar statement for Hilbert space-valued random variables. This is proved in the following using the definition above.

Lemma 2.9. *Let $\mathcal{H}$ be a Hilbert space and $(X_n)_{n\in\mathbb{N}}$ a sequence of $\mathcal{H}$-valued Gaussian random variables converging in $L^2(\Omega)\otimes\mathcal{H}$ to X, i.e.*

$$\|X_n - X\|_2^2 = \mathbb{E}(\| X_n - X \|^2) \longrightarrow 0$$

for $n \to +\infty$. Then X is Gaussian distributed.

Proof. Using Definition 2.7 it suffices to show that for all $h \in \mathcal{H}$ the real random variable (h, X) is Gaussian. Therefore we can reduce the problem to the real case and apply Theorem A.7 of [70]. All we have to check is the convergence of (h, X_n) to (h, X) in $L^2(\Omega)$. This is done in the following using the Cauchy-Schwarz inequality.

$$0 \le \mathbb{E}(((h, X_n) - (h, X))^2) = \mathbb{E}((h, X_n - X)^2) \le \| h \|^2 \mathbb{E}(\| X_n - X \|^2) \longrightarrow 0$$

for $n \to +\infty$ by the assumptions on X. Therefore the claim is proved. □

Definition 2.10. A U-valued stochastic process $W(t)$, $t \ge 0$ is called a *Q-Wiener process* on $[0, T]$, if

1. $W(0) = 0$,
2. W has continuous trajectories,
3. W has independent increments,
4. the increments $W(t) - W(s)$ are $\mathcal{N}(0, (t-s)Q)$ distributed for $t \ge s \ge 0$.

An important result for simulations is the following proposition.

Proposition 2.11. *Assume that W is a Q-Wiener process with* $\operatorname{tr} Q < +\infty$. *Then the following statements hold:*

1. *W is a Gaussian process on U and*

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

2. *For arbitrary t, W has the series expansion*

$$W(t) = \sum_{j=1}^{\infty} \sqrt{\lambda_j}\beta_j(t)e_j, \tag{2.2}$$

where

$$\beta_j(t) = \frac{1}{\sqrt{\lambda_j}}\langle W(t), e_j\rangle, \quad j = 1, 2, \ldots,$$

and the elements β_j are independent real-valued Brownian motions. The series in (2.2) is convergent in $L^2(\Omega, \mathcal{F}, P)$.

Definition 2.12. Let $\{\mathcal{F}_t\}_{t\geq 0}$ be a normal filtration in $\mathcal{F}$, i.e.

1. $\mathcal{F}_0$ contains all $A \in \mathcal{F}$ such that $P(A) = 0$,
2. $\mathcal{F}_t = \mathcal{F}_{t^+}$, $\forall\, t \in T$, where $\mathcal{F}_{t^+}$ is the intersection of all $\mathcal{F}_s$ where $s > t$.

Moreover assume that for $t \geq 0$

1. $W(t)$ is $\mathcal{F}_t$-measurable,
2. $W(t+h) - W(t)$ is independent of $\mathcal{F}_t$, $\forall\, h \geq 0$,

then W is called a *Q-Wiener process with respect to $\{\mathcal{F}_t\}_{t\geq 0}$*.

Definition 2.13. Let $\mathcal{P}_\infty$ be the σ-field generated by sets of the form

$$(s,t] \times F, \quad 0 \leq s < t < +\infty,\ F \in \mathcal{F}_s \qquad \text{and } \{0\} \times F, \quad f \in \mathcal{F}_0.$$

This σ-field is called a *predictable σ-field* and its elements *predictable sets*. The restriction of the σ-field $\mathcal{P}_\infty$ to $[0,T] \times \Omega$ will be denoted by $\mathcal{P}_T$. an arbitrary measurable mapping from $([0,+\infty) \times \Omega, \mathcal{P}_\infty)$ or $([0,T] \times \Omega, \mathcal{P}_T)$ into $(E, \mathcal{B}(E))$ is called a *predictable process*.

Lemma 2.14. *Let A be a linear closed operator and for all $t \geq 0$ let $W(t) \in D(A)$. Assume that* $\operatorname{tr} Q^{1/2}A^2Q^{1/2} < +\infty$. *Then $AW(t)$ is a Gaussian process with mean zero and covariance $tQ^{1/2}A^2Q^{1/2}$.*

Proof. To show that $AW(t)$ is Gaussian we first remark that

$$\mathbb{E}\big(\,\big\|\sum_{j=1}^{n} \sqrt{\lambda_j}\beta_j(t)e_j - \sum_{j=1}^{\infty} \sqrt{\lambda_j}\beta_j(t)e_j\big\|^2\big) \leq t \sum_{j=n+1}^{\infty} \lambda_j \,\|\, Ae_j \,\|^2 \downarrow 0$$

because $\lambda_j \geq 0$ for all $j \in \mathbb{N}$ and the process is in the domain of A. Applying Lemma 2.9, we have that $AW(t)$ is Gaussian.

Next we observe that because of linearity

$$A\sum_{j=1}^{n} \sqrt{\lambda_j}\beta_j(t)e_j = \sum_{j=1}^{n} \sqrt{\lambda_j}\beta_j(t)Ae_j,$$

and the closedness of A implies

$$A\sum_{j=1}^{\infty}\sqrt{\lambda_j}\beta_j(t)e_j=\sum_{j=1}^{\infty}\sqrt{\lambda_j}\beta_j(t)Ae_j.$$

In order to interchange the converging sum and the expectation we observe that

$$\|\,f_n\,\|:=\big\|\sum_{j=1}^{n}\sqrt{\lambda_j}\beta_j(t)e_j\big\|\le\|\,W(t)\,\|=:\|\,f\,\|$$

by Bessel's inequality and

$$\Big\|\int_\Omega(f_n-f)\,d\mu\Big\|\le\int_\Omega\|\,f_n-f\,\|\;d\mu\le\int_\Omega\|\,f_n\,\|+\|\,f\,\|\;d\mu\le 2\int_\Omega\|\,f\,\|\;d\mu<+\infty.$$

Now we can apply the dominated convergence theorem [5, Thm. 15.6] and we have

$$\mathbb{E}(AW(t))=\mathbb{E}\big(\sum_{j=1}^{\infty}\sqrt{\lambda_j}\beta_j(t)e_j\big)=\sum_{j=1}^{\infty}\sqrt{\lambda_j}\mathbb{E}(\beta_j(t))e_j=0.$$

Finally, to calculate the covariance operator we use the independence of the Brownian motions. This leads to

$$\begin{aligned}\mathrm{Cov}\,(A(W(t))&=\mathbb{E}((AW(t))^{\otimes 2})=\mathbb{E}\big(\sum_{j,k=1}^{\infty}\beta_j(t)\beta_k(t)A\sqrt{\lambda_j}e_j\otimes A\sqrt{\lambda_k}e_k\big)\\&=\sum_{j,k=1}^{\infty}\mathbb{E}(\beta_j(t)\beta_k(t))AQ^{1/2}e_j\otimes AQ^{1/2}e_k)=\sum_{j=1}^{\infty}t(AQ^{1/2}e_j)^{\otimes 2}\\&=t\sum_{j=1}^{\infty}e_j\otimes Q^{1/2}A^2Q^{1/2}e_j=t\;Q^{1/2}A^2Q^{1/2},\end{aligned}$$

because Q is positive definite and the convergence follows from the assumptions. □

Next we will define the stochastic integral for elementary processes. An $L=L(U,H)$-valued process $\Phi(t)$, $t\in[0,T]$, is said to be *elementary*, if there exist a sequence $0=t_0<t_1<\cdots<t_k=T<+\infty$ and a sequence $\Phi_0,\Phi_1,\ldots,\Phi_{k-1}$ of L-valued random variables taking only a finite number of values such that Φ_m is $\mathcal{F}_{t_m}$-measurable and

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

and $\Phi(0)=\Phi_0$.

Definition 2.15. The *stochastic integral* for elementary processes Φ is defined by the formula

$$\int_0^t\Phi(s)\,dW(s):=\sum_{m=0}^{k-1}\Phi_m(W(t_{m+1}\wedge t)-W(t_m\wedge t))$$

and denoted by $\Phi\cdot W(t)$, $t\in[0,T]$.

Define another Hilbert space $U_0 := Q^{1/2}(U) \subset U$ with inner product

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

where $\langle\cdot,\cdot\rangle$ is the inner product on the Hilbert space U. Moreover let $L_2^0 = L_2(U_0, H)$ be the space of all Hilbert–Schmidt operators from U_0 into H which can be proved to be a separable Hilbert space equipped with the norm

$$\| \Psi \|_{L_2^0}^2 = \sum_{h,k=1}^{\infty} |\langle \Psi g_h, f_k\rangle|^2 = \sum_{h,k=1}^{\infty} |\langle \Psi e_h, f_k\rangle|^2 = \|\Psi\, Q^{1/2}\|^2 = \mathrm{tr}[\Psi\, Q\, \Psi^\star],$$

where $\{g_j := \sqrt{\lambda_j} e_j\}$, $\{e_j\}$, and $\{f_j\}$, $j = 1, 2, \ldots$, are complete orthonormal bases in U_0, U, and H, respectively. For a measurable L_2^0-valued process $\Phi(t)$, $t \in [0, T]$, define the norm

$$|\!|\!| \Phi |\!|\!|_t := \mathbb{E}\Big(\int_0^t \| \Phi(s) \|_{L_2^0}^2 \, ds\Big)^{1/2} = \mathbb{E}\Big(\int_0^t \mathrm{tr}(\Phi(s)Q^{1/2})(\Phi(s)Q^{1/2})^\star\Big)^{1/2}, \qquad t \in [0, T].$$

After these definitions, we can now formulate a Hilbert space-valued Itô isometry for elementary processes.

Proposition 2.16. *If a process Φ is elementary and $|\!|\!| \Phi |\!|\!|_T < +\infty$, then the process $\Phi \cdot W$ is a continuous, square integrable H-valued martingale on $[0, T]$ and*

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

To extend the definition of the stochastic integral to more general processes, it is convenient to regard integrands as random variables defined on the product space $\Omega_T = [0, T] \times \Omega$ equipped with the product σ-field $\mathcal{B}([0, T]) \otimes \mathcal{F}$. The product of the Lebesgue measure on $[0, T]$ and the probability measure P is denoted by P_T. Finally Proposition 4.7 and Lemma 4.8 in [16] allow to extend the definition of the stochastic integral to all L_2^0 predictable processes Φ with $|\!|\!| \Phi |\!|\!|_T < +\infty$. These processes form a Hilbert space which will be denoted by $\mathcal{N}_W^2(0, T; L_2^0) = \mathcal{N}_W^2(0, T) = \mathcal{N}_W^T$, and the elementary processes form a dense set in this space. By a *localization procedure* the definition of the stochastic integral can be extended to L_2^0 predictable processes satisfying the weaker condition

$$P\Big(\int_0^T \| \Phi(s) \|_{L_2^0}^2 \, ds < +\infty\Big) = 1.$$

These processes are called *stochastically integrable* on $[0, T]$. The linear space formed by these processes is denoted by $\mathcal{N}_W(0, T; L_2^0) = \mathcal{N}_W(0, T) = \mathcal{N}_W$.

Two important properties of the stochastic integral are formulated in the the following:

Proposition 2.17. *Assume that $\Phi_1, \Phi_2 \in \mathcal{N}_W^2(0, T; L_0^2)$. Then*

$$\mathbb{E}(\Phi_i \cdot W(t)) = 0, \qquad \mathbb{E}(\| \Phi_i \cdot W(t) \|^2) < +\infty, \qquad t \in [0, T],\ i = 1, 2.$$

Moreover, the covariance operators are given by the formula

$$\mathrm{Cov}\,(\Phi_1 \cdot W(t), \Phi_2 \cdot W(s)) = \mathbb{E}\Big(\int_0^{t\wedge s} (\Phi_2(r)Q^{1/2})(\Phi_1(r)Q^{1/2})^\star \, dr\Big).$$

Corollary 2.18. *Under the hypotheses of the previous proposition, we have*

$$\mathbb{E}(\langle \Phi_1 \cdot W(t), \Phi_2 \cdot W(s)\rangle) = \mathbb{E}\Big(\int_0^{t\wedge s} \operatorname{tr}[(\Phi_2(r)Q^{1/2})(\Phi_1(r)Q^{1/2})^\star]\,dr\Big).$$

Using this integration theory, we will next state some results on the existence and uniqueness of solutions for linear SPDEs. Therefore let $W(t)$ be a Q-Wiener process as defined above. Consider the linear equation

$$\begin{cases} dX(t) = [AX(t) + f(t)]\,dt + B\,dW(t), \\ X(0) = \xi, \end{cases} \tag{2.3}$$

where $A : D(A) \subset H \to H$ and $B : U \to H$ are linear operators, f is an H-valued stochastic process.

Assumption 2.19. Assume that the deterministic Cauchy problem

$$\begin{cases} u'(t) = Au(t), \\ u(0) = x \in H \end{cases}$$

is uniformly well posed and that B is bounded, that is

1. A generates a strongly continuous semigroup $S(\cdot)$ in H, which is explained in Appendix A,
2. $B \in L(U, H)$.

Assumption 2.20. Furthermore we require that

1. f is a predictable process with Bochner integrable trajectories on an arbitrary finite interval $[0, T]$,
2. ξ is $\mathcal{F}_0$-measurable.

Remark 2.21. If W is a Q-Wiener process in U, then $W_1 = BW$ is a $BQB^\star$-Wiener process in H. So we could assume, without loss of generality, that $U = H$.

Definition 2.22. An H-valued predictable process $X(t)$, $t \in [0, T]$, is said to be a *strong solution* to Equation (2.3) if X takes values in $D(A)$, P_T-a.s., $\int_0^T |AX(s)|\,ds < +\infty$, P-a.s., and for $t \in [0, T]$,

$$X(t) = x + \int_0^t [AX(s) + f(s)]\,ds + BW(t), \qquad P\text{-a.s.}$$

This definition is only meaningful, if BW is a U-valued process and therefore requires that $\operatorname{tr} BQB^\star < +\infty$.

Moreover, an H-valued predictable process $X(t)$, $t \in [0, T]$, is said to be a *weak solution* of Equation (2.3) if the trajectories of $X(\cdot)$ are P-a.s. Bochner integrable and if for all $\zeta \in D(A^\star)$ and all $t \in [0, T]$ we have

$$\langle \zeta, X(t)\rangle = \langle \zeta, x\rangle + \int_0^t [\langle A^\star\zeta, X(s)\rangle + \langle \zeta, f(s)\rangle]\,ds + \langle \zeta, BW(t)\rangle, \qquad P\text{-a.s.}.$$

Theorem 2.23. *Let the* stochastic convolution W_A *be given by*

$$W_A(t) := \int_0^t S(t-s)B\,dW(s),$$

let Assumption 2.19 hold, and

$$\int_0^T \| S(r)B \|_{L_2^0}^2\,dr = \int_0^T \operatorname{tr}[S(r)BQB^\star S^\star(r)]\,dt < +\infty. \tag{2.4}$$

Then

1. *the process $W_A(\cdot)$ is Gaussian, continuous in mean square and has a predictable modification,*
2. *it holds*

$$\operatorname{Cov}(W_A(t)) = \int_0^t S(r)BQB^\star S^\star(r)\,dr, \qquad t \in [0,T].$$

Theorem 2.24. *Let Assumption 2.19, Assumption 2.20, and Equation (2.4) hold. Then Equation (2.3) has exactly one weak solution which is given by the formula*

$$X(t) = S(t)\xi + \int_0^t S(t-s)f(s)\,ds + \int_0^t S(t-s)B\,dW(s), \qquad t \in [0,T].$$

Lemma 2.25. *Let A be the generator of a semigroup. Assume* $\operatorname{tr} Q < +\infty$ *and set*

$$Y(t) := \int_0^t S(t-s)W(s)\,ds, \quad t \geq 0.$$

Then $Y(\cdot)$ belongs to $C^1([0,\infty); D(A))$, P-a.s., and

$$\begin{aligned} W_A(t) &= W(t) + A\int_0^t S(t-s)W(s)\,ds \\ &= \frac{\partial}{\partial t}\int_0^t S(t-s)W(s)\,ds = \frac{\partial}{\partial t}Y(t), \quad t \geq 0. \end{aligned}$$

Lemma 2.26. *Let A be a linear closed generator of a semigroup. For fixed $k \in \mathbb{N}$ let*

$$\int_0^t \left\| A^{k+1}S(t-s)W(s) \right\|_2 ds < +\infty,$$

where $W(s) \in D(A^i)$, $i = 1,\ldots,k+1$ and $S(t-s)W(s) \in D(A^{k+1})$. Then

$$A^k \int_0^t S(t-s)\,dW(s) = A^k W(t) + \int_0^t S(t-s)A^{k+1}W(s)\,ds.$$

Proof. First an application of Lemma 2.25 and the linearity of A lead to

$$A^k \int_0^t S(t-s)\,dW(s) = A^k(W(t)+A\int_0^t S(t-s)W(s)\,ds) = A^kW(t)+A^{k+1}\int_0^t S(t-s)W(s)\,ds.$$

Then Proposition 2.6 yields

$$A^k \int_0^t S(t-s)\,dW(s) = A^kW(t) + \int_0^t A^{k+1}S(t-s)W(s)\,ds.$$

And finally Theorem 1.2.4 in [74] implies

$$A^k \int_0^t S(t-s)\,dW(s) = A^kW(t) + \int_0^t S(t-s)A^{k+1}W(s)\,ds.$$

□

Theorem 2.27. *Assume that*

1. $\operatorname{tr} Q < +\infty$, $U = H$, $B = \mathbb{1}$, $\operatorname{Im} Q^{1/2} \subset D(A)$, *and* $\operatorname{tr}(AQ^{1/2}) < +\infty$),
2. $f \in C^1([0,T];H) \cap C([0,T];D(A))$, *P-a.s.*,
3. $\xi \in D(A)$, *P-a.s.*.

Then Problem (2.3) has a strong solution.

Theorem 2.28. *Assume that*

1. $(-A)^\beta \in L_2(H)$ *for some* $\beta \in (\frac{1}{2}, 1)$,
2. $f \in C^\alpha([0,T];H) \cap C([0,T];D_A(\alpha,\infty))$ *for some* $\alpha \in (0,1)$, *P-a.s.*,
3. $\xi \in D(A)$, *P-a.s.*.

Then Problem (2.3) has a strong solution.

Having considered linear SPDEs with additive noise, we will next look at results about those with multiplicative noise. Therefore let the following Cauchy problem be given

$$\begin{cases} dX(t) = [AX(t) + f(t)]\,dt + B(X(t))\,dW(t), \\ X(0) = \xi, \end{cases} \tag{2.5}$$

on a time interval $[0,T]$, where $A : D(A) \subset H \to H$ is the infinitesimal generator of a strongly continuous semigroup $S(\cdot)$, ξ is an H-valued, $\mathcal{F}_0$-measurable random variable, f is a predictable process with locally integrable trajectories, and $B : D(B) \to L_2^0$ is a linear operator.

Let $\{g_j\}$ be a complete orthonormal basis in $U_0 := Q^{1/2}U$, where U is a Hilbert space. Since for arbitrary $x \in D(B)$, B is a Hilbert–Schmidt operator from U_0 into H,

$$\sum_{j=1}^{\infty} |B(x)g_j|^2 < +\infty, \quad x \in D(B).$$

The operators

$$B_j x = B(x)g_j, \qquad x \in D(B),\ j = 1, 2, \dots$$

are linear and

$$B(x)u = \sum_{j=1}^{\infty} B_j x \langle u, g_j \rangle_{U_0}, \qquad x \in D(B),\ u \in U_0.$$

Consequently if

$$W(t) = \sum_{j=1}^{\infty} \beta_j(t) g_j,$$

then Equation (2.5) can equivalently be written as

$$\begin{cases} dX(t) = [AX(t) + f(t)]\, dt + \sum_{j=1}^{\infty} B_j X(t)\, d\beta_j(t), \\ X(0) = \xi, \end{cases}$$

Similar to Definition 2.22 for additive noise, we define strong, weak, and mild solutions.

Definition 2.29. A *strong solution* of Problem (2.5) is an H-valued predictable process $X(t)$, $t \in [0, T]$, which takes values in $D(A) \cap D(B)$, P_T-a.s. such that

$$P\Big(\int_0^T (|X(s)| + |AX(s)|)\, ds < +\infty\Big) = 1,$$
$$P\Big(\int_0^T \| B(X(s)) \|_{L_2^0}^2\, ds < +\infty\Big) = 1,$$

and, for arbitrary $t \in [0, T]$ and P-a.s.,

$$X(t) = \xi + \int_0^t [AX(s) + f(s)]\, ds + \int_0^t B(X(s))\, dW(s).$$

An H-valued predictable process $X(t)$, $t \in [0, T]$, is said to be a *weak solution* to Problem (2.5) if X takes values in $D(B)$, P_T-a.s.,

$$P\Big(\int_0^T |X(s)|\, ds < +\infty\Big) = 1, \tag{2.6}$$
$$P\Big(\int_0^T \| B(X(s)) \|_{L_2^0}^2\, ds < +\infty\Big) = 1, \tag{2.7}$$

and, for arbitrary $t \in [0, T]$ and $\zeta \in D(A^\star)$,

$$\langle \zeta, X(t) \rangle = \langle \zeta, \xi \rangle + \int_0^t [\langle A^\star \zeta, X(s) \rangle + \langle \zeta, f(s) \rangle]\, ds + \int_0^t \langle \zeta, B(X(s)) \rangle\, dW(s), \qquad P\text{-a.s..}$$

An H-valued predictable process $X(t)$, $t \in [0, T]$ is called a *mild solution* to (2.5), if X takes values in $D(B)$, P_T-a.s., (2.6) and (2.7) hold, and, for arbitrary $t \in [0, T]$,

$$X(t) = S(t)\xi + \int_0^t S(t - s) f(s)\, ds + \int_0^t S(t - s) B(X(s))\, dW(s).$$

Theorem 2.30. *Assume that $A : D(A) \subset H \to H$ is the infinitesimal generator of a C_0-semigroup $S(\cdot)$ in H. Then a strong solution is always a weak solution and a weak solution is a mild solution of Problem (2.5). Conversely if X is a mild solution of (2.5) and*

$$\mathbb{E}(\int_0^T \| B(X(s)) \|_{L_2^0}^2 \, ds) < +\infty,$$

then X is also a weak solution of (2.5).

Theorem 2.31. *Assume that A is the infinitesimal generator of a C_0-semigroup $S(\cdot)$ in H, $\mathbb{E}(|\xi|^2) < +\infty$, and $B \in L(H, L_2^0)$. Then Equation (2.5) has a unique mild solution $X \in \mathcal{N}_W^2(0, T; H)$, identical with a weak solution.*

Proposition 2.32. *Assume that the hypotheses of Theorem 2.31 hold, $\xi = x \in D(A)$, and $f \equiv 0$. Let moreover $0 \in \rho(A)$ and B_A, given by*

$$B_A(x)u := AB(A^{-1}x)u, \qquad x \in H, \ u \in U,$$

belong to $L(H, L_2^0)$, where $\rho(A)$ denotes the resolvent set of A. Then Equation (2.5) has a unique strong solution.

2.2. Numerical Discretization of SPDEs

For the implementation of an SPDE using a computer, we have to discretize the deterministic part of the equation in space and in time. Moreover we have to approximate operators and the Wiener process. The following subsections explain shortly how discretization is done in this thesis. We will just focus on simple schemes, but there exists a wide range of discretizations of SPDEs. The following methods can for instance be used for solving SPDEs numerically: finite difference methods [2, 17, 33, 34, 38, 75, 80, 88, 95, 96, 97], the splitting up method [6, 7, 22, 35, 37, 45, 69], Galerkin approximation [31, 63], finite element methods [28], spectral methods [88].

2.2.1. Discretization of PDEs

The approximation of deterministic partial differential equations has been studied a lot. References for possible discretization schemes are e.g. [19, 32, 53, 55, 56, 77, 80, 85].

In the examples of this chapter we will focus on explicit finite difference schemes for the deterministic part and we will compare the known bounds for the deterministic heat equation to those of the corresponding SPDEs with additive and multiplicative noise. It turns out that we get similar results but the constants are not exactly the same. For an introduction to finite difference methods see for instance [80]. On a finite interval $[0, b]$ we will use an explicit scheme in time and we will discretize the Laplacian by the three point formula, i.e.

$$\Delta u_n(k\triangle x) \approx \frac{u_n((k-1)\triangle x - 2u_n(k\triangle x) + u_n((k+1)\triangle x)}{(\triangle x)^2},$$

where $\triangle x$ is the grid size and $k\triangle x$ denotes the grid points. For the first and the last grid point the boundary conditions play an important role, i.e. they define $u_n(-\triangle x)$ and $u_n(b+\triangle x)$. In the

following we will use periodic boundary conditions for convenience. Therefore $u_n(-\triangle x) = u_n(b)$ and $u_n(b + \triangle x) = u_n(0)$. The approximation error of the Laplace approximation by centered differences, as denoted before, is of order $O((\triangle x)^2)$ if the fourth derivative of u in x exists and if it is bounded, see e.g. [85].

2.2.2. Simulation of Hilbert Space-Valued Stochastic Processes

In the following we will see how to combine the Hilbert space-valued stochastic processes of the previous section and the random field algorithms of Chapter 1. We first consider a Q-Wiener process W which was defined in Definition 2.12. Let $0 = t_0 < t_1 < \cdots < t_n = t$ be a partition of the interval $[0, t]$, then $W(t)$ can be expressed in the following way:

$$W(t) = W(t) - W(0) = \sum_{i=0}^{n-1} (W(t_{i+1}) - W(t_i)),$$

where the increments are independent with variance $t_{i+1} - t_i$ in time and covariance Q in space. In Chapter 1 we approximated random fields φ with covariance C in space. If we scale these by $\sqrt{t_{i+1} - t_i}$, the resulting GRFs will have the properties of the increments of a Q-Wiener process, also called Brownian motion. Thus a Q-Wiener process can be approximated and efficiently simulated by scaling the GRFs constructed in the previous chapter.

In simulations it might be interesting to simulate stochastic processes X with covariance Q in space, but another covariance in time. Then the calculations are similar to those of real-valued stochastic processes [52] but instead of a real-valued Brownian motion, the Q-Wiener process is used. The following example illustrates this construction.

Example 2.33. An *Ornstein-Uhlenbeck process* $\varphi(t, x)$ is either defined by the differential equation

$$\begin{cases} d\varphi(t, x) = -\gamma\varphi(t, x)\, dt + dW(t, x), \\ \varphi(0, x) = \varphi_0(x), \end{cases}$$

or by the solution of this SDE

$$\varphi(t, x) = e^{-\gamma t}\, \varphi_0(x) + \int_0^t e^{-\gamma(t-s)}\, dW(s, x),$$

where γ is a real-valued constant. These two definitions lead immediately to two different ways of implementing φ. The first is to simulate the SDE, i.e.

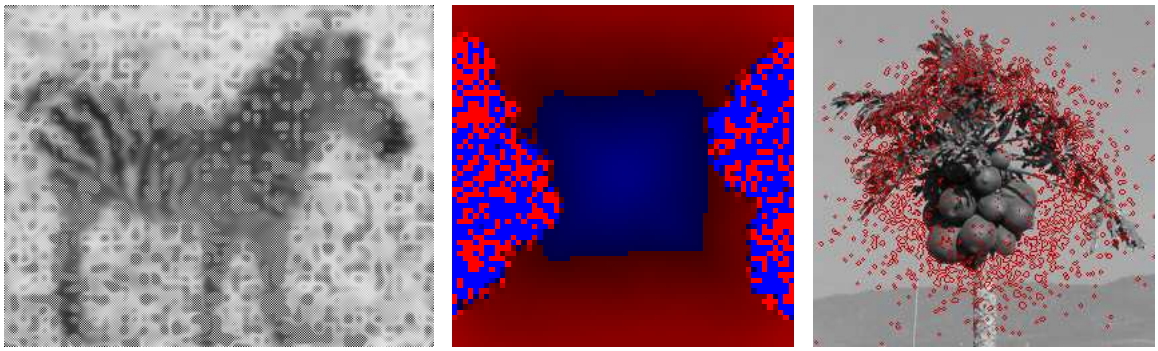
$$\varphi(t + \triangle t, x) = \varphi(t, x) - \triangle t\, \gamma\, \varphi(t, x) + (W(t + \triangle t, x) - W(t, x)),$$

and therefore the increments in each time step are approximated by

$$\varphi(t + \triangle t, x) - \varphi(t, x) = -\triangle t\, \gamma\, \varphi(t, x) + (W(t + \triangle t, x) - W(t, x)),$$

where the last expression is the increment of the Q-Wiener process which can be simulated as presented before.

The second possible construction is to construct φ directely using the solution of the SDE but the problem is that many exponential functions have to be calculated which is computationally expensive. Therefore I recommend to use the simulation of the SDE.



(a) heat equation, (b) motion by mean curvature, (c) segmentation,

Figure 2.1.: Examples of numerical instability when simulating SPDEs.

2.3. Numerical Stability of SPDEs

This section summarizes the work that has been done in the analysis of convergence of discretization schemes and in *numerical stability analysis*. Extensions of a definition of stability from PDEs [85] and one from SDEs [52] to SPDEs are made and studied. A general problem concerning numerical stability is that a big amount of different mathematical definitions exists. All theses definitions try to catch the phenomenon that during a simulation on a computer numerical errors explode and dominate the simulation such that the solution is not usable at all. Examples are shown in Figure 2.1 in the case of a heat equation, motion by mean curvature, and segmentation. In the following we will first look at what research has been done using different definitions. Afterwards we will generalize *stochastically numerical stability* of Kloeden and Platen in [52] and show that this definition does not seem to be useful for SPDEs. Afterwards we extend the definition of *numerical stability* of Sewell [85] from classical PDE theory to SPDEs. We will explicitely calculate the critical values for stability for the heat equation with additive and multiplicative noise using these two definitions. It turns out that the stability results are different.

2.3.1. Convergence in the Literature

Numerical stability of SDEs and SPDEs has rarely been studied in the literature so far. All that has been done are proofs for the convergence of discretization schemes. It turns out that many different definitions of stability and convergence are used and it is not easy to compare the different approaches. We will look at [35, 37, 39, 40, 44, 63, 86, 87, 88, 90] where we start with SDEs and proceed to SPDEs afterwards.

Papers written about SDEs are [39, 44, 90]. We start in alphabetical order with [39] written by Erika Hausenblas.

Error Analysis for Approximation of Stochastic Differential Equations Driven by Poisson Random Measures

This section summerizes [39] by Erika Hausenblas. Let X_t be a real-valued process and solution to

$$X_t(x_0) = x_0 + \int_0^t \int \sigma(X_{s-}, z)(\mu - \gamma)(dz, ds) + \int_0^t b(X_{s-})\, ds \tag{2.8}$$

where μ is a Poisson random measure satisfying certain conditions and γ is its compensator. Assume that $b : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ are Lipschitz continuous in x. The SDE admits a unique solution and the solution is a semimartingale. Let X_t^n be the approximation of X_t by the Euler scheme with step size $1/n$ defined by

$$\begin{cases} X_t^n = X_{[t]_n}^n + b(X_{[t]_n})(t - [t]_n) + \int_{[t]_n}^t \int \sigma(X_{[t]_n}^n, z)(\mu - \gamma)(dz, ds), \\ X_0^n = x_0, \end{cases}$$

where $[t]_n = [tn]/n$. The entity $\mathbb{E}(f(X_t))$, $f \in L^\infty$, will be approximated by a finite sum over a large number N of independent trajectories, i.e.

$$\mathbb{E}(f(X_T)) \approx \frac{1}{N} \sum_{i=1}^N f(X_t^N(\omega_i)).$$

The resulting error $e(n, N)$ depends on the sample size N and on the step size $1/n$, i.e.

$$e(n, N) \le \Big|\frac{1}{N} \sum_{i=1}^N f(X_t^N(\omega_i)) - \mathbb{E}(f(X_T^n))\Big| + \Big|\mathbb{E}(f(X_T^n)) - \mathbb{E}(f(X_T))\Big| =: \mathrm{I} + \mathrm{II}.$$

If the driving process has finite variance, an upper bound for (I) can be found by the central limit theorem or large deviation results. The main result of the paper is an error bound for the entity (II) under appropriate hypotheses for σ and b.

Let μ be a random measure generated by a Poisson point process whose characteristic measure is Lebesgue, γ is its compensator. Let X_t be a solution of (2.8).

Definition 2.34 (Bass and Cranston [4, p. 513])**.** $\sigma(x, z)$ is called *quasi-stable* of order k between the indices α^- and α^+ if there exist $0 \le z_0 < +\infty$ and $0 < c_1, c_2 < +\infty$ such that

$$c_1 |z|^{-\frac{1}{\alpha^-} - i} \le \big|\partial_z^i \sigma(x, z)\big| \le c_2 |z|^{-\frac{1}{\alpha^+} - i} \tag{2.9}$$

for $i = 0, \dots, k$, $|z| > z_0$, and all x.

Theorem 2.35. *Let X_t be the solution of SDE (2.8), where $\sigma(x, z)$ is quasi-stable of order five between the indices α^- and α^+, $0 < \alpha^- \le \alpha^+ < 2$, such that $\partial_z \sigma(x, z) = \sigma_z(x, z) \ge 0$. Moreover, assume that there exist constants $1 \le M < +\infty$ and $1 \le m_b \ll M$ such that σ and b satisfy the following hypotheses:*

(H0) For $0 < j \le 5$, $x \in \mathbb{R}$, and $i = 1, \dots, 5 - j$, either $\partial_x^j \sigma(x, z) = 0$ or estimate (2.9) holds for $\partial_x^j \sigma(x, z)$.

(H1) *For all x and z the quantities $|\partial_z^i \partial_x^j \sigma(x,z)|$ are uniformly bounded by M in z and x for all i and j, $i+j \leq 5$, $j \neq 0$.*

(H2) $\sup_x \left|\partial_x^i b(x)\right|$, $i = 1, \ldots, 5$, *is bounded by* m_b *and* $\sup_x \left|\partial_x^i \sigma(x,z)\right|$, $i = 1, \ldots, 5$, *is bounded by* $h_\sigma(z)$ *such that* $\| h_\sigma \|_p \leq M$ *for all* $p \geq 2$.

(H3) *Let* $\bar{z} = \sup_z \left\{|\sigma_z(s,z)| > \frac{1}{4} \text{ for all } x\right\}$. *Then the functions* $(\partial_x + \frac{\sigma_x(x,z)}{\sigma_z(x,z)}\partial_z)^i \sigma_x(x,z)$, $i = 1,2,3,4$, *are uniformly bounded by* M *in* x *for all* $|z| < \bar{z}$ *with the convention that* $0/0 = 0$.

If X_t is approximated by the Euler scheme, i.e. by X_t^n, then for $f \in L^\infty$

$$|\mathbb{E}(f(X_T)) - \mathbb{E}(f(X_T^n))| \leq C(T) \cdot n^{-1} \cdot M^{21}(1 + \exp(M^{16})).$$

The Optimal Discretization of Stochastic Differential Equations

The following paper [44] of Hofmann, Müller-Gronbach, and Ritter from 2001 studies the pathwise approximation of scalar stochastic differential equations and shows an optimal discretization in the sense of a mean squared L^2-error. Let

$$dX_t = a(t, X_t)\, dt + \sigma(t, X_t)\, dW_t, \qquad t \in [0,1] \tag{2.10}$$

be given with one-dimensional Brownian motion W, drift coefficient a, diffusion coefficient σ, and initial value X_0. A pathwise approximation to the solution X is a stochastic process $\bar{X}$ whose paths are close to the respective paths of X. The error $e(\bar{X})$ of the method $\bar{X}$ is defined by

$$e(\bar{X}) = \mathbb{E}(\left\|X - \bar{X}\right\|_2^2)^{1/2},$$

i.e. the quality of $\bar{X}$ is characterized globally on $[0,1]$ and not only on a finite number of points.

The authors answer the question of how much it costs to achieve an error of at most ε, i.e. they give an upper bound for specific methods and lower bounds for arbitrary methods. Thus they determine the smoothness of the solution at the point (t, X_t) by

$$\mathbb{E}((X_{t+\delta} - X_t)^2 | X_t) = \sigma^2(t, X_t) \cdot \delta + o(\delta),$$

and they call $|\sigma|\,(t, X_t)$ a *conditional Hölder constant*. Their method $\hat{X}_h^{**}$ works in the following way. For $h > 0$ take $k_h \in \mathbb{N}$ such that

$$\lim_{h \to 0} k_h \cdot h = 0$$

and

$$\lim_{h \to 0} k_h^2 \cdot h = +\infty.$$

Set

$$\tau_{i,0} = i/k_h, \qquad i = 0, \ldots, k_h - 1$$

and suppose that

$$x_i = \hat{X}_h^{**}(\tau_{i,0})$$

is already computed, where $x_0 = X(0)$. Put

$$\sigma_i = \sigma(\tau_{i,0}, x_i), \qquad a_i = a(\tau_{i,0}, x_i),$$

and define an adaptive discretization of the subinterval $(\tau_{i,0}, \tau_{i+1,0})$ by

$$\tau_{i,j+1} = \tau_{i,j} + h/\left|\sigma_i\right|$$

as long as the right-hand side is less than $\tau_{i+1,0}$. Here $\tau_{k_h,0} = 1$. On the subinterval $(\tau_{i,0}, \tau_{i+1,0})$ the Euler method is used

$$\hat{X}_h^{**}(\tau_{i,j+1}) = \hat{X}_h^{**}(\tau_{i,j}) + a_i \cdot (\tau_{i,j+1} - \tau_{i,j}) + \sigma_i \cdot (W(\tau_{i,j+1}) - W(\tau_{i,j})),$$

without updating the drift and diffusion coefficient. The approximated solution at the right endpoint $x_{i+1} = \hat{X}_h^{**}(\tau_{i+1,0})$ is computed by a single Milstein step of length $1/k_h$ starting at the left endpoint $\tau_{i,0}$ with initial value x_i. The Milstein step is based on $W(\tau_{i,j+1}) - W(\tau_{i,j})$. Globally piecewise linear interpolation is used. This method satisfies

$$\lim_{h\to 0} n(\hat{X}_h^{**})^{1/2} \cdot e(\hat{X}_h^{**}) = \mathbb{E}\left(\int_0^1 \left|\sigma\right|(t, X_t)\, dt\right)/\sqrt{6},$$

where $h > 0$ is the basic step size and $n(\hat{X}_h^{**})$ is the expected number of observations of W, and gives an asymptotic upper bound for the method. The matching asymptotic lower bound reads

$$\liminf_{N\to\infty} N^{1/2} \cdot e(\bar{X}_N) \geq \mathbb{E}\left(\int_0^1 \left|\sigma\right|(t, X_t)\, dt\right)/\sqrt{6},$$

for every sequence of methods $\bar{X}_N$ such that $n(\bar{X}_N) \leq N$. It follows that the method $\hat{X}_h^{**}$ is asymptotically optimal, that the best order of convergence is $1/2$ in terms of $n(\bar{X})$, and that the best asymptotic constant is given by the mean of the conditional Hölder constant in space and time.

Mean-Square Stability of Second-Order Runge-Kutta Methods for Stochastic Differential Equations

The paper [90] published in 2005 by Tocino shows numerical stability of second-order Runge-Kutta methods using test equations. The author considers the scalar Itô SDEs

$$\begin{cases} dX_t = a(t, X_t)\, dt + b(t, X_t)\, dW_t, & t_0 \leq t \leq T, \\ X_{t_0} = x_0, \end{cases}$$

where $a, b : [t_0, T] \times \mathbb{R} \to \mathbb{R}$ are the drift and diffusion coefficients, $\{W_t\}_{t_0 \leq t \leq T}$ represents the one-dimensional standard Wiener process and for simplicity $x_0 \in \mathbb{R}$ is nonrandom. In order to ensure existence and uniqueness of solutions as well as the existence of absolute moments, he assumes

1. (Lipschitz) There exists a constant $K_1 > 0$ such that

   $$|a(t, x) - a(t, y)| + |b(t, x) - b(t, y)| \leq K_1 |x - y|$$

   for all $t \in [t_0, T]$ and $x, y \in \mathbb{R}$.

2. (Linear growth bound) There exists a constant $K_2 > 0$ such that

$$|a(t,x)|^2 + |b(t,x)|^2 \leq K_2(1+|x|^2)$$

for all $t \in [t_0, T]$ and $x \in \mathbb{R}$.

The numerical schemes presented in the paper are all constructed along time discretizations $t_0 \leq t_1 \leq \cdots \leq t_N = T$ with constant step size

$$\Delta = \frac{T - t_0}{N} > 0,$$

therefore t_n denotes the n-th step point. Let C_P^β be the space of all $f : [t_0, T] \times \mathbb{R} \to \mathbb{R}$ for which f and all its partial derivatives up to order β have polynomial growth, i.e. there exist $K > 0$ and $r \in \mathbb{N}$, $r < +\infty$ such that $|f(t,x)| \leq K(1+|x|^r)$ for all $t \in [t_0, T]$ and $x \in \mathbb{R}$.

Definition 2.36. A discrete approximation $\bar{X} = \{\bar{X}_0, \bar{X}_1, \ldots, \bar{X}_N\}$ (based on a step size Δ) is said to *converge weakly with order* β towards the solution $X = \{X_t\}$ of the SDE if for each $g \in C_P^{2\beta+2}$ there exist constants $K_g > 0$ (not depending on Δ) and $\delta_0 > 0$ such that

$$\left|\mathbb{E}(g(\bar{X}_N) - g(X_t))\right| \leq K_g \Delta^\beta$$

for all $\Delta \in (0, \delta_0)$.

For weakly convergent approximations, the author proposes a second-order Runge-Kutta method but proves mean-square stability only for the test equation

$$dX_t = \lambda X_t\, dt + \mu X_t\, dW_t, \quad t > t_0,\ \lambda, \mu \in \mathbb{C}, \tag{2.11}$$

where mean-square stability is defined in the following way.

Definition 2.37. Let X_t be the solution of Equation (2.11) then

$$\lim_{t\to\infty} \mathbb{E}(|X_t|^2) = 0$$

is called *mean-square stability*.

Similarly, a numerical solution $\{X_n\}_{n\in\mathbb{N}}$ generated by a scheme with equidistant step size applied to test equation (2.11) is *mean-square stable* (*MS-stable*) if

$$\lim_{n\to\infty} \mathbb{E}(|X_n|^2) = 0.$$

Using this definition Tocino shows MS-stability for second-order Runge-Kutta methods and proves that the domain of MS-stability of a Runge-Kutta scheme depends on just one discretization parameter. He also confirms his theoretical result with numerical experiments.

This finishes the presentation of literature about convergence of discretization schemes of SDEs. More literature about stability analysis can be found in the references of the papers presented above and in [52]. Numerical stability of SDEs will be presented in Section 2.3.2. Next we will have a look at some work that has been done on the numerical stability analysis of SPDEs. We will start with an overview paper written by Gyöngy.

Approximations of Stochastic Partial Differential Equations

In [35] Gyöngy summarizes results of [37] and [95, 96]. The methods he uses are the splitting up method and a finite difference scheme. For the splitting up method the following equation is considered

$$\begin{aligned} du(t,x) &= \sum_{r=1}^{d_1}(L_r u(t,x) + f_r(t,x))\,dt \\ &\quad +(L_0 u(t,x) + f_0(t,x))\,dV_t^0 + (M_k u(t,x) + g_k(t,x))\,dY_t^k, \end{aligned}$$

where $t \in (0,T]$, $x \in \mathbb{R}^d$, $d_1 \in \mathbb{N}$ with initial condition $u(0,x) = u_0(x)$, the operators L_r and M_k are of the form

$$L_r = a_r^{ij}(t,x)D_{ij} + a_r^i(t,x)D_i + a_r(t,x), \quad M_k = b_k^i(t,x)D_i + b_k(t,x),$$

and Y is a d_0-dimensional continuous $\mathcal{F}_t$-martingale and V_0 is an $\mathcal{F}_t$-adapted continuous increasing process.

Assumption 2.38 (Smoothness of the coefficients)**.** All the coefficients $a_r^{ij}(t,x)$, $a_r^i(t,x)$, $a_r(t,x)$, $b_k^i(t,x)$, $b_k(t,x)$ are predictable for any $x \in \mathbb{R}^d$ and, for any $(\omega,t) \in \Omega \times (0,\infty)$, their derivatives up to order $m+3$ exist, are continuous, and by magnitude are bounded by K.

Assumption 2.39. For each $\omega \in \Omega$, the functions $f_r(t) = f_r(t,\cdot)$ are weakly continuous as H^{m+3}-valued functions, $g_k(t) = g_k(t,\cdot)$ are weakly continuous as H^{m+4}-valued functions, and $u_0 \in L^2(\Omega, \mathcal{F}_0; H^{m+3})$. Furthermore, f_r and g_k are predictable and

$$\mathbb{E}(\sup_{t\in[0,T]} \| f \|_{m+3}^p) + \mathbb{E}(\sup_{t\in[0,T]} \| g \|_{m+4}^p) + \mathbb{E}(\| u_0 \|_{m+3}^p) \le K,$$

where $\| f \|_{m+3}^2 = \sum_r \| f_r(t) \|_{m+3}^2$ and $\| g \|_{m+4}^2 = \sum_k \| g_k(t) \|_{m+4}^2$.

Assumption 2.40. The process V_t^0 is predictable, continuous, increasing, and starting at zero. We have $V_T^0 + \langle Y \rangle_T \le K$. The matrices (a_r^{ij}) are nonnegative and, for any $\omega \in \Omega$, $x, \lambda \in \mathbb{R}^d$, we have

$$2a_0^{ij}(t,x)\lambda^i\lambda^j\,dV_t^0 - b_k^i(t,x)b_r^j(t,x)\lambda^i\lambda^j\,d\langle Y^k, Y^r\rangle_t \ge 0$$

in the sense of measures on $[0,T]$.

Assumption 2.41. There exists a continuous $\mathcal{F}_t$-martingale

$$Z_t = (Z_t^1, \ldots, Z_T^{d_2})$$

and for any $x \in \mathbb{R}^d$ there exist bounded predictable functions

$$h_r(t,x) = (a_{\gamma r}^{ij}(t,x), a_{\gamma r}^i(t,x), a_{\gamma r}(t,x), f_{\gamma r}(t,x))$$

defined on $\Omega \times (0,T]$ for $r = 0, 1, \ldots, d_2$, such that

1. $d\langle Z \rangle_t \le dV_t$,

2.
$$h(t,x) = h(0,x) + \int_0^t h_0(s,x)\,dV_s + \sum_{r=1}^{d_2} \int_0^t h_r(s,x)\,dZ_s^r$$
for all ω and t, where, as usual, the summation in r is carried over all possible values, which in this case are $1, 2, \ldots, d_2$. Furthermore, all h_r are continuously differentiable with respect to x up to order $m+1$ and $\left|D^\beta h_r\right| \le K$ for $|\beta| \le m+1$.

Theorem 2.42 ([36]). *Under Assumptions 2.38, 2.39, 2.40, and 2.41 there is a constant N depending only on d, d_0, d_1, d_2, K, p, M, and T, such that*
$$\mathbb{E}\big(\max_{t\in T_n} \|u^{(n)}(t) - u(t)\|_m^p\big) \le N\, n^{-p}$$
for all integers $n \ge 1$.

Let $C^l = C^l(\mathbb{R}^d)$ denote the Banach space of functions $f = f(x)$, $x \in \mathbb{R}^d$, having continuous derivatives up to order l, such that $\| f \|_{C^l} := \sup_{x\in\mathbb{R}^d} \sum_{|\beta|\le l} \left|D^\beta f(x)\right| < +\infty$, then we have the following two corollaries.

Corollary 2.43 ([36]). *Assume the conditions of the previous theorem. If Assumptions 2.38, 2.39, and 2.41 hold with $m > l + d/2$ and nonnegative integer l, then for some N, where $N := N(d, d_0, d_1, d_2, K, p, m)$,*
$$\mathbb{E}(\max_{t\in T_n} \|u^{(n)}(t) - u(t)\|_{C^l}^p) \le N\, n^{-p}$$
for all $n \ge 1$.

Corollary 2.44 ([36]). *Assume the conditions of Theorem 2.42. If Assumption 2.39 holds with $p > \kappa$, for some $\kappa > 1$, then there exists a random variable ξ, such that almost surely*
$$\max_{t\in T_n} \|u^{(n)}(t) - u(t)\|_X \le \xi\, n^{-1+1/\kappa},$$
for all $n \ge 1$, where $X = H^m$, or $X := C^l$ can also be taken if $m > l + d/2$.

In the second part of the paper approximations and estimates on finite difference schemes are considered. The results summarized in this part are due to Yoo [95, 96, 97] and based on an L^2-theory for SPDEs. Therefore consider the SPDE
$$\begin{cases} du(t,x) = (D_i(a^{ij} D_j u(t,x) + f^i(t,x))\,dt + (D_i b_k^i(t,x) + g_k(t,x))\,dW_t^k, \\ u(0,x) = u_0(x), \qquad x \in \mathbb{R}^d, \end{cases} \tag{2.12}$$
where W is a d_1-dimensional Wiener process and a^{ij}, b_k^i, f^i, g_k are functions on $\Omega \times [0,T] \times \mathbb{R}^d$. Let T, K, and $\lambda > 0$ be fixed positive numbers and let $m \ge 0$ be a fixed integer. Assume that the following conditions are satisfied.

Assumption 2.45. The coefficients $a^{ij}(t,x)$, $b^i_k(t,x)$, are predictable for any $x \in \mathbb{R}^d$, and their derivatives up to order m are continuous and by magnitude are bounded by K. Furthermore, for $l \le m$ and

$$K_l(t) := \int_0^t \Big(\sum_i \| f^i(s) \|_l^2 + \sum_k \| g_k(s) \|_l^2 \Big)\, ds,$$

let

$$\mathbb{E}(\| u_0 \|_m^2) + \mathbb{E}(K_m(T)) < +\infty.$$

Assumption 2.46. For all $\omega \in \Omega$, $t \ge 0$, $x, z \in \mathbb{R}^d$

$$(2a^{ij}(t,x) - b^i_k b^j_k(t,x)) z^i z^j \ge \lambda \sum_i \left| z^i \right|^2.$$

These assumptions imply that the Cauchy problem (2.12) has a unique solution u, and

$$\begin{aligned} &\mathbb{E}\big(\sup_{0\le t\le T} \| u(t) \|_m^2 \big) + \mathbb{E}\big(\int_0^T \| u(t) \|_{m+1}^2 \, dt \big) \\ &\qquad \le N\Big(\mathbb{E}\big(\| u_0 \|_m^2 \big) + \mathbb{E}\big(\int_0^T \big(\sum_i \| f^i(t) \|_m^2 + \sum_k \| g_k(t) \|_m^2 \big)\, dt \big) \Big) \end{aligned}$$

with some constant N.

Let $\mathcal{Z}_h^d := \{ hz = (hz_1, hz_2, \dots, hz_d), z \in \mathcal{Z}^d \}$ be the grid for some fixed constant $h \in (0,1)$, where $\mathcal{Z}^d$ is a fixed grid. Define the finite difference operators δ_i^+ and δ_i^- by

$$\begin{aligned} \delta_i^+ v(x) &:= h^{-1}(v(x + he_i) - v(x)), \\ \delta_i^- v(x) &:= h^{-1}(v(x) - v(x - he_i)), \end{aligned}$$

where $e_1, e_2, \dots, e_d$ is the standard basis in $\mathbb{R}^d$, and consider the equation

$$\begin{aligned} du_h(t,x) = {} & \delta_i^-(a^{ij}(t,x)\delta_j^+ u_h(t,x) + f^i(t,x))\, dt \\ & + (b^i_k(T,x)\delta_i^+ u_h(t,x) + g_k(t,x))\, dW_t^k, \quad t \in [0,T], \quad x \in \mathcal{Z}_h^d. \end{aligned} \tag{2.13}$$

Define the discrete Sobolev space H_h^m to be the space of all real-valued functions f on $\mathcal{Z}_h^d$ with norm

$$\| f \|_{H_h^m} := \Big(\int_{[-\pi/h, \pi/h]^d} (1 + |y|^2)^m \left| \hat{f}(y) \right|^2 \, dy \Big)^{1/2} < +\infty,$$

where $\hat{f}(y) := (2\pi)^{-d} \sum_{z \in \mathcal{Z}_h^d} e^{ih(z,y)} f(z)$ is the discrete Fourier transform of f.

Theorem 2.47 ([95],[96])**.** *Let Assumptions 2.45, 2.46 hold with $m > d/2 + 1$ and let $m' \ge 0$ be such that $m < m' + d/2$. Then Equation (2.13) has a unique solution u_h, and*

$$\begin{aligned} &\mathbb{E}\big(\sup_{0\le t\le T} \| u_h(t) \|_{H_h^{m'}}^2 \big) + \mathbb{E}\big(\int_0^T \| e_h(t) \|_{H_h^{m'+1}}^2 \big)\, dt \\ &\qquad \le N\Big(\mathbb{E}\big(\| u_0 \|_{H_h^{m'}}^2 \big) + \mathbb{E}\big(\int_0^T \big(\sum_i \| f^i(t) \|_{H_h^{m'}}^2 + \sum_k \| g_k(t) \|_{H_h^{m'}}^2 \big)\, dt \big) \Big) \end{aligned}$$

with a constant N. Assume moreover that $m < m' + 2 + d/2$. Then for the error $e_h := u - u_h$ we have

$$\mathbb{E}\big(\sup_{0\le t\le T} \| e_h(t) \|^2_{H_h^{m'}}\big) + \mathbb{E}\big(\int_0^T \| u_h(t) \|^2_{H_h^{m'+1}}\big) \\ \le Nh^2\Big(\mathbb{E}\big(\| u_0 \|^2_m\big) + \mathbb{E}\big(\int_0^T \big(\sum_i \| f^i(t) \|^2_m + \sum_k \| g_k(t) \|^2_m\big)\, dt\big)\Big),$$

with some constant N, depending only on λ, K, m, and T.

Next Equation (2.13) is discretized in time along the partition $T_n := \{t_i = i\,T/n, i = 0, \dots, n\}$ of $[0, T]$ for a fixed integer $n \ge 1$. Then the Euler approximation applied to Equation (2.13) gives

$$\begin{cases} \triangle u_h^n(t,x) = \delta_i^-(a^{ij}(t,x)\delta_j^+ u_h^n(t,x) + f^i(t,x))\triangle t \\ \qquad\qquad +(b_k^i(t,x)\delta_i^+ u_h^n(t,x) + g_k(t,x))\ \triangle W_t^k, \quad t \in T_n, \quad x \in \mathcal{Z}_h^d, \\ u_h^n(0,x) = u(0,x), \quad x \in \mathcal{Z}_h^d. \end{cases}$$

In order to estimate the error $e_h^n := u_h - u_h^n$, one needs assumptions on the dependence of the coefficients a^{ij}, b_k^i, and the free terms f^i, g_k on t.

Assumption 2.48. For all $\omega \in \Omega$, $t \in [0, T]$, $x \in \mathbb{R}^d$, and for all multiindices $|\alpha| \le m$

$$\sum_{i,j} \big|\frac{\partial}{\partial t} D^\alpha a^{ij}\big| + \sum_{i,k} \big|\frac{\partial}{\partial t} D^\alpha b_k^i\big| \le K.$$

Moreover $\frac{\partial}{\partial t} f^i(t) \in H^m$, $\frac{\partial}{\partial t}\, g_k^i(t) \in H^m$ for all ω, t, i, k, and

$$\sum_i \mathbb{E}\big(\int_0^T \big\|\frac{\partial}{\partial t} f^i(t)\big\|^2_m\, dt\big) + \sum_{i,k} \mathbb{E}\big(\int_0^T \big\|\frac{\partial}{\partial t} g_k^i(t)\big\|^2_m\, dt\big) < +\infty.$$

Theorem 2.49 ([95],[96])**.** *Let Assumptions 2.45, 2.46, and 2.48 hold with $m > m' + 2 + d/2$ where m' is a nonnegative integer. Then there exists $0 < \theta = \theta(\lambda, K, m, d)$ such that if $\triangle t/h^2 \le \theta$ then*

$$\mathbb{E}\big(\sup_{t\in T_n} \| e_h^n(t) \|^2_{H_h^{m'}}\big) + \mathbb{E}\big(\sum_{k=0}^n \| u_h^n(t,x) \|^2_{H_h^{m'+1}}\big)\triangle t \\ \le N\triangle t\mathbb{E}\Big(\| u_0 \|^2_m + \int_0^T \big(\sum_i \| f^i(t) \|^2_m + \sum_k \| g_k(t) \|^2_m\big)\, dt \\ + \sum_i \int_0^T \big\|\frac{\partial}{\partial t} f^i(t)\big\|^2_m\, dt + \sum_{ik} \int_0^T \big\|\frac{\partial}{\partial t} g_k^i(t)\big\|^2_m\, dt\Big),$$

where N is a constant depending only on λ, d, T, K, and m.

And finally the error $u - u_h^n$ can be estimated in the following way.

Theorem 2.50 ([95], [96]). *Let Assumptions 2.45, 2.46, and 2.48 hold with $m > l + 2 + d$, where l is a nonnegative integer. Then there exists $0 < \theta = \theta(\lambda, K, m, d)$ such that if $\triangle t/h^2 \leq \theta$, then*

$$\mathbb{E}\big(\sup_{t\in T_n}\sup_{z\in\mathcal{Z}_h^d}|\delta^\alpha(u(t,x)-u_h^n(t,x))|^2\big)$$
$$\leq Nh^2\mathbb{E}\Big(\|u_0\|_m^2+\sum_i\int_0^T\big(\|f^i(t)\|_m^2+\|\frac{\partial}{\partial t}f^i(t)\|_m^2\big)\,dt$$
$$+\sum_{i,k}\int_0^T\big(\|g_k^i(t)\|_m^2+\|\frac{\partial}{\partial t}g_i^i(t)\|_m^2\big)\,dt\Big),$$

for all multiindices $|\alpha| \leq l$, where N is a constant depending only on λ, d, K, T, l, and m.

The last part of the presented paper applies the results to nonlinear filtering.

A Numerical Approximation of Parabolic Stochastic Partial Differential Equations Driven by a Poisson Random Measure

In the following paper [40] by Hausenblas and Marchis, the authors give estimates on the error of their approximation compared to the solution of the SPDE of the following form

$$\begin{cases} du(t) = Au(t)\,dtf(u(t-))\,dt \in_Z g(u(t-);z)\,\tilde{\eta}(dz,dt), \\ u(0) = u_0 \in V_{-\rho}, \end{cases} \tag{2.14}$$

where $f : E \to V_{-\delta_f}$ and $g : E \to L(Z; V_{-\delta_g})$ are not necessarily Lipschitz continuous functions and ρ, δ_f, δ_g are some non negative constants specified later. Furthermore $\tilde{\eta} = \eta - \gamma$ where η is a Poisson random measure over a probability space $(\Omega, \mathcal{F}, P)$ and γ its compensator, and for $\delta > 0$ V_δ is the domain of the fractional power of A with exponent δ and norm $\|\cdot\|_\delta := \|(I-A)^\delta \cdot\|$, i.e. $V_\delta := D((I-A)^\delta)$. For $n \in \mathbb{N}$ let V_{-n} be the completion of the Banach space E with respect to the norm $\|\cdot\|_{-n} := \|(-A)^{-n}\cdot\|$. If $\delta < 0$ such that $\delta = -n + \beta$ where $\beta \in [0,1)$ and $n \in \mathbb{N}$ then V_δ denotes the domain of the fractional power of A in V_{-n} with exponent β and norm $\|\cdot\|_\delta := \|(I-A)^\beta \cdot\|_{-n}$.

The space discretization will satisfy

$$\begin{cases} du_n(t) + A_n u_n(t_-)\,dt = P_n f(E_n u_n(t))\,dt + \int_{Z_n} P_n g(E_n u_n(t);z)\,\eta(dz;dt), \\ u_n(0) = P_n x, \end{cases}$$

where $A_n = P_n A E_n$ is a bounded operator on X_n approximating A, the operator P_n denotes a 'projection' from X into X_n, where X_n is finite dimensional, of dimension d_n. Let E_n be an 'interpolation' operator that associates with each element of X_n an element of X. The space Z is approximated by Z_n, where $Z_n := P_n Z$ and η is approximated by η_n. The authors assume that the space discretization satisfies the following assumptions.

Assumption 2.51.

1. X, X_1, X_2, X_3,... are all real- or all complex-valued Banach spaces. All norms will be denoted by $\|\cdot\|$.

2. P_n is a bounded linear operator, satisfying $\| P_n x \| \leq p \, \| x \|$ for all $n \geq 1$, $x \in X$ for some $p \geq 0$.

3. E_n is a bounded linear operator, satisfying $\| E_n x \| \leq q \, \| x \|$ for all $n \geq 1$, $x \in X$ and for some $q \geq 0$.

4. $P_n E_n x = x$ for all $n \geq 1$ and $x \in X_n$.

Moreover they assume the following stability condition.

Assumption 2.52. A_n is a bounded operator and there exist some $M < +\infty$ and $\omega \in \mathbb{R}$ such that

$$\left\| e^{A_n t} \right\| \leq M e^{\omega t} \qquad \text{for } t \geq 0, \ n \geq 1.$$

Assumption 2.53. Assume that there exist $\gamma > 0$ and some function $\varphi : \mathbb{N} \times (0, \gamma] \to [0, \gamma]$, with $\varphi(n, \delta) \to 0$ as $n \to +\infty$ and

$$\begin{aligned} \| (I - E_n P_n) x \| &\leq \varphi(n, \delta) \, \| x \|_\delta , \\ \| A(I - E_n P_n) x \| &\leq \varphi(n, \delta) \, \| A_n \| \, \| x \|_\delta , \qquad \text{for all } \delta \in (0, \gamma], \end{aligned}$$

for $x \in V_\alpha$, where α is given by Theorem 2.54.

The time discretization is either done by an explicit or an implicit Euler scheme. Let $v_n^k = u_n(k\tau_n)$ where τ_n is the step size corresponding to the space X_n. For the explicit scheme we have

$$\begin{cases} v_n^{k+1} = (1 + \tau_n A_n) v_n^k + \tau_n P_n f(E_n v_n^k) + \int_{k\tau_n}^{(k+1)\tau_n} \int_Z P_n g(E_n u_n(t); z) \, \eta_n(dz; dt), \\ v_n^0 = P_n u_0, \end{cases} \tag{2.15}$$

and for the implicit scheme

$$\begin{cases} v_n^{k+1} = (1 - \tau_n A_n)^{-1} \left(v_n^k + \tau_n P_n f(k\tau_n, v_n^k) + \xi_n^k(E_n u_n(t)) \right), \\ v_n^0 = P_n u_0, \end{cases} \tag{2.16}$$

where

$$\xi_n^k(E_n u_n(t)) := \int_{k\tau_n}^{(k+1)\tau_n} \int_Z P_n g(E_n u_n(t); z) \, \eta_n(dz; dt).$$

All these prerequisites are used in the following main result of the paper that gives an estimate on the error.

Theorem 2.54 ([40])**.** *Let X be a Banach space of M-type p — i.e. there exists a constant $C = C(X, p)$ such that for any X-valued discrete martingale $(M_0, M_1, M_2 \ldots)$ with $M_0 = 0$, the inequality*

$$\sup_{n \geq 1} \mathbb{E}(|M_n|^p) \leq C \sum_{n \geq 1} \mathbb{E}(|M_n - M_{n-1}|)$$

holds — $Z \subset X$, $\mathcal{X}$ and $\mathcal{Z}$ the Borel σ-fields. Let A be an unbounded operator on X with domain $D(A)$ generating an analytic semigroup on X. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ be the usual filtration and η a Poisson random measure with symmetric , p-integrable, characteristic measure $\nu \in \mathcal{L}(Z)$.

We assume that the Poisson random measure η is approximated by a Poisson random measure η^n with intensity ν^n such that there exists a function $\gamma : \mathbb{N} \to \mathbb{R}^+$ with

$$\int_Z |z|^p \, (\sigma - \sigma_n)\,(dz) \leq \gamma(n).$$

Let $1 < p \leq 2$ and $q = p^m$ for some $m \in \mathbb{N}$. Let δ_f and δ_g be two constants such that $\delta_f < 1 - q^{-1}$ and $\delta_g < p^{-1} - q^{-1}$. We assume that $f : V_{-\delta_f} \to X$ and $g : V_{-\delta_g} \times Z \to X$ satisfy the growth conditions

$$\begin{aligned} \| f(x) \|_{-\delta_f} &\leq C_f(1 + \| x \|), && x \in X, \\ \int_Z \| g(x;z) \|^{p^l}_{-\delta_g} \, \nu(dz) &\leq C_g(1 + \| x \|^{p^l}), && x \in X,\ l = 1, \ldots, m, \end{aligned}$$

and are globally Lipschitz, i.e.

$$\begin{aligned} \| f(x-y) \|_{-\delta_f} &\leq L_f \| x - y \|, && x, y \in X, \\ \int_Z \| g(x;z) - g(y;z) \|^{p^l}_{-\delta_g} &\leq G(\| x - y \|)^{p^l}, && x, y \in X, l = 1, \ldots, m, \\ \| g(x;z_1) - g(x;z_2) \|_{-\delta_g - \gamma} &\leq L(\| x \|) \| z_1 - z_2 \|_{-\gamma}, && x \in X, z_1, z_2 \in Z, 0 < \gamma < p^{-1} - q^{-1}, \end{aligned}$$

where $G, L : \mathbb{R} \to \mathbb{R}$ are globally Lipschitz and continuous. Assume there exists $0 < \delta < 1$ such that the initial condition u_0 belongs to V_δ. Assume furthermore that the space discretization satisfies the Assumptions 2.51, 2.52, 2.53 and let v_n^k be the approximation of (2.14) given by Equation (2.15) or (2.16), respectively. Then we have for all $1 \leq r \leq p$ and $0 < \varepsilon < \alpha$

$$\sup_{0 \leq k \leq N_n} \mathbb{E}(\|u(k\tau_n) - v_n^k\|^r)^{\frac{1}{r}} \leq \tfrac{C_1}{\varepsilon} \, \tau^{\alpha - \varepsilon} + \tfrac{C_2}{\varepsilon} \, \varphi(n, \alpha - \varepsilon) + C_3 \, (\varphi(n, \delta) + \tau_n^\delta),$$

where $\alpha = \min(1 - q^{-1} - \delta_f, p^{-1} - q^{-1} - \delta_g)$ and $T = \tau_n \, N_n < +\infty$ is fixed.

A Numerical Scheme for Stochastic PDEs with Gevrey Regularity

Lord and Rougemont consider in [63] the SPDE

$$du(t) = \Delta u(t) + F(u(t))\, dt + Q \, dW(t) \tag{2.17}$$

with periodic boundary conditions on $[0, 2\pi)$ and initial condition $u_0 \in C_{\mathrm{per}}([0, 2\pi], \mathbb{C})$, the space of continuous periodic functions. Here W is a cylindrical Wiener process and Q is a Hilbert–Schmidt operator on Gevrey space G_α, namely it has exponentially decaying Fourier coefficients. Therefore $Q\dot{W}(t)$ is given by

$$Q\dot{W}(t) = \sum_{n \in \mathbb{Z}} b_n \varphi_n \dot{\beta}_n(t), \qquad \text{with} \quad \sum_{n \in \mathbb{Z}} e^{2\alpha|n|}(1 + n^2) \, |b_n|^2 < +\infty,$$

where $b_n \in \mathbb{C}$, $\{\beta_n \in \mathbb{R},\, n \in \mathbb{Z}\}$ are mutually independent ordinary Brownian motions, and $\varphi_n(x) = \exp(inx)$. Let $b_0 = 0$. The authors use a Galerkin discretization to solve Equation (2.17) similarly to the approximations of deterministic PDEs in [62] and [89]. First, they take Fourier series and obtain the infinite system of coupled equations

$$u_n(t) = e^{-tn^2} u_n(0) + \int_0^t e^{-(t-s)n^2} F_n(u(x))\, ds + \int_0^t e^{-(t-s)n^2} b_n \, d\beta_n(x), \qquad n \in \mathbb{Z},$$

where $F_n(u)$ is the nth Fourier coefficient of $F(u)$. This leads to the discretization

$$u_n^N(t) = e^{-tn^2}u_n^N(0) + \sum_{k=1}^{[t/h]} e^{-(t-(k-1)h)n^2}(hF_n(u^N((k-1)h)) + b_n\sqrt{h}X_{k,n}) \tag{2.18}$$

where $X_{k,n} = h^{-1/2}(\beta_n(kh) - \beta_n((k-1)h))$.

Finally, in the following proposition, they prove an estimate on the error of the approximation.

Proposition 2.55 ([63]). *Equation (2.18) converges strongly in H^1 to a solution of (2.17) as $N \to +\infty$ and $h \to 0$. If $u_0 \in G_\alpha$, then the error can be estimated by*

$$\mathbb{E}(\sup_{\substack{t\in(0,T]\\ t/n\in\mathbb{N}}} \| u(t) - u^N(t) \|_{\mathcal{H}}) \leq K(T,m)(N^{-m} + h),$$

for all $m > 0$, K a function.

Weak Convergence of a Numerical Method for a Stochastic Heat Equation

In [87] Shardlow shows weak convergence for a finite difference scheme in space and time. Therefore consider the following stochastic heat equation on $[0,1]$ with homogeneous Dirichlet boundary conditions:

$$\begin{cases} du + Au\,dt = dW(t), \\ u(0) = U, \end{cases} \tag{2.19}$$

where the initial data U is in $L^2(0,1)$, $A := -\Delta$ denotes the Laplacian scaled to be positive definite with domain $H^2(0,1) \cap H_0^1(0,1)$, and $W(t)$ is a Wiener process with covariance Q. For simplicity, we suppose that Q has eigenvalues $\alpha_j \geq 0$ corresponding to the eigenfunctions $e_j(\cdot) := \sqrt{2}\sin(j\pi\cdot)$ of Δ; in other words,

$$W(t) = \sum_{j=1}^{\infty} \alpha_j^{1/2} e_j \beta_j(t),$$

where the elements $\beta_j(t)$ are independent and identically distributed Brownian motions. Then Equation (2.19) admits a unique mild solution for initial condition $U \in L^2(0,1)$. For further details see Section 2.1 or [16].

Shardlow uses the following discretization based on the θ method in time and the standard three point approximation of the Laplacian. Therefore we consider a time step $\triangle t$ and a grid size $\triangle x = 1/J$ for some $J \in \mathbb{N}$. The Wiener process is approximated by truncating its Fourier expansion to $J-1$ terms. Let $\mathcal{P}_{J-1}$ denote the operator taking f to its first $J-1$ modes,

$$\mathcal{P}_{J-1}f = 2\sum_{j=1}^{J-1} (f, \sin(j\pi\cdot))\sin(j\pi\cdot).$$

Define the approximation of the Wiener process by

$$dB_{\triangle t}(n) := \int_{n\triangle t}^{(n+1)\triangle t} \mathcal{P}_{J-1}\, dW(x).$$

This gives an $L^2(0,1)$ function. The numerical method evaluates this function at the grid points $j\triangle x$ for $j = 1, \ldots, J-1$. The chosen initial condition is $U_0 = \mathcal{P}_{J-1}U$. Then, for $0 \le \theta \le 1$, we iterate

$$u_{n+1} - u_n + \frac{\triangle t}{(\triangle x)^2} A_\Delta((1-\theta)u_n + \theta u_{n+1}) = \begin{pmatrix} dB_{\triangle t}(n)(\triangle x) \\ \vdots \\ dB_{\triangle t}(n)((J-1)\triangle x) \end{pmatrix} \tag{2.20}$$

where

$$A_\Delta = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}.$$

Let

$$\tilde{u}(n\triangle t; U) = \sum_{j=1}^{J-1} (\tilde{u})_j\, e_j, \tag{2.21}$$

where the elements $\tilde{u}_j$ are chosen such that $\tilde{u}(n\triangle t; U)(j\triangle x)$ equals the jth component of u_n for $j = 1, \ldots, J-1$.

The above numerical method has been studied by the author in [86] for the problem of space-time white noise, i.e. $Q = \mathbb{1}$ in strong convergence. In that paper he proves that for $T, \varepsilon > 0$, there exists K_ε such that

$$\mathbb{E}(\| u(n\triangle t; U) - \tilde{u}(n\triangle t; U) \|^2)^{1/2} \le K_\varepsilon (\triangle x)^{(1-\varepsilon)/2} (1 + \| U \|)(1 + (n\triangle t)^{\varepsilon - 1}),$$

$0 < n\triangle t \le T$, as $\triangle t, \triangle x \to 0$ with $\nu := \triangle t/(\triangle x)^2$ constrained by $\nu(1-\theta) \le 1/4$. The paper presented here [87] relaxes the condition to $\nu(1-2\theta) \le 1/2$.

The main result of [87] is the following theorem that deals with weak convergence.

Theorem 2.56 ([87])**.** *Let $u(t;U)$ (resp., $\tilde{u}(t;U)$) denote a solution of (2.19) (resp., the trigonometric interpolant of the numerical solution (2.20) defined in (2.21)) corresponding to initial data $U \in L^2(0,1)$. Suppose that $\sum_{j=1}^{\infty} \alpha_j j^r < +\infty$, for some $-2 < r \le 0$. For $\varepsilon, T > 0$ and a twice continuously bounded differentiable function $\phi : L^2(0,1) \to \mathbb{R}$, there exists a constant $K > 0$ such that*

$$\left|\mathbb{E}(\phi(u(n\triangle t; U))) - \mathbb{E}(\phi(\tilde{u}(n\triangle t; U)))\right| \le K(\triangle x)^{2+r-\varepsilon}(1 + \| U \|^2), \quad T = n\triangle t$$

as $\triangle t, \triangle x \to 0$ with $\nu = \triangle t/(\triangle x)^2$ fixed and $(1-2\theta)\nu < 1/2$.

2.3.2. Numerical Stability in the sense of Kloeden and Platen

Kloeden and Platen define in [52] when a discrete approximation of an SDE is called *stochastically numerically stable*. In this section we give some examples for stability and extend the definition to SPDEs. In the following let the solution of an SDE be denoted by X_t and the discrete approximation by Y_{n_t}.

Definition 2.57 ([52])**.** Let Y^δ denote a time discrete approximation with maximum step size $\delta > 0$ starting at time t_0 at Y_0^δ, and $\bar{Y}^\delta$ denotes the corresponding approximation starting at $\bar{Y}_0^\delta$. A discrete approximation Y^δ is *stochastically numerically stable* for a given SDE if for any finite interval $[t_0, T]$ there exists a positive constant Δ_0 such that for each $\varepsilon > 0$ and each $\delta \in (0, \Delta_0)$

$$\lim_{|Y_0^\delta - \bar{Y}_0^\delta| \to 0} \sup_{t_0 \le t \le T} P(|Y_{n_t}^\delta - \bar{Y}_{n_t}^\delta| \ge \varepsilon) = 0.$$

One observes that *stochastically numerically stable* has nothing to do with the original SDE and it is not important, if the approximation really approximates the solution of the SDE. For linear equations it is straightforward to show stability in the sense of Kloeden and Platen. The following examples of explicit approximation schemes give an idea of how to prove that an approximation is stochastically numerically stable.

Example 2.58 (Linear SDE with additive noise)**.** Let

$$dX_t = \lambda X_t \, dt + \sigma(t) dW_t$$

be a given SDE with $\lambda \in \mathbb{R}$, $\sigma : [t_0, T] \to \mathbb{R}$ and W_t is a Wiener process. Then the discretization

$$Y_{t_{n+1}} = Y_{t_n} + h \cdot \lambda Y_{t_n} + \sigma(t_n) \eta_{t_{n+1}},$$

where the interval $[t_0, T]$ is divided into equidistant steps with $h = t_{n+1} - t_n$ for all n and $\eta_{t_{n+1}} = \int_{t_n}^{t_{n+1}} dW_t$, is stochastically numerically stable for all $h > 0$.

This can be seen in the following. First we reduce

$$\begin{aligned} Y_{t_n} &= (1 + h\lambda) Y_{t_{n-1}} + \sigma(t_{n-1}) \eta_{t_n} \\ &= (1 + h\lambda)((1 + h\lambda) Y_{t_{n-2}} + \sigma(t_{n-2}) \eta_{t_{n-1}}) + \sigma(t_{n-1}) \eta_{t_n} \\ &= (1 + h\lambda)^n Y_{t_0} + \sum_{k=1}^{n} (1 + h\lambda)^{n-k} \sigma(t_{k-1}) \eta_{t_k}. \end{aligned}$$

Then for two different starting conditions Y_{t_0} and $\bar{Y}_{t_0}$ we have after n steps

$$\begin{aligned} \left| Y_{t_n} - \bar{Y}_{t_n} \right| &= |1 + h\lambda|^n \left| Y_{t_0} - \bar{Y}_{t_0} \right| \\ &\le e^{nh\lambda} \left| Y_{t_0} - \bar{Y}_{t_0} \right| \\ &= e^{t_n \lambda} \left| Y_{t_0} - \bar{Y}_{t_0} \right|. \end{aligned}$$

As $t_n \in [0,T]$, the expression $e^{t_n\lambda}$ is bounded by $e^{T|\lambda|}$ for all discretizations and by the calculations which were done before

$$\begin{aligned}
\lim_{|Y_0^\delta - \bar{Y}_0^\delta|\to 0} \sup_{t_0\le t\le T} P(|Y_{t_n}^\delta - \bar{Y}_{t_n}^\delta| \ge \varepsilon) &= \lim_{|Y_0^\delta - \bar{Y}_0^\delta|\to 0} \sup_{t_0\le t\le T} P(|1+h\lambda|^{t_n}\,|Y_0^\delta - \bar{Y}_0^\delta| \ge \varepsilon) \\
&\le \lim_{|Y_0^\delta - \bar{Y}_0^\delta|\to 0} \sup_{t_0\le t\le T} P(e^{t_n\lambda}|Y_0^\delta - \bar{Y}_0^\delta| \ge \varepsilon) \\
&\le \lim_{|Y_0^\delta - \bar{Y}_0^\delta|\to 0} \sup_{t_0\le t\le T} P(e^{-T|\lambda|}|Y_0^\delta - \bar{Y}_0^\delta| \ge \varepsilon) \\
&\le \lim_{|Y_0^\delta - \bar{Y}_0^\delta|\to 0} P(|Y_0^\delta - \bar{Y}_0^\delta| \ge \varepsilon\cdot e^{-T|\lambda|}) \\
&= 0.
\end{aligned}$$

This implies that the discretization scheme is stochastically numerically stable.

A second example for a stochastically numerically stable discretization of an SDE with multiplicative noise is the following.

Example 2.59 (Linear SDE with multiplicative noise)**.** Similar to Example 2.58 let

$$dX_t = \lambda X_t\,dt + X_t\,dW_t$$

be a given SDE with $\lambda \in \mathbb{R}$ and W_t a Wiener process. Then the discretization

$$Y_{t_{n+1}} = Y_{t_n} + h\cdot\lambda Y_{t_n} + Y_{t_n}\eta_{t_{n+1}},$$

where the interval $[t_0,T]$ is divided into equidistant steps with $h = t_{n+1} - t_n$ for all n and $\eta_{t_{n+1}} = \int_{t_n}^{t_{n+1}} dW_t$, is stochastically numerically stable for all $h > 0$.

The important observation in the following proof is the independence of the increments of the Wiener process. Let again Y_{t_0} and $\bar{Y}_{t_0}$ be two different starting conditions and $t_n \in [t_0,T]$, then

$$Y_{t_n} = (1+h\lambda+\eta_{t_n})Y_{t_{n-1}} = \prod_{i=1}^{n}(1+h\lambda+\eta_{t_i})Y_{t_0},$$

and therefore

$$\left|Y_{t_n} - \bar{Y}_{t_n}\right| = \Big(\prod_{i=1}^{n}|1+h\lambda+\eta_{t_i}|\Big)\left|Y_{t_0} - \bar{Y}_{t_0}\right|.$$

As the elements η_{t_i} are the independent increments of the Wiener process W_t, it holds $\mathbb{E}(\eta_{t_i}) = 0$ and $\mathbb{E}(\eta_{t_i}\eta_{t_j}) = \delta_{ij}h$. Before we finish the proof, we observe that

$$\begin{aligned}
\mathbb{E}((1+h\lambda+\eta_{t_i})^2) &= (1+h\lambda)^2 + 2(1+h\lambda)\mathbb{E}(\eta_{t_i}) + \mathbb{E}(\eta_{t_i}^2) \\
&= 1 + h(1+2\lambda+h\lambda^2) \\
&\le e^{h(1+2\lambda+h\lambda^2)}.
\end{aligned}$$

Using Chebyshev's inequality we finally get for all $\varepsilon > 0$

$$\begin{aligned}
P\big(|Y_{t_n} - \bar{Y}_{t_n}| \geq \varepsilon\big) &= P\big(\prod_{i=1}^{n} |1 + h\lambda + \eta_{t_i}| \, \big|Y_{t_0} - \bar{Y}_{t_0}\big| \geq \varepsilon\big) \\
&= P\big(\prod_{i=1}^{n} |1 + h\lambda + \eta_{t_i}| \geq \varepsilon \, \big|Y_{t_0} - \bar{Y}_{t_0}\big|^{-1}\big) \\
&\leq \frac{\big|Y_{t_0} - \bar{Y}_{t_0}\big|^2}{\varepsilon^2} \, \mathbb{E}\big(\prod_{i=1}^{n} (1 + h\lambda + \eta_{t_i})^2\big) \\
&= \big|Y_{t_0} - \bar{Y}_{t_0}\big|^2 \varepsilon^{-2} \prod_{i=1}^{n} \mathbb{E}((1 + h\lambda + \eta_{t_i})^2) \\
&\leq \big|Y_{t_0} - \bar{Y}_{t_0}\big|^2 \varepsilon^{-2} \exp(h(1 + 2\lambda + h\lambda^2))^n \\
&= \big|Y_{t_0} - \bar{Y}_{t_0}\big|^2 \varepsilon^{-2} \exp(t_n(1 + 2\lambda + h\lambda^2)).
\end{aligned}$$

The last expression is bounded because $t_n \in [t_0, T]$ for all n and therefore

$$\lim_{|Y^\delta_{t_0} - \bar{Y}^\delta_{t_0}| \to 0} \sup_{t_0 \leq t \leq T} P(|Y^\delta_{t_n} - \bar{Y}^\delta_{t_n}| \geq \varepsilon) \leq \lim_{|Y^\delta_{t_0} - \bar{Y}^\delta_{t_0}| \to 0} \sup_{t_0 \leq t \leq T} \big|Y_{t_0} - \bar{Y}_{t_0}\big|^2 \varepsilon^{-2} e^{t(1+2\lambda+h\lambda^2)} = 0.$$

This finishes the proof and shows that there are no restrictions on h.

In the next section we will see that the approximation of Example 2.59 will not be stable in the sense of Sewell and that for this definition of stability a higher order Milstein scheme is needed. Therefore the property of being *stochastically numerically stable* does not imply stability in the sense of Sewell.

Let us next extend the SDE definition of Kloeden and Platen to SPDEs.

Definition 2.60. Let $U^{(\delta,\kappa)}$ denote a time and space discrete approximation with maximum step size $\delta > 0$ in time and $\kappa > 0$ in space, starting at time t_0 at $U_0^{(\delta,\kappa)}$, and $\bar{U}^{(\delta,\kappa)}$ denotes the corresponding approximation starting at $\bar{U}_0^{(\delta,\kappa)}$.

A discrete approximation $U^{(\delta,\kappa)}$ is *stochastically numerically stable* for a given SPDE on a bounded region G, if for any finite interval $[t_0, T]$ there exist positive constants Δ_0 and K_0 such that for each $\varepsilon > 0$, each $\delta \in (0, \Delta_0)$, and each $\kappa \in (0, K_0)$,

$$\sup_{x \in G} \lim_{\left|U_0^{(\delta,\kappa)}(m_x) - \bar{U}_0^{(\delta,\kappa)}(m_x)\right| \to 0} \sup_{t_0 \leq t \leq T} P\big(|U_{n_t}^{(\delta,\kappa)}(m_x) - \bar{U}_{n_t}^{(\delta,\kappa)}(m_x)| \geq \varepsilon\big) = 0.$$

In the following we will study stability of the heat equation with additive and multiplicative noise. It turns out that for an explicit scheme there are no restrictions on the relation between the step size in time and in space. On the contrary for the corresponding deterministic problem there are bounds of the form $\delta/\kappa^2 \leq$ const. and Figure 2.2 shows that also with noise this scheme is not stable if $\delta = 1.5$ and $\kappa = 0.1$. Therefore the extension of Kloeden and Platen's definition of stability does not seem to be a good choice.

(a) original image, (b) instable simulation,

Figure 2.2.: Instable simulation of the heat equation with additive noise and parameters $\triangle t = 1.5$ and $\triangle x = 0.1$

Example 2.61 (Heat equation with additive noise)**.** Let

$$du(t) = \frac{1}{2}\Delta u(t)\,dt + \sigma(t)\,dW(t),$$

be the heat equation on a finite interval I with periodic boundary conditions in a finite time interval $[t_0, T]$ where $W(t)$ is an infinite dimensional Wiener process on I with given covariance operator Q and $\sigma : \mathbb{R}_+ \times I \to \mathbb{R}$. Then the following approximation scheme is stochastically numerically stable for all choices $h, k > 0$ where h is the step size in time and k is the step size in space, $n \in \mathbb{N}$ denotes the discretization steps in time, $m \in \mathbb{N}$ those in space,

$$U_{n+1}(m) = U_n(m) + \tfrac{1}{2}hk^{-2}(U_n(m+1) - 2U_n(m) + U_n(m-1)) + \sigma_n(m) \cdot \eta_{n+1}(m),$$

where $\sigma_n(m) := \sigma(t_n, x_m)$ and $\eta_{n+1}(m)$ is the approximation of $\int_{t_n}^{t_{n+1}} dW(s, x_m)$.

To prove this let U_0 and $\bar{U}_0$ be two different starting conditions, $U_n(m)$ and $\bar{U}_n(m)$ the corresponding approximations, and $\mathcal{U}_n := \sup_m \left|U_n(m) - \bar{U}_n(m)\right|$, then we have for every $h, k > 0$ fixed, $\alpha := h/k^2$,

$$\begin{aligned}
\left|U_n(m) - \bar{U}_n(m)\right| &\le \left|1 - \alpha\right| \left|U_{n-1}(m) - \bar{U}_{n-1}(m)\right| \\
&\quad + \tfrac{1}{2}\,\alpha \left(\left|U_{n-1}(m+1) - \bar{U}_{n-1}(m+1)\right| + \left|U_{n-1}(m-1) - \bar{U}_{n-1}(m-1)\right| \right) \\
&\le \left(\left|1 - \alpha\right| + \alpha\right) \mathcal{U}_n \\
&\le \left(\left|1 - \alpha\right| + \alpha\right)^n \mathcal{U}_0.
\end{aligned}$$

As for fixed h and k, n is bounded by $(T - t_0)/h$, the expression $(|1 - \alpha| + \alpha)^n$ is bounded, and therefore for $\mathcal{U}_0 \to 0$ implies stability in the sense of Kloeden and Platen.

Again we can show a similar stability result for the heat equation with multiplicative noise using Chebyshev's inequality:

Example 2.62 (Heat equation with multiplicative noise)**.** Let

$$du(t) = \frac{1}{2}\Delta u(t)\,dt + \sigma(t)u(t)\,dW(t),$$

be the heat equation on a finite interval I with periodic boundary conditions on a finite time interval $[t_0, T]$ where $W(t)$ is an infinite dimensional Wiener process on I with given covariance operator Q and $\sigma : \mathbb{R}_+ \times I \to \mathbb{R}$ bounded in space and time. Then the following approximation scheme is stochastically numerically stable for all choices $h, k > 0$ where h is the step size in time and k is the step size in space:

$$U_{n+1}(m) = U_n(m) + \tfrac{1}{2}\,hk^{-2}(U_n(m+1) - 2U_n(m) + U_n(m-1)) + \sigma_n(m)U_n(m)\cdot\eta_{n+1}(m)$$

where $\sigma_n(m) := \sigma(t_n, x_m)$ and $\eta_{n+1}(m)$ is the approximation of $\int_{t_n}^{t_{n+1}} dW(s, x_m)$.

This can be shown in the following way. Let $\varepsilon > 0$, U_0 and $\bar{U}_0$ two different starting conditions, $U_n(m)$ and $\bar{U}_n(m)$ the corresponding approximations, $\mathcal{U}_n := \sup_m \left\|U_n(m) - \bar{U}_n(m)\right\|_2$, and $\tilde{\sigma} := \sup_{n,m} |\sigma_n(m)|$, then it holds

$$\begin{aligned}
P(|U_n(m) - \bar{U}_n(m)| \geq \varepsilon) &\leq \varepsilon^{-2}\,\mathbb{E}(\left|U_n(m) - \bar{U}_n(m)\right|^2) \\
&= \varepsilon^{-2}\left\|U_n(m) - \bar{U}_n(m)\right\|_2^2 \\
&\leq \varepsilon^{-2}\big(|1-\alpha|\left\|U_{n-1}(m) - \bar{U}_{n-1}(m)\right\|_2 \\
&\qquad + \tfrac{1}{2}\,\alpha\left(\left\|U_{n-1}(m+1) - \bar{U}_{n-1}(m+1)\right\|_2\right. \\
&\qquad\qquad \left. + \left\|U_{n-1}(m-1) - \bar{U}_{n-1}(m-1)\right\|_2\right) \\
&\qquad + |\sigma_n(m)|\,\mathbb{E}(\eta_{n+1}(m)^2)^{1/2}\left\|U_{n-1}(m) - \bar{U}_{n-1}(m)\right\|_2\big)^2 \\
&\leq \varepsilon^{-2}\big(|1-\alpha| + \alpha + \sqrt{h}\,Q(m,m)^{1/2}\,\tilde{\sigma}\big)\,\mathcal{U}_{n-1}^2 \\
&\leq \varepsilon^{-2}\big(|1-\alpha| + \alpha + \sqrt{h}\,Q(m,m)^{1/2}\,\tilde{\sigma}\big)^{2n}\,\mathcal{U}_0^2
\end{aligned}$$

Similarly to Example 2.61 the first part of the last expression is bounded for every fixed h, k. And therefore for $\mathcal{U}_0 \to 0$ we have

$$P(|U_n(m) - \bar{U}_n(m)| \geq \varepsilon) \to 0.$$

2.3.3. Numerical Stability in the sense of Sewell

Another approach to numerical stability for SDEs and SPDEs is to extend the definition which is used by Sewell in [85] for deterministic ODEs and PDEs.

Definition 2.63. Let X_t be the solution of an SDE and Y_t an approximation with approximation scheme $Y_{t+h} = D(Y_t, Y_{t+h})$ and step size h. Then the approximation scheme is *consistent* if for all $t \geq 0$

$$\lim_{h\to 0} \|T_{t+h}\|_2 := \lim_{h\to 0} \|X_{t+h} - D(X_t, X_{t+h})\|_2 = 0,$$

where T_t is called the *truncation error* and $\|T_t\|_2^2 = \mathbb{E}(T_t^2)$. An approximation is called *numerically stable* if it is consistent and if the size of the *error* $\|e_t\|_2 := \|Y_t - X_t\|_2 \to 0$ for $\|T_t\|_2 \to 0$ for all $t \in [t_0, T]$.

To get an idea of how to work with this kind of stability, we will look at similar examples as in the previous section.

Example 2.64 (Linear SDE with additive noise)**.** Let

$$\begin{cases} dX_t = \lambda X_t\, dt + \sigma(t) dW_t, \\ X_0 = \xi, \end{cases}$$

be a given Cauchy problem in $\mathbb{R}$ with $\lambda \in \mathbb{R}$, $\sigma : [t_0, T] \to \mathbb{R}$ bounded with bounded derivative, and W_t a one-dimensional Wiener process. Then the discretization

$$Y_{t_{n+1}} = Y_{t_n} + h \cdot \lambda Y_{t_n} + \sigma(t)\eta_{t_{n+1}}$$

is numerically stable for all $h > 0$, where h is the step size and $\eta_{t_{n+1}} = \int_{t_n}^{t_{n+1}} dW_t$.

To prove this, we first have to prove consistency and therefore we have to give estimates on the truncation error. The solution of the SDE is

$$X_t = e^{\lambda t}\xi + \int_0^t e^{\lambda(t-s)}\sigma(s)\, dW_s,$$

and we can express X_{t+h} in terms of X_t in the following way:

$$X_{t+h} = e^{\lambda h} X_t + \int_0^h e^{\lambda(h-s)}\sigma(t+s)\, dW_{t+s}.$$

This leads to the truncation error

$$\begin{aligned} T_{t+h} &= X_{t+h} - D(X_t, X_{t+h}) \\ &= X_{t+h} - (1+h\lambda)X_t - \sigma(t)\eta_{t+h} \\ &= e^{h\lambda}X_t + \int_0^h e^{\lambda(h-s)}\sigma(t+s)\, dW_{t+s} - (1+h\lambda)X_t - \sigma(t)\eta_{t+h} \\ &= (e^{h\lambda} - (1+h\lambda))X_t + \int_0^h e^{\lambda(h-s)}\sigma(t+s) - \sigma(t)\, dW_{t+s} \\ &= \int_0^h \lambda e^{s\lambda} - \lambda\, ds\, X_t + \int_0^h e^{\lambda(h-s)}(\sigma(t+s) - \sigma(t)) + (e^{\lambda(h-s)} - 1)\sigma(t)\, dW_{t+s} \\ &= \lambda^2 \int_0^h \int_0^s e^{r\lambda}\, dr\, ds\, X_t + \int_0^h e^{\lambda(h-s)} \int_0^s \sigma'(t+r)\, dr\, dW_{t+s} \\ &\quad + \int_0^h \int_0^{h-s} \lambda\, e^{r\lambda}\sigma(t)\, dr\, dW_{t+s}. \end{aligned}$$

Next we calculate bounds on the summands.

$$\begin{aligned} \left\| \lambda^2 \int_0^h \int_0^s e^{r\lambda}\, dr\, ds\, X_t \right\|_2 &\leq \tfrac{1}{2}\, h^2 e^{|\lambda| h}\, \|X_t\|_2 \\ &= O(h^2), \end{aligned}$$

if $X_t \in L^2$ which is given by the assumptions on σ and ξ. For the second expression we use the Itô isometry and get

$$\begin{aligned}\left\| \int_0^h e^{\lambda(h-s)} \int_0^s \sigma'(t+r)\, dr\, dW_{t+s} \right\|_2^2 &= \int_0^h e^{2\lambda(h-s)} \big(\int_0^s \sigma'(t+r)\, dr \big)^2 ds \\ &\leq e^{2|\lambda|h} \sup_{s\in[t,t+h]} \left|\sigma'(s)\right|^2 \int_0^h s^2\, ds \\ &= O(h^3),\end{aligned}$$

as σ' is bounded. And finally we have using again the Itô isometry

$$\begin{aligned}\left\| \int_0^h \int_0^{h-s} \lambda\, e^{r\lambda} \sigma(t)\, dr\, dW_{t+s} \right\|_2^2 &= \int_0^h \big(\int_0^{h-s} \lambda\, e^{r\lambda} \sigma(t)\, dr \big)^2 ds \leq \lambda^2 e^{2|\lambda|h} \left|\sigma(t)\right|^2 \int_0^h (h-s)^2\, ds \\ &= O(h^3).\end{aligned}$$

Therefore the truncation error is bounded by

$$\begin{aligned}\|T_{t+h}\|_2 &\leq \left\| \lambda^2 \int_0^h \int_0^s e^{r\lambda}\, dr\, ds X_t \right\|_2 + \left\| \int_0^h e^{\lambda(h-s)} \int_0^s \sigma'(t+r)\, dr\, dW_{t+s} \right\|_2 \\ &\quad + \left\| \int_0^h \int_0^{h-s} \lambda e^{r\lambda} \sigma(t)\, dr\, dW_{t+s} \right\|_2 \\ &= O(h^{3/2})\end{aligned}$$

and the approximation is consistent with the SDE.

To prove stability we look at the error

$$e_{t_{n+1}} - (1+h\lambda) e_{t_n} = Y_{t_{n+1}} - D(Y_{t_n}, Y_{t_{n+1}}) - (X_{t_{n+1}} - D(X_{t_n}, X_{t_{n+1}})) = -T_{t_{n+1}}.$$

If we set $T_{\max} = \sup_n \|T_{t_n}\|_2$, we have

$$\|e_{t_n}\|_2 \leq |1+h\lambda| \left\|e_{t_{n-1}}\right\|_2 + \|T_{t_n}\|_2 \leq e^{\lambda t_n} (\|e_0\|_2 + n \cdot T_{\max}) \leq e^{\lambda t_n} t_n O(h^{1/2})$$

because $t_n = n \cdot h$, and numerical stability is proved.

In the case of multiplicative noise, a solution does not have to be given explicitly. Therefore, although an explicit solution is known, we will prove stability in the following example using integral equations. This is supposed to show in an easy way how to generalize this result to linear SPDEs later in this section.

Example 2.65 (Linear SDE with multiplicative noise)**.** Let

$$\begin{cases} dX_t &= \lambda X_t\, dt + X_t dW_t, \\ X_0 &= \xi, \end{cases}$$

be a given Cauchy problem in $\mathbb{R}$ with $\lambda \in \mathbb{R}$ and W_t is a real-valued Wiener process. Then the Milstein approximation

$$Y_{t_{n+1}} = Y_{t_n} + h \cdot \lambda Y_{t_n} + Y_{t_n} \eta_{t_{n+1}} + Y_{t_n} \tfrac{\eta_{t_{n+1}}^2 - h}{2} = (1 + \lambda h + \eta_{t_{n+1}} + \tfrac{\eta_{t_{n+1}}^2 - h}{2}) Y_{t_n}$$

is stochastically numerically stable for all $h > 0$ with $h = t_{n+1} - t_n$ for all n and $\eta_{t_{n+1}} = \int_{t_n}^{t_{n+1}} dW_t$.

Similarly to the previous example we first prove consistency. Therefore we observe that the SDE can be rewritten as the integral equation

$$X_t = X_0\, e^{\lambda t} + \int_0^t e^{\lambda(t-s)} X_s\, dW_s.$$

We could solve this SDE explicitly and give a similar proof to Example 2.64 but as explicit solutions are not known for all SDEs, we will choose this easy example to show how in general stability may be proved. Let us observe next that

$$X_{t+h} = e^{\lambda h} X_t + \int_t^{t+h} e^{\lambda(t+h-s)} X_s\, dW_s,$$

and for all $s > t$

$$X_s = e^{\lambda(s-t)} X_t + \int_t^s e^{\lambda(s-r)} X_r\, dW_r.$$

Then the truncation error can be calculated in the following way

$$\begin{aligned}
T_{t+h} &= X_{t+h} - (1 + \lambda h + \eta_{t+h} + \tfrac{\eta_{t+h}^2 - h}{2}) X_t \\
&= (e^{\lambda h} - (1+\lambda h)) X_t + \int_t^{t+h} e^{\lambda(t+h-s)} X_s\, dW_s - (\eta_{t+h} + \tfrac{\eta_{t+h}^2 - h}{2})\, X_t \\
&= \big(\int_0^h \lambda(e^{\lambda s} - 1)\, ds\big)\, X_t + \int_t^{t+h} e^{\lambda(t+h-s)} e^{(s-t)} X_t\, dW_s \\
&\quad + \int_t^{t+h} \int_t^s e^{\lambda(s-r)} X_r\, dW_r\, dW_s - (\eta_{t+h} + \tfrac{\eta_{t+h}^2 - h}{2})\, X_t \\
&= \big(\int_0^h \int_0^s \lambda^2 e^{\lambda h}\, dr\, ds\big)\, X_t + (e^{\lambda h} - 1) X_t \eta_{t+h} \\
&\quad + \int_t^{t+h} \int_t^s e^{\lambda(t+h-r)} e^{\lambda(r-t)} X_t\, dW_r\, dW_s - \tfrac{\eta_{t+h}^2 - h}{2} X_t \\
&\quad + \int_t^{t+h} \int_t^s \int_t^r e^{\lambda(t+h-r)} e^{\lambda(r-p)} X_p\, dW_p\, dW_r\, dW_s \\
&= \big(\int_0^h \int_0^s \lambda^2 e^{\lambda h}\, dr\, ds + \int_0^h \lambda e^{\lambda s}\, ds\, (\eta_{t+h} + \tfrac{\eta_{t+h}^2 - h}{2})\big)\, X_t \\
&\quad + \int_t^{t+h} \int_t^s \int_t^r e^{\lambda(t+h-p)} X_p\, dW_p\, dW_r\, dW_s.
\end{aligned}$$

This leads to the estimates

$$\begin{aligned}
\|T_{t+h}\|_2 &\le (O(h^2) + O(h)(\sqrt{h} + \frac{h}{\sqrt{2}}))\, \|X_t\|_2 + (\int_t^{t+h} \int_t^s \int_t^r e^{2\lambda(t+h-p)}\, \|X_p\|_2^2\, dp\, dr\, ds)^{1/2} \\
&\le O(h^{3/2})(\|X_t\|_2 + \max_{s\in[t,t+h]} \|X_s\|_2) \\
&\le O(h^{3/2}) \max_{s\in[t,t+h]} \|X_s\|_2
\end{aligned}$$

and therefore if the solution is in L^2, the approximation is consistent. Next we will check that the solution is in L^2. Let $s \in [0,T]$. The solution of the SDE is given by

$$X_s = X_0\, e^{W_s+(\lambda-1/2)s}$$

and this implies

$$\|X_s\|_2 = \|X_0\|_2\, e^{(\lambda-1/2)s}\, \left\|e^{W_s}\right\|_2 = \|\xi\|_2\, e^{(\lambda-1/2)s+s} = \|\xi\|_2\, e^{(\lambda+1/2)s}.$$

So for every finite time interval the solution is in L^2 for $\xi \in L^2$.

Finally we show stability by calculating the error

$$\begin{aligned} e_{t_n} &= (1+\lambda h+\eta_{t_n}+\tfrac{\eta_{t_n}^2-h}{2}\, e_{t_{n-1}}) - T_{t_n} \\ &= \prod_{i=1}^{n}(1+\lambda h+\eta_{t_i}+\tfrac{\eta_{t_i}^2-h}{2})\, e_{t_0} - \sum_{k=1}^{n}\prod_{i=k+1}^{n}(1+\lambda h+\eta_{t_i}+\tfrac{\eta_{t_i}^2-h}{2})\, T_{t_k} \\ &= -\sum_{k=1}^{n}\prod_{i=k+1}^{n}(1+\lambda h+\eta_{t_i}+\tfrac{\eta_{t_i}^2-h}{2})\, T_{t_k}, \end{aligned}$$

and we use the independence of the increments of the Wiener process which leads to

$$\begin{aligned} \|e_{t_n}\|_2 &\le \Big(\sum_{k=1}^{n}\prod_{i=k+1}^{n}\mathbb{E}\big((1+\lambda h+\eta_{t_i}+\tfrac{\eta_{t_i}^2-h}{2})^2\big)\,\mathbb{E}(T_{t_k}^2)\Big)^{1/2} \\ &\le O(h^{3/2})\Big(\sum_{k=1}^{n}\big(1+(1+2\lambda+(\lambda^2+\tfrac{1}{2})\,h)\,h\big)^{n-k}\Big)^{1/2} \\ &\le O(h^{3/2})\Big(\sum_{k=0}^{n-1}\exp\big((1+2\,|\lambda|+(\lambda^2+\tfrac{1}{2})\,h)\,hk\big)\Big)^{1/2} \\ &\le O(h^{3/2})\,n\,\exp\big((1+2\,|\lambda|+(\lambda^2+\tfrac{1}{2})\,h)\,h(n-1)/2\big) \\ &\le O(h^{1/2})\,t_n\,\exp\big((1+2\,|\lambda|+(\lambda^2+\tfrac{1}{2})\,h)\,t_n/2\big). \end{aligned}$$

This expression tends to zero for $T_{t_i} \to 0$ and the Milstein scheme is stable in the sense of Sewell.

Remark 2.66. In Example 2.65 a higher order Milstein scheme is needed to show stability while at the corresponding Example 2.59 a simpler Euler scheme was sufficient for stability in the sense of Kloeden and Platen. This scheme would not have been good enough for numerical stability in the sense of Sewell because in the calculations of the truncation error we had

$$\left\|\frac{\eta_{t_n}^2-h}{2}\right\|_2 = \frac{h}{\sqrt{2}}$$

and therefore the convergence of the truncation error would have been no better then $O(h)$. In the error estimates by recurrency, a factor of n appears which compensates the $O(h)$ to t_n, the fixed time where stability is checked. Therefore the final estimate is independent of h and for $h \to 0$ the error does not converge to zero. We remark that because of this, in all studied SDEs and SPDEs, the convergence of the truncation error had to be at least of order $O(h^{1+\varepsilon})$ with $\varepsilon > 0$ in order to get convergence of the error.

Next we will generalize Definition 2.63 to SPDEs. We could look at stability either pointwise or in terms of functions on a bounded region.

Definition 2.67. Let $u(t)$, $t \in [t_0, T]$, be the solution of an SPDE on a bounded region G with boundary conditions and $U(t)$ an approximation of the solution with approximation scheme $U(t+h) = D(U(t), U(t+h))$ and step size $h > 0$ in time and $\triangle x$ in space. Then the approximation scheme is *consistent* if for all $t \geq 0$

$$\lim_{h, \triangle x \to 0} \sup_{x \in G} \|T_{t+h}(x)\|_2 := \lim_{h, \triangle x \to 0} \sup_{x \in G} \|(u(t+h) - D(u(t), u(t+h)))(x)\|_2 = 0,$$

where $T_t(x)$ is called the *truncation error* at time t and point x.

An approximation is called *numerically stable*, if it is consistent and if for all $t \in [t_0, T]$, $x \in G$

$$\lim_{\|T_t(x)\|_2 \to 0} \|e_t(x)\|_2 := \lim_{\|T_t(x)\|_2 \to 0} \|(U(t) - u(t))(x)\|_2 = 0.$$

In the following, we will prove numerical stability of the heat equation with noise similarly to Example 2.61 and 2.62.

For the remainder of this section we will make the following assumptions and definitions. Let

$$W(t, x) = \sum_{k=0}^{\infty} \sqrt{a_k}\, e_k(x) \beta_k(t),$$

be a Hilbert space-valued Brownian motion with covariance operator Q and $a_k = (m^l + k^l)^{-n}$, where $m \in \mathbb{R}_+$, $l/2, n \in \mathbb{N}$, $l \geq 2$, and $\{e_k\}$ is a Hilbert space basis on the interval $[0, 2\pi)$ with periodic boundary conditions. The elements $\beta_k(t)$ are independent one-dimensional Brownian motions. Moreover let the approximation of $W(t)$ be given by

$$\eta_{t+h} = \sum_{k=0}^{(T/\triangle x)-1} \sqrt{a_k}\, e_k(x) \left(\beta_k(t+h) - \beta_k(t)\right),$$

which is computed by one of the algorithms given in Chapter 1. The stochastic process

$$\sum_{k=0}^{(T/\triangle x)-1} \sqrt{a_k}\, e_k(x) \beta_k(t),$$

converges almost surely to $W(t)$ for $\triangle x \to 0$ [12]. Furthermore let the Laplacian $\frac{1}{2}\Delta$ be discretized in the following way:

$$D\,U(t, x) := \tfrac{1}{2} (\triangle x)^{-2} \left(U(t, x + \triangle x) - 2\,U(t, x) + U(t, x - \triangle x)\right),$$

and calculations in x are done modulo 2π.

Proposition 2.68 (Heat equation with additive noise)**.** *Let*

$$\begin{cases} du(t) &= \frac{1}{2}\Delta u(t)\, dt + \lambda\, dW(t), \\ u(0) &= u_0, \end{cases}$$

be the heat equation with additive noise on the interval $[0, 2\pi)$ *with periodic boundary conditions. Let* $\lambda \in \mathbb{R}$ *and* $l \cdot n \geq 14$. *Assume that* $\left\|\Delta^2 u_0\right\|_2 < +\infty$ *for all* $t \in [0, T]$. *Then the explicit approximation scheme*

$$U(t+h) = (1 + hD)U(t) + \lambda\, \eta_{t+h}$$

is consistent and numerically stable for $h \leq (\triangle x)^2$.

Proof. We start with the calculation of the truncation error. Therefore we observe that the solution of the SPDE is given by

$$u(t) = P_t\, u_0 + \int_0^t P_{t-s}\, \lambda\, dW(s)$$

where $P_\cdot$ is the semigroup generated by $\frac{1}{2}\Delta$. Then $u(t+h)$ can be expressed recursively by

$$u(t+h) = P_h\, u(t) + \int_0^h P_{h-s}\, \lambda\, dW(t+s).$$

Thus the truncation error can be calculated similarly to Example 2.64.

$$\begin{aligned}
T_{t+h} &= u(t+h) - (1+hD)u(t) - \lambda\, \eta_{t+h} \\
&= (P_h - (1+hD))\, u(t) + \int_0^h P_{h-s}\, \lambda\, dW(t+s) - \lambda\, \eta_{t+h} \\
&= \int_0^h (P_s \tfrac{1}{2}\Delta - D)u(t)\, ds + \int_0^h (P_{h-s} - 1)\lambda\, dW(t+s) + \lambda(W(t+h) - W(t) - \eta_{t+h}) \\
&= \tfrac{1}{4}\int_0^h \int_0^s P_r\, \Delta^2 u(t)\, dr\, ds + h(\tfrac{1}{2}\Delta - D)u(t) \\
&\quad + \int_0^h \int_0^s P_r \tfrac{1}{2}\Delta\lambda\, dr\, dW(t+s) + \lambda(W(t+h) - W(t) - \eta_{t+h})
\end{aligned}$$

Next, the norm of the summands will be approximated.

1. $$\left\|\int_0^h \int_0^s P_r \Delta^2 u(t)\, dr\, ds\right\|_2 \leq \int_0^h \int_0^s \left\|\Delta^2 u(t)\right\|_2\, dr\, ds = O(h^2)\left\|\Delta^2 u(t)\right\|_2,$$

2. $$\left\|h(\tfrac{1}{2}\Delta - D)u(t)\right\|_2 \leq h \cdot O((\triangle x)^2)\left\|u(t)\right\|_2 = O(h(\triangle x)^2)\left\|u(t)\right\|_2.$$

3. Lemma 2.14, which can be applied as $W(t) \in D(\Delta)$ [48], and the Itô isometry lead to

$$\begin{aligned}
\left\|\int_0^h \int_0^s P_r \Delta\lambda\, dr\, dW(t+s)\right\|_2^2 &= |\lambda|^2 \left\|\int_0^h \int_0^s P_r\, dr\, d(\Delta W(s))\right\|_2^2 \\
&\leq |\lambda|^2 \int_0^h (\int_0^s \|\, P_r \,\|\, dr)^2\, \mathrm{tr}(\Delta^2 Q)\, ds \\
&\leq |\lambda|^2 \int_0^h s^2\, ds\, \mathrm{tr}(\Delta^2 Q) = O(h^3).
\end{aligned}$$

4. Finally,

$$\begin{aligned}\|\lambda(W(t+h)-W(t)-\eta_{t+h})\|_2 &= |\lambda| \Big\| \sum_{k=T/\triangle x}^{\infty} \sqrt{a_k}\, e_k(x)(\beta_k(t+h)-\beta_k(t))\Big\|_2 \\ &\le |\lambda| \sqrt{h} \sum_{k=T/\triangle x}^{\infty} \sqrt{a_k} \\ &\le |\lambda| \sqrt{h} \left(\frac{\triangle x}{T}\right)^5 \cdot \text{const.} = O(\sqrt{h}\,(\triangle x)^5),\end{aligned}$$

where the last inequality follows by the assumptions on the elements a_k.

This implies for the size of the truncation error

$$\|T_{t+h}\|_2 \le \left\|\Delta^2 u(t)\right\|_2 O(h^2) + O(h^{3/2}) + \|u(t)\|_2\, O(h(\triangle x)^2) + O(\sqrt{h}\,(\triangle x)^5)$$

and if all summands are bounded, consistency follows. The boundedness of the summands follows from the assumptions and only has to be checked for $\left\|\Delta^2 u(t)\right\|_2$ in the following.

$$\begin{aligned}\left\|\Delta^2 u(t)\right\|_2 &\le \left\|\Delta^2 P_t u_0\right\|_2 + \left\|\lambda\, \Delta^2 \int_0^t P_{t-s}\, dW(s)\right\|_2 \\ &\le \left\|\Delta^2 u_0\right\|_2 + |\lambda| \left\|\Delta^2 W(t) + \int_0^t P_{t-s}\Delta^3 W(s)\, ds\right\|_2 \\ &\le \left\|\Delta^2 u_0\right\|_2 + |\lambda|\,(t\cdot \operatorname{tr}(\Delta^4 Q) + t\cdot \operatorname{tr}(\Delta^6 Q))\end{aligned}$$

follows by applying Lemma 2.25 and the assumptions made above. The assumptions on u_0 and Q done in this example finally imply consistency.

Stability follows by setting $E_t = \sup_{x\in[0,2\pi)} \|e_t(x)\|_2$, $T_t = \sup_{x\in[0,2\pi)} \|T_t(x)\|_2$, $T\max = \sup_{t\in[0,T]} T_t$, and $\alpha = h/(\triangle x)^2$ and by the equation

$$e_{t+h}(x) = (1-\alpha)e_t(x) + \tfrac{1}{2}\alpha\,(e_t(x+\triangle x) + e_t(x-\triangle x)) - T_{t+h}(x).$$

For all $x \in [0, 2\pi)$ and $t_n \in [0, T]$, we have

$$\begin{aligned}\|e_{t_n}(x)\|_2 &\le \left|1-\alpha\right| \left\|e_{t_{n-1}}(x)\right\|_2 + \tfrac{1}{2}\alpha\, \left(\|e_t(x+\triangle x)\|_2 + \|e_t(x-\triangle x)\|_2\right) + \|T_t(x)\|_2 \\ &\le \left(\left|1-\alpha\right| + \alpha\right) E_{t_{n-1}} + T_{t_n} \\ &\le \left(\left|1-\alpha\right| + \alpha\right)^n E_{t_0} + \sum_{k=1}^{n} \left(\left|1-\alpha\right| + \alpha\right)^{n-k} T_{t_k} \\ &\le \left(\left|1-\alpha\right| + \alpha\right)^n n\, T\max.\end{aligned}$$

This is only bounded for $|1-\alpha| + \alpha \le 1$ and therefore $h \le (\triangle x)^2$ leads to

$$\|e_{t_n}(x)\|_2 \le n\, O(h^{3/2}) = t_n\, O(\sqrt{h})$$

and the approximation scheme is numerically stable for $h \le (\triangle x)^2$. □

Remark 2.69. The proof of Proposition 2.68 was done with semigroup theory. One can also prove stability with Fourier methods. Appendix B shows the corresponding calculations which are done in a slightly informal way but which give an idea of how to get stability without abstract mathematical theory at least in some cases. On the other hand, it is noticeable that the Fourier calculations are much longer and harder to read.

We extend the previous proposition to the case where the stochastic integral is of the form $\sigma(t)\,dW(t)$.

Proposition 2.70 (Heat equation with additive noise and time dependent coefficient)**.** *Let*

$$\begin{cases} du(t) &= \frac{1}{2}\Delta u(t)\,dt + \sigma(t)\,dW(t), \\ u(0) &= u_0, \end{cases}$$

be the heat equation with additive noise on the interval $[0, 2\pi)$ *and periodic boundary conditions. Let* $\sigma : \mathbb{R} \to \mathbb{R}$*, for all* $t \in [0,T]$ $|\dot{\sigma}(t)|$ *is bounded and* $l \cdot n \geq 14$*. Assume that* $\left\|\Delta^2 u_0\right\|_2 < +\infty$*. Then the explicit approximation scheme*

$$U(t+h) = (1+hD)U(t) + \sigma(t)\eta_{t+h}$$

is consistent and numerically stable for $h \leq (\triangle x)^2$*.*

Proof. This claim can be proved with the previous example. We first remark that $\sigma(t)\,dW(t) = d(\sigma(t)W(t))$ because of linearity. Therefore the SPDE is similar to Example 2.68 but the Q-Wiener process is transformed to a $(\sigma^2 Q)$-Wiener process which is not stationary.

Again we start calculating the truncation error. Therefore we observe that the solution of the SPDE is given by

$$u(t) = P_t u_0 + \int_0^t P_{t-s}\sigma(s)\,dW(s),$$

where $P_\cdot$ is the semigroup generated by $\frac{1}{2}\Delta$. Then $u(t+h)$ can be expressed recursively by

$$u(t+h) = P_h u(t) + \int_0^h P_{h-s}\sigma(t+s)\,dW(t+s).$$

Thus the truncation error can be calculated similarly to Proposition 2.68

$$\begin{aligned} T_{t+h} &= u(t+h) - (1+hD)u(t) - \sigma(t)\eta_{t+h} \\ &= (P_h - (1+hD))u(t) + \int_0^h P_{h-s}\,\sigma(t+s)\,dW(t+s) - \sigma(t)\eta_{t+h} \\ &= \int_0^h (P_s\,\tfrac{1}{2}\Delta - D)u(t)\,ds + \int_0^h P_{h-s}(\sigma(t+s) - \sigma(t))\,dW(t+s) \\ &\quad + \int_0^h (P_{h-s} - 1)\sigma(t)\,dW(t+s) + \sigma(t)(W(t+h) - W(t) - \eta_{t+h}) \\ &= \tfrac{1}{4}\int_0^h \int_0^s P_r \Delta^2 u(t)\,dr\,ds + h(\tfrac{1}{2}\Delta - D)u(t) \\ &\quad + \int_0^h \int_0^s P_{h-s}\,\dot{\sigma}(t+r)\,dr\,dW(t+s) + \sigma(t)\int_0^h \int_0^s P_r\,\tfrac{1}{2}\Delta\,dr\,dW(t+s) \\ &\quad + \sigma(t)(W(t+h) - W(t) - \eta_{t+h}). \end{aligned}$$

Next, the norm of the summands is approximated.

1. $$\Big\|\int_0^h \int_0^s P_r \Delta^2 u(t)\, dr\, ds\Big\|_2 \le \int_0^h \int_0^s \left\|\Delta^2 u(t)\right\|_2 dr\, ds = O(h^2) \left\|\Delta^2 u(t)\right\|_2,$$

2. $$\left\|h(\tfrac{1}{2}\Delta - D)u(t)\right\|_2 \le h \cdot O((\triangle x)^2) \left\|u(t)\right\|_2 = O(h(\triangle x)^2) \left\|u(t)\right\|_2,$$

3. the Itô isometry yields

$$\begin{aligned}\Big\|\int_0^h \int_0^s P_{h-s}\dot\sigma(t+r)\, dr\, dW(t+s)\Big\|_2^2 &\le \int_0^h \Big(\int_0^s |\dot\sigma(t+r)|\, dr\Big)^2 \operatorname{tr} Q\, ds \\ &\le \max_{s\in[t,t+h]} |\dot\sigma(s)|^2 \operatorname{tr} Q\, O(h^3),\end{aligned}$$

4. Lemma 2.14, which can be applied as $W(t) \in D(\Delta)$ [48] and the Itô isometry imply

$$\begin{aligned}\Big\|\sigma(t)\int_0^h \int_0^s P_r \Delta\, dr\, dW(t+s)\Big\|_2^2 &= \Big\|\sigma(t)\int_0^h \int_0^s P_r\, dr\, d(\Delta W(t+s))\Big\|_2^2 \\ &\le |\sigma(t)|^2 \operatorname{tr}(\Delta^2 Q) O(h^3),\end{aligned}$$

5. $$\begin{aligned}\|\sigma(t)(W(t+h) - W(t) - \eta_{t+h})\|_2 &= |\sigma(t)| \Big\| \sum_{k=T/\triangle x}^{\infty} \sqrt{a_k}\, e_k(x)(\beta_k(t+h) - \beta_k(t))\Big\|_2 \\ &\le |\sigma(t)| \sqrt{h} \sum_{k=T/\triangle x}^{\infty} \sqrt{a_k} \\ &\le |\sigma(t)| \sqrt{h} \left(\triangle x\, T^{-1}\right)^5 \cdot \text{const.} \\ &= |\sigma(t)|\, O(\sqrt{h}(\triangle x)^5),\end{aligned}$$

   where the last inequality follows by the assumptions on the elements a_k.

This implies for the size of the truncation error

$$\begin{aligned}\|T_{t+h}\|_2 \le \left\|\Delta^2 u(t)\right\|_2 O(h^2) &+ \big(\max_{s\in[t,t+h]} |\dot\sigma(s)| + |\sigma(t)|\big)\, O(h^{3/2}) \\ &+ \|u(t)\|_2\, O(h(\triangle x)^2) + |\sigma(t)|\, O(\sqrt{h}(\triangle x)^5),\end{aligned}$$

and if all summands are bounded, consistency follows. The boundedness of the summands follows from the assumptions and only has to be checked for $\left\|\Delta^2 u(t)\right\|_2$ in the following

$$\begin{aligned}\left\|\Delta^2 u(t)\right\|_2 &\le \left\|\Delta^2 P_t u_0\right\|_2 + \Big\|\Delta^2 \int_0^t P_{t-s}\sigma(s)\, dW(s)\Big\|_2 \\ &\le \left\|\Delta^2 u_0\right\|_2 + \Big\|\Delta^2 \int_0^t P_{t-s}\, d(\sigma(s)W(s))\Big\|_2.\end{aligned}$$

If we define $\tilde{W}(t) = \sigma(t)W(t)$, where $\tilde{W}(t)$ is Gaussian with mean zero and covariance $t\,\sigma(t)^2 Q$, we can use the proof of Example 2.68 and reduce the boundedness of the solution to

$$\left\|\Delta^2 u_0\right\|_2 + |\lambda|\left(t \cdot \mathrm{tr}(\Delta^4(\sigma(t)^2 Q)) + t \cdot \mathrm{tr}(\Delta^6(\sigma(t)^2 Q)\right) < +\infty.$$

Using the assumptions done in this example, consistency is proved.

Stability finally follows, setting $E_t = \sup_{x\in[0,2\pi)} \|e_t(x)\|_2$, $T_t = \sup_{x\in[0,2\pi)} \|T_t(x)\|_2$, $T\max = \sup_{t\in[0,T]} T_t$, and $\alpha = h/(\triangle x)^2$ and observing

$$e_{t+h}(x) = (1-\alpha)e_t(x) + \tfrac{1}{2}\alpha\left(e_t(x+\triangle x) + e_t(x-\triangle x)\right) - T_{t+h}(x).$$

For all $x \in [0, 2\pi)$ and $t_n \in [0, T]$, we have

$$\begin{aligned}
\|e_{t_n}(x)\|_2 &\le \left|1-\alpha\right| \left\|e_{t_{n-1}}(x)\right\|_2 + \tfrac{1}{2}\alpha\left(\|e_t(x+\triangle x)\|_2 + \|e_t(x-\triangle x)\|_2\right) + \|T_t(x)\|_2 \\
&\le \left(\left|1-\alpha\right| + \alpha\right) E_{t_{n-1}} + T_{t_n} \\
&\le \left(\left|1-\alpha\right| + \alpha\right)^n E_{t_0} + \sum_{k=1}^{n} \left(\left|1-\alpha\right| + \alpha\right)^{n-k} T_{t_k} \\
&\le \left(\left|1-\alpha\right| + \alpha\right)^n n\, T_{\max}.
\end{aligned}$$

This is only bounded for $|1-\alpha| + \alpha \le 1$ and therefore $h \le (\triangle x)^2$ leads to

$$\|e_{t_n}(x)\|_2 \le n\, O(h^{3/2}) = t_n\, O(\sqrt{h})$$

and the approximation scheme is numerically stable for $h \le (\triangle x)^2$. □

A combination of Example 2.64 and Proposition 2.68 leads to the following result.

Proposition 2.71 (Heat equation with multiplicative noise)**.** *Let*

$$\begin{cases} du(t) &= \frac{1}{2}\Delta u(t)\,dt + u(t)\,dW(t), \\ u(0) &= u_0, \end{cases}$$

be the heat equation with multiplicative noise on the interval $[0, 2\pi)$ with periodic boundary conditions and $l \cdot n \ge 10$. Assume that $\|\Delta^2 u_0\|_2 < +\infty$ and $\|\Delta^2 u\|_2$ exists with $\|\Delta^2 u\|_2 < +\infty$. Then the explicit approximation scheme

$$U(t+h) = (1+hD)\,U(t) + \left(\eta_{t+h} + \tfrac{1}{2}(\eta_{t+h}^2 - h)\right) U(t)$$

is consistent and numerically stable for $h < (\triangle x)^2$.

Proof. First we observe that the solution of the SPDE can be expressed as

$$u(t) = P_t u_0 + \int_0^t P_{t-s} u(s)\,dW_s$$

and

$$\begin{aligned}
u(t+h) &= P_h u(t) + \int_t^{t+h} P_{t+h-s} u(s)\, dW_s \\
&= P_h u(t) + \int_t^{t+h} P_{t+h-s}(P_{s-t}u(t) + \int_t^s P_{s-r} u(r)\, dW_r)\, dW_s \\
&= P_h \left(1 + (W_{t+h} - W_t) + \tfrac{1}{2}((W_{t+h} - W_t)^2 - h)\right)\, u(t) \\
&\quad + \int_t^{t+h} \int_t^s \int_t^r P_{t+h-p}\, u(p)\, dW_p\, dW_r\, dW_s,
\end{aligned}$$

where $P_\cdot$ is the semigroup generated by $\frac{1}{2}\Delta$. Then it is easy to calculate the truncation error using results from Example 2.68

$$\begin{aligned}
T_{t+h} &= u(t+h) - \left(1 + hD + \eta_{t+h} + \tfrac{1}{2}(\eta_{t+h}^2 - h)\right)\, u(t) \\
&= \Big(P_h\big(1 + (W_{t+h} - W_t) + \tfrac{1}{2}((W_{t+h} - W_t)^2 - h)\big) - \big(1 + hD + \eta_{t+h} + \tfrac{1}{2}(\eta_{t+h}^2 - h)\big)\Big)\, u(t) \\
&\quad + \int_t^{t+h} \int_t^s \int_t^r P_{t+h-p} u(p)\, dW_p\, dW_r\, dW_s \\
&= \tfrac{1}{4} \int_0^h \int_0^s P_r \Delta^2 u(t)\, dr\, ds + h(\tfrac{1}{2}\Delta - D)u(t) \\
&\quad + \int_0^h P_s \tfrac{1}{2}\Delta\big((W_{t+h} - W_t) + \tfrac{1}{2}((W_{t+h} - W_t)^2 - h)\big)\, u(t)\, ds \\
&\quad + (W_{t+h} - W_t) - \eta_{t+h} + \tfrac{1}{2}((W_{t+h} - W_t)^2 - h) - \tfrac{1}{2}(\eta_{t+h}^2 - h) \\
&\quad + \int_t^{t+h} \int_t^s \int_t^r P_{t+h-p}\, u(p)\, dW_p\, dW_r\, dW_s
\end{aligned}$$

Assume that $\left\|\Delta^2 u(t)\right\|_2 < +\infty$. This can be shown by standard contraction methods and expressed as restriction on the coefficients of the covariance of the Brownian motion and of the starting condition but this proof goes beyond the scope of this thesis. Consistency holds by the following estimates.

1. First observe that

$$\begin{aligned}
&\int_0^h P_s\, \Delta\Big[\big((W_{t+h} - W_t) + \tfrac{1}{2}((W_{t+h} - W_t)^2 - h)\big)\, u(t)\Big]\, ds \\
&\qquad = \int_0^h P_s \Big[(\Delta(W(t+h) - W(t)) + \Delta\tfrac{1}{2}((W_{t+h} - W_t)^2 - h)\Big] u(t)\, ds \\
&\qquad\quad + \int_0^h P_s \Big[(W_{t+h} - W_t) + \tfrac{1}{2}((W_{t+h} - W_t)^2 - h)\Big] \Delta u(t)\, ds \\
&\qquad\quad + 2\int_0^h P_s \Big[(\nabla\,(W(t+h) - W(t)) + \nabla\,\tfrac{1}{2}((W_{t+h} - W_t)^2 - h)\Big] \nabla\, u(t)\, ds.
\end{aligned}$$

Then we have to estimate the properties of the coefficients. This includes

$$\|\Delta(W(t+h) - W(t))\|_2^2 = h \cdot \mathrm{tr}(\Delta^2 Q),$$

using Lemma 2.14 and

$$\begin{aligned}
&\left\|\Delta(W(t+h)-W(t))^2\right\|_2^2 \\
&\qquad = \Big\|\Delta\Big(\sum_{k=1}^{\infty}\sqrt{a_k}(\beta_k(t+h)-\beta_k(t))e_k\Big)^2\Big\|_2^2 \\
&\qquad = \Big\|\sum_{i,j=1}^{\infty}\sqrt{a_i a_j}(\beta_i(t+h)-\beta_i(t))(\beta_j(t+h)-\beta_j(t))(i^2+2ij+j^2)e_i e_j\Big\|_2^2 \\
&\qquad \le \sum_{i\neq j} a_i a_j (i^2+2ij+j^2)^2 \mathbb{E}((\beta_i(t+h)-\beta_i(t))^2)\mathbb{E}((\beta_j(t+h)-\beta_j(t))^2) \\
&\qquad\quad + \sum_i a_i^2\, 16\, i^4\, \mathbb{E}((\beta_i(t+h)-\beta_i(t))^4) \\
&\qquad = h^2\Big(\sum_{i\neq j}(i^2+2ij+j^2)^2 a_i a_j + 48\sum_i i^4 a_i^2\Big) = O(h^2)
\end{aligned}$$

which follows by the convergence of the sums using the properties of the elements a_k and simple calculus. Similarly we have

$$\|\nabla\,(W(t+h)-W(t))\|_2 = h\cdot\operatorname{tr}(\Delta Q)$$

and

$$\begin{aligned}
&\left\|\nabla\,(W(t+h)-W(t))^2\right\|_2^2 \\
&\qquad = \Big\|\sum_{i,j=1}^{\infty}\sqrt{a_i a_j}(\beta_i(t+h)-\beta_i(t))(\beta_j(t+h)-\beta_j(t))(i+j)e_i e_j\Big\|_2^2 \\
&\qquad \le \sum_{i\neq j} a_i a_j (i+j)^2 \mathbb{E}((\beta_i(t+h)-\beta_i(t))^2)\mathbb{E}((\beta_j(t+h)-\beta_j(t))^2) \\
&\qquad\quad + \sum_i a_i^2\, 4\, i^2\, \mathbb{E}((\beta_i(t+h)-\beta_i(t))^4) \\
&\qquad = h^2\Big(\sum_{i\neq j}(i+j)^2 a_i\, a_j + 12\sum_i i^2 a_i^2\Big) = O(h^2).
\end{aligned}$$

So finally we have

$$\begin{aligned}
&\left\|\int_0^h P_s\Delta\left((W_{t+h}-W_t)+\tfrac{1}{2}((W_{t+h}-W_t)^2-h)\right)u(t)\,ds\right\|_2 \\
&\le h\Big((O(\sqrt{h}+O(h))\,\|u(t)\|_2 + (O(\sqrt{h}+O(h))\,\|\Delta u(t)\|_2 + (O(\sqrt{h})+O(h))\,\|\nabla\, u(t)\|_2)\Big) \\
&= O(h^{3/2})\left(\|u(t)\|_2 + \|\nabla\, u(t)\|_2 + \|\Delta u(t)\|_2\right).
\end{aligned}$$

2. The second estimate that has not been done before is

$$\Big\|\int_t^{t+h}\int_t^s\int_t^r P_{t+h-p}u(p)\,dW_p\,dW_r\,dW_s\Big\|_2 \le \max_{s\in[t,t+h]}\|u(s)\|_2\, O(h^{3/2}).$$

So finally

$$\begin{aligned}\|T_{t+h}\|_2 \leq \left\|\Delta^2 u(t)\right\|_2 O(h^2) + \|u(t)\|_2 O(h(\triangle x)^2) + O(\sqrt{h}(\triangle x)^5) \\ + \left(\|\Delta u(t)\|_2 + \|\nabla u(t)\|_2 + \max_{s\in[t,t+h]} \|u(t)\|_2\right) O(h^{3/2}),\end{aligned}$$

and therefore the the approximation scheme is consistent.

Next we show stability. Set $E_t = \sup_{x\in[0,2\pi)} \|e_t(x)\|_2$, $T_t = \sup_{x\in[0,2\pi)} \|T_t(x)\|_2$, $T\max = \sup_{t\in[0,T]} T_t$, and $\alpha = h/(\triangle x)^2$. Then

$$e_{t+h}(x) = \left(1 - \alpha + \eta_{t+h} + \tfrac{1}{2}(\eta_{t+h}^2 - h)\right) e_t(x) + \tfrac{1}{2}\alpha\left(e_t(x+\triangle x) + e_t(x-\triangle x)\right) - T_{t+h}(x).$$

For all $x \in [0, 2\pi)$ and $t_n \in [0, T]$, we have

$$\begin{aligned}\|e_{t_n}(x)\|_2 &\leq \left(\left\|1 - \alpha + \eta_{t_{n-1}} + \tfrac{1}{2}(\eta_{t_{n-1}}^2 - h)\right\|_2 + \alpha\right) E_{t_{n-1}} + \left\|T_{t_{n-1}}(x)\right\|_2 \\ &\leq \sum_{k=1}^{n} \prod_{j=k+1}^{n} \left(\left\|1 - \alpha + \eta_{t_j} + \tfrac{1}{2}(\eta_{t_j}^2 - h)\right\|_2 + \alpha\right) \|T_{t_k}(x)\|_2 \\ &\leq T\max \sum_{k=1}^{n} \prod_{j=k+1}^{n} \left(\sqrt{(1-\alpha)^2 + h + \tfrac{1}{2}h^2} + \alpha\right) \\ &\leq T\max \cdot n \cdot \left(\sqrt{(1-\alpha)^2 + h + \tfrac{1}{2}h^2} + \alpha\right)^n\end{aligned}$$

Finally we have to show for which choices of h and $\triangle x$ $a_n := (\sqrt{(1-\alpha)^2 + h + \frac{1}{2}h^2} + \alpha)^n$ is bounded for all n, i.e. when the sequence converges. Therefore let $h = t_n/n$, where t_n is the fixed time where the error is calculated, and $0 < \alpha = h/(\triangle x)^2 < 1$. Then a_n converges if $n \cdot (a_n - 1)$ converges. For α defined as above, we have

$$\begin{aligned}n\left(\sqrt{(1-\alpha)^2 + \frac{t_n}{n} + \frac{t_n^2}{2n^2}} + \alpha - 1\right) &= (1-\alpha)\, n \left(\sqrt{1 + \frac{t_n}{n(1-\alpha)^2}(1 + \frac{t_n}{2n})} - 1\right) \\ &\leq (1-\alpha)\, n \left(1 + \frac{t_n}{2n(1-\alpha)^2}\left(1 + \frac{t_n}{2n}\right) - 1\right) \\ &= \frac{t_n(1 + t_n/2n)}{2(1-\alpha)} \\ &\to \frac{t_n}{2(1-\alpha)} \qquad \text{for } n \to +\infty.\end{aligned}$$

So finally for $t_n < T$ and $0 < \alpha < 1$ we have

$$\|e_{t_n}(x)\|_2 \leq O(h^{3/2}) \cdot n \cdot \exp\left(\frac{t_n}{2 - 2\alpha}\right) \leq O(h^{1/2})\, t_n \exp\left(\frac{t_n}{2-2\alpha}\right) \leq O(h^{1/2})\, T \exp\left(\frac{T}{2-2\alpha}\right)$$

and numerical stability follows.

□

This concludes the second chapter. We have seen some theory on Hilbert space-valued SDEs. Numerical discretizations of theses SDEs were discussed and different definitions of numerical stability were introduced. The heat equation with additive and multiplicative noise was used to compare different numerical stability definitions. It turned out that numerical stability in the sense of Kloeden and Platen and in the sense of Sewell lead to different stability results.

3. Segmentation

3.1. Introduction to Level Sets

This section gives an overview about level sets. A good introduction to level set methods and its applications can be found in [72, 84]. The first applications in computer vision with geodesic active contours were published by Casellas et al. in [9] and Chan and Vese in [10].

Level sets are a possibility to represent curves in the plane. The goal is to find a representation that leads to easy and quick calculations of evolving curves, e.g. the motion by mean curvature. We will use level sets because this method represents curves in implicit form and therefore the representation is independent of the parameterization of the curve. That is why they are easier to handle than parameterized curves, especially when the curve splits up or is reunified.

Definition 3.1. Let $\Lambda \subset \mathbb{R}^2$ be a region and $\phi : \Lambda \to \mathbb{R}$ a continuous map. The *zero level set* Γ of ϕ is given by

$$\Gamma := \{(x, y) \in \Lambda,\, \phi(x, y) = 0\}.$$

Any level set Γ_α for a given element $\alpha \in \mathbb{R}$ is defined by

$$\Gamma_\alpha := \{(x, y) \in \Lambda,\, \phi(x, y) = \alpha\}.$$

In most cases we are given a closed curve. This curve can be interpreted as the zero level set Γ. Therefore in the beginning ϕ is just defined on Γ. There are many ways of how ϕ can be continued outside of Γ. For a smooth continuation we define the *signed distance function*.

Definition 3.2. Let Γ be an embedded simple closed curve in the region $\Lambda \subset \mathbb{R}^2$. The *signed distance function* $\phi : \Lambda \to \mathbb{R}$ is given by

$$\phi(x, y) \mid_{inside(\Gamma)} := d((x, y), \Gamma)$$

if $(x, y) \in \Lambda$ is inside of Γ and

$$\phi(x, y) \mid_{outside(\Gamma)} := -d((x, y), \Gamma)$$

else, where $d((x, y), \Gamma)$ is the Euclidean distance between the point (x, y) and the curve Γ. Depending on the simulation that should be done, the minus sign can be swapped, i.e. $\phi(x, y) < 0$ if (x, y) inside of Γ and $\phi(x, y) > 0$ if (x, y) outside of Γ. A general embedded closed curve can be seen as the union of simple closed curves. Then the signed distance function can be defined similarly to the one for simple closed curves.

An example of a signed distance function is shown in Figure 3.1. Note that the signed distance function is just one possibility to initialize ϕ. There are other possibilities as well. Evans and Spruck showed in [20] that if ϕ is initialized as the signed distance function, the solution of motion by mean curvature is the same as in classical motion by mean curvature with a parameterized curve. Moreover they showed in [20, Section 5] that the solution of the motion of the curve is independent of the initial choice of ϕ. The same was also shown independently by Chen et al. [11, Thm. 7.1].

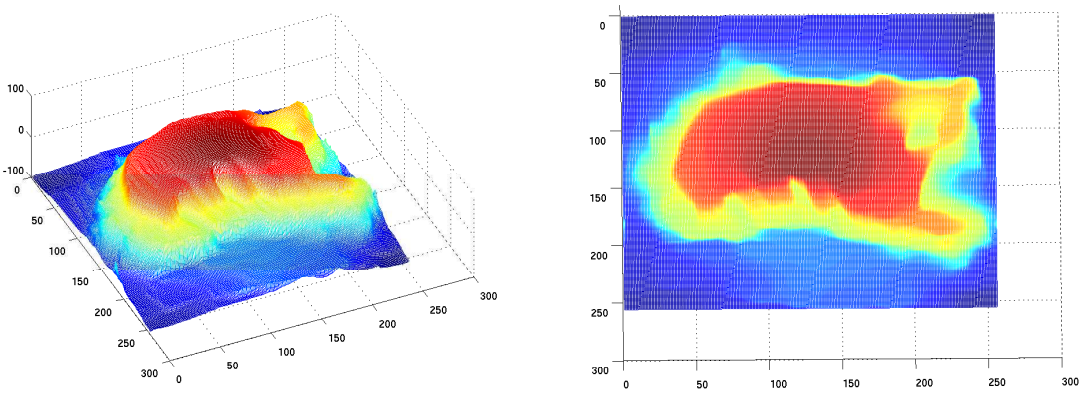


Figure 3.1.: Signed distance function of the segmentation result of a black bear.

3.2. Mathematical Background of Calculus of Variations

In the following, calculus of variations is used to derive the Euler–Lagrange equations. The resulting PDEs will be studied in the following sections, especially in the context of segmentation of images. Most of this section is due to [71]. We will work in the Euclidean space $\mathbb{R}^d$. More general results on smooth manifolds can be found in [1] and [91]. The explicit calculation of the first variation for the special case of $\mathbb{R}^2$ is due to [14].

In the following let $\Lambda \subset \mathbb{R}^d$ an open, connected set with smooth boundary $\partial\,\Lambda$.

Definition 3.3. A *variational problem* consists of finding the extrema (maxima or minima) of the functional

$$\mathcal{L}[u] = \int_\Lambda L(x, u^{(n)})\, dx$$

in some class of functions $u(x) = f(x)$ with $f : \Lambda \to \mathbb{R}^q$. The integrand $L(x, u^{(n)})$, called the *Lagrangian* of the variational problem $\mathcal{L}$, is a smooth function of x, u, and various derivatives of u.

In the following we denote by $\langle \cdot, \cdot \rangle$ the L^2 inner product

Definition 3.4. Let $\mathcal{L}[u]$ be a variational problem. The *variational derivative* of $\mathcal{L}$ is the unique q-tuple

$$\delta\mathcal{L}[u] = (\delta_1\mathcal{L}[u], \ldots, \delta_q\mathcal{L}[u]),$$

with the property that

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathcal{L}[f + \varepsilon\eta] = \int_\Lambda \delta\mathcal{L}[f(x)] \cdot \eta(x)\, dx \tag{3.1}$$

whenever $u(x) = f(x)$ is a smooth function defined on Λ, and $\eta(x) = (\eta^1(x), \ldots, \eta^q(x))$ is a smooth function with compact support in Λ, so that $f + \varepsilon\eta$ still satisfies any boundary conditions that might be entailed on the space of functions over which we are minimizing $\mathcal{L}$. The component $\delta_\alpha\mathcal{L} = \delta\mathcal{L}/\delta u^\alpha$ is the *variational derivative* of $\mathcal{L}$ with respect to u^α, $\alpha \in \{1, \ldots, q\}$.

Proposition 3.5 ([71, Prop. 4.2]). *If $u(x) = f(x)$ is an extremal of $\mathcal{L}[u(x)]$, then*

$$\delta\mathcal{L}[f(x)] = 0, \qquad x \in \Lambda. \tag{3.2}$$

For the case that $u = f : \mathbb{R}^2 \to \mathbb{R}$, $\delta\mathcal{L}$ can be calculated in the following way: Let $x = (x_1, x_2)$ and let $\mathcal{L}[u]$ just depend on u and the first partial derivatives. For the calculations the abbreviation $u_{x_i} := \frac{d}{dx_i} u$ is used. We search for an extremum of

$$\mathcal{L}[u] = \int_\Lambda L(x_1, x_2, u, u_{x_1}, u_{x_2})\, dx.$$

Therefore we have to calculate

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathcal{L}[f + \varepsilon\eta] = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \int_\Lambda L(x_1, x_2, f + \varepsilon\eta, f_{x_1} + \varepsilon\eta_{x_1}, f_{x_2} + \varepsilon\eta_{x_2})\, dx.$$

As we assume L to be smooth, we can change the order of integration and differentiation.

$$\begin{aligned}\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathcal{L}[f + \varepsilon\eta] &= \int_\Lambda \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} L(x_1, x_2, f + \varepsilon\eta, f_{x_1} + \varepsilon\eta_{x_1}, f_{x_2} + \varepsilon\eta_{x_2})\, dx \\ &= \int_\Lambda (\eta\, L_f + \eta_{x_1} L_{f_{x_1}} + \eta_{x_2} L_{f_{x_2}})\, dx.\end{aligned}$$

In order to get an expression of the form $\delta\mathcal{L} \cdot \eta$ we have to transform the partial derivatives of η. Therefore we need *Gauss's Theorem*, also known as *Green's Theorem* or *Divergence Theorem*, see e.g. [13], [18], [23] or in a very general form in [54]. A classical version is given in the following.

Theorem 3.6 (Gauss's Theorem [13, p. 320]). *Let f and g be smooth functions with smooth derivative mapping elements from a region $\Lambda \subset \mathbb{R}^2$ to $\mathbb{R}$, then*

$$\iint_\Lambda f_x(x, y) + g_y(x, y)\, dx\, dy = \int_{\partial\Lambda} (f(x, y)\, dy - g(x, y)\, dx)$$

where the right hand side is a way of writing down the line integral along $\partial\Lambda$.

If we apply Theorem 3.6 to our problem, we set $\tilde{f}(x, y) = \eta\, L_{f_x}$ and $g(x, y) = \eta\, L_{f_y}$. Then

$$\tilde{f}_x(x, y) = \eta_x\, L_{f_x} + \eta\, \frac{\partial}{\partial x} L_{f_x} g_y(x, y) = \eta_y\, L_{f_y} + \eta\, \frac{\partial}{\partial y} L_{f_y}$$

and Theorem 3.6 yields

$$\iint_\Lambda \eta_x\, L_{f_x} + \eta_y\, L_{f_y}\, dx\, dy = \int_{\partial\Lambda} \eta\, (L_{f_x}\, dy + L_{f_y}\, dx) - \iint_\Lambda \eta\, \big(\frac{\partial}{\partial x} L_{f_x} + \frac{\partial}{\partial y} L_{f_y}\big)\, dx\, dy.$$

As η was defined to have compact support and therefore should be zero on the boundary, the equation transforms to

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathcal{L}[f + \varepsilon\eta] = \int_\Lambda (L_f - \frac{\partial}{\partial x} L_{f_{x_1}} - \frac{\partial}{\partial y} L_{f_{x_2}}) \cdot \eta\, dx.$$

Comparing this expression with Equation (3.1), we conclude that

$$\delta\mathcal{L}[f(x)] = L_f - \frac{\partial}{\partial x}L_{f_{x_1}} - \frac{\partial}{\partial y}L_{f_{x_2}}$$

and by applying Proposition 3.5 it holds that if $u(x) = f(x)$ is an extremal of $\mathcal{L}[u(x)]$, then

$$L_f - \frac{\partial}{\partial x}L_{f_{x_1}} - \frac{\partial}{\partial y}L_{f_{x_2}} = 0.$$

In the following we state a theorem for the more general case of arbitrary dimensions. Therefore we first need another definition.

Definition 3.7. For $1 \leq \alpha \leq q$, the α-th *Euler operator* is given by

$$\mathrm{E}_\alpha = \sum_J (-D)_J \frac{\partial}{\partial u_J^\alpha},$$

the sum extending over all multi-indices $J = (j_1, \ldots, j_k)$ with $1 \leq j_k \leq d$, $k \geq 0$ and

$$(-D)_J := (-1)^k D_J = (-D_{j_1})(-D_{j_2}) \cdots (-D_{j_k}),$$

where D_i denotes the total derivative in the i-th direction. Note that to apply E_α to any given function $L(x, u^{(n)})$ of u and its derivatives, only finitely many terms in the summation are required, since L depends on only finitely many derivatives u_J^α.

Theorem 3.8 (Euler–Lagrange equations [71, Thm. 4.4])**.** *If $u(x) = f(x)$ is a smooth extremal of the variational problem $\mathcal{L}[u] = \int_\Lambda L(x, u^{(n)})\, dx$, then it must be a solution of the* Euler–Lagrange equations

$$E_\nu(L) = 0, \qquad \nu = 1, \ldots, q.$$

3.3. Active Contours

This section presents different energy functionals and the corresponding Euler–Lagrange equations for segmentation. These functionals seem to be good choices in order to get segmentation results that coincide with the intuitively expected results. By parameterizing the descent direction with an artificial time, PDEs are derived which can be simulated and will be modified to SPDEs later in this chapter. The calculations in this section are done informally because there is no rigorous mathematical proof so far and this extends the scope of this thesis.

3.3.1. An Energy Functional by Chan and Vese

The standard motion by mean curvature shrinks all curves to a circle and finally they vanish [26, 30]. Our goal is to use a similar kind of motion for segmentation. Therefore the evolution has to stop at an extremum. In other words we define an energy functional and try to minimize it by the corresponding Euler–Lagrange equation. The first approach to a variational level set method is due to Zhao et al. [99]. In the following we will use the energy functional proposed by Chan and Vese in [10] and do the calculation of the Euler–Lagrange equation explicitely. Let $\Lambda \subset \mathbb{R}^2$ be a region.

Definition 3.9. The *Heaviside function* H is defined by

$$H : \begin{cases} \Lambda & \to & \mathbb{R} \\ x & \mapsto & \begin{cases} 0 & \text{if} \quad x \leq 0 \\ 1 & \text{if} \quad x > 0. \end{cases} \end{cases}$$

Moreover the *Delta function* δ is defined as the distributional derivative of the Heaviside function, i.e.

$$\delta(x) := H'(x).$$

Similarly to Section 3.1, let Γ be a curve given by the zero level set of a function ϕ initialized as signed distance function. Denote in the following the evolving curve by C, depending on space and time and C is the zero level set of $\phi(x, y)$. Then the length of the curve is given by

$$\mathrm{L}(C) = \iint_C \sqrt{\phi_x^2 + \phi_y^2}\, dx\, dy \quad = \quad \int_\Lambda \delta(\phi(x,y))\, |\nabla\, \phi(x,y)|\, dx\, dy.$$

Next we calculate the area inside and outside of the curve denoted by *in* and *out*.

$$\begin{aligned} \mathrm{A}(in(C)) &= \int_{in(C)} 1\, dx\, dy \;= \int_\Lambda H(\phi(x,y))\, dx\, dy, \\ \mathrm{A}(out(C)) &= \int_{out(C)} 1\, dx\, dy \;= \int_\Lambda (1 - H(\phi(x,y)))\, dx\, dy. \end{aligned}$$

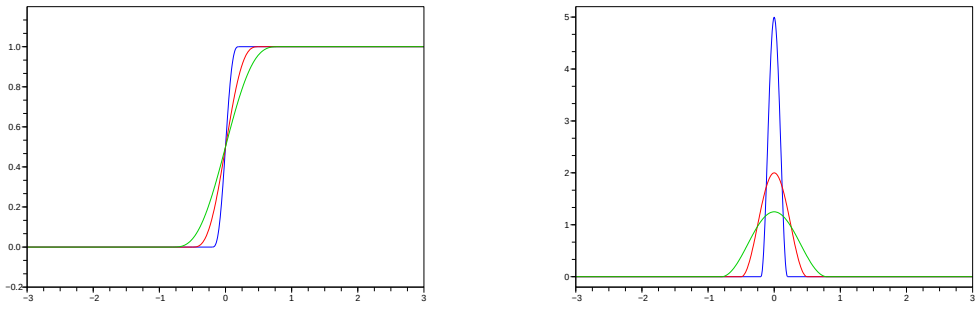
In order to detect edges in a picture we have to understand what "edge" means in the mathematical language. We can think of it as a place with high gradient in normal direction. If we think in terms of energy, our goal is to achieve that the difference of the mean value and the value at every point inside the curve is as small as possible and the same holds for the part outside the curve. Let $u_0(x, y)$ be the value of the picture at (x, y), e.g. the gray value at the point (x, y). Define c_1 to be the average value inside the curve and c_2 the average value outside the curve. If ϕ is fixed, these values are calculated

$$\begin{aligned} c_1(\phi) &= \int_{in(C)} u_0(x,y)\, dx\, dy \,/\, \mathrm{A}(in(C)) \\ &= \int_\Lambda u_0(x,y) H(\phi(x,y))\, dx\, dy \,/ \int_\Lambda H(\phi(x,y))\, dx\, dy, \\ c_2(\phi) &= \int_{out(C)} u_0(x,y)\, dx\, dy \,/\, \mathrm{A}(out(C)) \\ &= \int_\Lambda u_0(x,y)\, (1 - H(\phi(x,y)))\, dx\, dy \,/ \int_\Lambda (1 - H(\phi(x,y)))\, dx\, dy. \end{aligned}$$

Now we are able to formulate the energy $\tilde{J}$ that should be minimized. To this end let

$$\tilde{J}(c_1, c_2, \phi) := \int_\Lambda |u_0(x,y) - c_1|^2\, H(\phi(x,y))\, dx\, dy + \int_\Lambda |u_0(x,y) - c_2|^2\, (1 - H(\phi(x,y)))\, dx\, dy.$$

It turns out that the minimum of $\tilde{J}$ might not be the curve that gives the desired segmentation result. The modified functional proposed in [10] also depends on the length of the curve



(a) Heaviside function H_ε, (b) delta function δ_ε,

Figure 3.2.: The regularized Heaviside function and delta function for $\varepsilon = 0.2, 0.5, 0.8$.

and the area inside the curve. Another modeling question that arises is how to weight the different curve properties in the resulting energy functional. Denote the weighting coefficients by μ, ν, λ_1, λ_2 and define the energy functional by

$$\begin{aligned} J(c_1, c_2, \phi) := \mu \int_\Lambda \delta(\phi(x,y)) \left|\nabla \phi(x,y)\right| \, dx\, dy \\ + \ \nu \int_\Lambda H(\phi(x,y)) \, dx\, dy \\ + \ \lambda_1 \int_\Lambda \left|u_0(x,y) - c_1\right|^2 H(\phi(x,y)) \, dx\, dy \\ + \ \lambda_2 \int_\Lambda \left|u_0(x,y) - c_2\right|^2 (1 - H(\phi(x,y))) \, dx\, dy. \end{aligned}$$

The existence of minimizers of this functional can be shown by lower-semicontinuity of the total variation and classical arguments of calculus of variations. Similar equations were derived in [67, 68].

Next the Euler–Lagrange equation is calculated but therefore we first take a regularization of the Heaviside function and denote the regularized functional of J by J_ε. Define H_ε to be the regularized version of the Heaviside function [99] by

$$H_\varepsilon(x) := \begin{cases} 0 & \text{if} \quad x < -\varepsilon \\ 2^{-1}(1 + x/\varepsilon + \pi^{-1} \sin(\pi x/\varepsilon)) & \text{if} \quad -\varepsilon \le x \le \varepsilon \\ 1 & \text{if} \quad x > \varepsilon, \end{cases}$$

and similarly to the standard delta function the regularized version δ_ε by

$$\delta_\varepsilon(x) := H'_\varepsilon(x).$$

Figure 3.2 shows a plot of H_ε and δ_ε for $\varepsilon = 0.2, 0.5, 0.8$.

The regularized energy functional is given by

$$\begin{aligned} J_\varepsilon(c_1, c_2, \phi) := \int_\Lambda & \mu\, \delta_\varepsilon(\phi(x,y))\, |\nabla\, \phi(x,y)| + \ \nu\, H_\varepsilon(\phi(x,y)) \\ & + \ \lambda_1\, |u_0(x,y) - c_1|^2\, H_\varepsilon(\phi(x,y)) \\ & + \ \lambda_2\, |u_0(x,y) - c_2|^2\, (1 - H_\varepsilon(\phi(x,y)))\ dx\, dy. \end{aligned}$$

In order to calculate the Euler–Lagrange equations we observe that the Lagrangian is given by

$$\begin{aligned} F(x, y, c_1, c_2, \phi, \phi_x, \phi_y) = \mu\, & \delta_\varepsilon(\phi(x,y))\, |\nabla\, \phi(x,y)| + \ \nu\, H_\varepsilon(\phi(x,y)) \\ & + \ \lambda_1\, |u_0(x,y) - c_1|^2\, H_\varepsilon(\phi(x,y)) \\ & + \ \lambda_2\, |u_0(x,y) - c_2|^2\, (1 - H_\varepsilon(\phi(x,y))). \end{aligned}$$

The partial derivatives necessary for the usage of Theorem 3.8 are:

$$\begin{aligned} &\frac{\partial}{\partial \phi} |\nabla\, \phi| = 0, && \frac{\partial}{\partial \phi} |u_0 - c_1|^2 = 0, && \frac{\partial}{\partial \phi} |u_0 - c_2|^2 = 0, \\ &\frac{\partial}{\partial \phi} H_\varepsilon(\phi) = \delta_\varepsilon(\phi), && \frac{\partial}{\partial \phi} \delta_\varepsilon(\phi) = 0, && \text{for } \varepsilon \text{ small enough.} \end{aligned}$$

This implies

$$F_\phi = \nu\, \delta_\varepsilon(\phi) + \lambda_1\, |u_0 - c_1|^2\, \delta_\varepsilon(\phi) + \lambda_2\, |u_0 - c_2|^2\, \delta_\varepsilon(\phi).$$

Moreover we calculate the following partial derivatives

$$\begin{aligned} \frac{\partial}{\partial \phi_x} \delta_\varepsilon(\phi) = 0 &= \frac{\partial}{\partial \phi_y} \delta_\varepsilon(\phi), \\ \frac{\partial}{\partial \phi_x} H_\varepsilon(\phi) = 0 &= \frac{\partial}{\partial \phi_y} H_\varepsilon(\phi), \\ \frac{\partial}{\partial \phi_x} |u_0 - c_1|^2 = 0 &= \frac{\partial}{\partial \phi_y} |u_0 - c_1|^2, \\ \frac{\partial}{\partial \phi_x} |u_0 - c_2|^2 = 0 &= \frac{\partial}{\partial \phi_y} |u_0 - c_2|^2. \end{aligned}$$

The only derivatives with respect to ϕ_x and ϕ_y that are not equal to zero are

$$\begin{aligned} \frac{\partial}{\partial \phi_x} |\nabla\, \phi| &= \frac{\partial}{\partial \phi_x} \sqrt{\phi_x^2 + \phi_y^2} = \frac{\phi_x}{|\nabla\, \phi|}, \\ \frac{\partial}{\partial \phi_y} |\nabla\, \phi| &= \frac{\partial}{\partial \phi_y} \sqrt{\phi_x^2 + \phi_y^2} = \frac{\phi_y}{|\nabla\, \phi|}. \end{aligned}$$

So we have

$$\begin{aligned} F_{\phi_x} &= \mu\, \delta_\varepsilon(\phi)\, \frac{\phi_x}{|\nabla\, \phi|}, \\ F_{\phi_y} &= \mu\, \delta_\varepsilon(\phi)\, \frac{\phi_y}{|\nabla\, \phi|}. \end{aligned}$$

Finally for ε small enough we have

$$\frac{\partial}{\partial x}\,\delta_\varepsilon(\phi(x,y)) = 0,$$
$$\frac{\partial}{\partial y}\,\delta_\varepsilon(\phi(x,y)) = 0,$$

so that the last two terms of the Euler–Lagrange equation can be calculated.

$$\frac{\partial}{\partial x}F_{\phi_x} = \mu\,\delta_\varepsilon(\phi)\frac{\phi_{xx}\phi_y^2 - \phi_{xy}\phi_x\phi_y}{|\nabla\,\phi|^3},$$
$$\frac{\partial}{\partial y}F_{\phi_y} = \mu\,\delta_\varepsilon(\phi)\frac{\phi_{yy}\phi_x^2 - \phi_{xy}\phi_x\phi_y}{|\nabla\,\phi|^3}$$

Putting everything together and applying Proposition C.6 and Proposition C.7 we get

$$\begin{aligned}
&\frac{\partial}{\partial x}F_{\phi_x}+\frac{\partial}{\partial y}F_{\phi_y} - F_\phi \\
&\quad= \delta_\varepsilon(\phi)\left(\mu\,\frac{\phi_{xx}\phi_y^2 - 2\phi_{xy}\phi_x\phi_y + \phi_{yy}\phi_x^2}{|\nabla\,\phi|^3} - \nu - \lambda_1\ |u_0-c_1|^2 + \lambda_2\ |u_0-c_2|^2\right) \\
&\quad= \delta_\varepsilon(\phi)\left(\mu\,\kappa(\phi) - \nu - \lambda_1\ |u_0-c_1|^2 + \lambda_2\ |u_0-c_2|^2\right) = 0.
\end{aligned}$$

If we parameterize the descent direction by an artificial time t, i.e. the greater the time the smaller the energy, we obtain the following non-linear partial differential equation

$$\frac{\partial\phi}{\partial t} = \delta_\varepsilon(\phi)\,(\mu\,\kappa(\phi) - \nu - \lambda_1\,|u_0-c_1|^2 + \lambda_2\,|u_0-c_2|^2) = 0,$$

where $\kappa(\phi) = \nabla\cdot\frac{\nabla\,\phi}{|\nabla\,\phi|}$ denotes motion by mean curvature as introduced in Appendix C. As initial condition we want $\phi(x,y,0)$ to be the signed distance function initialized by the given curve Γ. In [10] the authors also require that

$$\frac{\delta_\varepsilon(\phi)}{|\nabla\,\phi|}\frac{\partial\phi}{\partial n} = 0$$

on the boundary $\partial\Lambda$ where $\frac{\partial\phi}{\partial n}$ is the normal derivative and the normal n is pointing outwards. Therefore they require a Neumann boundary condition, if the curve is on the boundary.

In the following we will use periodic boundary conditions because in the simulations we would like to extend this PDE to an SPDE and therefore we would like to use the random fields with periodic boundary conditions that were generated in Chapter 1. For more details on curvature see also Appendix C.

3.3.2. Energy Functionals Based on Gaussian Distributions

In the following a single Gaussian model by Rousson and Deriche [83] is summarized. A similar summary can be found in [47]. In their unsupervised segmentation framework [83], the authors model each region of a gray-valued or color image I by a single Gaussian distribution

of unknown mean μ_i and variance Q_i, where $\mu_i \in \mathbb{R}^m$ and Q_i is an $m \times m$-matrix, where m is the number of different image features, e.g. the color channels red, green, and blue. The case of two region segmentation turns into minimizing the following energy:

$$E(\Gamma, \mu_1, Q_1, \mu_2, Q_2) = \int_{in(\Gamma)} e_1(x)\, dx + \int_{out(\Gamma)} e_2(x)\, dx + \nu L(\Gamma),$$

where the notations are similar to those of Section 3.3.1 and $e_i(x) = -\log p_{\mu_i Q_i}(I(x))$ with

$$p_{\mu_i Q_i}(I(x)) = ((2\pi)^m \det(Q_i))^{-1/2} \exp(-\tfrac{1}{2}\langle I(x) - \mu_i, Q_i^{-1}(I(x) - \mu_i)\rangle)$$

being the m-dimensional Gaussian density for a given value $I(x)$ with respect to the hypothesis (μ_i, Q_i). The parameters (μ_i, Q_i), estimated from the pixel currently inside and outside Γ, are functions of Γ. Thus, the energy is a function just depending on Γ. Its Euler–Lagrange equation is not obvious but finally simplifies to the minimization dynamics

$$\beta_c = e_2(x) - e_1(x) + \nu\, \kappa(u),$$

where $\kappa(u)$ denotes motion by mean curvature as introduced in Appendix C. This approach leads to successful segmentations of two regions, even when they have the same mean but only different variances. But also the PDE resulting of this approach can easily get stuck in "false" local minimum.

In [47], Juan et al. extend this model to region statistics modeled by a mixture of Gaussian distributions of parameters $\Theta_i = (\pi_i^1, \mu_i^1, Q_i^1, \ldots, \pi_i^{n_i}, \mu_i^{n_i}, Q_i^{n_i})$ with $\sum_{j=1}^{n_i} \pi_i^j = 1$, $\pi_i^j \in [0, 1]$, in order to give an example where the Euler–Lagrange equation cannot be computed. Then the Gaussian density for a given value $I(x)$ is given by

$$p_{\Theta_i}(I(x)) = \sum_{j=1}^{n_i} \pi_i^j p_{\mu_i^j Q_i^j}(I(x)).$$

The next question that arises is how to estimate the number of Gaussian distributions to be chosen. Juan et al. suggest to estimate these either at the initial time step or dynamically using a minimum description length criterion [82] or the minimal message length method [92]. For the problem of estimating Θ_i from input samples they finally chose the K-means algorithm by MacQueen [64] which seemed to be the best for their application after they had tested several algorithms to estimate the parameters. The new segmentation problem consists of minimizing the energy previously defined in [83] with $e_i = -\log p_{\Theta_i}(I(x))$. The dependency of Θ_i with respect to Γ is a complex problem. The learning algorithm that estimates Θ_i acts as a "black box" implementing $\Gamma \mapsto \Theta_i(\Gamma)$. As a consequence, the Euler–Lagrange equation of the energy $E(\Gamma, \Theta_1(\Gamma), \Theta_2(\Gamma)) = E(\Gamma)$ cannot be computed explicitly. If, similar to the case of one Gaussian distribution, the evolution is driven by $\beta_c = e_2 - e_1 + \nu\kappa$, where κ is the mean curvature, the deterministic algorithm might also get stuck in local minima because the evolution is not driven by the exact coefficients but just by a guessed one. In order to overcome "false" local minima in a better way, stochastic active contours are introduced in the next section.

3.4. Stochastic Active Contours

This section is devoted to generalize the resulting PDEs of the last section to different types of SPDEs. First we will review a paper by Juan et al. [47] on which this thesis is based. Afterwards possible modifications of the stochastic modeling are presented.

3.4.1. Idea of Juan, Keriven, and Postelnicu

This paragraph summarizes the paper "Stochastic Motion and the Level Set Method in Computer Vision: *Stochastic Active Contours*" by Juan, Keriven, and Postelnicu [47] which was the starting point of this thesis. The paper suggests using stochastics based on recent work by Lions and Souganidis [57, 58, 60, 61] to overcome "false" local minima and to get better segmentation results.

Let $F = F(D^2u, Du, x, t)$ be a deterministic driving force, $W(t)$ a one-dimensional Brownian motion, and the SPDEs

$$du = F(D^2u, Du, x, t)\,dt + \varepsilon\,|Du| \circ dW(t), \tag{3.3}$$

$$du = F(D^2u, Du, x, t)\,dt + |Du|\,\dot{\xi}_\alpha(t) \tag{3.4}$$

be given, where $\varepsilon \geq 0$ and ξ_α is a family of smooth functions $\xi_\alpha : \mathbb{R}_+ \to \mathbb{R}$. Then the properties of these equations are proved in the papers of Lions and Souganidis and summarized in the following theorem which can be found in [47].

Theorem 3.10 ([57, 58, 60, 61])**.** *The following hold a.s. in ω.*

1. *There exists a unique solution to Equation (3.3).*
2. *Let $\{\xi_\alpha(t)\}_{\alpha>0}$ and $\{\eta_\beta(t)\}_{\beta>0}$ be two families of smooth functions such that as α and $\beta \to 0$, ξ_α and η_β converge to W uniformly in $t \in [0,T]$ for any T and a.s. in ω. Let $\{u_\alpha\}_{\alpha>0}$ and $\{v_\beta\}_{\beta>0}$ in $BUC(\mathbb{R}_+ \times \mathbb{R}^N)$ (bounded uniformly continuous) be the corresponding solutions of Equation (3.4). If $\lim_{\alpha,\beta\to 0} \| u_\alpha(\cdot,0) - v_\beta(\cdot,0) \|_{C(\mathbb{R}^N)} = 0$, then, for all $T > 0$, $\lim_{\alpha,\beta\to 0} \| u_\alpha - v_\beta \|_{C([0,T]\times\mathbb{R}^N)} = 0$. In particular, any smooth approximation of W produces solutions converging to the unique stochastic viscosity solution of Equation (3.3).*
3. *As $\varepsilon \to 0$, the solution u_ε of Equation (3.3) converges in $C(\mathbb{R}_+ \times \mathbb{R}^N)$ to the solution of Equation (3.3) with $\varepsilon = 0$.*

Remark 3.11. For an introduction to viscosity solutions for deterministic PDEs, the reader is referred to the "User's Guide" by Crandall, Ishii, and Lions [15], the book of Barles [3], and the book of Fleming and Soner [21]. An extension of this theory to SPDEs was done by Lions and Souganidis [57, 58, 60, 61].

The evolution of a curve is simulated as the zero level set of an N-dimensional region Λ. The authors use the Stratonovich convention because then the result is independent of the initialization of all points except the zero level set. This can easily be seen by the Itô formula. Therefore let $\alpha : \mathbb{R} \to \mathbb{R}$ be a smooth increasing function with $\alpha(0) = 0$. If $u(t,x)$ is the solution of the SPDE

$$du = |Du|\ dW(t), \tag{3.5}$$

given some initial condition $u_0(x)$, then $\alpha(u(t,x))$ is not the solution of Equation (3.5) with initial condition $\alpha(u_0(x))$. Moreover, the solution of Equation (3.5) with initial condition $\alpha(u_0(x))$ has not the same zero level set as $u(t,x)$. Therefore the evolution depends on the implicit representation of the initial curve. This can be seen in the following calculations pointed out in [59] and mentioned in [47]. Consider a function $\alpha : \mathbb{R} \to \mathbb{R}$ such that $\alpha' > 0$ and $\alpha(0) = 0$, and the initial value problem (3.5) with initial condition $u(0,\cdot) = u_0(\cdot)$. If we consider the solution of (3.5), then $v = \alpha(u)$ should verify the same dynamics but with a different initial condition $v(0,\cdot) = \alpha(u_0(\cdot))$ as is the case in the deterministic framework. Using the Itô formula, we get

$$dv = d\alpha(u) = \alpha'(u)\, du + \frac{1}{2}\alpha''(u)(du)^2 = |Dv|\ dW(t) + \frac{1}{2}\alpha''(u)\,|Du|^2\ dt,$$

and the assertion is not verified because there is the additional term $\frac{1}{2}\alpha''(u)\,|Du|^2\ dt$, . Hence the problem is ill-posed from a level sets point of view. This problem can be overcome by using the Stratonovich convention because it obeys the chain rule known from real analysis [49], i.e. for the initial value problem

$$\begin{cases} du = |Du| \circ dW(t), \\ u(0,\cdot) = u_0(\cdot), \end{cases} \tag{3.6}$$

it holds for a smooth increasing function $\alpha : \mathbb{R} \to \mathbb{R}$ with $\alpha(0) = 0$, that if u is a solution of Problem (3.6), then it is also a solution of

$$\begin{cases} d\alpha(u) = |D\alpha(u)| \circ dW(t), \\ \alpha(u(0,\cdot)) = \alpha(u_0(\cdot)). \end{cases}$$

Therefore the solution is independent of the initialization of the implicit function, if only the zero level set is given. For the Itô case we finally observe that the equation is invariant in case that α is a linear transformation.

The authors of [47] simulate a nonlinear equation of the form

$$du(t,x) = F(D^2u(t,x), Du(t,x), x, t)\, dt + |Du(t,x)| \sum_{i=1}^{m} \phi_i(x) \circ dW_i(t),$$

where $W(t) = (W_1(t), \ldots, W_m(t))$ is an m-dimensional Brownian motion, F is the diffusion coefficient, and the elements $\phi_i : \mathbb{R}^N \to \mathbb{R}$ are smooth functions with compact support, i.e. there is a certain finite number of *noise sources* x_i and

$$\phi_i(x) = \phi(x - x_i)$$

where ϕ is some convenient regular function and for all i, j it holds that $\phi_i(x_j) = \delta_{ij}$. The authors modify the elements ϕ_i to $\widetilde{\phi}_i(x) := \phi_i(x)/(\sum_{j=1}^{m} \phi_j^2(x))$ in order to have a constant unit variance in the whole simulation area. We remark that the existence of a viscosity solution for this equation is still an open problem. As implicit schemes for nonlinear equations can not

be simulated efficiently (see also Section 3.5.2), the authors transform their equation by the Wong-Zakai theorem [94] to the following Itô SDE

$$du = F\,dt + |Du| \sum_{i=1}^{m} \phi_i(x)\,dW_i(t) + \frac{1}{2}\Big(\big(\sum_{i=1}^{m} \phi_i^2(x)\big) D^2 u\, \frac{Du}{|Du|} + \big(\sum_{i=1}^{m} \phi_i(x) D\phi_i(x)\big) \cdot Du \Big)\, dt.$$

To make the expression better readable, we omitted the arguments of the functions. These are similar to the equation written with Stratonovich convention. For the numerical simulations they use a WENO3 scheme [46]. The convergence of this scheme is verified by statistical tests. The invariance of the Stratonovich convention is tested by setting $\alpha(x) = e^x - 1$ which conserves the zero level sets.

For the application of this model to computer vision which is called *stochastic active contours*, we first have to introduce the simulated annealing algorithm. This algorithm is based on the work of Metropolis et al. [66] and was first mentioned by Kirkpatrick et al. [51] as an application of statistical physics to optimization problems. The algorithm applied to computer vision works in the following way.

Algorithm 3.12. Given some computer vision problem in a variational framework, where we have to find the region Γ described by a function u that minimizes an energy $E(u)$, the following simulated annealing decision scheme is used:

1. Start from some initial guess u_0.
2. Compute u_{n+1} from u_n using some dynamics, e.g. $du = |Du| \sum_{i=1}^{m} \phi_i(x) \circ dW_i(t)$.
3. Compute the energy $E(u_{n+1})$.
4. Accept u_{n+1},
   - if $E(u_{n+1}) < E(u_n)$,
   - otherwise, accept u_{n+1} with probability $\exp(-(E(u_{n+1}) - E(u_n))/T(n))$.
5. Loop back to step 2, until some stopping condition is fulfilled, e.g. the energy does not change for a by the user specified time.

Here $T(n)$ is a time-dependent function that plays the same role as a decreasing temperature. Its choice is not obvious. A classical choice is $T(n) = T_0/\sqrt{n}$, where T_0 is a constant.

The classical way to solve the minimization problem of Algorithm 3.12 is to use a gradient descent method. Therefore the Euler–Lagrange equation is computed, leading to some evolution $\partial u/\partial t = \beta_c |Du|$ in the level set framework. The authors use the classical motion as heuristics that drives the evolution faster towards a minimum, and they replace the dynamics of step 2 by

$$du = \beta_c |Du|\ dt + |Du| \sum_{i=1}^{m} \phi_i(x) \circ dW_i(t).$$

The convergence of this algorithm towards a global minimum is still an open problem. The authors claim that practical simulations indicate that Algorithm 3.12 is more likely to overcome

Figure 3.3.: Segmentation of two regions modeled by two unknown Gaussian distributions. From left to right: the initial curve, the final time step of the classical method, again stuck in a local minimum and the final step of the method presented in [47].

local minima than the classical approach. For their simulations, the energy functionals presented in Section 3.3.2 are used although one could also use the functionals from Section 3.3.1 or those of Casellas et al. [9]. One of their segmentation results is the zebra shown in Figure 3.3 which was also successfully segmented by the authors of [73] with supervised texture segmentation.

3.4.2. Segmentation Using Different Noise Terms

The noise used in the previous section has the disadvantage that there is a fixed finite set of points that is perturbed and all other points are connected to this set by a deterministic function. Therefore it might be an interesting modification to use Q-Wiener processes and the corresponding numerical simulations presented in Section 2.2.2. The existence of solutions can be found in [16]. The question, when a discretization of a stochastic differential equation of the form

$$du(t,x) = F(D^2u, Du, x, t)\,dt + |Du(t,x)|\,dW(t,x)$$

is consistent and numerically stable, is still open. A first idea how these properties can be checked is given in Section 2.3 in the case of the heat equation with additive and multiplicative noise. Concerning the question whether to use the Itô or the Stratonovich convention, simulations on both were done which are presented in the next section.

3.5. Simulations

This section summarizes the simulation of SPDEs for segmentation with different numerical schemes and compares them. To show the effect of different covariances, a heat equation on an image and motion by mean curvature will be used. Moreover the differential equations from Section 3.3 and the stochastic ones of the previous section will be implemented and tested. The implementation was done in C++ and the graphical user interface with Qt. The random numbers were generated using Marsaglia's 'mother of all random number generators", which is a multiply-with-carry generator and which was published on the web [65]. These equally distributed random numbers were transformed by a Box-Muller algorithm to standard normally distributed random numbers. Most of the algorithms used in the implementation,

Figure 3.4.: Two hibiscus blossoms of size 256×256 pixels.

e.g. the solution of linear and nonlinear algebraic equations, are due to [76]. The discretization of differential operators was always done by finite differences. In our case, it seems reasonable to choose finite grids and this simple type of discretization because images are represented with rectangular regions subdivided into pixels. All work was done with grayscale pictures. If the pictures would have been segmented as color images, the segmentation results would probably have been better in some cases.

3.5.1. Noise Effects

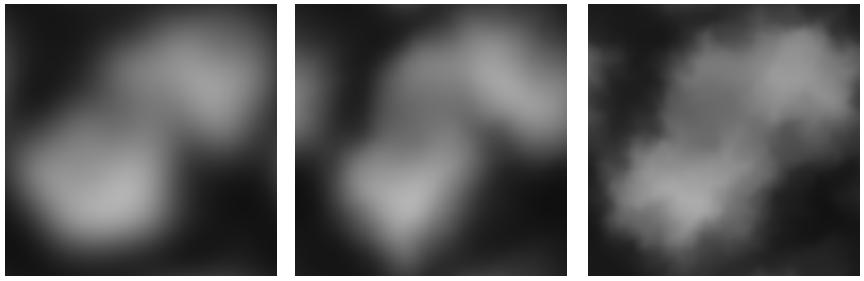
To give an idea what type of noise might be a good choice for different simulations and how to choose a proper covariance function, a heat equation and motion by mean curvature will be used in the following. These SPDEs give a good visualization of the effects due to the noise. Therefore we discretize

$$du(t) = \tfrac{1}{2}\Delta u(t) + |Du| \; dW(t)$$

and choose a grayscale picture of two hibiscus blossoms, see Figure 3.4. The gray value is used as temperature and dark parts represent cold regions while the white ones are the hot parts of the image. Figure 3.5 shows the result after 50 simulation steps with $\triangle t = 1$ and $\triangle x = 0.5$. The used space covariance for the simulations with noise is given by the Fourier transform of

$$f(p_1, p_2) = (m^{2kn} + (p_1^{2k} + p_2^{2k})^n)^{-l},$$

where $n = 2$, $k = 1$, and $l = 1$ were chosen. The value of m was varied in order to get different strengths of correlations. Details on the effect of this function on the resulting random fields can be found in Section 1.6. In the following, the differences of the images in Figure 3.5 will be described. The first image shows the heat equation without any disturbance and therefore the contours become diffuse but the shape stays about the same. From Pictures (b) to (f) the exponential decay of the covariance gets stronger and this entails that perturbations appear more and more localized. While the whole shape is corrupted in Picture (b), the shape of the blossoms is more and more similar to the unperturbed image and the size of the perturbation of the borders decreases. If we make use of these observations in segmentation, this will mean that



(a) no noise (b) $m = 1$ (c) $m = 5$

(d) $m = 10$ (e) $m = 20$ (f) $m = 50$

Figure 3.5.: Heat equation with gradient coupled noise and different covariances after 50 time steps using Figure 3.4, $n = 2, k = 1, l = 1$.

m should be chosen large, when the curve has to be perturbed only very locally. Otherwise, if the object to be found is not very delicate, m has to be chosen small enough.

So far, everything was fixed on the covariance in space and it has just been looked at processes that are white in time. One example of a heat equation perturbed with gradient coupled noise in space and an Ornstein-Uhlenbeck process in time is shown in Figure 3.6. This process was introduced in Example 2.33. In this simulation the covariance is the same as in Figure 3.5(c) and all simulations were done with the same random numbers. We observe from Example 2.33 that regions that were strongly perturbed up to time t will also be perturbed in the following time step where the strength depends on γ. The simulation results show that a small value $\gamma = 0.2$ smoothes the noise while a relatively large value $\gamma = 0.9$ frays out the contours. Depending on the effect that should be achieved, an Ornstein-Uhlenbeck process considered in theory might be a good choice. This is especially the case if in a segmentation problem the curve is pushed into a corner that would be smoothed out by the corresponding deterministic equation. Then in the next simulation step the curve will be perturbed at the same place and therefore hopefully forced deeper into the corner. In the simulations done for this thesis however, it turns out that this approach does not lead to remarkable better results.

Similar observations to those with the stochastic heat equation are made when implementing motion by mean curvature. Figure 3.7 shows in (a) the initial state where the zero level set is the border between red and blue. The unperturbed PDE shrinks the square to a decreasing circle as shown in (b) and finally vanishes. If the PDE is perturbed with gradient coupled noise, the square also shrinks. In (c) the covariance is very rapidly decreasing to zero and therefore

(a) no noise, (b) $\gamma = 0$, (c) $\gamma = 0.2$,

(d) $\gamma = 0.5$, (e) $\gamma = 0.7$, (f) $\gamma = 0.9$,

Figure 3.6.: Heat equation with gradient coupled noise, fixed covariance in space (see Figure 3.5(c)), and Ornstein-Uhlenbeck process in time with different coefficients γ after 50 time steps using Figure 3.4.

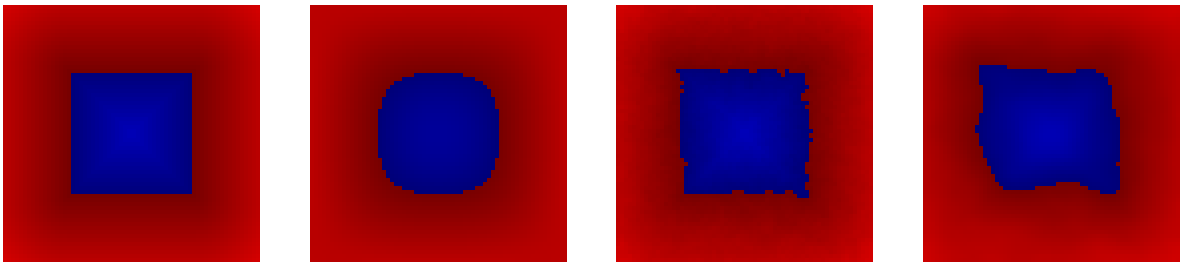
the general motion of the shape is about the same as without noise but single pixels on the border are perturbed, while in (d) the noise is smoother and therefore the whole shape changes and one observes that the curve is moving all over the image and even breaking into several parts. These impressions should help the reader to find the right noise with given covariance for his own applications and simulations. Next we move on to segmentation with different types of noise, especially differently coupled to the PDE.

3.5.2. Implicit Schemes

One possible implementation of SPDEs with Stratonovich convention is to use implicit schemes, i.e.

$$\int_t^{t+\triangle t} |\nabla u(s,x)| \circ dW(s,x) \approx \tfrac{1}{2} \left(|\nabla u(t+\triangle t,x)| + |\nabla u(t,x)| \right) \left(W(t+\triangle t,x) - W(t,x) \right).$$

This leads to nonlinear algebraic equations that can only be solved numerically. The existing algorithms need a good guess of the solution in order to converge. Possible choices for this guess are the solution of the SPDE of the previous step or the solution that would come from an explicit scheme. In the simulations, the problem occurred that these guesses were not good enough to take a large time step size $\triangle t$. If a time step size that was not numerically stable in the corresponding explicit scheme, was taken, the algorithm for solving the system of nonlinear equations would not converge. Therefore these equations do not have the advantage that a



(a) initial condition, (b) without noise, (c) $f(p) = (1 + |p|^2)^{-1}$, (d) $f(p) = (1 + |p|^2)^{-2}$,

Figure 3.7.: Simulation of motion by mean curvature perturbed with different types of gradient coupled noise.

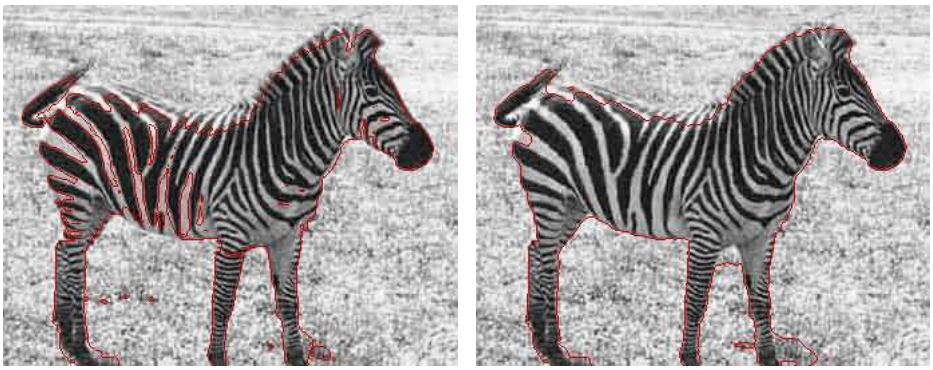
larger step size can be chosen than for the explicit schemes but the simulation is a lot slower. The segmentation results are similar to those with explicit scheme. It seems that the zero level sets are a bit smoother in comparison to the explicit solutions and therefore the implicit schemes avoid the formation of cusps.

3.5.3. Explicit Schemes

The fastest and easiest way to implement SPDEs is to use explicit schemes and finite differences. This kind of implementation also leads to fast simulations because no algebraic equation systems have to be solved. Numerical stability is achieved for $\triangle x = 0.5$ and $\triangle t \leq 0.001$, if the energy functional by Chan and Vese (see Section 3.3.1) is used. In the following we have to distinguish between five different equations: without noise, with additive noise, with multiplicative noise, with gradient coupled noise, and with curvature coupled noise, i.e. the following SPDEs were implemented:

$$du(t,x) = F(D^2u, Du, x, t)\,dt + \begin{cases} 0, \\ \alpha\, dW(t,x), \\ \alpha\, u(t,x)\, dW(t,x), \\ \alpha\, |Du(t,x)|\, dW(t,x), \\ \alpha \left(1 + \operatorname{div}\left(\dfrac{\nabla u(t,x)}{|\nabla u(t,x)|}\right)\right) dW(t,x), \end{cases}$$

where $\alpha \in \mathbb{R}_+$ is a constant to regulate the strength of the noise. For the deterministic equation, a full time simulation was used, i.e. the PDE was simulated up to time T, while the SPDEs where embedded into Algorithm 3.12. In most cases the results do not differ much. One example where the gradient coupled noise led to a much better result than the deterministic equation can be found in Figure 3.8. Before showing other segmentation results, the meaning of the different types of noise is discussed. Additive and multiplicative noise hardly change anything. Additive noise disturbs the whole function everywhere and the simulation results give the impression that if perturbed, the curve is destroyed at the wrong places. Using



(a) without noise (b) with gradient coupled noise

Figure 3.8.: Segmentation of a zebra.

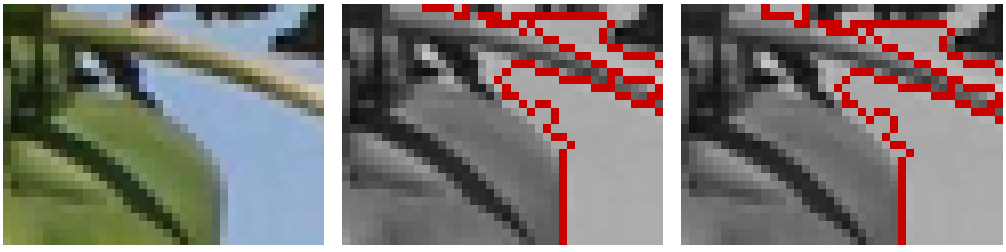
multiplicative noise means that the function is not perturbed at the zero level set but at all other places. This means the further a pixel is away from the curve, the stronger it is perturbed. As the goal is to get better segmentation results and therefore a different shape of the curve than without noise, this approach does not seem to be a good choice. If the noise is weighted with the size of the gradient, one observes the best results. It is also noticeable that the signed distance function is not too much perturbed even without reinitialization. This can be explained in the following way. The noise is weighted with the size of the gradient which is equal to one for the signed distance function. If the function becomes much steeper, i.e. the gradient is much larger than one, these parts will undergo stronger perturbations and therefore they will probably be pushed back into the direction of the signed distance function. The gradient coupled noise seems to be the best choice to get small details. The idea of the curvature coupled noise was to find small details because parts with high curvature will be perturbed stronger than straight lines. In simulations it turned out that this is better achieved by gradient coupled noise.

A typical segmentation result with a fixed deterministic equation of Chan and Vese type and different noise is shown in Figure 3.9. Looking at the complete segmentation, there are hardly any differences but in Figure 3.10 we compare the details. These show that the gradient coupled noise is the best to find the edges of the leaves. Therefore it overcomes the problem that the approach by Chan and Vese [10] smoothes out details if the mean curvature term dominates the PDE. Finally, Figure 3.11 shows that the segmentation of the zebra shown in Figure 3.3 also gives good results, if the deterministic PDE presented in Section 3.3.2 is used, but with stochastics the segmentation result is even better. If the noise developed in this thesis is applied to segmentation, one observes that the results are comparable to other segmentation methods with and without stochastics. An advantage of the infinite-dimensional noise approach presented in Section 1.4 and used here is that the random field generation is fast. Therefore the segmentation of the images did not take longer than five minutes on a standard PC. Still, the simulation without noise is the fastest of all versions. We also remark that the speed of the noise generation depends highly on the size of the image. This is due to the fact that Algorithm 1.18 uses FFT implemented via *fftw3* [25] which is the fastest for powers of two.

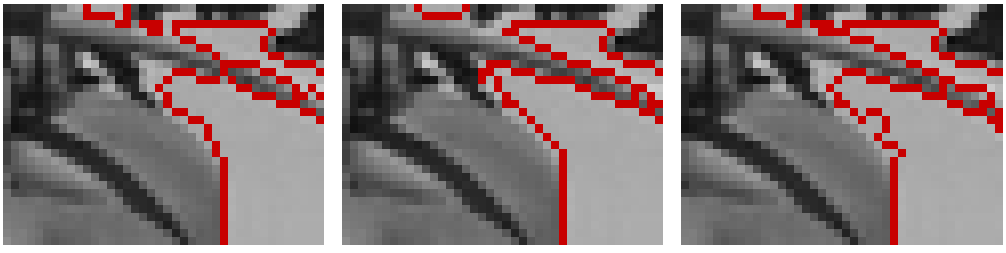
(a) original image, (b) without noise, (c) additive noise,

(d) multiplicative noise, (e) gradient coupled noise, (f) curvature coupled noise,

Figure 3.9.: Segmentation of a papaya tree with different types of noise.

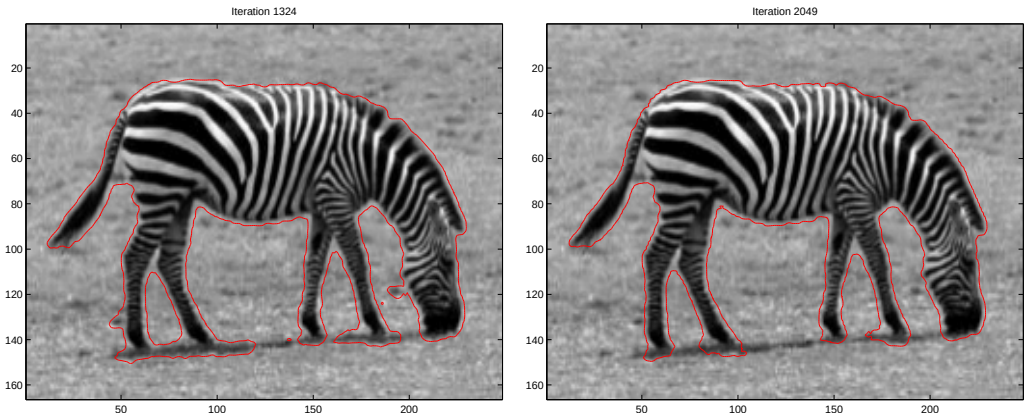


(a) original image, (b) without noise, (c) additive noise,



(d) multiplicative noise, (e) gradient coupled noise, (f) curvature coupled noise,

Figure 3.10.: Detail of Figure 3.9 showing the differences between various noise types in detail but with the same deterministic equation.



(a) without noise, (b) gradient coupled noise,

Figure 3.11.: Segmentation of Figure 3.3 using Gaussian distributions.

Appendix A.

Basic Properties of Strongly Continuous Semigroups

This appendix summerizes some important properties of strongly continuous semigroups. For more details the reader is referred to [43, 50, 74, 98]. In the following let E be a Banach space with norm $\|\cdot\|$.

Definition A.1. A one parameter family $S(t)$, $0 \leq t < +\infty$, of bounded linear operators from E into E is a *semigroup of bounded linear operators* on E if

1. $S(0) = \mathbb{1}$, where $\mathbb{1}$ is the identity operator on E,
2. $S(t+s) = S(t)S(s)$ for every $t, s \geq 0$.

A semigroup of bounded linear operators $S(\cdot)$ is *uniformly continuous* if

$$\lim_{t\downarrow} \| S(t) - \mathbb{1} \| = 0.$$

The linear operator A defined by

$$D(A) = \left\{ x \in E,\ \lim_{t\downarrow 0} \frac{S(t)x - x}{t} \text{ exists} \right\}$$

and

$$Ax = \lim_{t\downarrow 0} \frac{S(t)x - x}{t} \qquad \text{for } x \in D(A)$$

is the *infinitesimal generator* of the semigroup $S(\cdot)$, $D(A)$ is the *domain* of A.

Definition A.2. A semigroup $S(t)$, $0 \leq t < +\infty$, of bounded linear operators on E is a *strongly continuous semigroup* of bounded linear operators if

$$\lim_{t\downarrow 0} S(t)x = x \qquad \text{for every } x \in E.$$

A strongly continuous semigroup of bounded linear operators on E is called a *semigroup of class* C_0 or simply a C_0 *semigroup.*

Corollary A.3. *If $S(\cdot)$ is a C_0 semigroup then for every $x \in E$, $t \mapsto S(t)x$ is a continuous map from the nonnegative real line into E.*

Theorem A.4. *Let $S(\cdot)$ be a C_0 semigroup and let A be its infinitesimal generator.*

1. *For $x \in E$,*

$$\lim_{h\to 0} \frac{1}{h} \int_t^{t+h} S(s)x\, ds = S(t)x.$$

2. *For $x \in E$,*

$$\int_0^t S(s)x\,ds \in D(A) \quad \textit{and} \quad A\left(\int_0^t S(s)x\,ds\right) = S(t)x - x.$$

3. *For $x \in D(A)$, $S(t)x \in D(A)$ and*

$$\frac{d}{dt}S(t)x = AS(t)x = S(t)Ax.$$

4. *For $x \in D(A)$,*

$$S(t)x - S(s)x = \int_s^t S(\tau)Ax\,d\tau = \int_s^t AS(\tau)x\,d\tau.$$

Definition A.5. A linear operator $A : D(A) \subset E \to E$ is *closed* if its graph

$$\mathcal{G}_A := \{(x, y) \in E \times E,\ x \in D(A),\ Ax = y\}$$

is closed in $E \times E$.

If A is closed, the domain $D(A)$ can be endowed with the *graph norm*

$$\| x \|_{D(A)} := \| x \| + \| Ax \|$$

which can be shown to be indeed a norm.

Corollary A.6. *If A is the infinitesimal generator of a C_0 semigroup $S(\cdot)$ then $D(A)$, the domain of A, is dense in E and A is a closed linear operator.*

Theorem A.7. *Let A be the infinitesimal generator of the C_0 semigroup $S(\cdot)$. If $D(A^n)$ is the domain of A^n, then $\cap_{n=1}^{\infty} D(A^n)$ is dense in E.*

Definition A.8. A complex number λ belongs to the *resolvent set* $\rho(A)$ of A if $\lambda\mathbb{1} - A$ is one-to-one and onto. If $\lambda \in \rho(A)$, we set

$$R(\lambda, A) := (\lambda\mathbb{1} - A)^{-1},$$

and call $R(\lambda.A)$ the *resolvent operator* of A. The complement of $\rho(A)$ in $\mathbb{C}$ is called the *spectrum* of A.

Theorem A.9 (Hille–Yosida)**.** *Let $A : D(A) \subset E \to E$ be a linear closed operator on E. Then the following statements are equivalent:*

1. *A is the infinitesimal generator of a C_0 semigroup $S(\cdot)$ such that*

$$\| S(t) \| \leq Me^{\omega t}, \qquad \textit{for all } t \geq 0.$$

2. *$D(A)$ is dense in E, the resolvent set $\rho(A)$ contains the interval $(\omega, +\infty)$ and the following estimates hold*

$$\left\| R^k(\lambda, A) \right\| \leq \frac{M}{(\lambda - \omega)^k}, \qquad \textit{for } k \in \mathbb{N}.$$

Moreover if either 1. or 2. holds then

$$R(\lambda, A)x = \int_0^\infty e^{-\lambda t} S(t)x \, dt, \qquad x \in E,\ \lambda > \omega.$$

Finally

$$S(t)x = \lim_{n \to +\infty} e^{tA_n} x, \qquad \text{for all } x \in E,$$

where $A_n := nAR(n, A)$ *and the following estimate holds*

$$\left\| e^{tA_n} \right\| \leq M e^{\frac{\omega n t}{n - \omega}}, \qquad \text{for all } t \geq 0,\ n > \omega.$$

The operators $A_n = AJ_n$ where $J_n := nR(n, A)$, $n > \omega$, are called the *Yosida approximations* of A. The following properties of Yosida approximations are frequently used.

Proposition A.10. *Let* $A : D(A) \subset E \to E$ *be the infinitesimal generator of a* C_0 *semigroup. Then*

$$\lim_{n \to +\infty} nR(n, A)x = x, \qquad \text{for all } x \in E,$$
$$\lim_{n \to +\infty} A_n x = Ax, \qquad \text{for all } x \in D(A).$$

Appendix B.

Stability Using Fundamental Solutions and Fourier Methods

In the following we will derive Proposition 2.68 in the language of fundamental solutions and Fourier methods in an informal way instead of semigroup theory. We assume that we are doing simulations on $\mathbb{R}$ and that we do not need boundary conditions. Moreover we assume that the stochastic integral can be approximated without error. For the calculations we will first shortly introduce the necessary basics on fundamental solutions. These can be found in [24].

Let

$$\left(\frac{\partial}{\partial t} - \frac{1}{2}\triangle\right) u(t,x) = 0$$

be the homogeneous heat equation on $\mathbb{R}$ with the corresponding Cauchy problem $u(0,x) = \varphi(x)$ where φ is a continuous function on $\mathbb{R}$ with $\int_{\mathbb{R}} \varphi(x)\,dx < +\infty$.

This PDE has the fundamental solution

$$K(t,x) = \frac{1}{\sqrt{2\pi t}}\, e^{-\frac{x^2}{2t}}.$$

The solution of the heat equation with initial condition $u(0,x) = \varphi(x)$ is given by

$$u(t,x) = (K(t,\cdot) * \varphi)(x) = \int_{\mathbb{R}} K(t,x-y)\varphi(y)\,dy.$$

If we extend the heat equation with an inhomogeneity, i.e.

$$\left(\frac{\partial}{\partial t} - \frac{1}{2}\triangle\right) u(t,x) = f(t,x)$$

where f is continuous and locally Hölder continuous, the solution is given by

$$u(t,x) = (K * f)(t,x) \quad = \quad \int_0^t \int_{\mathbb{R}} K(t-s,x-y) f(s,y)\,dy\,ds.$$

For the corresponding Cauchy problem with $u(0,x) = \varphi(x)$, the solution is given by

$$u(t,x) = (K(t,\cdot) * \varphi)(x) + (K * f)(t,x).$$

Proposition B.1 (Heat equation with additive noise)**.** *Let*

$$du(t,x) = \frac{1}{2}\triangle u(t,x)\,dt + dW(t,x),$$

be a stochastic partial differential equation where $W(t,x)$ is a stochastic process with given space covariance

$$C(x,y) = \int_{\mathbb{R}} e^{-2\pi i(x-y)p} f(p)\,dp,$$

where $f(p)$ is symmetric under rotation, such that $C(x,y)$ is exponentially decreasing for $|x-y|$ large and $\int_{\mathbb{R}} p^k f(p)\,dp < +\infty, k \in \{0,\dots,6\}$, and the covariance in time is given by a process $dX_t = b(X_t)\,dt + \sigma(X_t)\,dB_t$ where $\sigma(X_t)$ is bounded, i.e. $0 < \varepsilon < \sigma(X_t) < b < +\infty$ and B_t denotes Brownian motion. Possible examples for $W(t,x)$ are Brownian motion and Ornstein-Uhlenbeck processes. Assume that K_{tt} and K_{xxxx} are bounded where K denotes the fundamental solution. Then the approximation method

$$U(t+h,x) = U(t,x) + \frac{h}{2(\triangle x)^2}(U(t,x+k) - 2U(t,x) + U(t,x-k)) + \eta(t,x)$$

is consistent and stable for $h \le (\triangle x)^2$, $t \in (0,T]$ where $\eta(t,x) = W(t+h,x) - W(t,x)$

Proof. First, we have to calculate the truncation error to prove consistency. Therefore we set $\alpha := h/(\triangle x)^2$ in order to make notations easier.

$$\begin{aligned}
&T_{t+h}(x)\\
&= u(t+h,x) - (1 - \frac{h}{(\triangle x)^2})u(t,x) - \frac{h}{2(\triangle x)^2}(u(t,x+\triangle x) + u(t,x-\triangle x)) - \eta(t,x)\\
&= \int_{\mathbb{R}} K(t+h,x-y)\varphi(y)\,dy + \int_{\mathbb{R}}\int_0^{t+h} K(t+h-s,x-y)\,dW(s,y)\,dy - \int_t^{t+h} dW(s,x)\\
&\quad - (1-\alpha)(\int_{\mathbb{R}} K(t,x-y)\varphi(y)\,dy + \int_{\mathbb{R}}\int_0^t K(t-s,x-y)\,dW(s,y)\,dy)\\
&\quad - \frac{\alpha}{2}(\int_{\mathbb{R}} K(t,x+\triangle x-y)\varphi(y)\,dy + \int_{\mathbb{R}}\int_0^t K(t-s,x+\triangle x-y)\,dW(s,y)\,dy)\\
&\quad - \frac{\alpha}{2}(\int_{\mathbb{R}} K(t,x-\triangle x-y)\varphi(y)\,dy + \int_{\mathbb{R}}\int_0^t K(t-s,x-\triangle x-y)\,dW(s,y)\,dy)\\
&= \int_{\mathbb{R}} (K(t+h,x-y) - (1-\alpha)K(t,x-y) \qquad \text{(B.1)}\\
&\qquad - \frac{\alpha}{2}(K(t,x+\triangle x-y) + K(t,x-\triangle x-y)))\varphi(y)\,dy\\
&\quad + \int_{\mathbb{R}}\int_t^{t+h} K(t+h-s,x-y)\,dW(s,y)\,dy - \int_t^{t+h} dW(s,x) \qquad \text{(B.2)}\\
&\quad + \int_{\mathbb{R}}\int_0^t (K(t+h-s,x-y) - (1-\alpha)K(t-s,x-y) \qquad \text{(B.3)}\\
&\qquad - \frac{\alpha}{2}(K(t-s,x+\triangle x-y) + K(t-s,x-\triangle x-y)))\,dW(s,y)\,dy
\end{aligned}$$

In the following we will calculate the L^2 norm of the three integrals above separately and show that

$$\|T(t,x)\|_2 \le O(h^{3/2})$$

and therefore that the truncation error converges to zero in L^2 as h tends to zero with at least rate $O(h^{3/2})$.

The estimates on the first integral are the following: We assume $t > 0$, because otherwise the integral is equal to $\varphi(x)$. We also notice that for $t > 0$ the derivatives in space and time exist and that the integral over these is finite if φ satisfies $\int_{\mathbb{R}} \varphi(y)\,dy < +\infty$.

$$\begin{aligned}
&\int_{\mathbb{R}} (K(t+h, x-y) - (1-\alpha)K(t, x-y) - \frac{\alpha}{2}(K(t, x+\triangle x - y) + K(t, x-\triangle x - y)))\varphi(y)\,dy \\
&\quad = \int_{\mathbb{R}} (K(t, x-y) + hK_t(t, x-y) + \frac{h^2}{2}K_{tt}(\tilde{t}, x-y) - (1-\alpha)K(t, x-y)) \\
&\qquad - \frac{\alpha}{2}(K(t, x-y) + \triangle x K_x(t, x-y) + \frac{(\triangle x)^2}{2}K_{xx}(t, x-y) \\
&\qquad\quad + \frac{(\triangle x)^3}{6}K_{xxx}(t, x-y) + \frac{(\triangle x)^4}{24}K_{xxxx}(t, \tilde{x} - y)) \\
&\qquad - \frac{\alpha}{2}(K(t, x-y) - \triangle x K_x(t, x-y) + \frac{(\triangle x)^2}{2}K_{xx}(t, x-y) \\
&\qquad\quad - \frac{(\triangle x)^3}{6}K_{xxx}(t, x-y) + \frac{(\triangle x)^4}{24}K_{xxxx}(t, \tilde{\tilde{x}} - y)))\varphi(y)\,dy \\
&\quad = \int_{\mathbb{R}} h(K_t(t, x-y) - \frac{1}{2}K_{xx}(t, x-y))\varphi(y)\,dy + \frac{h^2}{2}\int_{\mathbb{R}} K_{tt}(\tilde{t}, x-y)\varphi(y)dy \\
&\qquad - \frac{h(\triangle x)^2}{48}\int_{\mathbb{R}} (K_{xxxx}(t, \tilde{x} - y) + K_{xxxx}(t, \tilde{\tilde{x}} - y))\varphi(y)\,dy \\
&\quad = \frac{h^2}{2}\int_{\mathbb{R}} K_{tt}(\tilde{t}, x-y)\varphi(y)dy - \frac{h(\triangle x)^2}{48}\int_{\mathbb{R}} (K_{xxxx}(t, \tilde{x} - y) + K_{xxxx}(t, \tilde{\tilde{x}} - y))\varphi(y)\,dy.
\end{aligned}$$

Therefore as K_{tt} and K_{xxxx} are bounded by assumption and $\int_{\mathbb{R}} \varphi(y)\,dy < +\infty$, we get

$$\begin{aligned}
\Big\|\frac{h^2}{2}\int_{\mathbb{R}} K_{tt}(\tilde{t}, x-y)\varphi(y)dy - \frac{h(\triangle x)^2}{48}\int_{\mathbb{R}} (K_{xxxx}(t, \tilde{x} - y) + K_{xxxx}(t, \tilde{\tilde{x}} - y))\varphi(y)\,dy\Big\|_2 \\
\leq O(h^2) + O(h(\triangle x)^2).
\end{aligned}$$

If $h \sim (\triangle x)^2$, we have that the first integral is bounded by $O(h^2)$.

For the estimates on the second integral we have

$$\begin{aligned}
&\Big\|\int_{\mathbb{R}}\int_t^{t+h} K(t+h-s, x-y)\,dW(s,y)\,dy - \int_t^{t+h} dW(s,x)\Big\|_2^2 \\
&\quad = \mathbb{E}((\int_{\mathbb{R}}\int_t^{t+h} K(t+h-s, x-y)\,dW(s,y)\,dy)^2) + \mathbb{E}((\int_t^{t+h} dW(s,x))^2) \\
&\qquad - 2\cdot\mathbb{E}(\int_{\mathbb{R}}\int_t^{t+h} K(t+h-s, x-y)\,dW(s,y)\,dy \int_t^{t+h} dW(s,x)) \\
&\quad = \int_t^{t+h}\int_{\mathbb{R}\times\mathbb{R}} K(t+h-s, x-y)C(y,z)K(t+h-s, x-z)\,dy\,dz\,ds \\
&\qquad - 2\int_t^{t+h}\int_{\mathbb{R}} K(t+h-s, x-y)C(x,y)\,dy\,ds + \int_t^{t+h} C(x,x)\,ds
\end{aligned}$$

$$
\begin{aligned}
&= \int_t^{t+h} \mathcal{F}(\mathcal{F}^{-1}(K(t+h,x)) \cdot \mathcal{F}^{-1}(C(x,x)) \cdot \mathcal{F}^{-1}(K(t+h,x)))\, ds \\
&\quad - 2\int_t^{t+h} \mathcal{F}(\mathcal{F}^{-1}(K(t+h,x)) \cdot \mathcal{F}^{-1}(C(x,x)))\, ds + hC(x,x) \\
&= \int_t^{t+h} \int_{\mathbb{R}} e^{2\pi i(x-x)q} e^{-2(t+h-s)\pi^2 q^2} f(q) e^{-2(t+h-s)\pi^2 q^2}\, dq\, ds \\
&\quad - 2\int_t^{t+h} \int_{\mathbb{R}} e^{2\pi i(x-x)q} e^{-2(t+h-s)\pi^2 q^2} f(q)\, dq\, ds + \int_{\mathbb{R}} hf(q)\, dq \\
&= \int_{\mathbb{R}} (e^{-4h\pi^2 q^2} \frac{(e^{4h\pi^2 q^2}-1)}{4\pi^2 q^2} - e^{-2\pi^2 q^2} \frac{2(e^{2\pi^2 q^2}-1)}{2\pi^2 q^2} - h) f(q)\, dq \\
&\leq \int_{\mathbb{R}} (\frac{1}{4\pi^2 q^2}(1-(1-h4\pi^2 q^2 + \frac{h^2}{2}(4\pi^2 q^2)^2 - \frac{h^3}{6}(4\pi^2 q^2)^3)) \\
&\quad - \frac{2}{2\pi^2 q^2}(1-(1-h2\pi^2 q^2 + \frac{h^2}{2}(2\pi^2 q^2)^2)) + h) f(q)\, dq \\
&= h^3 \frac{4}{3}\pi^4 \int_{\mathbb{R}} q^4 f(q)\, dq \quad = \quad O(h^3),
\end{aligned}
$$

using in the second step the Itô isometry. Therefore it follows that the second integral is bounded by $O(h^{3/2})$.

Finally the norm in L^2 of the third integral can be estimated in the following way.

$$
\begin{aligned}
&\Big\| \int_{\mathbb{R}} \int_0^t (K(t+h-s,x-y) - (1-\alpha)K(t-s,x-y) \\
&\qquad -\frac{\alpha}{2}(K(t-s,x+k-y) + K(t-s,x-k-y)))\, dW(s,y)\, dy \Big\|_2^2 \\
&= \int_0^t \int_{\mathbb{R}\times\mathbb{R}} (K(t+h-s,x-y) - (1-\alpha)K(t-s,x-y) \\
&\qquad -\frac{\alpha}{2}(K(t-s,x+k-y) + K(t-s,x-k-y))) \\
&\qquad C(y,z)(K(t+h-s,x-z) - (1-\alpha)K(t-s,x-z) \\
&\qquad -\frac{\alpha}{2}(K(t-s,x+k-z) + K(t-s,x-k-z)))\, dy\, dz\, ds \\
&= \int_0^t \mathcal{F}(\mathcal{F}^{-1}(K(t+h-s,x) - (1-\alpha)K(t-s,x) - \frac{\alpha}{2}(K(t-s,x+k) + K(t-s,x-k))) \\
&\qquad \mathcal{F}^{-1}(C(x,x))\mathcal{F}^{-1}(K(t+h-s,x) - (1-\alpha)K(t-s,x) \\
&\qquad -\frac{\alpha}{2}(K(t-s,x+k) + K(t-s,x-k)))\, ds \\
&= \int_0^t \int_{\mathbb{R}} e^{2\pi i(x-x)q} \Big(e^{-2(t+h-s)\pi^2 q^2} - (1-\alpha)e^{-2(t-s)\pi^2 q^2} \\
&\qquad -\frac{\alpha}{2}\left(e^{2\pi ikq} e^{-2(t-s)\pi^2 q^2} + e^{-2\pi ikq} e^{-2(t-s)\pi^2 q^2} \right) \Big)^2 f(q)\, dq\, ds \\
&= \int_{\mathbb{R}} \left(e^{-2h\pi^2 q^2} - 1 + \alpha(1-\cos(2\pi kq)) \right)^2 \int_0^t e^{-4(t-s)\pi^2 q^2}\, ds\, f(q)\, dq
\end{aligned}
$$

$$\leq \int_{\mathbb{R}} \frac{1-e^{-4t\pi^2q^2}}{4\pi^2q^2}\Big(1-h2\pi^2q^2+\frac{h^2}{2}(2\pi^2q^2)^2-1+\alpha(1-1+\frac{(\triangle x)^2}{2}4\pi^2q^2)\Big)^2 f(q)\,dq$$
$$\leq \int_{\mathbb{R}} 1\cdot h^4\pi^6q^6 f(q)\,dq = O(h^4)$$

The last equation follows because $\int_{\mathbb{R}} q^6 f(q)\,dq < +\infty$. Therefore we conclude that the third integral is bounded by $O(h^2)$. So in conclusion we have proved that $\|T(t,x)\|_2 \leq O(h^{3/2})$ and that the approximation method is consistent.

We finish the proof by calculating the error and showing that the method is stable. The equation

$$e(t+h,x)-(1-\alpha)e(t,x)-\frac{\alpha}{2}(e(t,x+k)+e(t,x-k)) = -T(t,x)$$

implies that

$$\|e(t+h,x)\|_2 \leq |1-\alpha|\,\|e(t,x)\|_2+\frac{\alpha}{2}(\|e(t,x+k)\|_2+\|e(t,x-k)\|_2)+\|T(t,x)\|_2\,.$$

Let the time steps in the simulation be given by $t_n = h\cdot n, 0\leq t_n \leq T$ and we set

$$\begin{aligned} E(t_n) &= \max_x \|e(t_n,x)\|_2\,,\\ T(t_n) &= \max_x \|T(t_n,x)\|_2\,,\\ T_{\max} &= \max_{t_n} T(t_n). \end{aligned}$$

$E(t_n)$, $T(t_n)$, and $T_{\max}$ exist because an error is just defined on the finite grid of our simulation in space and time. Then for $h \leq (\triangle x)^2$ we get

$$\begin{aligned} E(t_{n+1}) &\leq (|1-\alpha|+\alpha)\,E(t_n)+T(t_n)\\ &\leq E(t_n)+T(t_n)\\ &\leq E(t_0)+n\cdot T_{\max}\\ &\leq 0+nh\,O(\sqrt{h})\\ &= t_n O(\sqrt{h}). \end{aligned}$$

Therefore $E(t_{n+1})$ tends to zero if $T_{\max}$ goes to zero and the used approximation method is stable for $h \leq (\triangle x)^2$.

□

Appendix C.

Curvature of a Curve

There are three different ways how to represent a curve in the plane: in the parametric form, explicitely, or implicitely. In the following, we will see the calculations and transformations of the curvature in the different representations. Note that we use curves in the plane and therefore that there exists just one curvature value and the mean curvature is equal to the principal curvature and the curvature.

Definition C.1. Let $r : [a, b) \to \mathbb{R}^2$ be the embedding of a curve given by $r(s) = (x(s), y(s))$. The curve $r(s)$ is called *parameterized by arclength* if the tangent vector $\mathbf{t}(s) = r'(s)$ satisfies $|\mathbf{t}(s)| = 1$ for all $s \in [a, b)$.

Moreover let $\mathbf{n}(s)$ be the normal vector of the curve. The direction of $\mathbf{n}(s)$ is given by

$$\mathbf{n}(s) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \mathbf{t}(s)^{\mathrm{t}}.$$

The orientation is made clear in Figure C.1. Now we are able to define the curvature of a curve in the plane.

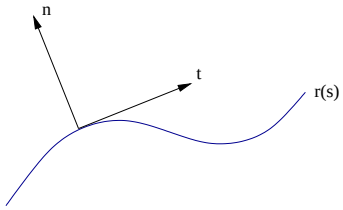


Figure C.1.: A given curve $r(s)$ with tangent vector t and normal n.

Definition C.2. Let r be an embedded curve as defined above and $\mathbf{t}(s)$ the tangent vector at the point $r(s)$. Furthermore let $\varphi(s)$ be the angle between the tangent vector $\mathbf{t}(s)$ and the positive x-axis. Then the *curvature* of the curve is given by

$$k(s) := \frac{d\,\varphi(s)}{d\,s}.$$

The *vector of curvature* is defined by

$$\mathbf{k}(s) := k(s) \cdot \mathbf{n}(s).$$

Lemma C.3. *Given an embedded curve r with $r(s) = (x(s), y(s))$ and let the curve be parameterized by arclength with s. Let t be an arbitrary parameterization. Then*

$$\frac{d\,s(t)}{d\,t} = \dot{s}(t) \;=\; \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}. \tag{C.1}$$

Proof. Let $s(t)$ be the arclength of the curve between t_0 and t. Then $s(t)$ is given by

$$s(t) = \int_{t_0}^{t} \sqrt{\dot{x}(\tilde{t})^2 + \dot{y}(\tilde{t})^2}\; d\,\tilde{t}.$$

Calculating the derivative of $s(t)$ with respect to t

$$\dot{s}(t) = \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}$$

finishes the proof. □

Proposition C.4. *If r is an embedded curve with an arbitrary parameterization given by $r(t) = (x(t), y(t))$, then the curvature $k(t)$ is given by*

$$k(t) = \frac{\dot{x}\,\ddot{y} - \ddot{x}\,\dot{y}}{(\dot{x}^2 + \dot{y}^2)^{3/2}}, \tag{C.2}$$

where $x := x(t)$ and $y := y(t)$ and $\dot{x}$ is the derivative of x with respect to t.

Proof. Let φ be the angle between the tangent vector and the positive x-axis, then

$$\tan\varphi = \frac{d\,y}{d\,x} = \frac{d\,y}{d\,t} \cdot \left(\frac{d\,x}{d\,t}\right)^{-1} = \frac{\dot{y}}{\dot{x}}$$

and therefore φ is given by

$$\varphi = \tan^{-1}\left(\frac{\dot{y}}{\dot{x}}\right).$$

Finally we calculate the derivative using Equation (C.1).

$$\frac{d\,\varphi}{d\,s} = \frac{d\,\varphi}{d\,t} \cdot \frac{d\,t}{d\,s} = \left(1 + \left(\frac{\dot{y}}{\dot{x}}\right)^2\right)^{-1} \frac{d}{d\,t}\left(\frac{\dot{y}}{\dot{x}}\right) \cdot (\dot{x}^2 + \dot{y}^2)^{-1/2} = \frac{\dot{x}\,\ddot{y} - \ddot{x}\,\dot{y}}{(\dot{x}^2 + \dot{y}^2)^{3/2}}$$

□

Next let the curve be represented explicitely, i.e. it is given by $y = f(x)$ where f is a map with $f : \mathbb{R} \to \mathbb{R}$.

Proposition C.5. *If the curve in the plane is given explicitely by $y = f(x)$, the curvature of the curve at a point $(x, f(x))$ is*

$$k(x) = \frac{f''(x)}{(1 + f'(x)^2)^{3/2}}, \tag{C.3}$$

where f' denotes the derivative of f with respect to x.

Proof. Let the parameterization be given by $x = t$ and $y = f(t)$, then by using Equation (C.2) we finish the proof.

$$k(x) = k(t) \;=\; \frac{\dot{x}\,\ddot{y} - \ddot{x}\,\dot{y}}{(\dot{x}^2 + \dot{y}^2)^{3/2}} = \frac{1 \cdot f''(t) - 0 \cdot f'(t)}{(1^2 + f'(t)^2)^{3/2}} = \frac{f''(x)}{(1 + f'(x)^2)^{3/2}}$$

□

Finally let the curve Γ be given by the zero level set of a function $F : \mathbb{R}^2 \to \mathbb{R}$, i.e. $\Gamma = \{(x, y),\, F(x, y) = 0\}$. Then the curvature is given by the following proposition.

Proposition C.6. *The curvature of a curve given implicitely, i.e. by the zero level set of a function $F : \mathbb{R}^2 \to \mathbb{R}$, is given by*

$$k(x, y) = -\frac{F_{xx}F_y^2 - 2 \cdot F_{xy}F_xF_y + F_{yy}F_x^2}{(F_x^2 + F_y^2)^{3/2}} \tag{C.4}$$

where F_x denotes the derivative in x-direction and similarly F_y the derivative in y-direction.

Proof. First we assume the $F_y \neq 0$. Then in a neighborhood of each point the curve can be represented explicitely by $y = f(x)$. This holds by the implicit function theorem. By the definition of the zero level set we have $F(x, f(x)) = 0$ and therefore

$$\frac{\partial}{\partial x} F(x, f(x)) = F_x + F_y f'(x) = 0$$

and

$$\frac{\partial^2}{\partial x^2} F(x, f(x)) = F_{xx} + F + xyf'(x) + F_{xy}f'(x) + F_y y f'(x)^2 + F_y f''(x) = 0.$$

As $F_y \neq 0$ we can transform these equations to

$$\begin{aligned} f'(x) &= -F_x/F_y \\ f''(x) &= F_y^{-1} \cdot (-F_{xx} + 2F_{xy}F_xF_y^{-1} - F_{yy}F_x^2F_y^{-2}). \end{aligned}$$

Finally we use Equation (C.3) and get the claim. The denominator is

$$(1 + f'(x)^2)^{3/2} = F_y^{-3}(F_x^2 + F_y^2)^{3/2}.$$

And therefore the final equation is given by

$$k(x, y) = \frac{f''(x)}{(1 + f'(x)^2)^{3/2}} = -\frac{F_{xx}F_y^2 - 2 \cdot F_{xy}F_xF_y + F_{yy}F_x^2}{(F_x^2 + F_y^2)^{3/2}}.$$

□

We conclude this section by showing an easy expression that is equal to the curvature of an implicit curve.

Proposition C.7. *The curvature k of a curve represented implicitely is given by the formula*

$$k(x,y) = -\nabla \cdot \frac{\nabla F(x,y)}{|\nabla F(x,y)|},$$

where we will denote in the following

$$\kappa(F) = \nabla \cdot \frac{\nabla F}{|\nabla F|} = \operatorname{div}\left(\frac{\nabla F}{|\nabla F|}\right) = \operatorname{div}\left(\frac{DF}{|DF|}\right),$$

which are other notations common in the literature.

Proof. In order to prove this proposition we just calculate $\nabla \cdot \frac{\nabla F}{|\nabla F|}$.

$$\begin{aligned}
\nabla \cdot \frac{\nabla F}{|\nabla F|} = \operatorname{div}\left(\frac{(F_x, F_y)}{\sqrt{F_x^2 + F_y^2}}\right) &= \frac{\partial}{\partial x}\left(\frac{F_x}{\sqrt{F_x^2 + F_y^2}}\right) + \frac{\partial}{\partial y}\left(\frac{F_y}{\sqrt{F_x^2 + F_y^2}}\right) \\
&= \frac{F_{xx} \cdot \sqrt{F_x^2 + F_y^2} - F_x \cdot (F_{xx}F_x + F_{xy}F_y) \cdot (F_x^2 + F_y^2)^{-1/2}}{F_x^2 + F_y^2} \\
&\quad + \frac{F_{yy} \cdot \sqrt{F_x^2 + F_y^2} - F_y \cdot (F_{yy}F_y + F_{xy}F_x) \cdot (F_x^2 + F_y^2)^{-1/2}}{F_x^2 + F_y^2} \\
&= \frac{F_{xx}F_y^2 - 2 \cdot F_{xy}F_xF_y + F_{yy}F_x^2}{(F_x^2 + F_y^2)^{3/2}} = -k,
\end{aligned}$$

where we used Equation (C.4) in the last step. □

Bibliography

[1] I. M. Anderson, *Introduction to the variational bicomplex*, in Mathematical aspects of classical field theory (Seattle, WA, 1991), vol. 132 of Contemp. Math., Amer. Math. Soc., Providence, RI, 1992, pp. 51–73.

[2] D. Barkley, *A model for fast computer simulation of waves in excitable media*, Phys. D, 49 (1991), pp. 61–70.

[3] G. Barles, *Viscosity solutions of Hamilton-Jacobi equations. (Solutions de viscosité des équations de Hamilton-Jacobi.)*, Mathématiques & Applications (Paris). 17. Paris: Springer. ix, 194 p., 1994.

[4] R. Bass and M. Cranston, *The Malliavin calculus for pure jump processes and applications to local time*, Ann. Probab., 14 (1986), pp. 490–532.

[5] H. Bauer, *Measure and integration theory. (Maß- und Integrationstheorie.) 2., überarb. Aufl.*, de Gruyter Lehrbuch. Berlin: Walter de Gruyter. xviii, 260 p., 1992.

[6] A. Bensoussan and R. Glowinski, *Approximation of Zakaï equation by the splitting up method*. Stochastic systems and optimization, Proc. 6th IFIP WG 7.1 Work. Conf., Warsaw/Pol. 1988, Lect. Notes Control Inf. Sci. 136, 257–265, 1989.

[7] A. Bensoussan, R. Glowinski, and A. Răşcanu, *Approximation of some stochastic differential equations by the splitting up method*, Appl. Math. Optimization, 25 (1992), pp. 81–106.

[8] S. Bochner, *Harmonic analysis and the theory of probability*, (California Monographs in Mathematical Sciences.) Berkeley: University of California Press. viii, 176 p., 1955.

[9] V. Caselles, R. Kimmel, and G. Sapiro, *Geodesic Active Contours*, Int. J. of Comput. Vision, 22 (1997), pp. 61–79.

[10] T. Chan and L. Vese, *Active contours without edges*, IEEE Trans. Image Processing, 10 (2001), pp. 266–277.

[11] Y. G. Chen, Y. Giga, and S. Goto, *Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations*, J. Differential Geom., 33 (1991), pp. 749–786.

[12] P.-L. Chow, *Stochastic partial differential equations*, Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series. Boca Raton, FL: Chapman & Hall/CRC. ix, 281 p., 2007.

[13] R. Courant, *Vorlesungen über Differential- und Integralrechnung. Zweiter Band: Funktionen mehrerer Veränderlicher*, Vierte Auflage, Berlin – Heidelberg – New York: Springer. xi, 468p., 1972.

[14] R. Courant and D. Hilbert, *Methoden der Mathematischen Physik I*, Heidelberger Taschenbücher. Berlin – Heidelberg – New York: Springer. xv, 469p., 1968.

[15] M. G. Crandall, H. Ishii, and P.-L. Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Am. Math. Soc., New Ser., 27 (1992), pp. 1–67.

[16] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, Encyclopedia of Mathematics and Its Applications. 44. Cambridge: Cambridge University Press. xviii, 454 p., 1992.

[17] A. Davie and J. Gaines, *Convergence of numerical schemes for the solution of parabolic stochastic partial differential equations.*, Math. Comput., 70 (2001), pp. 121–134.

[18] F. Erwe, *Differential- und Integralrechnung. Band II: Integralrechnung*, Bibliographisches Institut, Mannheim, 1964.

[19] G. Evans, J. Blackledge, and P. Yardley, *Numerical methods for partial differential equations*, Springer Undergraduate Mathematics Series. Berlin: Springer. xii, 290 p., 2000.

[20] L. C. Evans and J. Spruck, *Motion of level sets by mean curvature I*, J. Differential Geom., 33 (1991), pp. 635–681.

[21] W. H. Fleming and H. Soner, *Controlled Markov processes and viscosity solutions*, Applications of Mathematics. 25. New York: Springer. xv, 428 p., 1993.

[22] P. Florchinger and F. Le Gland, *Time-discretization of the Zakai equation for diffusion processes observed in correlated noise*, Stochastics Stochastics Rep., 35 (1991), pp. 233–256.

[23] O. Forster, *Analysis. 3*, vol. 52 of Vieweg Studium: Aufbaukurs Mathematik [Vieweg Studies: Mathematics Course], Friedr. Vieweg & Sohn, Braunschweig, 1981. Integralrechnung im $\mathbf{R}^n$ mit Anwendungen. [Integral calculus in $\mathbf{R}^n$ with applications].

[24] A. Friedman, *Partial differential equations of parabolic type*, Englewood Cliffs, N.J.: Prentice-Hall, Inc. xiv, 347 p., 1964.

[25] M. Frigo and S. G. Johnson, *The Design and Implementation of FFTW3*, Proceedings of the IEEE, 93 (2005), pp. 216–231. special issue on "Program Generation, Optimization, and Platform Adaptation".

[26] M. Gage and R. S. Hamilton, *The heat equation shrinking convex plane curves*, J. Differential Geom., 23 (1986), pp. 69–96.

[27] I. M. Gel'fand and N. Y. Vilenkin, *Generalized functions. Vol. 4: Applications of harmonic analysis*, Translated by Amiel Feinstein, Academic Press. xiv, 384 p., New York, 1964.

[28] A. Germani and M. Piccioni, *Semi-discretization of stochastic partial differential equations on R^d by a finite-element technique*, Stochastics, 23 (1988), pp. 131–148.

[29] J. Glimm and A. Jaffe, *Quantum physics. A functional integral point of view*, New York – Heidelberg – Berlin: Springer. xx, 417 p., 1981.

[30] M. A. Grayson, *The heat equation shrinks embedded plane curves to round points*, J. Differential Geom., 26 (1987), pp. 285–314.

[31] W. Grecksch and P. E. Kloeden, *Time-discretised Galerkin approximations of parabolic stochastic PDEs*, Bull. Austral. Math. Soc., 54 (1996), pp. 79–85.

[32] C. Grossmann and H.-G. Roos, *Numerics of partial differential equations. (Numerik partieller Differentialgleichungen.)*, Teubner Studienbücher: Mathematik. Stuttgart: B. G. Teubner. 477 p., 1992.

[33] I. Gyöngy, *Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise. I*, Potential Anal., 9 (1998), pp. 1–25.

[34] ——, *Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise. II*, Potential Anal., 11 (1999), pp. 1–37.

[35] ——, *Approximations of stochastic partial differential equations.* Da Prato, Giuseppe (ed.) et al., Stochastic partial differential equations and applications. New York, NY: Marcel Dekker. Lect. Notes Pure Appl. Math. 227, 287–307 (2002)., 2002.

[36] I. Gyöngy and N. Krylov, *On splitting up method and stochastic partial differential equations.* IMA Preprint Series, No 1737, December 2000.

[37] ——, *On the splitting-up method and stochastic partial differential equations*, Ann. Probab., 31 (2003), pp. 564–591.

[38] I. Gyöngy and D. Nualart, *Implicit scheme for stochastic parabolic partial differential equations driven by space-time white noise*, Potential Anal., 7 (1997), pp. 725–757.

[39] E. Hausenblas, *Error analysis for approximation of stochastic differential equations driven by Poisson random measures*, SIAM J. Numer. Anal., 40 (2002), pp. 87–113.

[40] E. Hausenblas and I. Marchis, *A numerical approximation of parabolic stochastic partial differential equations driven by a Poisson random measure*, BIT, 46 (2006), pp. 773–811.

[41] T. Hida, *Brownian motion. Transl. by the author and T. P. Speed*, Applications of Mathematics, 11. New York – Heidelberg – Berlin: Springer. xvi, 325 p., 1980.

[42] T. Hida, H.-H. Kuo, J. Potthoff, and L. Streit, *White noise. An infinite dimensional calculus*, Mathematics and its Applications (Dordrecht). 253. Dordrecht: Kluwer Academic Publishers. xiii, 516 p., 1993.

[43] E. Hille and R. S. Phillips, *Functional analysis and semi-groups*, Colloquium Publications, 31. Providence, R.I.: American Mathematical Society (AMS). xii, 808 p., 1981.

[44] N. Hofmann, T. Müller-Gronbach, and K. Ritter, *The optimal discretization of stochastic differential equations*, J. Complexity, 17 (2001), pp. 117–153.

[45] K. Itô and B. Rozovskii, *Approximation of the Kushner equation for nonlinear filtering*, SIAM J. Control Optimization, 38 (2000), pp. 893–915.

[46] G.-S. Jiang and D. Peng, *Weighted ENO schemes for Hamilton-Jacobi equations*, SIAM J. Sci. Comput., 21 (2000), pp. 2126–2143.

[47] O. Juan, R. Keriven, and G. Postelnicu, *Stochastic Motion and the Level Set Method in Computer Vision: Stochastic Active Contours*, International Journal of Computer Vision, 69 (2006), pp. 7–25.

[48] J.-P. Kahane, *Some random series of functions. 2nd ed.*, Cambridge Studies in Advanced Mathematics, 5. Cambridge: Cambridge University Press. xiii, 305 p., 1985.

[49] I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus. 2nd ed.*, Graduate Texts in Mathematics, 113. New York: Springer. xxiii, 470 p., 1991.

[50] T. Kato, *Perturbation theory for linear operators. Corr. printing of the 2nd ed.*, Grundlehren der mathematischen Wissenschaften, 132. Berlin – Heidelberg – New York: Springer. xxi, 619 p., 1980.

[51] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, *Optimization by Simulated Annealing*, Science, 220 (1983), pp. 671–680.

[52] P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, Applications of Mathematics. 23. Berlin: Springer. xxxv, 632 p., 1992.

[53] P. Knabner and L. Angermann, *Numerical methods for elliptic and parabolic partial differential equations. Translation from the German*, Texts in Applied Mathematics 44. New York: Springer. xv, 424 p., 2003.

[54] H. König, *Ein einfacher Beweis des Integralsatzes von Gauss*, Jahresber. Deutsch. Math.-Verein., 66 (1963/1964), pp. 119–138.

[55] H. P. Langtangen, *Computational partial differential equations. Numerical methods and Diffpack programming. 2nd ed.*, Texts in Computational Science and Engineering 1. Berlin: Springer. xxvi, 855 p., 2003.

[56] S. Larsson and V. Thomée, *Partial differential equations with numerical methods*, Texts in Applied Mathematics. 45. Berlin: Springer. ix, 259 p., 2003.

[57] P.-L. Lions and P. E. Souganidis, *Fully nonlinear stochastic partial differential equations*, C. R. Acad. Sci. Paris Sér. I Math., 326 (1998), pp. 1085–1092.

[58] ——, *Fully nonlinear stochastic partial differential equations: Non-smooth equations and applications*, C. R. Acad. Sci. Paris Sér. I Math., 327 (1998), pp. 735–741.

[59] ——, *Équations aux dérivées partielles stochastiques nonlinéaires et solutions de viscosité*, in Seminaire: Équations aux Dérivées Partielles, 1998–1999, Sémin. Équ. Dériv. Partielles, École Polytech., Palaiseau, 1999, pp. Exp. No. I, 15.

[60] ——, *Fully nonlinear stochastic PDE with semilinear stochastic dependence*, C. R. Acad. Sci. Paris Sér. I Math., 331 (2000), pp. 617–624.

[61] ——, *Uniqueness of weak solutions of fully nonlinear stochastic partial differential equations*, C. R. Acad. Sci. Paris Sér. I Math., 331 (2000), pp. 783–790.

[62] G. J. Lord and J. Rougemont, *Topological and ε-entropy for large volume limits of discretized parabolic equations*, SIAM J. Numer. Anal., 40 (2002), pp. 1311–1329.

[63] ——, *A numerical scheme for stochastic PDEs with Gevrey regularity*, IMA J. Numer. Anal., 24 (2004), pp. 587–604.

[64] J. MacQueen, *Some methods for classification and analysis of multivariate observations*, in Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, University of California, 1967, pp. 281–297.

[65] G. Marsaglia, *Yet another rng*. Posted to electronic bulletin board sci.stat.math, August 1994.

[66] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *Equation of State Calculations by Fast Computing Machines*, The Journal of Chemical Physics, 21 (1953), pp. 1087–1092.

[67] J.-M. Morel and S. Solimini, *Segmentation of images by variational methods: a constructive approach*, Rev. Mat. Univ. Complut. Madrid, 1 (1988), pp. 169–182.

[68] D. Mumford and J. Shah, *Optimal approximations by piecewise smooth functions and associated variational problems*, Comm. Pure Appl. Math., 42 (1989), pp. 577–685.

[69] N. Nagase, *Remarks on nonlinear stochastic partial differential equations: An application of the splitting-up method*, SIAM J. Control Optimization, 33 (1995), pp. 1716–1730.

[70] B. Øksendal, *Stochastic differential equations. An introduction with applications. 6th ed.*, Universitext. Berlin: Springer. xxiii, 360 p., 2003.

[71] P. J. Olver, *Applications of Lie groups to differential equations*, vol. 107 of Graduate Texts in Mathematics, Graduate Texts in Mathematics. New York – Berlin – Heidelberg: Springer. xxviii, 513 p., New York, second ed., 1993.

[72] S. Osher and R. Fedkiw, *Level set methods and dynamic implicit surfaces*, vol. 153 of Applied Mathematical Sciences, Applied Mathematical Sciences, 153. New York – Berlin – Heidelberg: Springer. xiv, 273 p., New York, 2003.

[73] N. Paragios and R. Deriche, *Geodesic active regions and level set methods for supervised texture segmentation*, Int. J. Comput. Vis., 46 (2002), pp. 223–247.

[74] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Applied Mathematical Sciences, 44. New York – Berlin – Heidelberg – Tokyo: Springer. viii, 279 p., 1983.

[75] M. Piccioni, *Convergence of implicit discretization schemes for linear differential equations with application to filtering.* Stochastic partial differential equations and applications, Proc. Conf., Trento/Italy 1985, Lect. Notes Math. 1236, 208–229, 1987.

[76] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C++*, Cambridge: Cambridge University Press. xxviii, 1002 p., 2002.

[77] A. Quarteroni and A. Valli, *Numerical approximation of partial differential equations*, Springer Series in Computational Mathematics. 23. Berlin: Springer. xvi, 543 p., 1994.

[78] M. Reed and B. Simon, *Methods of modern mathematical physics. I: Functional analysis*, New York – London: Academic Press, Inc. xvii, 325 p., 1972.

[79] ——, *Methods of modern mathematical physics. II: Fourier analysis, self- adjointness*, New York – San Francisco – London: Academic Press, a subsidiary of Harcourt Brace Jovanovich, Publishers. xv, 361 p., 1975.

[80] R. D. Richtmyer and K. Morton, *Difference methods for initial-value problems. Repr. of the 2nd ed. 1967*, Melbourne, FL: Krieger Publishing Company. xiv, 405 p., 1994.

[81] B. D. Ripley, *Stochastic simulation*, Wiley Series in Probability and Mathematical Statistics. Applied Probability and Statistics. New York: John Wiley & Sons. xi, 237 p., 1987.

[82] J. Rissanen, *Stochastic complexity in statistical inquiry*, Series in Computer Science. 15. Singapore etc.: World Scientific. iii, 177 p., 1989.

[83] M. Rousson and R. Deriche, *A Variational Framework for Active and Adaptative Segmentation of Vector Valued Images*, in MOTION '02: Proceedings of the Workshop on Motion and Video Computing, Washington, DC, USA, 2002, IEEE Computer Society, p. 56.

[84] J. Sethian, *Level set methods and fast marching methods. Evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*, Cambridge Monographs on Applied and Computational Mathematics. 3. Cambridge: Cambridge University Press. xx, 378 p., 1999.

[85] G. Sewell, *The numerical solution of ordinary and partial differential equations. 2nd ed.*, Pure and Applied Mathematics. A Wiley-Interscience Series of Texts, Monographs, and Tracts. Hoboken, NJ: John Wiley & Sons. xii, 333 p., 2005.

[86] T. Shardlow, *Numerical methods for stochastic parabolic PDEs*, Numer. Funct. Anal. Optimization, 20 (1999), pp. 121–145.

[87] ——, *Weak convergence of a numerical method for a stochastic heat equation*, BIT, 43 (2003), pp. 179–193.

[88] ——, *Numerical simulation of stochastic PDEs for excitable media*, J. Comput. Appl. Math., 175 (2005), pp. 429–446.

[89] P. Takáč and A. Jüngel, *A nonstiff Euler discretization of the complex Ginzburg-Landau equation in one space dimension*, SIAM J. Numer. Anal., 38 (2000), pp. 292–328.

[90] A. Tocino, *Mean-square stability of second-order Runge-Kutta methods for stochastic differential equations*, J. Comput. Appl. Math., 175 (2005), pp. 355–367.

[91] A. M. Vinogradov, *On the algebro-geometric foundations of Lagrangian field theory*, Dokl. Akad. Nauk SSSR, 236 (1977), pp. 284–287.

[92] C. Wallace and D. Boulton, *An information measure for classification*, Comput. J., 11 (1968), pp. 185–194.

[93] D. Werner, *Functional analysis. (Funktionalanalysis.) 4. überarbeitete Aufl.*, Springer-Lehrbuch. Berlin: Springer. xiii, 503 p., 2002.

[94] E. Wong and M. Zakai, *Riemann-Stieltjes approximations of stochastic integrals*, Z. Wahrscheinlichkeitstheor. Verw. Geb., 12 (1969), pp. 87–97.

[95] H. Yoo, *An analytic approach to stochastic differential equations and its applications*, PhD thesis, University of Minnesota, 1998.

[96] ——, *On L_2-theory of discrete stochastic evolution equations and its application to finite difference approximations of stochastic PDEs.* Preprint, 1998.

[97] ——, *Semi-discretization of stochastic partial differential equations on $\mathbb{R}^1$ by a finite-difference method*, Math. Comput., 69 (2000), pp. 653–666.

[98] K. Yosida, *Functional analysis. 6th ed.*, Grundlehren der mathematischen Wissenschaften, 123. Berlin – Heidelberg – New York: Springer. xii, 501 p., 1980.

[99] H.-K. Zhao, T. Chan, B. Merriman, and S. Osher, *A variational level set approach to multiphase motion*, J. Comput. Phys., 127 (1996), pp. 179–195.