

Chapter 1

Introduction

Bifurcation theory for stochastic partial differential equations (SPDEs) is still not fully developed. In contrast to that stochastic ordinary differential equations (SDEs) especially with one-dimensional phase space are widely studied, and there are numerous results on bifurcation. For instance, for one-dimensional SDEs a detailed classification of possible bifurcation scenarios is established. See for example [Ste00] or [CIS99]. But already with a two-dimensional phase space there are open problems.

An interesting feature of stochastic bifurcation theory is that there are several approaches describing bifurcations in SDEs, which sometimes yield completely different results. Numerous articles are devoted to the study of relations and differences of these concepts. Some examples of concepts are:

- Phenomenological bifurcation (or P -bifurcation), which characterises a bifurcation by using qualitative changes in the unique invariant measure for the corresponding Markov semigroup given by the law of stationary solutions.
- D -bifurcation (sometimes also called dynamical bifurcation), which is characterised by the changes in the set of random invariant measures for random dynamical systems.

For a D -bifurcation one could simply look at the change in the structure of the random attractor, which is a compact random set attracting all orbits or all bounded sets. See for instance [Arn98] for the theory of random dynamical systems and for examples of possible bifurcations (see Chapter 9 of [Arn98]). The connection between random attractors and invariant measures is for instance studied in [Sch99; Cra01]. Furthermore first steps are provided towards an abstract bifurcation theory in the spirit of catastrophe-theory for dynamical systems. See for example [DW06a; DW05] or [DTZ05b; DTZ05a].

Phenomenological Bifurcation

If we consider phenomenological bifurcation for equations with additive noise, then it is usually easy to derive the uniqueness of invariant measures under very mild non-

degeneracy conditions on the noise. See [DPZ96] for a textbook and for examples of SPDEs with mild degeneracy conditions see [EM01] or [EH01] and the references therein. Thus on the level of Markov-processes the long-time evolution is well described by this invariant measure.

Furthermore, one can study the invariant measure as the solution of the well known Kolmogorov equation, which is extensively studied (see for example [Cer01] or [DPZ02] or [RS04; RS06; KRZD99]). But in contrast to that there are hardly any examples describing qualitatively the structure of the invariant measures.

In Sections 3.1 and Theorem 4.4 we describe the results obtained in [BH04] and [BHP05]. These are two applications, where one actually can describe the structure of the invariant measure near a bifurcation. Hence, it is possible to detect phenomenological bifurcations in these models.

Random Attractors

There are many examples deriving the existence of random attractors for SPDEs. See for example [CF94; CDF97] for the definition of random attractors and for instance [FS96; FS99; DKS01] for some applications. But there are only a few results on the fine structure of the random attractors. For instance, upper bounds for the dimension are well known, but for many examples with additive noise, the random attractor is just a single random point.

For multiplicative noise the structure of a random attractor near a deterministic fixed point was studied in [CLR01]. It was shown that the random attractor actually undergoes a bifurcation, where the attractor changes from a single point to a random set. This was achieved by looking at the local dimension of the corresponding stable and unstable invariant manifolds in a small neighbourhood near the fixed point. For the general theory of invariant manifolds for SPDEs see [DLS03] or [MZZ07], but these results focus mainly on existence.

Random attractors for SDEs or SPDEs have been the topic of intense research. But when we look at the structure of a random attractor, it is just a single random point for many examples with additive noise. See for instance the celebrated work of Crauel and Flandoli [CF98] showing that for a one-dimensional SDE with additive noise there is no bifurcation on the level of random attractors. This was extended to monotone SPDEs in [CS04]. Similar results for non-monotone systems were derived by [Tea06b; Tea06a]. These results show that different concepts of bifurcation yield different results, as most of these examples do exhibit a phenomenological bifurcation. But the bifurcation can only be seen from the probability distribution or the dynamics of the random fixed point in time.

The problem of the attractor being a single random point is in general open. Even for simple two-dimensional phase space, there are only results for very small or very large noise or monotone SDEs.

Here we take a somewhat different approach. Instead of looking at the bifurcation of objects for time to ∞ , we look at the typical transient behaviour of solutions

of SPDEs. Here transient means possibly large, but still finite, time intervals. But even for SDEs the typical transient dynamics is not very well described in many examples. One exception being the work of Berglund and Gentz (see for instance [BG02a], [BG03], or the book [BG06]), where a detailed description of dynamics was derived for low-dimensional phase-spaces. Our results allow to carry some of these results immediately over to SPDEs. See for example [Blö03].

Freidlin-Wentzell Theory

We focus on equations with small noise in a neighbourhood of a change of stability. A similar but slightly different approach is the celebrated theory of Freidlin and Wentzell on large deviation effects. See [FW98] or Section 12 of [DPZ92]. One topic in this theory is to point out the effect of very small noise on the exit from domains of attraction of the deterministic model. Nevertheless, these effects occur only on time-scales, which are exponentially large in the noise strength, as in the small noise limit the solution follows the deterministic model for most of the time with very high probability. Thus, as the noise is induced for instance by very small thermal fluctuations, all these effects are obviously difficult to detect in experiments.

For Freidlin-Wentzell theory of SPDEs see [DPZ92; OV05]. For application to Cahn-Hilliard and Allen-Cahn see [Gaw06; Bra91; BGW07].

Our point of view is different from classical Freidlin-Wentzell Theory. We consider not only small noise, but also small parameters in the equation. This reflects the fact, that we are interested in experiments, where there is on one hand a very small source of noise (e.g. thermal fluctuations), which is usually described as the small noise limit of noise intensity to 0. But on the other hand one also tries to get the parameters sufficiently close to the change of stability, in order to see the effects of this bifurcation on the dynamics of the model, which is a second limit to 0.

One of our interest is to have a coupling of these two limits. In the context of Freidlin-Wentzell theory, this was only studied by using first the small noise limit, and then later on the limit of the parameter to 0 in the action functional, but not both limits at the same time. See for example [BDP05; OKRVE06; RK06].

Multi-Scale Analysis

The approach presented here relies on the fact that near a change of stability the dynamics of (stochastic) PDEs is driven by some dominant modes. These are exactly the eigenfunctions of the linearised operator, with eigenvalues close to 0. All other modes are subject to strong linear damping. Amplitude equations describe these essential dynamics. This is well known in the physics community and rigorously established for the deterministic PDEs. See for example [CE90; KSM92; Sch94; MS95] or for a recent review on bifurcation theory on unbounded domains [Sch01]. For PDEs on unbounded domains these amplitude equations are also known as modulation equations.

On a formal level, as already mentioned, there are numerous results using non-rigorous multiple scale analysis to derive reduced descriptions for the dynamics. See for example [CH93] or [Hak83] both containing plenty of such formal approximations. An interesting example is [Kus03] for highly oscillating solutions of an SDE.

First rigorous results are possible do to the theorem of Kurtz [Kur73] but, based on abstract semi-group theory for generators of Markov-diffusions, it provides no error bound, just the weak convergence of the approximation. Furthermore, it is a result for SDEs and not for SPDEs. See for example [WSPS04; MTVE01] for applications.

Pattern Formation Below Threshold of Stability

If the distance from the change of stability is sufficiently small, then the influence of small noise is detected in experiments. See for example the work of Ahlers, Rehberg et.al. on pattern formation below threshold of instability in convection problems. In Bénard-type models for electro-convection of liquid crystals [SA02; SR94] and in Rayleigh-Bénard convection for fluids [OdZSA04; OA03], it was verified experimentally, that near a change of stability there is a significant impact of thermal noise on the dynamics, leading to the formation of pattern in otherwise deterministically stable equations. This was long conjectured (see for example [CH93; HS92] and the references therein). The main difficulty of the experiment was to stabilise the control parameters (for example the temperature in Rayleigh-Bénard convection) to the precision of the noise strength, which is extremely small in case of thermal fluctuations in fluids. This is the main reason, why the effect was seen in electro-convection first.

The non-rigorous explanation of the experimentally verified effects relies on a formal expansion of the solution in the noise strength and separation between slow and fast components, in order to derive an effective equation for the amplitudes of the dominating pattern. This is also known as multi-scale expansion. We present results using this type of formal calculation in Section 1.1.

Amplitude Equation

The main object of this book is to rigorously establish estimates for the approximation of SPDEs via amplitude equations, and to answer questions on how noise influences the dynamics in systems near a change of stability. This has a lot of interesting applications, like pattern formation below criticality (see Section 3.2) or the structure of invariant measures for SPDEs (see Section 3.1), which gives insight into phenomenological bifurcations in SPDEs.

On bounded domains for SPDEs the approximation via amplitude equations was first rigorously verified in [BMPS01] for a simple Swift-Hohenberg model, and later on extended in [Blö03; Blö05a; BH04]. We outline typical results in Section

1.2 in a non-technical way. Here the amplitude equation for the dominant modes is given by an SDE. A typical example is

$$\partial_T \mathcal{A} = \nu \mathcal{A} - c |\mathcal{A}|^2 \mathcal{A} + \dot{\beta},$$

where $\mathcal{A}(T) \in \mathbb{R}^n$, $\nu, c \in \mathbb{R}$, and $\dot{\beta}$ is noise in \mathbb{R}^n .

In Chapter 2 we review these results in detail and present the theory applied to a simple model with multiplicative noise, also known as parameter noise. In that model rigorous results were not considered before.

While all results for amplitude equations are mainly limited to transient behaviour, it was in [BH04] also possible to approximate the long-time behaviour in terms of the structure of invariant measures for the corresponding Markov-semigroup. See also Section 3.1 or [BH05].

The case of unbounded or just very large domains is significantly different. The amplitudes of the dominant modes are subject to a long-range modulation in space, and hence not given by an SDE, but an SPDE instead. The celebrated example is an SPDE of Ginzburg-Landau type:

$$\partial_T A = \partial_X^2 A + \nu A - c |A|^2 A + \xi, \quad (1.1)$$

where $A(T, X) \in \mathbb{C}$ and ξ is space-time white noise.

The case of large, but still bounded, domains is discussed in [BHP05]. See also [MSZ00] for the deterministic equation. In both cases the domain is bounded, but it scales with respect to the distance from bifurcation. A review of the stochastic result can be found in Chapter 4. As already mentioned before, there is also a very large literature for deterministic equations on unbounded domains, but this seems to be out of reach for SPDEs.

In all of these articles for the stochastic case, we consider noise that is on one hand sufficiently small, as given from the experiment, but on the other hand sufficiently large compared to the distance from the bifurcation. As discussed before, the viewpoint of the experiment is that we try to adjust the bifurcation parameter sufficiently close to the bifurcation, in order to see both stochastic effects and the small linear stability or instability in the amplitude equation. Of course, we could look at different scalings between these two small quantities, losing either the noise or the linear (in)stability in the approximation. Depending on the type of equation the scaling of the coupling between the noise strength and the distance from bifurcation may change, in order to get an interesting stochastic amplitude equation.

The main difference between small and large domains is the existence of a large spectral gap of order $\mathcal{O}(1)$ in the linearised operator of the PDE. On bounded domains, we have a finite number $e = (e_1, \dots, e_n)$ of modes (or eigenfunctions) such that the corresponding eigenvalues change sign at the change of stability. If we are close to the bifurcation, all other eigenvalues are negative and sufficiently far away from 0. Formal arguments (see for instance Section 1.1.1) show, that the

amplitudes $\mathcal{A} \in \mathbb{R}^n$ of the dominating modes are given by the so called amplitude equation, while the solution u of the SPDE is well approximated by

$$u(t, x) = \varepsilon \mathcal{A}(\varepsilon^2 t) \cdot e(x) + \mathcal{O}(\varepsilon^2),$$

where ε^2 is the typical scale for the distance from bifurcation.

On unbounded or just very large domains this picture changes completely. Even very close to the bifurcation a large number of modes are near or already above the threshold of stability, but still small. In this case the amplitude \mathcal{A} is also a function in x that is concentrated in Fourier space near the dominant modes. Hence, \mathcal{A} is subject to slow modulations in space, taking into account the large number of weakly (un)stable modes. In Section 1.1.4 we see that the solution u is in this case given by

$$u(t, x) = \varepsilon \mathcal{A}(\varepsilon^2 t, \varepsilon x) \cdot e(x) + \mathcal{O}(\varepsilon^2)$$

and \mathcal{A} fulfils a (stochastic) PDE, which is called *amplitude or modulation equation*. The typical example being a stochastic Ginzburg-Landau equation, where for instance $e = (\sin, \cos)$, and we can thus write

$$u(t, x) = \varepsilon A(\varepsilon^2 t, \varepsilon x) e^{ix} + c.c. + \mathcal{O}(\varepsilon^2)$$

where the complex amplitude $A(T, X) \in \mathbb{C}$ solves a Ginzburg-Landau equation similar to (1.1).

A key point is that the amplitude equation for \mathcal{A} is in a certain sense independent of ε . Although the paths of the process $\{\mathcal{A}(T)\}_{T \geq 0}$ might depend explicitly on ε , this dependence disappears on the level of probability measures due to the scaling properties of the noise. This is explained in more detail in Remark 1.2. Thus any result for the amplitude equation immediately carries over to a rescaled statement for the original equation, provided that $\varepsilon > 0$ is small. This is for instance important in the results on pattern formation. See Section 3.2. One could also refer to this fact as a type of normal form theory, as the amplitude equation for many examples will always be of the same Ginzburg-Landau type.

In the remainder of the introduction, we first present results on the formal derivation of the amplitude equation in Section 1.1. Cubic nonlinearities are treated in Section 1.1.1, where we also take higher order corrections into account. Quadratic nonlinearities exhibit special properties, as in many examples dominant modes are not mapped to dominant modes by the nonlinearity. This will be addressed in Section 1.1.3, while in Section 1.1.4 the effect of large domains is studied. Section 1.2 presents the general method of proof, which is similar for all cases, and states typical results. While the final Section 1.3 of the introduction presents some examples of SPDEs, where the general theory applies.

1.1 Formal Derivation of Amplitude Equations

In this section, we discuss the formal derivation of amplitude equations and higher order corrections. Therefore, we use multiple scale analysis to reduce the equation to the essential dynamics, which involves the expansion of all terms in a small parameter. This is well known for many examples. Here we present results described in more detail for quadratic nonlinearities in [Blö05a] and for cubic nonlinearities in [BH04]. For large domains we summarise results of [BHP05] in Section 1.1.4.

Let us consider parabolic semilinear SPDEs or systems of SPDEs perturbed by additive forcing near a change of stability. Let us suppose, that the noise is of the order of the distance from the bifurcation. The use of additive noise is mainly for simplicity of presentation, and it is not very restrictive. We comment on multiplicative noise later in several occasions in Chapter 2. A large body of the research papers are on additive noise, which we will summarise later. In this book simple multiplicative noise is used to present a self-contained introduction to the topic.

The general prototype is an equation of the type

$$\partial_t u = Lu + \varepsilon^2 Au + \mathcal{F}(u) + \varepsilon^2 \xi, \quad (1.2)$$

where

- L is a symmetric non-positive differential operator (e.g. $1 + \partial_x^2$) with non-zero finite dimensional kernel (or null-space),
- $\varepsilon^2 Au$ is a small (linear) deterministic perturbation,
- \mathcal{F} is some nonlinearity, for instance a stable cubic like $-u^3$.
- $\xi = \xi(t, x)$ is a Gaussian noise in space and time

We later give examples of the noise, which is taken to be white in time and can be either white or coloured in space. To be more precise, suppose that ξ is a generalised Gaussian process such that for mean and correlation

$$\mathbb{E}\xi(t, x) = 0 \quad \text{and} \quad \mathbb{E}\xi(t, x)\xi(s, y) = \delta(t - s)q(x - y),$$

for some suitable spatial correlation function (or distribution) q . If q is the Delta-distribution δ , too, then we call ξ space-time white noise. In this case $\xi = \partial_t W$ is the generalised derivative of a cylindrical Wiener-process $\{W(t)\}_{t \geq 0}$ in a suitable Hilbert space. This means

- $W(0) = 0$
- The increments $W(t) - W(s)$ are independent for disjoint intervals (s, t) .
- The path $t \mapsto W(t)$ is continuous with probability 1.
- $W(t)$ is normal with covariance operator $t \cdot Id$.

For general q we can always write

$$\xi = \partial_t QW$$

for some symmetric linear operator Q . We will state details later when necessary. For a detailed discussion see [DPZ92] or Section 2.5.

For the formal calculation in the bounded domain case, we rely mainly on the scaling properties of the noise in time. To be more precise, we need that $\varepsilon^{-1}\xi(\varepsilon^{-2}T, x)$ and $\xi(T, x)$ are for all $\varepsilon > 0$ versions of the same noise process, which means that they are different ε -dependent processes, but their law is the same. This is easy to verify on the level of correlation functions, using the scaling properties of the Delta-distribution, or on a more rigorous level, one can use the scaling properties of W (cf. Assumption 1.1).

Remark 1.1 *For the rigorous results, we work with an integrated version of (1.2), using mild solutions in some Hilbert space X . The precise definition of this is given for instance in Proposition 2.1. Nevertheless, for the formal calculation we directly work with the SPDE. This is usual for formal calculations in physics. The standard mathematical way of writing the SPDE is*

$$du = (Lu + \varepsilon^2 Au + \mathcal{F}(u))dt + \varepsilon^2 dQW,$$

where one uses Itô-differentials.

1.1.1 Cubic Nonlinearities

One interesting example of an equation with cubic nonlinearity is the Swift-Hohenberg equation, which was first used as a toy model for the convective instability in the Rayleigh-Bénard problem (see [SH77] or Section 1.3).

On a formal level for the Swift-Hohenberg equation the derivation of the amplitude equation is well known, see for instance (4.31) or (5.11) in the comprehensive review article [CH93] and references therein. The amplitude equation for (1.3) was already treated in [BMPS01]. But here we follow the presentation from [BH04], taking into account second order corrections.

The formal SPDE is

$$\partial_t u = -(1 + \Delta)^2 u + \varepsilon^2 \nu u - u^3 + \varepsilon^2 \partial_t QW. \quad (1.3)$$

It is obviously of the type of (1.2) with $L = -(1 + \Delta)^2$, $A = \nu I$ for some $\nu \in [-1, 1]$, and $\mathcal{F}(u) = -u^3$. We can for instance consider periodic boundary conditions on the domain $[0, 2\pi l]^d$ for dimension $d \in \mathbb{N}$ and integer length $l > 0$. This is mainly for convenience to ensure that the change of stability is exactly at $\nu = 0$. After slight modifications we can also treat non-integer length $l > 0$ or non-squared domains.

For the formal derivation in this section we consider an equation of the type (1.2) or (1.3) and assume:

Assumption 1.1 *Let $\{QW(t)\}_{t \geq 0}$ be a Q -Wiener process. This implies especially that $\{W(t)\}_{t \geq 0}$ and $\{\varepsilon W(\varepsilon^{-2}t)\}_{t \geq 0}$ are in law the same process.*

Furthermore, let \mathcal{F} be cubic (i.e. $\mathcal{F}(u) = \mathcal{F}(u, u, u)$ is trilinear).

Denote the kernel (or nullspace) of L by \mathcal{N} and the orthogonal projection onto \mathcal{N} by P_c . Define $P_s = I - P_c$.

Then we make the following ansatz:

$$u(t) = \varepsilon a(\varepsilon^2 t) + \varepsilon^2 b(\varepsilon^2 t) + \varepsilon^2 \psi(t) + \mathcal{O}(\varepsilon^3), \quad (1.4)$$

with $a, b \in \mathcal{N}$ and $\psi \in \mathcal{S} := \mathcal{N}^\perp$ the orthogonal complement of \mathcal{N} in X .

This ansatz is motivated by the fact that, due to the linear damping of order one in \mathcal{S} , the modes in \mathcal{S} are expected to evolve on time scales of order one, whereas the modes in \mathcal{N} are expected to evolve on much slower time scales of order ε^{-2} , as the linear operator is of order ε^2 . This is mainly due to the fact that we have a well defined spectral gap of order $\mathcal{O}(1)$ between 0 and the first non-zero eigenvalue together with a small linear perturbation of order ε^2 .

We do not use lower order terms, as we expect that small solutions stay small. Furthermore, using linear and nonlinear stability, it is possible to verify a priori estimates that rigorously verify that the typical scaling of a solution corresponds to the one prescribed by the ansatz (1.4). The statement is called the *attractivity* result (cf. Section 1.2).

Let us now come back to the formal derivation. Plugging the ansatz (1.4) into (1.2), rescaling to the *slow time-scale* $T = \varepsilon^2 t$ and expanding in orders of ε , we obtain by collecting all terms of order ε^3 in \mathcal{N}

$$\partial_T a(T) = A_c a(T) + \mathcal{F}_c(a(T)) + \partial_T \beta(T). \quad (1.5)$$

Here,

$$\beta(T) = \varepsilon P_c QW(\varepsilon^{-2} T), \quad T \geq 0$$

is a Wiener process in \mathcal{N} with law independent of ε , due to the scaling properties of the Wiener process. We used

$$A_c = P_c A \quad \text{and} \quad \mathcal{F}_c = P_c \mathcal{F}$$

for short.

This approximating equation in (1.5) is called *amplitude equation*, as it can be rewritten to an SDE for the amplitudes of an expansion of a with respect to a basis in \mathcal{N} . Results like this well known for many examples in the physics and applied mathematics literature (for example [CH93, (4.31),(5.11)]). Moreover, there are numerous variants of this method. However, most of these results are non-rigorous approximations using this type of formal multi-scale analysis.

Remark 1.2 *One key point is that (1.5) is, at least in law, completely independent of the small parameter $\varepsilon > 0$, as the amplitude equation is in general ε -independent. Therefore, we do not use an index ε for a . Although the paths of the Brownian motion β in the amplitude equation (1.5) depend obviously on ε due to the rescaling in time, the law of β is independent of ε . This is due to the scaling properties of*

a Wiener process, which state that the process $\{\varepsilon W(\varepsilon^{-2}t)\}_{t \geq 0}$ is a version of the Wiener process $\{W(t)\}_{t \geq 0}$ for all $\varepsilon > 0$. Therefore, by the weak uniqueness the law of solutions of (1.5) is independent of ε . Thus we also neglect the dependence of a on ε in the notation, as it is only path-wise.

Higher Order Corrections

We already verified, that the slow modes (or amplitudes of the dominant modes) decouple from the fast modes, at least approximatively to the first order in ε . We now see that this remains true for the second order corrections, too.

Collecting terms of order ε^2 in \mathcal{S} yields for ψ the linear equation

$$\partial_t \psi(t) = L_s \psi(t) + P_s \xi(t), \tag{1.6}$$

where we defined $L_s = P_s L$.

One can show furthermore that $b = 0$. Therefore, we first rescale ψ to the slow time-scale to obtain formally that in law

$$\psi(T\varepsilon^{-2}) = \varepsilon L_s^{-1} P_s \xi(T) + \text{“higher order terms”}.$$

Thus the term $\psi(T\varepsilon^{-2})$, when viewed on the slow time-scale, gives a contribution of order ε . Using this, we obtain for terms of order ε^4 in \mathcal{N}

$$\partial_T b = A_c b + 3\mathcal{F}_c(a, a, b).$$

Since the initial condition for the equation is $u(0) = \varepsilon a(0) + \varepsilon^2 \psi(0)$ by the ansatz, one has $b(0) = 0$, and therefore b vanishes identically.

It is not easy to proceed to higher order corrections, as they do involve couplings of fast and slow modes, which will make the approximating equations much more complicated. Moreover, we formally need to justify quadratic functionals of the noise term, which is highly nontrivial. See [BH04] for a detailed discussion.

1.1.2 Other Types of Nonlinearities

Cubic nonlinearities are not very special, we can extend the simple idea of the previous section, using scaling and projection, to a lot of different types of nonlinearities. If we look at suitable scalings of the noise and the linear (in)stability we obtain in all cases interesting results. If we do not adapt the scaling, we either loose the linear instability or the noise in the amplitude equation.

Suppose for this section that $\mathcal{F}^{(n)}$ is some multi-linear nonlinearity, which is homogeneous of degree $n \in \mathbb{N}$ with $n \geq 2$ (i.e. for $\alpha > 0$, $\mathcal{F}^{(n)}(\alpha u) = \alpha^n \mathcal{F}^{(n)}(u)$). Then the noise strength in the SPDE (1.2) should be changed to $\sigma^2 = \varepsilon^{(n+1)/(n-1)}$ instead of ε^2 . Thus the equation reads in the interesting scaling

$$\partial_t u = Lu + \varepsilon^2 Au + \mathcal{F}^{(n)}(u) + \varepsilon^{(n+1)/(n-1)} \xi. \tag{1.7}$$

Now with the ansatz

$$u(t) = \varepsilon^{2/(n-1)} a(\varepsilon^2 t) + \varepsilon^{(n+1)/(n-1)} \psi(t) + \mathcal{O}(\varepsilon^{2n/(n-1)}) \quad (1.8)$$

and a similar formal calculation as in the previous section, we derive the same type of amplitude equation. First collecting all terms of order $\varepsilon^{2n/(n-1)}$ in \mathcal{N} yields

$$\partial_T a = P_c A a + P_c \mathcal{F}^{(n)}(a) + \partial_T \beta, \quad (1.9)$$

which now contains a nonlinearity which is homogeneous of degree n . The second order correction is exactly the same (cf. (1.6)) as in the cubic case, but now it contains all terms in \mathcal{S} of order $\varepsilon^{(n+1)/(n-1)}$.

We will not focus on rigorous results for this type of equations, as they are very similar to the cubic case. After minor modifications one can easily transfer all results to the general case.

Remark 1.3 *Let us finally comment on the distance from the bifurcation, that is necessary in order to see the stochastic effects in the amplitude equation. The order of magnitude of the distance in terms of the noise strength σ^2 is*

$$\varepsilon^2 = \sigma^{4(n-1)/(n+1)} \approx \begin{cases} \sigma^4 & \text{for } n \rightarrow \infty \\ 1 & \text{for } n \rightarrow 1 \end{cases}.$$

Thus, the higher the order of the nonlinearity, the nearer we have to go to the bifurcation, in order to see the influence of the noise on the dominant modes, at least on the time-scale under consideration. Let us finally change the point of view and express the noise in terms of $\varepsilon > 0$. Now

$$\sigma^2 = \varepsilon^{(n+1)/(n-1)} \approx \begin{cases} \varepsilon & \text{for } n \rightarrow \infty \\ 0 & \text{for } n \rightarrow 1 \end{cases}.$$

1.1.3 Quadratic Nonlinearities

An interesting feature of quadratic nonlinearities $B(u) = B(u, u)$ is that in many examples $P_c B(a) \equiv 0$ for all $a \in \mathcal{N}$. In this case, the ansatz (1.8) yields only the linearisation. See (1.9). This means that we still look at solutions that are too small to capture any of the nonlinear effects present in the equation. In order to obtain a nonlinear amplitude equation, we either consider larger noise, or we look at a parameter regime where we are nearer to the change of stability.

To illustrate this problem, we briefly discuss a one-dimensional Burgers' equation, which is given by

$$\partial_t u = \partial_x^2 u + \mu_\varepsilon u - u \partial_x u + \sigma_\varepsilon \xi.$$

Let ξ be space-time white noise for simplicity.

Example 1.1 For periodic boundary conditions on $[0, 2\pi]$ and $\mu_\varepsilon = \mathcal{O}(\varepsilon^2)$, we readily obtain for the space of dominant modes $\mathcal{N} = \text{span}\{1\}$. All other eigenfunctions of ∂_x^2 are spanned by $\sin(kx)$ and $\cos(kx)$ with $k \in \mathbb{N}$ and eigenvalues $-k^2 \leq -1 < 0$. But now for $B(u) = u\partial_x u$ we have $B(1) = 0$, obviously.

Example 1.2 Consider Dirichlet boundary conditions on $[0, \pi]$, then the linear instability arises for $\mu_\varepsilon = 1 + \mathcal{O}(\varepsilon^2)$. Furthermore, $\mathcal{N} = \text{span}\{\sin\}$, as all other eigenfunctions are spanned by $\sin(kx)$ with eigenvalues again bounded by -1 . But now $B_c(\sin) = 0$, where we used again the short-hand notation $B_c = P_c B$ and $B_s = P_s B$.

Hence, in both examples with the ansatz (1.4) we only derive the linearisation as the amplitude equation.

There are numerous examples in the physics literature of equations with quadratic nonlinearities and the same property, as described above. One model is the growth of rough amorphous surfaces. See for example [BGR02] and the references therein. Another example is the Kuramoto-Sivashinsky equation, but the probably most important one is the Rayleigh-Bénard problem which is the paradigm of pattern formation in convection problems. All equations are described briefly in Section 1.3.

If we want to take into account nonlinear effects, then we have to look at the coupling of the slow dominant modes to the fast modes. In the following, we will follow the presentation of [Blö05a], in order to briefly comment on the formal derivation of the amplitude equation in this case.

Consider an equation of the type

$$\partial_t u = Lu + \varepsilon^2 Au + B(u, u) + \varepsilon^2 \xi \tag{1.10}$$

with

- B a symmetric and bilinear operator with $B_c(a, a) = 0$ for $a \in \mathcal{N}$.
- L, A , and ξ as in (1.2)

We make the following ansatz, which is significantly different from (1.8):

$$u(t) = \varepsilon a(\varepsilon^2 t) + \varepsilon^2 \psi(t) + \mathcal{O}(\varepsilon^3),$$

with $a \in \mathcal{N}$ and $\psi \in \mathcal{S}$, in order to take into account both nonlinear and noise terms in the amplitude equation. Our new ansatz yields in lowest order in ε the following system of formal approximations. First of order $\mathcal{O}(\varepsilon^2)$ on the fast time-scale t in \mathcal{S} .

$$\partial_t \psi(t) = L_s \psi(t) + B_s(a(\varepsilon^2 t), a(\varepsilon^2 t)) + P_s \xi(t). \tag{1.11}$$

Secondly of order ε^3 in \mathcal{N} on the slow time-scale $T = \varepsilon^2 t$

$$\partial_T a(T) = A_c a(T) + 2B_c(a(T), \psi(\varepsilon^{-2} T)) + \partial_T \beta(T), \tag{1.12}$$

where the rescaled projection of the Wiener process β was defined after (1.5).

These equations (1.11) and (1.12) are a coupled system of equations. On one hand there is a dominating equation (1.12) on a slow time-scale, which is similar to the amplitude equation (1.5). This is coupled to an equation (1.11) on the fast time-scale. Equations with a similar structure are rigorously treated in [BG03; BG06] for SDEs, or in [PS03; KPS04] where tracers in a fast moving random velocity field are considered. There is also a review [GKS04] and numerous other references.

The aim is now to get an effective equation for the slow component completely independent of the fast modes. First rescale (1.11) to the slow time-scale $T = \varepsilon^2 t$ by $\psi(t) = \Phi(\varepsilon^2 t)$. Thus

$$\varepsilon^2 \partial_T \Phi = L_s \Phi + B_s(a, a) + \varepsilon P_s \hat{\xi},$$

where $\hat{\xi}(T) = \varepsilon^{-1} \xi(\varepsilon^{-2} T)$ is a rescaled version of the noise ξ . Now neglect all terms, which are not $\mathcal{O}(1)$. This yields the stationary problem

$$0 = L_s \Phi + B_s(a, a).$$

As L_s is invertible on \mathcal{S} , we get in lowest order of ε that

$$\Phi = -L_s^{-1} B_s(a, a).$$

This together with (1.12) establishes a single approximation equation.

$$\partial_T a = A_c a - 2B_c \left(a, L_s^{-1} B_s(a, a) \right) + \partial_T \beta, \quad (1.13)$$

Surprisingly, this equation involves a cubic nonlinearity, although the nonlinearity in the original equation was quadratic.

There is a whole zoo of other effects appearing, if we consider (1.10) with noise of strength ε , but not action on the dominant modes \mathcal{N} directly. See [Rob03] for the first observations and [BHP06] for the first rigorous treatment.

1.1.4 Large or Unbounded Domains

For unbounded domains the results are very different. First of all, we do not have a spectral gap, and near the change of stability a whole band of eigenvalues gets unstable. The same effect already occurs, if we consider large domains, which are at least of the size $\mathcal{O}(\varepsilon^{-1})$. In Figure 1.1 we briefly sketch the eigenvalue curve $k \mapsto -P(-k)$ with the corresponding eigenvalues of the Swift-Hohenberg operator $-P(i\partial_x) = -(1 + \partial_x^2)^2$. For the deterministic PDE this somewhat intermediate step was already discussed in [MSZ00]. The stochastic case is treated in [BHP05], but we present a different formal derivation here. This is closer to usual physical reasoning, and more in the spirit of [KSM92].

Consider as an example a one-dimensional version of the Swift-Hohenberg equation, which was first used as a toy-model for the convective instability in the

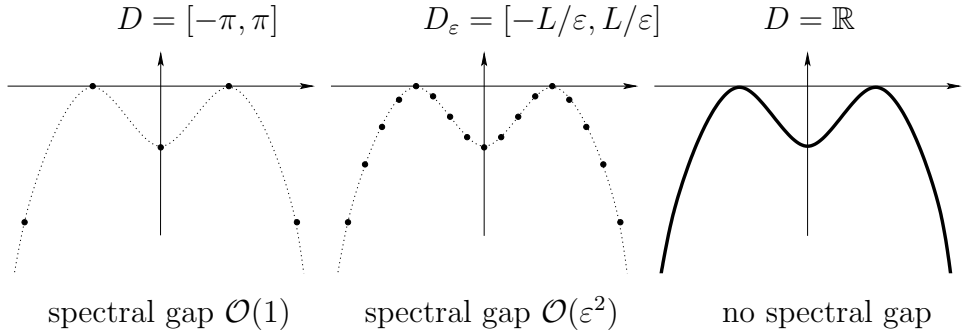


Fig. 1.1 Comparison between the spectrum of $-P(i\partial_x)$ for different domains. The dashed line is $k \mapsto -P(-k) = -(1 - k^2)^2$, with dots, provided the corresponding e^{ikx} is an eigenfunction of $-P(i\partial_x)$.

Rayleigh-Bénard problem (see [SH77]). Here

$$u(t, x) \in \mathbb{R}, \quad \text{for } t > 0, x \in D_\varepsilon = L\varepsilon^{-1} \cdot [-1, 1]$$

fulfils

$$\partial_t u = -P(i\partial_x)u + \varepsilon^2 \nu u - u^3 + \varepsilon^{\frac{3}{2}} \xi \tag{1.14}$$

subject to periodic boundary conditions. Note that we prescribe a scaling between the noise strength and the distance from bifurcation, that differs from the one used in the bounded domain case.

The linear operator is given by

$$P(\zeta) = (1 - \zeta^2)^2.$$

The complex eigenfunctions of the linear operator $P(i\partial_x)$ are $x \mapsto \exp\{ik\varepsilon\pi x/L\}$ with corresponding eigenvalue $P(k\varepsilon\pi/L)$ for $k \in \mathbb{Z}$. For simplicity, let ξ be space-time white noise in the following formal calculation. We rely on scaling properties for the noise, which are not that easy to formulate for coloured noise. See also Section 4.2. To be more precise, we use that ξ and $\hat{\xi}$ are versions of the same noise, when we define

$$\hat{\xi}(T, X) = \varepsilon^{-3/2} \xi(T\varepsilon^{-2}, X\varepsilon^{-1}). \tag{1.15}$$

We expect a linear instability at $e^{\pm ix}$, as $P(\pm 1) = 0$ and $P(x) > 0$ for $x \neq \pm 1$, but due to the boundedness of the domain $e^{\pm ix}$ is in general not an eigenfunction. The nearest eigenfunction is $e^{i\rho_c(\varepsilon/L)x}$, where

$$\rho_c(\varepsilon/L) := \frac{\varepsilon\pi}{L} \cdot \left[\frac{L}{\varepsilon\pi} \right],$$

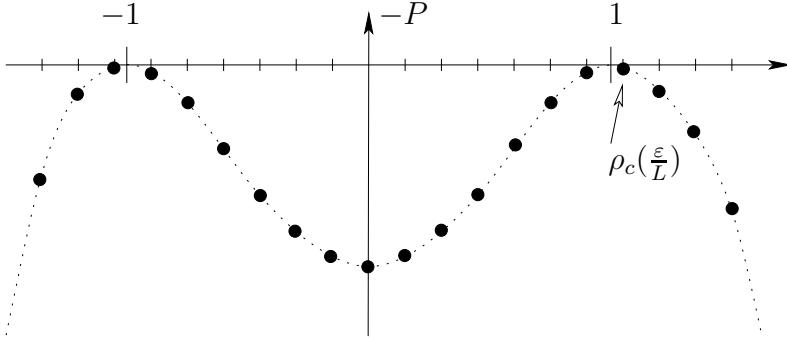


Fig. 1.2 Sketch of the eigenvalue curve for large domains of size $\mathcal{O}(\varepsilon^{-1})$. Note that $e^{\pm ix}$ is in general not an eigenfunction. The eigenfunctions with eigenvalue nearest to 0 are $e^{\pm i\rho_c x}$.

and $[\frac{L}{\varepsilon\pi}]$ is the nearest integer to $L/\varepsilon\pi$. See Figure 1.2, where we give a sketch of the eigenvalue curve of $-P(i\partial_x)$. Note that obviously

$$\left| \frac{1 - \rho_c(\varepsilon/L)}{\varepsilon} \right| \leq \frac{\pi}{2} \cdot \frac{1}{L}.$$

Leaving out the error term for simplicity of presentation, we make the following ansatz:

$$u(t, x) = \varepsilon A(\varepsilon^2 t, \varepsilon x) e^{i\rho_c(\varepsilon/L)x} + \varepsilon^3 B(\varepsilon^2 t, \varepsilon x) e^{3i\rho_c(\varepsilon/L)x} + c.c. , \tag{1.16}$$

where $c.c.$ denotes the complex conjugate. Here $\Re(z) = \frac{1}{2}(z + c.c.)$ denotes the real part of $z \in \mathbb{C}$. The term involving B just simplifies the formal calculation. It has no impact on the final result. A similar idea was used in [KSM92] for the deterministic equation on the unbounded domain.

For the deterministic equation the effect of bounded but large domains was first discussed in [MSZ00]. They used the ansatz with $\rho_c \equiv 1$, and for the Swift-Hohenberg equation they obtain the following amplitude equation for A :

$$\partial_T A = 4\partial_X^2 A + \nu A - 3|A|^2 A , \tag{1.17}$$

but subject to ε -dependent boundary conditions on $[-L, L]$, which take into account that the ansatz as a solution should be L/ε -periodic, which $e^{\pm ix}$ is in general not. We use the other ansatz (1.16) to see that all ε -dependent terms in the amplitude equation are actually uniformly small in L , and they vanish for $L \in \varepsilon\pi\mathbb{N}$. Moreover, due to (1.16) the formal calculation yields for the amplitude equation standard periodic boundary conditions, which are more familiar.

Plugging the ansatz (1.16) into (1.14) and using

$$-P(i\partial_x)[f e^{ikx}] = [-P(i\partial_x - k)f] e^{ikx} ,$$

which is easy to verify, we get in lowest order (which is ε^3)

$$\begin{aligned} \partial_T A e^{i\rho_c x} + c.c. = & \left[4\partial_X^2 A - 4i \frac{1 - \rho_c^2}{\varepsilon} \partial_X A - \frac{(1 - \rho_c^2)^2}{\varepsilon^2} A + \nu A \right] \cdot e^{i\rho_c x} \\ & - (1 - 9\rho_c^2)^2 B e^{3i\rho_c x} - A^3 e^{3i\rho_c x} - 3A|A|^2 e^{i\rho_c x} + c.c. + \hat{\xi}. \end{aligned}$$

Here $\hat{\xi}$, as in (1.15), is a rescaled version of ξ . In order to get rid of the terms depending on e^{3ix} , we choose

$$B = -(1 - 9\rho_c^2)^{-2} A^3.$$

Furthermore, we use

$$1 - \rho_c^2 = 2(1 - \rho_c) + \mathcal{O}\left(\frac{\varepsilon^2}{L^2}\right).$$

Finally, we have to rewrite the noise. We will see below, that one can define a complex-valued space-time white noise η with law independent of ε , such that we obtain the following amplitude equation

$$\partial_T A = 4\left(\partial_X A - 2i \frac{1 - \rho_c}{\varepsilon}\right)^2 A + \nu A - 3A|A|^2 + \eta, \tag{1.18}$$

subject to periodic boundary conditions on $[-L, L]$. Note that $(1 - \rho_c)/\varepsilon$ is small for large L uniformly with respect to ε . Moreover, it vanishes for $L \in \varepsilon\pi\mathbb{N}$.

Discussion on the Noise

On the formal level there is no trivial argument available, why we can rewrite the noise ξ in order to obtain η . Furthermore, one has to be extremely careful with the formal analysis at this point. For example, consider the argument that there is a complex valued space-time white noise $\tilde{\eta}$ such that

$$\hat{\xi}(T, X) = \frac{1}{\sqrt{2}} [\tilde{\eta}(T, X) e^{ix} + c.c.] .$$

This is obviously true, by calculation the corresponding correlation functions. But the rigorous proofs of Chapter 4 show that this simple argument is wrong. Using $\tilde{\eta}$ we obtain a factor $1/\sqrt{2}$ in front of the noise in the amplitude equation, but this is not present in the rigorous result.

The main fault is that space-time white noise is of order $\mathcal{O}(\varepsilon)$, when we restrict it to modes in Fourier space of order larger than ε^{-1} . Especially, if we look at the linear equation, then it is possible to argue that modes are actually small, once their wavenumber is far away from the instability $\pm\rho_c$. Thus we neglect all modes that are sufficiently far away. Note that the instability $\pm\rho_c$ corresponds to modes e^{ikx} with wavenumber k of order $\pm\rho_c\pi/\varepsilon L$.

Thus after neglecting the modes far away from the instability, the right decomposition of ξ is now to separate it into two parts. One contains all the remaining

Fourier modes with positive wavenumber, and the other one all modes with negative wave-numbers. Then we can shift these modes in Fourier space by $\pm\rho_c\pi/\varepsilon L$ by pulling a factor $e^{\pm i\rho_c x}$ out of the Fourier series of the noise. Now we put these transformed noise processes into the amplitude equation, or its complex conjugate version.

Finally, we fill up these noise processes with noise on high wave numbers, in order to get space-time white noise again. The fact that we can fill up the high modes of the noise is by no means obvious from the PDE. The main reason is that we have a strong linear damping on high modes present in the Ginzburg-Landau equation (1.18), which arises from the linear part. See Section 4.3 or [BHP05]. It will be easier to see later, when we later change to the mild formulation.

1.2 General Structure of the Approach

It is not the aim of this section to present rigorous results. Instead it highlights the key steps in a non-technical way. For all our results in the stochastic case, the general method of proof already dates back to [BMPS01]. Furthermore it was already used for amplitude equations for deterministic equations, for instance, in [KSM92] and [Sch94].

For simplicity of presentation we focus on the case of bounded domains. The case of large or unbounded domains is similar, but it exhibits many additional technical difficulties. Furthermore, we stick to cubic nonlinearities with additive noise. This was discussed in Section 1.1.1. The method of proof for other types of equations is very similar, only the formulation and the technical details differ.

Due to the lack of regularity, we cannot proceed analogous to the deterministic setting. This is one of the main issues for SPDEs, as the approach for deterministic PDEs relies on bounds for solutions of the amplitude equations in spaces with sufficiently high regularity. But especially on large domains for SPDEs this is never the case. See Section 4.3 or Remark 4.1.

In order to give SPDEs like (1.2) a meaning, we use the concept of mild solutions. These are stochastic processes with continuous paths that fulfil the following variation of constants formula

$$u(t) = e^{tL}u(0) + \int_0^t e^{(t-\tau)L}[\varepsilon^2 Au + \mathcal{F}(u)](\tau)d\tau + \varepsilon^2 W_L(t) \quad (1.19)$$

for $t \leq t^*$, where $t^* > 0$ is some stopping time. Here $\{e^{tL}\}_{t \geq 0}$ denotes the semigroup of operators generated by the differential operator L . For a detailed definition see [Paz83; Hen81; Lun95] or Section 2.5.1. The main point here is that $w(t) = e^{tL}w_0$ solves $\partial_t w = Lw$ with $w(0) = w_0$, and thus $\partial_t e^{tL} = Le^{tL}$.

For the definition of the stochastic convolution

$$W_L(t) = \int_0^t e^{(t-\tau)L} dQW(\tau), \quad t \geq 0$$

see [DPZ92]. Formally differentiating (1.19) yields immediately that $u(t)$ solves (1.2).

Here $\partial_t QW = \xi$ in a generalised sense, and W is some cylindrical Wiener process in some Hilbert space (see Assumption 2.8 and the discussion below that). For the connection between the noise ξ and Q -Wiener processes see [Blö05b]. For a different approach using the Brownian sheet and an explicit representation of the semigroup e^{tL} via the Green function see [Wal86].

We use the projection P_c onto the kernel \mathcal{N} of L and $P_s = I - P_c$, which were defined before (cf. Section 1.1.1). Now we project the equation to \mathcal{N} and \mathcal{S} .

Definition 1.1 We call $u_s(t) = P_s u(t) \in \mathcal{S}$ *fast modes*, as they are subject to a deterministic exponential decay on a time-scale of order $\mathcal{O}(1)$. Moreover, $u_c(t) = P_c u(t) \in \mathcal{N}$ are the *slow modes*, as they change only on the slow time-scale $T = \varepsilon^2 t$.

For simplicity we assume that P_c , and hence P_s , commutes with L and therefore with e^{tL} , too. Moreover, $e^{tL} P_c = P_c = P_c e^{tL}$. Projecting (1.19), we derive

$$u_c(t) = u_c(0) + \int_0^t [\varepsilon^2 A_c(u_c + u_s) + \mathcal{F}_c(u_c + u_s)](\tau) d\tau + \varepsilon^2 P_c QW(t), \quad (1.20)$$

and

$$u_s(t) = e^{tL} u_s(0) + \int_0^t e^{(t-\tau)L} [\varepsilon^2 A_s(u_c + u_s) + \mathcal{F}_s(u_c + u_s)](\tau) d\tau + \varepsilon^2 P_s QW_L(t). \quad (1.21)$$

If we now use the scaling

$$u_c = \mathcal{O}(\varepsilon) \quad \text{and} \quad u_s = \mathcal{O}(\varepsilon^2),$$

as indicated in the ansatz (1.4), then we immediately see that (1.20) decouples from (1.21). Here we just neglect all terms of higher order and rescale to the slow time-scale. Again we derive the amplitude equation (1.5), but now in integrated form.

The Impact of Noise

We see in (1.20) and (1.21) that the influence of the noise on the slow and fast component is very different. On the time-scale t of order 1 both are ε^2 , but this changes significantly, when using the slow time-scale $T = \varepsilon^2 t$. In (1.20) for the slow variable u_c , we see that the noise acts like some Brownian motion in \mathcal{N} , which is isomorphic to $\mathbb{R}^{\dim(\mathcal{N})}$. Due to the scaling properties of W it is possible to verify that

$$\sup_{t \in [0, T_0 \varepsilon^{-2}]} \|P_c QW(t)\| = \varepsilon^{-1} \cdot \sup_{T \in [0, T_0]} \|P_c QW(T)\| \quad (\text{in law}).$$

Hence,

$$\sup_{t \in [0, T_0 \varepsilon^{-2}]} \|P_c QW(t)\| = \mathcal{O}(\varepsilon^{-1}).$$

On the other hand, in (1.21) for the fast component u_s the noise enters as an infinite dimensional Ornstein-Uhlenbeck process. Here it is possible to verify using the celebrated factorisation method (see [DPZ92] or the proof of Lemma A.4) that

$$\sup_{t \in [0, T_0 \varepsilon^{-2}]} \|P_s W_L(t)\| = \mathcal{O}(\varepsilon^{-\kappa}) \quad \text{for arbitrary } \kappa > 0.$$

See for instance Theorem 5.1 of [BMPS01]. Note that this is not the optimal bound. The right scaling should be logarithmic in ε .

From the previous discussion we see that the noise acting on (1.21) is about one order of magnitude smaller than the noise acting on (1.20). Using the stability of the equations a long calculation leads to a priori estimates. These establish different scalings for u_c and u_s , which are the ones already used in the ansatz (1.4). This is the so called *attractivity result* (see Theorem 1.1), which justifies the ansatz for the formal calculation. Moreover, it provides the typical structure of the initial condition necessary to start an *approximation result*. This means a rigorous error bound between the true solution, and the approximation via amplitude equations. See Theorem 1.3.

1.2.1 Meta Theorems

The first result presented here is the attractivity. It justifies the scaling of ansatz (1.4) used for the formal derivation. It heavily relies on the structure of the equation. Sometimes we rely on global nonlinear stability and sometimes we only use linear stability on the non-dominant modes. A typical statement would be:

Theorem 1.1 (Attractivity) *There is a time $t_\varepsilon = \mathcal{O}(\ln(\varepsilon^{-1}))$ such that for all solutions u of (1.19) with initial conditions $u(0)$ of order $\mathcal{O}(\varepsilon)$ we have $u_s(t_\varepsilon) = \mathcal{O}(\varepsilon^2)$ and $u_c(t_\varepsilon) = \mathcal{O}(\varepsilon)$. This means the solution looks at the time t_ε like ansatz (1.4). To be more precise $u(t_\varepsilon) = \varepsilon a_\varepsilon + \varepsilon^2 \psi_\varepsilon$ with $a_\varepsilon \in \mathcal{N}$ and $\psi_\varepsilon \in \mathcal{S}$ both of order $\mathcal{O}(1)$.*

If we assume additionally global nonlinear stability for the equation, then there is a time $T_\varepsilon = \mathcal{O}(\varepsilon^{-2})$ such that $u(T_\varepsilon) = \mathcal{O}(\varepsilon)$ independent of the initial condition.

This theorem is rigorously stated in Theorems 2.7 or 2.8. We will give a detailed discussion of these results for multiplicative noise in Theorems 2.1 and 2.4 for cubic and quadratic nonlinearities. A sketch of the typical dynamics for the local attractivity result is given in Figure 1.3.

Remark 1.4 *Depending on the assumptions the statement $g_\varepsilon = \mathcal{O}(f_\varepsilon)$ can have two different meanings. Depending on the context, we either use that for all $p > 0$ there is a constant $C > 0$ such that $\mathbb{E}\|g_\varepsilon\|^p \leq C f_\varepsilon^p$ for all $\varepsilon \in (0, 1]$. This is typically*

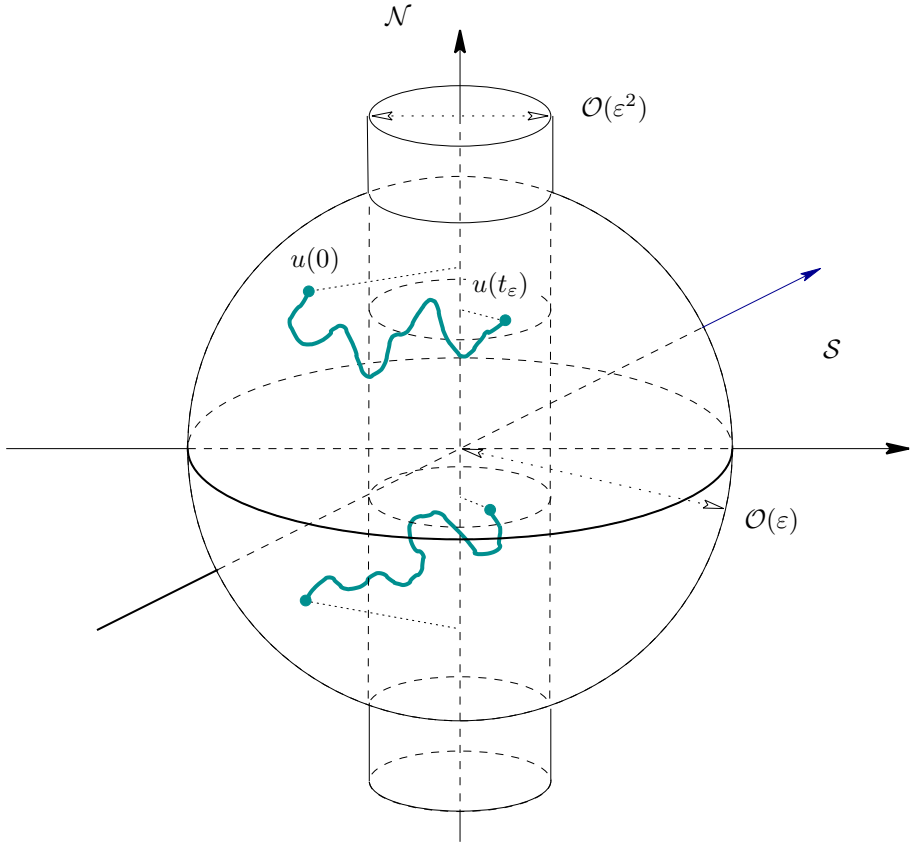


Fig. 1.3 Two typical trajectories for the local attractivity result from Theorem 1.1. For the global result relying on nonlinear stability the $\mathcal{O}(\varepsilon)$ -ball would be attractive, too.

only valid for nonlinear stable equations, where we can actually bound moments. In case of, for instance quadratic nonlinearities, where in general we do not have control on moments of solutions, we also use the somewhat weaker meaning that for some constant $C > 0$, we have $\mathbb{P}(\|g_\varepsilon\| \geq C f_\varepsilon)$ uniformly small for all $\varepsilon \in (0, 1]$. Sometimes we also give explicit convergence rates of this probability for $\varepsilon \rightarrow 0$.

For a solution a of (1.5) and ψ of (1.6) we define first the approximations εw_k of order k by

$$\varepsilon w_1(t) := \varepsilon a(\varepsilon^2 t) \quad \text{and} \quad \varepsilon w_2(t) := \varepsilon a(\varepsilon^2 t) + \varepsilon^2 \psi(t). \tag{1.22}$$

In our setting the *residual* of εw is defined by

$$\begin{aligned} \text{Res}(\varepsilon w)(t) &= -\varepsilon w(t) + e^{tL} \varepsilon w(0) + \varepsilon^2 W_L(t) \\ &\quad + \int_0^t e^{(t-\tau)L} [\varepsilon^3 A w + \mathcal{F}(\varepsilon w)](\tau) d\tau. \end{aligned} \tag{1.23}$$

In order to show that εw_j is a good approximation of a solution u of (1.19), the key step is to control the residual, which measures the distance of the approximation εw_j from being a solution. Obviously, $\text{Res}(\varepsilon w_j) = 0$, if and only if εw_j is a solution of (1.19).

Note that in general no additional information on the solution u is needed. We mainly rely on a priori estimates for the solutions of the amplitude equation, which are much easier to obtain. A typical statement is:

Theorem 1.2 (Residual) *For any $T_0 > 0$, and approximations given by (1.22) with initial conditions $a(0)$ and $\psi(0)$ of order $\mathcal{O}(1)$, we have*

$$\sup_{t \in [0, T_0 \varepsilon^{-2}]} \|\text{Res}(\varepsilon w_k)(t)\| = \mathcal{O}(\varepsilon^{1+k}).$$

A detailed discussion of the residual in the multiplicative noise case can be found in Theorem 2.2 for cubic and in Theorem 2.5 for quadratic nonlinearities.

Using the results for the residual, it is usually straightforward to derive the approximation result. Here we sometimes have to use additional assumptions on the nonlinearity. Again no bounds on solutions are necessary, but for simplicity we sometimes rely on them, when they are easy to establish.

Theorem 1.3 (Approximation) *For all $T_0 > 0$, all solutions u of (1.19), and all approximations given by (1.22) with $\|u(0) - \varepsilon w_k(0)\| = \mathcal{O}(\varepsilon^{1+k})$ we have*

$$\sup_{t \in [0, T_0 \varepsilon^{-2}]} \|u(t) - \varepsilon w_k(t)\| = \mathcal{O}(\varepsilon^{1+k}).$$

This result is rigorously stated in Theorems 2.9 and 2.10. A detailed discussion for multiplicative noise is given in Theorems 2.3 or 2.6. We sketch a typical trajectory from the approximation result in Figure 1.4. Note finally that the condition imposed on the initial conditions in Theorem 1.3 are usually trivial, in case we consider a solution given by the attractivity result.

1.3 Examples of Equations

In the literature there are numerous examples of equations where the abstract theorems do apply. In this section we focus mainly on additive noise. For instance, for cubic nonlinearities the well known Ginzburg-Landau equation (see [DE00] for a standard proof of existence of unique solutions)

$$\partial_t u = \Delta u + \nu u - u^3 + \sigma \xi$$

and the Swift-Hohenberg equation (see [CH93] for numerous references)

$$\partial_t u = -(\Delta + 1)^2 u + \nu u - u^3 + \sigma \xi$$

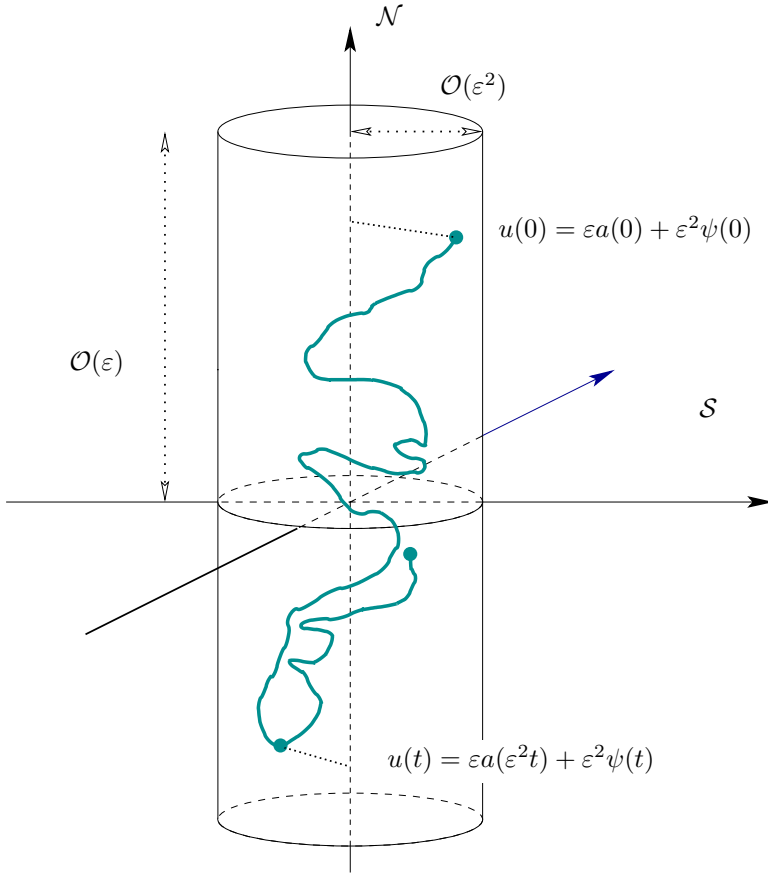


Fig. 1.4 A typical trajectory for the approximation result.

fall into the scope of our work, in case the parameters ν and σ are small and of comparable order of magnitude. Both equations are considered on bounded domains with suitable boundary conditions (e.g. periodic, Dirichlet, Neumann, etc.). The Swift-Hohenberg equation is a toy model for the convective instability in the Rayleigh-Bénard convection. A formal derivation of the equation from the Boussinesq approximation of fluid dynamics can be found in [SH77].

Another example arising in the theory of surface growth is

$$\partial_t u = -\Delta^2 u - \mu \Delta u + \nabla \cdot (|\nabla u|^2 \nabla u) + \sigma \xi, \tag{1.24}$$

subject to periodic boundary conditions and moving frame $\int_G u \, dx = 0$, where one rescales the mean growth of u out of the equation, in order to ensure a Poincaré type inequality. This model was first suggested by [LDS91]. The deterministic equation was rigorously treated in [KSW03]. For a review on surface growth see for example [BS95] or [HHZ95]. For this model we can consider $\mu = \mu_0 + \epsilon^2$ and $\sigma = \mathcal{O}(\epsilon^2)$,

where μ_0 is such that $L = -\Delta^2 u - \mu_0 \Delta u$ is a non-positive operator with non-zero kernel. We will see later on, that all examples presented up to now exhibit a stable nonlinearity in the sense of Assumption 2.2.

There are also numerous examples in the physics literature of equations with quadratic nonlinearities where our theory does apply. One example is the growth of rough amorphous surfaces. See for example [BGR02], [BG02b], or [BH04] and the references therein. The equation is of the type

$$\partial_t u = -\Delta^2 u - \mu \Delta u - \Delta |\nabla u|^2 + \sigma \xi, \quad (1.25)$$

for instance subject to periodic boundary conditions on $[0, L]^d$ and the moving frame condition. In the context of surface growth, this equation was first used in [SP94]. Later on it was used to model a special class of growth of amorphous surfaces (See for example [RML⁺00; RLH00] and the references therein). The local existence of unique mild solutions using standard fixed point methods is verified in [BG04] or for the deterministic equation in [SW05]. The uniqueness of global solutions is still open, even for $d = 1$. For the existence of global solutions with Markov property see [BFR06].

Other examples that fall into the scope of our work are first the Kuramoto-Sivashinsky equation

$$\partial_t u = -\Delta^2 u + \mu \Delta u + |\nabla u|^2 + \sigma \xi.$$

This was first used to model propagation of flames and recently as a model for surface growth via ion sputtering (cf. [CB95; FBK⁺02; LV05]). The second example is the Burgers' equation which was already discussed in Section 1.1.3. Another related model is the celebrated KPZ-equation.

But the probably most important example is the Rayleigh-Bénard problem (see for example [Get98; Wal97] or [CH93]) which is the paradigm of pattern formation in convection problems. In [Blö05a] we discuss in detail the amplitude equation for the stochastic two-dimensional Bénard problem in a strip, where a fluid is heated from below. The three dimensional problem in a box is treated similarly, but the notation is much more involved.

In the following denote by (v, w) the velocity field of the fluid in $(y, z) \in D := [0, 2\pi] \times [0, \pi]$, where z is the vertical direction. Hence, the fluid is heated at $z \equiv 0$. Let p be the pressure and θ the normalised temperature, which means that $\theta \equiv 0$ and $(v, w) \equiv 0$ is heat transport without motion.

In dimensionless quantities the governing Navier-Stokes and heat equation are given by (see e.g. [Get98] or [Wal97])

$$\partial_t(v, w) + ((v, w) \cdot \nabla)(v, w) = -\nabla p + (0, 1) \frac{R}{P} \theta + \Delta(v, w) \quad (1.26)$$

$$\partial_t \theta - v + ((v, w) \cdot \nabla) \theta = \frac{1}{P} \Delta \theta + \varepsilon^2 \xi \quad (1.27)$$

$$\operatorname{div}(v, w) = 0. \quad (1.28)$$

We suppose periodic boundary conditions in y both for θ and (v, w) . Moreover, $\partial_z v = w = \theta = 0$ for $z = 0$ and $z = \pi$. The noise ξ corresponds to fluctuations in the temperature. We could also incorporate fluctuations in the velocity field, but we neglect this for simplicity of presentation.

In order to rule out motion of the whole fluid in the y -direction, we suppose vanishing mean flux $\int_0^\pi v dz$. We use the following constants:

- R the Rayleigh number
- P the Prantl number
- $\rho = R/P$ the Reynolds number

The Rayleigh number is a dimensionless measure of the heat difference between top and bottom of the strip, while the Prantl number depends only on the properties of the fluid.

For a rigorous verification of the amplitude equation for the Bénard problem see [Blö05a]. We give a summary of all these results in Section 2.6. Note that in this case one of the major difficulties is that the linear operator L is not self-adjoint. However it exhibits a complete non-orthogonal basis of eigenfunctions. For simplicity of presentation, we do not focus on that technical point (cf. Assumption 2.1). For a detailed discussion see [Blö05a].

In this example L and A do not commute, and hence $P_c A \neq A P_c$, which is in contrast to most of the other examples stated above, where we have equality. This does not cause major technical difficulties, and we allow for quite general A in the abstract setting (cf. Assumptions 2.2 or 2.4).