Normal forms for stochastic differential equations

Ludwig Arnold
Institut für Dynamische Systeme
Universität Bremen
28334 Bremen, Germany
email:arnold@mathematik.uni-bremen.de

Peter Imkeller
Institut für Stochastik
Humboldt–Universität zu Berlin
10099 Berlin, Germany
email:imkeller@mathematik.hu-berlin.de

October 13, 1997

Abstract

We address the following problem from the intersection of dynamical systems and stochastic analysis: Two SDE $dx_t = \sum_{j=0}^{m} f_j(x_t) \circ dW_t^j$ and $dx_t = \sum_{j=0}^{m} g_j(x_t) \circ dW_t^j$ in $\mathbb{R}^d$ with smooth coefficients satisfying $f_j(0) = g_j(0) = 0$ are said to be smoothly equivalent if there is a smooth random diffeomorphism (coordinate transformation) $h(\omega)$ with $h(\omega, 0) = 0$ and $Dh(\omega, 0) = \mathrm{id}$ which conjugates the corresponding local flows,

$$\varphi(t, \omega) \circ h(\omega) = h(\theta_t \omega) \circ \psi(t, \omega),$$

where $\theta_t \omega(s) = \omega(t+s) - \omega(t)$ is the (ergodic) shift on the canonical Wiener space. The normal form problem for SDE consists in finding the "simplest–possible" member in the equivalence class of a given SDE, in particular in giving conditions under which it can be linearized ($g_j(x) = Df_j(0)x$).

We develop a mathematically rigorous normal form theory for SDE which justifies the engineering and physics literature on that problem. It is based on the multiplicative ergodic theorem and uses a uniform (with respect to a spatial parameter) Stratonovich calculus which allows the handling of non–adapted initial values and coefficients in the stochastic version of the cohomological equation. As a by–product, we prove a general theorem on the existence of a stationary solution of an anticipative affine SDE.

The study of the Duffing–van der Pol oscillator with small noise concludes the paper.

Key words and phrases: stochastic differential equation; stochastic normal form; stochastic cohomological equation; anticipative calculus; random dynamical system; multiplicative ergodic theory; affine stochastic differential equation; noisy Duffing–van der Pol oscillator.

AMS 1990 subject classification: primary 60 H 10, 34 F 05; secondary 58 F 11, 93 E 03.

1 Introduction

Normal form theory was initiated by Poincaré in 1892 and is a technique of fundamental importance for dynamical systems, in particular for bifurcation theory. It aims at simplifying a nonlinear deterministic or random dynamical system in the neighborhood of a reference solution by means of a smooth change of coordinates. In this paper, we consider dynamical systems in $\mathbb{R}^d$, and the reference solution is assumed to be the fixed point $x = 0$.

We briefly recall those facts from deterministic normal form theory relevant to the stochastic case. For recent presentations of the deterministic theory see e.g. Anosov and V. I. Arnold [1], Vanderbauwhede [23], or Katok and Hasselblatt [14].

Two smooth vector fields f and g in $\mathbb{R}^d$ with $f(0) = g(0) = 0$ are called smoothly equivalent, if the local flows φ and ψ generated by $\dot{x} = f(x)$ and $\dot{x} = g(x)$ are smoothly equivalent, i. e., if there is a $\mathcal{C}^\infty$ diffeomorphism (also called coordinate transformation)[1] $h : \mathbb{R}^d \to \mathbb{R}^d$ with $h(0) = 0$ for which

$$\varphi(t) \circ h = h \circ \psi(t) \quad \text{locally.}$$

Differentiating this with respect to t gives the equivalent infinitesimal form

$$f \circ h = Dh\, g. \tag{1.1}$$

Normal form theory for vector fields now seeks an h for which g is "simplest–possible", in other words: we look for the "simplest–possible" element in the $\mathcal{C}^\infty$ equivalence class of f.

Since (1.1) implies $B = Dh(0)^{-1} A Dh(0)$, where $A = Df(0)$ and $B = Dg(0)$ are the Jacobians of f and g at $x = 0$, respectively, and since a linear mapping is considered simplest–possible if its matrix representation is in Jordan canonical form (which requires the choice of a basis in $\mathbb{R}^d$) it is reasonable to assume

[1] Deterministic as well as stochastic normal form theory makes statements about *germs* of $\mathcal{C}^\infty$ diffeomorphisms or vector fields, i. e., equivalence classes of $\mathcal{C}^\infty$ diffeomorphisms or vector fields which coincide in a neighborhood of 0. However, for ease of presentation we ignore this point.

without loss of generality that A is in Jordan canonical form, and h is *near–identity*, i.e., $Dh(0) = \mathrm{id}$. Then $B = A$ for any such h. The normal form of f at the fixed point 0 is thus the natural generalization of the Jordan canonical form to the nonlinear case.

The ultimate aim of normal form theory is the linearization of f, i.e., to find an h such that $g(x) = Ax$, or $f \sim A$, against which there are obstructions in the form of "resonances".

Formal normal form theory now expands f, g and h into formal Taylor series at 0, $f(x) \sim Ax+\sum_{n=2}^{\infty} f_n(x)$, $g(x) \sim Ax+\sum_{n=2}^{\infty} g_n(x)$, $h(x) \sim x+\sum_{n=2}^{\infty} h_n(x)$, inserts those expansions into equation (1.1), equates coefficients of order $n \geq 2$ (which are homogeneous polynomials), and tries to successively determine h_n such that g_n is zero, or at least "simplest–possible". For this purpose we introduce spaces of homogeneous polynomials.

Let for $n \in \mathbb{N}$

$$\mathbb{N}_n^d := \{\tau = (\tau_1, \ldots, \tau_d) \in (\mathbb{Z}^+)^d : |\tau| := \sum_{i=1}^{d} \tau_i = n\}$$

be the set of multi–indices of length n. Denote by $x^\tau := x_1^{\tau_1} x_2^{\tau_2} \cdots x_d^{\tau_d}$ the scalar monomial in d variables of degree $|\tau| = n$. Then

$$H_{n,d} := H_{n,d}(\mathbb{R}^d) = \{f = \sum_{\tau \in \mathbb{N}_n^d} x^\tau f_\tau : f_\tau \in \mathbb{R}^d\}$$

is the vector space of homogeneous polynomials of degree n in d variables with values in $\mathbb{R}^d$. We also write $|\tau| = n$ for $\tau \in \mathbb{N}_n^d$. Observe that

$$\Delta := \Delta(n,d) = \#\mathbb{N}_n^d = \binom{n+d-1}{n},$$

so that

$$D = \dim H_{n,d}(\mathbb{R}^d) = \Delta \cdot d,$$

in particular

$$\Delta = \dim H_{n,d}(\mathbb{R}^1).$$

A basis $F = (u_1, \ldots, u_d)$ of $\mathbb{R}^d$ and the basis $(x^\tau)_{|\tau|=n}$ of $H_{n,d}(\mathbb{R}^1) \cong \mathbb{R}^\Delta$ give a basis $(x^\tau F) := (x^\tau u_i)_{i=1,\ldots,d;|\tau|=n}$ of $H_{n,d}(\mathbb{R}^d)$, and

$$H_{n,d} \ni f = \sum_{|\tau|=n} \sum_{i=1}^{d} f_{i,\tau} x^\tau u_i \cong k_F(f) = (k_F(f_\tau))_{|\tau|=n} = (f_{i,\tau}) \in \mathbb{R}^D$$

(column vectors, ordered lexicographically) identifies $H_{n,d}$ with $\mathbb{R}^D$, where k_F is the mapping which assigns F coordinates (respectively $(x^\tau F)$ coordinates) to an element of $\mathbb{R}^d$ (respectively $H_{n,d}$). We can identify $H_{n,d}(\mathbb{R}^d)$ with the tensor product of $H_{n,d}(\mathbb{R}^1)$ and $\mathbb{R}^d$,

$$H_{n,d}(\mathbb{R}^d) = H_{n,d}(\mathbb{R}^1) \otimes \mathbb{R}^d \cong \mathbb{R}^\Delta \otimes \mathbb{R}^d,$$

where the above choice of bases induces the basis with elements $x^\tau \otimes u_i$ in $H_{n,d}(\mathbb{R}^d)$, and the isomorphism induced by the coordinate mappings maps this basis to the standard basis $f_j \otimes e_i \in \mathbb{R}^\Delta \otimes \mathbb{R}^d$.

Inserting the above Taylor expansions into (1.1) and equating coefficients gives the **cohomological equations**

$$g_n(x) = (\mathrm{ad}_n A)h_n(x) + k_n(x), \quad n \geq 2,$$

where the linear operator

$$\mathrm{ad}_n A : H_{n,d} \to H_{n,d}, \quad h_n \mapsto (\mathrm{ad}_n A)h_n(x) := Ah_n(x) - Dh_n(x)Ax. \tag{1.2}$$

is called **cohomological operator**. Further,

$$k_n = f_n + P_n(f_k, g_k, h_k, 2 \leq k \leq n-1),$$

where

$$P_n = T_n\{S_{n-1}(f - A) \circ S_{n-1}(h - \mathrm{id}) - DS_{n-1}(h - \mathrm{id})S_{n-1}(g - A)\}, \tag{1.3}$$

is a polynomial of the lower order terms f_k, g_k, h_k, $2 \leq k \leq n-1$, of f, g, and h. Here $T_n(f)$ denotes the term of order n of f, $S_n(f)$ is the n–jet (Taylor polynomial of order n) of f. For example, $k_2 = f_2$, $k_3 = f_3 - Dh_2\, g_2$, $k_4 = f_4 + f_2 \circ g_2 - Dh_2\, g_3 - Dh_3\, g_2$, etc.

The operator $\mathrm{ad}_n A$ depends linearly on the entries of A, and its $D \times D$ matrix representation is

$$\mathrm{ad}_n A = I_1 \otimes A - T(A)_n \otimes I_2, \tag{1.4}$$

where I_1 and I_2 are $\Delta \times \Delta$ and $d \times d$ unit matrices, respectively, $T(A)_n$ is the $\Delta \times \Delta$ matrix describing the linear mapping on $H_{n,d}(\mathbb{R}^1) \cong \mathbb{R}^\Delta$ given by

$$h = \sum_{|\tau|=n} h_\tau x^\tau \mapsto \sum_{|\tau|=n} \sum_{j,k=1}^{d} h_\tau \frac{\partial(x^\tau)}{\partial x_j} a_{jk} x_k =: T(A)_n(h),$$

and $A \otimes B = (a_{ij}B)$ is the Kronecker product of A and B.

1.1. Remark. One easily checks that if A is diagonal, then so is $\mathrm{ad}_n A$ for any $n \geq 2$. ∎

We will now present a treatment of the normal form problem for stochastic differential equations (SDE). It deals with the basic question of how many "essentially different" (modulo a $\mathcal{C}^\infty$ near–identity coordinate transformation) SDE exist.

Our effort is motivated by the fact that many dynamical systems in engineering and physics are perturbed by noise. The desire to simplify those systems prompted numerous publications (see, e. g., Coullet, Elphick and Tirapegui [12], Nicolis and Nicolis [15], Schöner and Haken [20][21], Sri Namachchivaya and Lin [22], and the references therein). All those authors work with a smallness parameter multiplying the noise terms, hence obtaining the stochastic normal form as a small perturbation of the deterministic one.

Normal form theory without any smallness assumption was developed for random diffeomorphisms by Arnold and Xu in [8], and for random differential equations in [9] and [10].

However, normal form theory for SDE has so far defied rigorous analysis for the following technical reason: When solving the stochastic version of the cohomological equation one is forced to consider coefficients as well as solutions which at time t are not adapted to the natural forward filtration $\mathcal{F}_{-\infty}^t$ of the (two–sided) Wiener process. This fact, which is a "conflict" between (multiplicative) ergodic theory and classical stochastic analysis, has been clearly seen but not rigorously handled by the pioneers of stochastic normal form theory quoted above. It is the aim of this paper to dissolve this "conflict" and present a rigorous treatment on the basis of multiplicative ergodic theory and of a uniform Stratonovich and anticipative calculus. We also show by way of an example the usefulness of the stochastic normal form.

2 The stochastic cohomological equation

We consider the SDE

$$dx_t = \sum_{j=0}^{m} f_j(x_t) \circ dW_t^j, \quad t \in \mathbb{R}, \quad f_j(0) = 0, \tag{2.1}$$

in $\mathbb{R}^d$, where as usual, dW_t^0 stands for dt, $f_j \in \mathcal{C}^\infty$ for $0 \leq j \leq m$, $(\Omega, \mathcal{F}, \mathbb{P})$ is the canonical two–sided Wiener space, $W_t = (W_t^1, \ldots, W_t^m)$, $t \in \mathbb{R}$, the canonical two–sided Wiener process, $\mathcal{F}_{-\infty}^t$ the σ–algebra generated by W_s, $s \leq t$, completed by $\mathbb{P}$–null sets, and $\theta_t\omega(s) := \omega(t+s) - \omega(t)$, $s, t \in \mathbb{R}$. In the following ω will denote an arbitrary element of Ω, whenever it appears in the argument of a random variable.

Equation (2.1) uniquely generates a local $\mathcal{C}^\infty$ random dynamical system (RDS) (or cocycle) φ over the ergodic dynamical system $(\Omega, \mathcal{F}, \mathbb{P}, (\theta_t)_{t\in\mathbb{R}})$, (see Arnold and Scheutzow [6], Arnold [2] (Theorem 2.3.40), for a survey see Arnold [3]). The domain $D(t, \omega)$ and range $R(t, \omega)$ of $\varphi(t, \omega) : D(t, \omega) \to R(t, \omega)$ are neighborhoods of 0. How many different such local RDS do exist modulo a smooth random conjugacy?

The linear cocycle $\Phi_1(t,\omega) := \Phi(t,\omega) = D\varphi(t,\omega,0)$ on $T_0\mathbb{R}^d \cong \mathbb{R}^d$ is then generated by the linearized SDE

$$dv_t = \sum_{j=0}^{m} A_j v_t \circ dW_t^j, \quad A_j := Df_j(0), \quad 0 \le j \le m. \tag{2.2}$$

The multiplicative ergodic theorem holds for (2.2) (without any further integrability assumptions, see [2], section 6.2), giving the Lyapunov spectrum

$$\Sigma(\Phi_1) := \Sigma(\theta, A_0, \dots, A_m, W) = \{\Lambda_1 \ge \dots \ge \Lambda_d\},$$

and the splitting $\mathbb{R}^d = E_1(\omega) \oplus \dots \oplus E_p(\omega)$ which is invariant, $\Phi(t,\omega)E_i(\omega) = E_i(\theta_t\omega)$, $t \in \mathbb{R}$, and in which the different Lyapunov exponents $\lambda_1 > \dots > \lambda_p$ from the spectrum are realized as exponential growth rates forward and backward in time,

$$v \in E_i(\omega) \setminus \{0\} \iff \lim_{t\to\pm\infty} \frac{1}{t} \log \|\Phi(t,\omega)v\| = \lambda_i,$$

and for which $\dim E_i(\omega) = d_i$, d_i the multiplicity of λ_i, $1 \le i \le p$.

2.1. Definition (Random coordinate transformation). A measurable mapping $h : \Omega \times \mathbb{R}^d \to \mathbb{R}^d$ is called **(near–identity) random coordinate transformation** if

(i) $h(\omega,\cdot) \in \mathrm{Diff}^{\infty}(\mathbb{R}^d)$,
(ii) $h(\omega,0) = 0$,
(iii) $Dh(\omega,0) = \mathrm{id}$.

The diffeomorphism $h(\omega,\cdot)$ is also denoted by $h(\omega)$. ∎

We now conjugate the local RDS φ generated by (2.1) with another local RDS ψ by means of a random coordinate transformation,

$$\varphi(t,\omega) \circ h(\omega) = h(\theta_t\omega) \circ \psi(t,\omega) \quad \text{(locally)}, \tag{2.3}$$

where ψ is generated by an SDE

$$dx_t = \sum_{j=0}^{m} g_j(\theta_t\cdot, x_t) \circ dW_t^j, \tag{2.4}$$

and h is chosen such that the SDE (2.4) makes sense, and its coefficients g_j are "as simple as possible", the ultimate aim being linearization, i. e.,

$$dx_t = \sum_{j=0}^{m} A_j x_t \circ dW_t^j.$$

It is important to realize that, while the diffeomorphisms $\varphi(t,\omega)$ and $\psi(t,\omega)$ map the fiber over ω to the fiber over $\theta_t\omega$, the coordinate transformation h is "static" and maps each fiber to itself, hence the $h(\omega)$ on the left-hand side , and the $h(\theta_t\omega)$ on the right-hand side of (2.3).

Although it will turn out that the transformation $\bar{h}(t,\cdot,x) = h(\theta_t\cdot,x)$ to be applied at time t is in general *not* adapted to the filtration of W at t, we proceed formally (and justify later): Applying the Stratonovich lemma to (2.3) gives

$$d\varphi_t = dh_t + Dh(\theta_t\cdot,\psi_t)\circ d\psi_t, \tag{2.5}$$

where dh_t denotes the t–differential of $h(\theta_t\cdot,x)$, and $d\varphi_t$ the t–differential of $\varphi(t,\cdot)$, etc. Inserting the differentials of φ_t and ψ_t into (2.5) yields

$$\sum_{j=0}^{m} f_j(h(\theta_t\cdot,x_t)\circ dW_t^j = dh_t + \sum_{j=0}^{m} Dh(\theta_t\cdot,x_t)g_j(\theta_t\cdot,x_t)\circ dW_t^j, \tag{2.6}$$

equivalently

$$\begin{aligned} dx_t &= \sum_{j=0}^{m} g_j(\theta_t\cdot,x_t)\circ dW_t^j \qquad (2.7)\\ &= -Dh(\theta_t\cdot,x_t)^{-1}dh_t + \sum_{j=0}^{m} Dh(\theta_t\cdot,x_t)^{-1} f_j(h(\theta_t\cdot,x_t))\circ dW_t^j. \end{aligned}$$

This is an equation for h and the g_j, where the choice of h is made such that the g_j are simplest–possible.

We now make the simplifying assumption that we choose the canonical basis in $\mathbb{R}^d$. In other words, we leave the linearized SDE (2.2) untouched and refrain from making it "as simple as possible" by choosing an appropriate (necessarily non–adapted) random basis. See the treatment of the resonant case in subsection 4.1.

As in the deterministic case, we make the formal Taylor series Ansatz

$$f_j(x) \sim A_j x + \sum_{n=2}^{\infty} f_{j,n}(x), \quad j=0,\dots,m,$$

$$g_j(\omega,x) \sim A_j x + \sum_{n=2}^{\infty} g_{j,n}(\omega,x), \quad j=0,\dots,m,$$

$$h(\omega,x) \sim x + \sum_{n=2}^{\infty} h_n(\omega,x),$$

where $f_{j,n}\in H_{n,d}$, while $g_{j,n}(\cdot)$ and $h_n(\cdot)\in \mathcal{H}_{n,d}$ are $H_{n,d}$–valued random variables.

Plugging this into equation (2.6) and equating coefficients yields (calculations are as in the deterministic case) an identity for the linear part, and for $n \geq 2$

$$\begin{aligned}\sum_{j=0}^{m} g_{j,n}(\theta_t\cdot) \circ dW_t^j &= \sum_{j=0}^{m} (\mathrm{ad}_n A_j) h_n(\theta_t\cdot) \circ dW_t^j - dh_n(\theta_t\cdot) \\ &\quad + \sum_{j=0}^{m} k_{j,n}(\theta_t\cdot) \circ dW_t^j, \end{aligned} \tag{2.8}$$

where $\mathrm{ad}_n A_j : H_{n,d} \to H_{n,d}$ is the linear operator defined in (1.2) by $h_n \mapsto (\mathrm{ad}_n A_j)h_n(x) := A_j h_n(x) - Dh_n(x) A_j x$, in matrix form $\mathrm{ad}_n A_j = I_1 \otimes A_j - T(A_j)_n \otimes I_2$ on $H_{n,d}(\mathbb{R}^d) \cong H_{n,d}(\mathbb{R}^1) \otimes \mathbb{R}^d$, and

$$k_{j,n} = f_{j,n} + P_n(f_{j,k}, g_{j,k}, h_k,\ 2 \leq k \leq n-1), \quad j = 0, \ldots, m. \tag{2.9}$$

Here P_n is a deterministic polynomial of the lower order terms of $f_{j,k}$, $g_{j,k}$ and h_k, $2 \leq k \leq n-1$, which is independent of j and explicitly given in (1.3).

We now define the **stochastic cohomological operator** by

$$dL_n(h_n) := dh_n - \sum_{j=0}^{m} (\mathrm{ad}_n A_j) h_n \circ dW_t^j, \tag{2.10}$$

acting on those $H_{n,d}$–valued stationary stochastic processes $h_n(\theta_t\omega)$ for which the expression in (2.10) makes sense. With this definition, (2.8) turns into the system of **stochastic cohomological equations** (suppressing the $(\theta_t\omega)$ argument)

$$dL_n(h_n) = \sum_{j=0}^{m} (k_{j,n} - g_{j,n}) \circ dW_t^j =: dK_n - dG_n, \tag{2.11}$$

which have to be solved successively for $n = 2, \ldots$, for dK_n known from previous steps, and h_n choosen to make dG_n simplest–possible. The most desirable choice is $dG_n \equiv 0$, resulting in the task of solving the cohomological equations

$$dL_n(h_n) = dK_n, \quad n \geq 2,$$

or, equivalently, looking for stationary solutions of the affine SDE

$$dh_n = \sum_{j=0}^{m} ((\mathrm{ad}_n A_j) h_n + k_{j,n}(\theta_t\cdot)) \circ dW_t^j, \quad n \geq 2, \tag{2.12}$$

where $k_{j,n}$ depends on the solutions of (2.12) of lower order $2 \leq k \leq n-1$. We have $k_{j,2} = f_{j,2}$ deterministic, but for $n \geq 3$, $k_{j,n}$ is typically random and not adapted to the filtration of W.

Let Φ_n be the linear cocycle on $H_{n,d}$ generated by the SDE

$$dv_t = \sum_{j=0}^{m} (\mathrm{ad}_n A_j) v_t \circ dW_t^j, \quad n \geq 2. \tag{2.13}$$

The multiplicative ergodic theorem holds for Φ_n (again without additional assumptions) and gives the spectrum

$$\Sigma(\Phi_n) = \{\Lambda_i - (\Lambda, \tau) : \Lambda_i \in \Sigma(\Phi_1), |\tau| = n\},$$

where $(\Lambda, \tau) := \sum_{i=1}^{d} \Lambda_i \tau_i$. This can be deduced from the form of $\mathrm{ad}_n A_j$ given in (1.4) and the following facts: If the linear cocycles Ψ_1 and Ψ_2 are generated by the linear SDE $d\Psi_i = \sum_{j=0}^{m} A_j^{(i)} \Psi_i \circ dW_t^j$, $i = 1, 2$, then the linear cocycle $\Psi_1 \otimes \Psi_2$ is generated by the linear SDE

$$d(\Psi_1 \otimes \Psi_2)_t = \sum_{j=0}^{m} (A_j^{(1)} \otimes I_2 + I_1 \otimes A_j^{(2)})(\Psi_1 \otimes \Psi_2)_t \circ dW_t^j,$$

the spectrum of $\Psi_1 \otimes \Psi_2$ is

$$\Sigma(\Psi_1 \otimes \Psi_2) = \Sigma(\Psi_1) + \Sigma(\Psi_2)$$

([2], Theorem 5.4.2), and, finally, the spectrum of the cocycle generated by $dv_t = \sum_{j=0}^{m} (-T(A_j)_n) v_t \circ dW_t^j$ is $\{-(\Lambda, \tau) : |\tau| = n\}$ (see Arnold and Xu [10], Theorem 3.1).

It turns out that we have a chance of finding a (unique) stationary solution of (2.12) provided the linear SDE (2.13) is hyperbolic, i.e., $0 \notin \Sigma(\Phi_n)$. We call the linear cocycle $\Phi = \Phi_1$ generated by (2.2) **nonresonant of order n** if $0 \notin \Sigma(\Phi_n)$, **resonant of order n** otherwise, $n \geq 2$.

3 The nonresonant case

The crucial property of the stationary processes $k_{j,n}(\theta_t \cdot)$ in the cohomological equation (2.12) to ensure the existence of a stationary solution in case the linear SDE (2.13) is hyperbolic, and which is in fact inherited by the solution, is the property of temperedness.

3.1. Definition (Tempered random vector). A random vector X with values in $\mathbb{R}^d$ is called **tempered** with respect to the metric dynamical system $(\Omega, \mathcal{F}, \mathbb{P}, (\theta_t)_{t \in \mathbb{R}})$ if

$$\lim_{t \to \pm\infty} \frac{1}{|t|} \log^+ |X(\theta_t \omega)| = 0. \tag{3.1}$$

■

The assumption (3.1) excludes the case $\limsup_{t\to\pm\infty} \frac{1}{|t|} \log^+ |X(\theta_t\omega)| = \infty$, and is implied by the integrability condition $\sup_{0\le t\le 1} \log^+ |X(\theta_t\cdot)| \in L^1(\mathbb{P})$. We will also make use of the fact that the set of real–valued tempered random variables is a commutative ring with unit element.

3.2. THEOREM (FORMAL LINEARIZATION OF AN SDE). *Given the SDE*

$$dx_t = \sum_{j=0}^{m} f_j(x_t) \circ dW_t^j, \quad f_j(0) = 0, \quad 0 \le j \le m,$$

and assume that the linear cocycle generated by

$$dv_t = \sum_{j=0}^{m} A_j v_t \circ dW_t^j, \quad A_j := Df_j(0), \quad 0 \le j \le m,$$

is nonresonant of any order $n \ge 2$. Then there is a random coordinate transformation h whose formal Taylor series $h(\omega, x) \sim x + \sum_{n\ge 2} h_n(\omega, x)$ is uniquely determined and has tempered coefficients h_n, such that h formally linearizes the above SDE.

3.3. Remark. (i) It is quite remarkable that at the beginning and at the end of the above procedure we have two bona fide classical SDE, while the transformation converting the solutions of the first into the solutions of the second is anticipative.

(ii) With the above procedure, infinitely flat terms of the f_j at 0 cannot be detected as they do not appear in formal power series. A random version of Sternberg's linearization theorem would assert that in the above situation the nonlinear SDE (2.1) and its linearization (2.2) are indeed *smoothly* (and not just formally) conjugate by a random coordinate transformation $h(\omega, x)$. However, such a theorem is still lacking. ■

The proof of Theorem 3.2 is quite complicated and is divided into three steps, of which the first two are of technical nature, while the third one (on invariant measures of affine SDE) is of independent interest. The result will be a formal random Taylor series for h. To find a random coordinate transformation corresponding to this formal random Taylor series is accomplished by the following theorem.

3.4. THEOREM (BOREL'S THEOREM). *Given $h_0 = 0$, $h_1(\omega, x) = x$, h_n $H_{n,d}$–valued random variables, $n \ge 2$. Then there is a random coordinate transformation h whose formal random Taylor series expansion has the coefficients h_n.*

The proof is just an ω–wise version of the deterministic proof given by Vanderbauwhede [23], page 142, and is thus omitted.

3.1 Step 1: Boundedness of moments of solutions of a hierarchical system of affine SDE

Our main task here will consist in proving that all processes in our hierarchical system of affine SDE obtained by solving the cohomological equations (2.12) step by step for *fixed* initial conditions satisfy the conditions of the following lemma.

3.5.LEMMA. *Let $u = (u(t,x))_{t\in[0,1],x\in\mathbb{R}^d}$ be an $\mathbb{R}^k$–valued stochastic process which for fixed x is $\mathbb{P}$-a. s. continuous with respect to t, is adapted to $(\mathcal{F}^t_{-\infty})$, $0 \le t \le 1$, and satisfies the following two conditions which from now on are called*

Conditions (C): *For any $p \ge 2$ and any compact set $K \subset \mathbb{R}^d$ there exist constants c_p and $c_{p,K} \in \mathbb{R}^+$ such that*

$$\mathbb{E}\left(\sup_{0\le t\le 1} |u(t,0)|^p\right) \le c_p,$$

$$\mathbb{E}\left(\sup_{0\le t\le 1} |u(t,x) - u(t,y)|^p\right) \le c_{p,K}|x-y|^{p/2} \quad \textit{for all } x, y \in K.$$

Then u is $\mathbb{P}$-a. s. jointly continuous with respect to (t,x), and for $p > d$ and any compact set $K \subset \mathbb{R}^d$ there exist constants $C_p \in \mathbb{R}^+$ and $q \ge 1$ such that

$$\mathbb{E}\left(\sup_{x\in K}\sup_{0\le t\le 1} |u(t,x)|^p\right) \le C_p(\operatorname{diam} K)^q. \tag{3.2}$$

Proof. The joint continuity of u follows from (C) by Kolmogorov's continuity criterion, applied to the $\mathcal{C}[0,1]$–valued process $x \mapsto u(\cdot, x)$, and with $p/2 > d$.

(3.2) is a well–known implication of the fundamental continuity lemma of Garsia, Rodemich and Rumsey, see for example Barlow and Yor [11], formula (3.b), or Arnold and Imkeller [5]. □

We now prove that conditions (C) are passed on from u to processes obtained from u by reasonable operations.

3.6.LEMMA. *Let u satisfy conditions (C), and let*

$$v(t,x) = \int_0^t u(s,x)\, dW_s, \quad w(t,x) = \int_0^t u(s,x) ds,$$

where W is a scalar Wiener process. Then v and w satisfy conditions (C).

Proof. This is an immediate consequence of Burkholder's and Hölder's inequalities. □

The following considerations will be crucial for the algorithm by which we solve our hierarchical system of affine SDE. Let $d_1, d_2 \in \mathbb{N}$, and for $0 \le i \le m$ suppose that $(u_i(t,x_1))_{t\in[0,1],x_1\in\mathbb{R}^{d_1}}$ is a parametrized semimartingale with decomposition

$$u_i(t,x_1) = \int_0^t w_i(s,x_1)ds + \sum_{j=1}^m \int_0^t v_i^j(s,x_1)\, dW_s^j, \tag{3.3}$$

with values in $\mathbb{R}^{d_2}$. Let, moreover, B_0, $B_1,\ldots,B_m$ be $d_2 \times d_2$ matrices, and denote by $(\Phi(t))_{t\in\mathbb{R}}$ the linear flow in $\mathbb{R}^{d_2}$ generated by the linear SDE

$$dy_t = \sum_{j=0}^m B_j y_t \circ dW_t^j. \tag{3.4}$$

We now consider the SDE

$$dx_t = \sum_{j=0}^m (B_j x_t + u_j(t,x_1)) \circ dW_t^j, \quad x_0 = x_2 \in \mathbb{R}^{d_2}. \tag{3.5}$$

Let $(\varphi(t,x_1))_{t\in\mathbb{R}}$ be the flow generated by (3.5), the value of which at $x_2 \in \mathbb{R}^{d_2}$ will be written as $\varphi(t,x_1)x_2$.

3.7. LEMMA. *Let $u_i = (u_i(t,x_1))_{t\in[0,1],x_1\in\mathbb{R}^{d_1}}$ be given by (3.3). Assume that w_i and v_i^j (hence u_i), $0 \le i \le m$, $1 \le j \le m$, satisfy conditions (C). Let*

$$\varphi(t,x_1)x_2 = \int_0^t \rho(s,x_1,x_2)ds + \sum_{j=1}^m \int_0^t \psi^j(s,x_1,x_2)dW_s^j$$

be the semimartingale decomposition of the solution flow of (3.5). Then φ, ρ, and ψ^j, $1 \le j \le m$, satisfy conditions (C).

Proof. (i) We only prove the second of the conditions (C) for φ.

Let us start by writing the Itô form of (3.5). We have

$$\begin{aligned} dy_t &= \sum_{j=1}^m (B_j y_t + u_j(t,x_1))dW_t^j \qquad (3.6)\\ &\quad + \left((B_0 + \frac{1}{2}\sum_{j=1}^m B_j^2) y_t + \frac{1}{2}\sum_{j=1}^m v_j^j(t,x_1) \right) dt. \end{aligned}$$

Now fix $p \ge 2$ and a compact set $K \subset \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. Then for any (x_1,x_2), $(y_1,y_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, Burkholder's and Jensen's inequalities as well as the hypothesis yield, with suitable constants c_1, c_2, $c_{3,K}$ and with the abbreviation

$$f(t) := \mathbb{E}\left(\sup_{s\in[0,t]} |\varphi(s,x_1)x_2 - \varphi(s,y_1)y_2|^p \right)^{1/p},$$

$$
\begin{aligned}
f(1) \quad \leq \quad & c_1 \Bigg\{ |x_2 - y_2| \\
& + \sum_{j=1}^m \mathbb{E} \left(\sup_{t\in[0,1]} | \int_0^t B_j(\varphi(s,x_1)x_2 - \varphi(s,y_1)y_2) dW_s^j|^p \right)^{1/p} \\
& + \sum_{j=1}^m \mathbb{E} \left(\sup_{t\in[0,1]} | \int_0^t (u_j(s,x_1) - u_j(s,y_1)) dW_s^j|^p \right)^{1/p} \\
& + \mathbb{E} \left(\sup_{t\in[0,1]} | \int_0^t (B_0 + \frac{1}{2}\sum_{j=1}^m B_j^2)(\varphi(s,x_1)x_2 - \varphi(s,y_1)y_2) ds|^p \right)^{1/p} \\
& + \sum_{j=1}^m \mathbb{E} \left(\sup_{t\in[0,1]} | \int_0^t (v_j^j(s,x_1) - v_j^j(s,y_1)) ds|^p \right)^{1/p} \Bigg\} \\
\leq \quad & c_2 \left(|x_2 - y_2| + \int_0^t f(s)ds + |x_1 - y_1|^{1/2} \right) \\
\leq \quad & c_{3,K} \left(|(x_1,x_2) - (y_1,y_2)|^{1/2} + \int_0^1 f(t)dt \right). \qquad (3.7)
\end{aligned}
$$

To (3.7) we have to apply Gronwall's lemma to obtain with a suitable constant $c_{p,K} \in \mathbb{R}^+$

$$
\mathbb{E} \left(\sup_{t\in[0,1]} |\varphi(t,x_1)x_2 - \varphi(t,y_1)y_2|^p \right) \leq c_{p,K} |(x_1,x_2) - (y_1,y_2)|^{p/2}. \qquad (3.8)
$$

This is the second of the conditions (C) for φ.

(ii) Next observe that according to (3.6) for any $t \in [0,1]$ and $(x_1,x_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$

$$
\psi^j(t,x_1,x_2) = B_j\varphi(t,x_1)x_2 + u_j(t,x_1),
$$

$$
\rho(t,x_1,x_2) = \left(B_0 + \frac{1}{2}\sum_{j=1}^m B_j^2 \right) \varphi(t,x_1)x_2 + \frac{1}{2}\sum_{j=1}^m v_j^j(t,x_1).
$$

Hence by our hypotheses and (3.8), ψ^j and ρ satisfy the conditions (C) as well. □

Next we show that conditions (C) are inherited from u and its characteristics to a polynomial of u and its characteristics.

3.8. LEMMA. *Let*

$$u(t,x) = \sum_{j=1}^{m} \int_0^t v^j(s,x) dW_s^j + \int_0^t w(s,x) ds, \quad t \in [0,1], x \in \mathbb{R}^{d_1},$$

take values in $\mathbb{R}^{d_2}$*, and assume that* u*,* v^j *and* w *satisfy conditions (C). Let* p *be a polynomial in the variable* $y \in \mathbb{R}^{d_2}$*. Then, if*

$$(p \circ u)(t,x) = \sum_{j=1}^{m} \int_0^t q^j(s,x) dW_s^j + \int_0^t r(s,x) ds,$$

also $p \circ u$*,* q^j *and* r *satisfy conditions (C).*

Proof. First note that, by Itô's formula, q^j and r are of the same structure as $p \circ u$. Hence it suffices to prove the conditions for $p \circ u$.

Due to Hölder's inequality, it is evidently enough to consider the case of a real–valued u and a polynomial p of the form $p(y) = y^l$ for some $l \in \mathbb{N}$. Then for $y_1, y_2 \in \mathbb{R}$

$$p(y_1) - p(y_2) = (y_1 - y_2) \sum_{k=0}^{l-1} y_1^k y_2^{l-1-k}. \tag{3.9}$$

Now observe that by conditions (C) for any $x \in \mathbb{R}^{d_1}$

$$\mathbb{E} \left(\sup_{t \in [0,1]} |u(t,x)|^p \right) \le c_p \tag{3.10}$$

for a suitable constant $c_p \in \mathbb{R}^+$. Then (3.9) and an application of Hölder's inequality obviously allow us to deduce conditions (C) for $p \circ u$ from (3.10) and the condition (C) for u. □

3.9. LEMMA. *Let* $u = (u(t,x))_{t \in [0,1], x \in \mathbb{R}^d}$ *with values in* $\mathbb{R}^d$ *satisfy conditions (C), let* $(\Phi(t))_{t \in \mathbb{R}}$ *be the linear flow generated by (3.4) for* $d_2 = d$*, and let* W *be a scalar Wiener process. Then the processes*

$$v(t,x) = \int_0^t \Phi(s)^{-1} u(s,x)\, dW_s, \quad w(t,x) = \int_0^t \Phi(s)^{-1} u(s,x) ds$$

satisfy conditions (C).

Proof. $\Phi(t)^{-1}$ satisfies the SDE

$$d\Phi(t)^{-1} = -\sum_{j=0}^{m} \Phi(t)^{-1} B_j \circ dW_t^j, \quad \Phi(0) = I.$$

Hence for any $p \geq 1$ we have

$$\mathbb{E}\left(\sup_{t\in[0,1]} \|\Phi(t)^{-1}\|^p\right) < \infty. \tag{3.11}$$

Then it is clear that Burkholder's inequality, Hölder's inequality and (3.11) imply that v and w satisfy conditions (C). □

We finally come to the announced hierarchical system of affine SDE. Let $(d_n)_{n\in\mathbb{N}}$ be a sequence of integers. For $0 \leq j \leq m$ and each $n \in \mathbb{N}$, let $A_{j,n}$ be a $d_n \times d_n$ matrix, and let $p_{j,n}$ be a polynomial in the variables $(x_1,\dots,x_{n-1}) \in \mathbb{R}^{d_1} \times \dots \times \mathbb{R}^{d_{n-1}}$, with $p_{j,1} = b_j \in \mathbb{R}^{d_1}$ a fixed vector. Then our **hierarchical system of affine SDE** is as follows:

$$\begin{aligned}
dx_t^1 &= \sum_{j=0}^m (A_{j,1}x_t^1 + p_{j,1}) \circ dW_t^j, \quad x_0^1 = x_1 \in \mathbb{R}^{d_1}, &(3.12)\\
dx_t^2 &= \sum_{j=0}^m (A_{j,2}x_t^2 + p_{j,2}(x_t^1)) \circ dW_t^j, \quad x_0^2 = x_2 \in \mathbb{R}^{d_2}, &(3.13)\\
\dots &\quad \dots\\
dx_t^n &= \sum_{j=0}^m (A_{j,n}x_t^n + p_{j,n}(x_t^1,\dots,x_t^{n-1})) \circ dW_t^j, &(3.14)\\
&\quad x_0^n = x_n \in \mathbb{R}^{d_n},\\
\dots &\quad \dots
\end{aligned}$$

The algorithm for successively solving this system is as follows: Let $(\Phi_n(t))_{t\in\mathbb{R}}$ be the linear cocycle in $\mathbb{R}^{d_n}$ generated by the linear SDE

$$dy_t^n = \sum_{j=0}^m A_{j,n} y_t^n \circ dW_t^j, \quad y_0^n = y_n \in \mathbb{R}^{d_n}. \tag{3.15}$$

We first solve the first affine SDE (3.12). Denote the resulting affine cocycle by $(\varphi_1(t))_{t\in\mathbb{R}}$ which by the variation of constants formula is given by

$$\varphi_1(t)x = \Phi_1(t)\left(x + \sum_{j=0}^m \int_0^t \Phi_1(s)^{-1} p_{j,1} \circ dW_s^j\right).$$

Then insert the cocycle $\varphi_1(t)x_1$ into the second equation (3.13) in place of x_t^1 which gives a non–autonomous affine SDE whose solution flow is denoted by $\varphi_2(t,x_1)$, etc. At step n insert $\varphi_1(t)x_1, \dots, \varphi_{n-1}(t,x_1,\dots,x_{n-2})x_{n-1}$ into the nth equation (3.14) which gives a non–autonomous affine SDE

$$\begin{aligned}
dx_t^n &= \sum_{j=0}^m (A_{j,n}x_t^n + p_{j,n}(\varphi_1(t)x_1,\dots,\varphi_{n-1}(t,x_1,\dots,x_{n-2})x_{n-1}) \circ dW_t^j,\\
&\quad x_0^n = x_n \in \mathbb{R}^{d_n}, &(3.16)
\end{aligned}$$

with solution flow $\varphi_n(t, x_1, \ldots, x_{n-1})$.

Note that the first n affine SDE considered as one equation generate the $\mathcal{C}^\infty$ RDS

$$\varphi(t)(x_1, \ldots, x_n) = (\varphi_1(t)x_1, \varphi_2(t, x_1)x_2, \ldots, \varphi_n(t, x_1, \ldots, x_{n-1})x_n). \quad (3.17)$$

3.10. PROPOSITION. *For $n \in \mathbb{N}$, represent the solution of the nth equation (3.16) as*

$$\begin{aligned}\varphi_n(t, x_1, \ldots, x_{n-1})x_n &= \int_0^t v^n(s, x_1, \ldots, x_n)ds \\ &\quad + \sum_{j=1}^m \int_0^t u_j^n(s, x_1, \ldots, x_n)dW_s^j.\end{aligned}$$

Then for any $n \in \mathbb{N}$, φ_n as well as v^n and u_j^n, $1 \leq j \leq m$, satisfy conditions (C).

Proof. We use induction on n.

The assertion holds for $n = 1$. This is an immediate consequence of Lemma 3.7, choosing $u_j = b_j$.

Let us now assume that the assertion holds for $\varphi_1, \ldots, \varphi_{n-1}$ and their characteristics. Consequently, Lemma 3.8 yields that all the components of $p_{j,n-1}(\varphi_1(t)x_1, \ldots, \varphi_{n-1}(t, x_1, \ldots, x_{n-2})x_{n-1})$ and their semimartingale characteristics satisfy conditions (C). Hence Lemma 3.7 applies and gives condition (C) for φ_n as well as for u_j^n and v^n. □

3.11. PROPOSITION. *Let for $t \in \mathbb{R}$*

$$X_1(t) = \sum_{j=0}^m \int_0^t \Phi_1(s)^{-1} b_j \circ dW_s^j,$$

and for $n \geq 2$ and $(x_1, \ldots, x_{n-1}) \in \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_{n-1}}$

$$X_n(t, x_1, \ldots, x_{n-1}) = \sum_{j=0}^m \int_0^t \Phi_n(s)^{-1} p_{j,n}(\varphi_1(t)x_1, \ldots, \varphi_{n-1}(t, x_1, \ldots, x_{n-2})x_{n-1}) \circ dW_s^j. \quad (3.18)$$

Then X_n satisfies conditions (C).

Proof. This is an immediate consequence of Proposition 3.10 and Lemmas 3.8 and 3.9. □

Here is our final and main result of step 1.

3.12. THEOREM. *Consider the hierarchical system of affine SDE introduced in (3.12) to (3.14). Then for any $n \geq 2$, any $p > d_1 + \cdots + d_{n-1}$, and any compact set $K \subset \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_{n-1}}$ there exist constants $c_{n,p} \in \mathbb{R}^+$ and $q \geq 1$ such that*

$$\mathbb{E}\left(\sup_{(x_1,\ldots,x_{n-1}) \in K} \sup_{t \in [0,1]} |X_n(t, x_1, \ldots, x_{n-1})|^p \right) \leq c_{n,p} (\operatorname{diam} K)^q, \tag{3.19}$$

where X_n is defined by (3.18).

Proof. Combine Proposition 3.11 with Lemma 3.5. □

3.2 Step 2: Inheritance of temperedness

We first study the inheritance of temperedness in case tempered vectors (see Definition 3.1) are inserted into random fields. Property (3.2) will play a crucial role hereby, as indicated by the following lemma.

3.13. LEMMA. *Let $(X(y))_{y \in \mathbb{R}^{d_1}}$ be a $\mathbb{P}$-a. s. continuous random field with values in $\mathbb{R}^{d_2}$ for which the following condition holds: For $p > d_1$ and for any compact set $K \subset \mathbb{R}^{d_1}$ there exist $c_p \in \mathbb{R}^+$ and $q \geq 1$ such that*

$$\mathbb{E}\left(\sup_{y \in K} |X(y)|^p \right) \leq c_p (\operatorname{diam} K)^q. \tag{3.20}$$

Let Y be a random vector with values in $\mathbb{R}^{d_1}$, and for $\varepsilon > 0$ and $m \in \mathbb{N}$ let

$$A_{\varepsilon,m} := \{\omega \in \Omega : |Y(\theta_t \omega)| \leq m \exp(t\varepsilon),\, t \in \mathbb{R}^+\}.$$

Then there exists a constant $c_{\varepsilon,m} \in \mathbb{R}^+$ such that for any $n \in \mathbb{Z}^+$

$$\mathbb{E}\left(1_{A_{\varepsilon,m}} \sup_{t \in [n,n+1]} |X(Y)(\theta_t \cdot)|^p \right) \leq c_{\varepsilon,m} \exp(n\varepsilon q),$$

where p and q are related by (3.20).

Proof. For $\omega \in A_{\varepsilon,m}$ and $t \in [n, n+1]$ we have

$$|Y(\theta_t \omega)| \leq m \exp(t\varepsilon) \leq m \exp(\varepsilon) \exp(n\varepsilon).$$

Hence due to (3.20)

$$\begin{aligned} \mathbb{E}\left(1_{A_{\varepsilon,m}} \sup_{t \in [n,n+1]} |X(Y)(\theta_t \cdot))|^p \right) &\leq \mathbb{E}\left(1_{A_{\varepsilon,m}} \sup_{y \in B_{\varepsilon,m,n}} |X(y)|^p \right) \\ &\leq c_p (\operatorname{diam} B_{\varepsilon,m,n})^q, \end{aligned} \tag{3.21}$$

where

$$B_{\varepsilon,m,n} := \{y \in \mathbb{R}^{d_1} \,:\, |y| \leq m \exp(\varepsilon) \exp(n\varepsilon)\}.$$

But $\operatorname{diam} B_{\varepsilon,m,n} \leq c_{d_1,\varepsilon,m} \exp(n\varepsilon)$, with a constant depending just on d_1, ε and m. Hence the desired inequality follows readily from (3.21). □

3.14. THEOREM. *Let $(X(y))_{y\in\mathbb{R}^{d_1}}$ be a $\mathbb{P}$-a. s. continuous random field with values in $\mathbb{R}^{d_2}$ for which the following condition holds: For $p > d_1$ and for any compact set $K \subset \mathbb{R}^{d_1}$ there exist a $c_p \in \mathbb{R}^+$ and a $q \geq 1$ such that*

$$\mathbb{E}\left(\sup_{y\in K}|X(y)|^p\right) \leq c_p(\operatorname{diam} K)^q. \tag{3.22}$$

Let Y be a tempered random vector with values in $\mathbb{R}^{d_1}$. Then the $\mathbb{R}^{d_2}$–valued random vector $X(Y)$ is tempered.

Proof. Define $A_{\varepsilon,m}$ as in Lemma 3.13. We shall prove that

$$\lim_{t\to\infty}\frac{1}{t}\log^+|X(Y)(\theta_t\omega))| = 0,$$

remarking that the behavior for $t \to -\infty$ can be treated similarly. Since Y is tempered, $A_{\varepsilon,m} \uparrow \Omega$ $(m \uparrow \infty)$ $\mathbb{P}$-a. s. for any $\varepsilon > 0$.

Let $\varepsilon > 0$ be given. We have to prove that there exists a $\delta(\varepsilon) > 0$ such that for any $m \in \mathbb{N}$ we have

$$1_{A_{\delta(\varepsilon),m}}\limsup_{t\to\infty}\frac{1}{t}\log^+|X(Y)(\theta_t\omega)| \leq \varepsilon, \quad \mathbb{P}\text{-a. s.} \tag{3.23}$$

(3.23) will indeed imply temperedness since $A_{\delta(\varepsilon),m} \uparrow \Omega$, $\mathbb{P}$-a. s. , for $m \uparrow \infty$ and any $\varepsilon > 0$.

To prove (3.23), let ε, $\delta > 0$, $m, n \in \mathbb{N}$ be given. Then by (3.22) and Lemma 3.13

$$\begin{aligned}
&\mathbb{P}(A_{\delta,m}\cap\{\sup_{t\in[n,n+1]}\frac{1}{t}\log^+|X(Y)(\theta_t\cdot)| > \varepsilon\}) \\
&\leq \mathbb{P}(A_{\delta,m}\cap\{\sup_{t\in[n,n+1]}|X(Y)(\theta_t\cdot)| > \exp(n\varepsilon)\}) \\
&\leq \mathbb{E}(1_{A_{\delta,m}}\exp(-np\varepsilon)\sup_{t\in[n,n+1]}|X(Y)(\theta_t\cdot))|^p) \\
&\leq c_{m,\delta}\exp(nq\delta - np\varepsilon) = c_{m,\delta}\exp(n(q\delta - p\varepsilon)).
\end{aligned}$$

Now choose $\delta < \frac{p\varepsilon}{q}$. Then for all $m \in \mathbb{N}$ the Borel Cantelli lemma yields

$$1_{A_{\delta,m}}\limsup_{n\to\infty}\sup_{t\in[n,n+1]}\frac{1}{t}\log^+|X(Y)(\theta_t\cdot))| \leq \varepsilon.$$

This clearly implies (3.23) and completes the proof. □

As a second issue we study the inheritance of temperedness by geometric series.

3.15. THEOREM. *Suppose X is an $\mathbb{R}^d$–valued tempered random vector. Let $(T_n)_{n\in\mathbb{N}}$ be a sequence of random linear operators in $\mathbb{R}^d$ with negative Lyapunov index, i. e., for some deterministic $\beta > 0$*

$$\limsup_{n\to\infty} \frac{1}{n} \log \|T_n\| \le -\beta.$$

Then

$$Y = \sum_{n=0}^{\infty} T_n(X \circ \theta(n))$$

is absolutely and geometrically convergent and tempered.

Proof. Choose $\varepsilon > 0$ such that $2\varepsilon < \beta$. Then by the assumptions, from a certain index n on, $\|T_n\|(X \circ \theta(n)) < \exp((2\varepsilon - \beta)n)$ from which the convergence statements follow. As to the temperedness of Y, there exists a random variable R_ε such that for $t \in \mathbb{R}$

$$\begin{aligned} |Y \circ \theta_t| &\le \sum_{n=0}^{\infty} \|T_n\| \, |X \circ \theta(n+t)| \\ &\le R_\varepsilon \sum_{n=0}^{\infty} \exp(\varepsilon n - \beta n) \exp(\varepsilon(n + |t|)) \\ &= R_\varepsilon \frac{1}{1 - \exp(2\varepsilon - \beta)} \exp(\varepsilon |t|), \end{aligned}$$

which clearly implies that Y is tempered. □

3.3 Step 3: Invariant measures of a hierarchical system of affine SDE

We return to the setting of a hierarchical system of affine SDE introduced above (see Theorem 3.12), and determine its invariant measures.

So let $(\Phi_n(t))_{t\in\mathbb{R}}$ be the linear cocycle in $\mathbb{R}^{d_n}$ generated by (3.15). The multiplicative ergodic theorem holds for Φ_n, giving its Lyapunov spectrum $\Sigma(\Phi_n)$. Having in mind our cohomological equations in the nonresonant case, we assume that all these cocycles are hyperbolic, i. e., $0 \notin \Sigma(\Phi_n)$ for all $n \in \mathbb{N}$. Then $\mathbb{R}^{d_n} = E_n^s(\omega) \oplus E_n^u(\omega)$, where $E_n^s(\omega) = \oplus_{\lambda_{i,n}<0} E_{i,n}(\omega)$ is the stable space, and $E_n^u(\omega) = \oplus_{\lambda_{i,n}>0} E_{i,n}(\omega)$ is the unstable space of Φ_n. Denote by π_n^u (π_n^s) the projection whose range is E_n^u (E_n^s), and whose kernel is E_n^s (E_n^u).

For convenience we recall the following facts (see [2], section 4.3).

3.16. LEMMA. *Assume that the cocycle Φ_n is hyperbolic. Then there exists a constant $\beta_n > 0$ such that*

$$\limsup_{k\to\infty} \frac{1}{k} \log \|\pi_n^u \Phi_n(k)^{-1}\| \le -\beta_n,$$

and

$$\limsup_{k\to\infty} \frac{1}{k} \log \|\pi_n^s \Phi_n(-k)^{-1}\| \leq -\beta_n.$$

We first consider the affine SDE of order one. We recall that a probability measure $\mu(d\omega, dx) = \mu_\omega(dx)\mathbb{P}(d\omega)$ on $(\Omega \times \mathbb{R}^d, \mathcal{F} \otimes \mathcal{B})$ with marginal $\mathbb{P}$ on $(\Omega, \mathcal{F})$ is called invariant for the RDS φ if $\varphi(t,\omega)\mu_\omega = \mu_{\theta_t\omega}$ $\mathbb{P}$-a. s.

3.17. THEOREM. *Consider the affine SDE in $\mathbb{R}^{d_1}$*

$$dx_t^1 = \sum_{j=0}^{m} (A_{j,1} x_t^1 + b_j) \circ dW_t^j, \tag{3.24}$$

and assume that the corresponding linear cocycle Φ_1 generated by $dy_t^1 = \sum_{j=0}^m A_{j,1} y_t^1 \circ dW_t^j$ is hyperbolic. Then the affine cocycle φ_1 generated by (3.24) has a unique invariant measure, namely the random Dirac measure $\mu_\omega = \delta_{\kappa_1(\omega)}$, i. e., we have $\varphi_1(t,\cdot)\kappa_1 = \kappa_1 \circ \theta_t$. Here $\kappa_1 = \kappa_1^s \oplus \kappa_1^u$, and

$$\kappa_1^s = \sum_{k=0}^{\infty} \pi_1^s \circ \Phi_1(-k)^{-1} (X_1^- \circ \theta_{-k}), \tag{3.25}$$

$$\kappa_1^u = -\sum_{k=0}^{\infty} \pi_1^u \circ \Phi_1(k)^{-1} (X_1^+ \circ \theta_k), \tag{3.26}$$

where

$$X_1^- := \sum_{j=0}^{m} \int_{-1}^{0} \Phi_1(t)^{-1} b_j \circ dW_t^j, \quad X_1^+ := \sum_{j=0}^{m} \int_{0}^{1} \Phi_1(t)^{-1} b_j \circ dW_t^j.$$

The sequences in (3.25) and (3.26) converge $\mathbb{P}$-a. s. absolutely and geometrically, and the limits κ_1^s and κ_1^u as well as $\kappa_1 = \kappa_1^s \oplus \kappa_1^u$ are tempered random vectors.

Proof. (i) We first prove the statements on temperedness. X_1^- and X_1^+ are tempered by Theorem 3.14 since a constant vector is tempered. Hence Theorem 3.15 and Lemma 3.16 yield that κ_1^s and κ_1^u exist and are tempered. Thus κ_1 is tempered.

(ii) We now check that δ_{κ_1} is invariant, i. e., that $\varphi_1(t,\cdot)\kappa_1 = \kappa_1 \circ \theta_t$. This is equivalent with $\pi_1^{s,u}(\theta_t \cdot)\varphi_1(t,\cdot)\kappa_1 = \kappa_1^{s,u}(\theta_t \cdot)$ for $t \in \mathbb{R}$. Using the variation of constants representation of φ_1 and the fact that for $k \in \mathbb{Z}^+$

$$\sum_{j=0}^{m} \int_{k}^{k+1} \Phi_1(s)^{-1} b_j \circ dW_s^j = \Phi_1(k)^{-1} (X_1^+ \circ \theta_k),$$

we obtain, after some rather lengthy, but elementary manipulations,

$$\begin{aligned}\varphi_1(t,\cdot)\kappa_1^u &= \lim_{N\to\infty}\varphi_1(t,\cdot)\left(-\sum_{k=0}^{N-1}\pi_1^u\circ\Phi_1(k)^{-1}(X_1^+\circ\theta_k)\right)\\ &= \kappa_1^u\circ\theta_t+\pi_1^s\sum_{j=0}^m\int_0^t\Phi_1(s)^{-1}b_j\circ dW_s^j.\end{aligned}$$

A similar argument holds for the stable component.

(iii) The uniqueness of the invariant measure is a consequence of hyperbolicity and is proved in [2], Theorem 5.6.2. □

Now suppose that $\kappa_1,\ldots,\kappa_{n-1}$ are tempered random vectors with values in $\mathbb{R}^{d_1},\ldots,\mathbb{R}^{d_{n-1}}$, respectively such that $\mu^i=\delta_{\kappa_i}$, $1\le i\le n-1$ is the unique invariant measure of the ith equation of our hierarchical system of affine SDE. We now study the nth SDE

$$dx_t^n=\sum_{j=0}^m(A_{j,n}x_t^n+p_{j,n}(\varphi_1(t)\kappa_1,\ldots,\varphi_{n-1}(t,\kappa_1,\ldots,\kappa_{n-2})\kappa_{n-1}))\circ dW_t^j \quad (3.27)$$

in $\mathbb{R}^{d_n}$. Note the fundamental fact that the validity of Lemma 3.5 entails the famous substitution rule for Stratonovich integrals (see Arnold and Imkeller [5], Corollary 1): If η is any random variable and W a scalar Wiener process, then $u(t,\eta)$ (though nonadapted) is Stratonovich integrable, and

$$\int_0^t u(s,\eta)\circ dW_s=\int_0^t u(s,x)\circ dW_s\Big|_{x=\eta},\quad t\in\mathbb{R}. \quad (3.28)$$

Hence the anticipative affine SDE (3.27) makes sense, and it generates an affine cocycle $\varphi_n(t,\kappa_1,\ldots,\kappa_{n-1})$. We now determine its invariant measure.

3.18. THEOREM. *Suppose all equations of the hierarchical system of affine SDE have a hyperbolic linear part. Then each of the equations has a unique invariant measure. This measure is a random Dirac measure (stationary solution) whose supporting random variable can be obtained step by step by determining κ_1 according to Theorem 3.17, inserting $\varphi_1(t)\kappa_1$ into the second SDE, etc., where the stable and unstable component of κ_n of the nth equation (3.27) is given by*

$$\kappa_n^s=\sum_{k=0}^\infty\pi_n^s\circ\Phi_n(-k)^{-1}(X_n^-\circ\theta_{-k}), \quad (3.29)$$

$$\kappa_n^u=-\sum_{k=0}^\infty\pi_n^u\circ\Phi_n(k)^{-1}(X_n^+\circ\theta_k), \quad (3.30)$$

where

$$X_n^- := \sum_{j=0}^m \int_{-1}^0 \Phi_n(t)^{-1} p_{j,n}(\varphi_1(t)\kappa_1, \ldots, \varphi_{n-1}(t,\kappa_1,\ldots,\kappa_{n-2})\kappa_{n-1}) \circ dW_t^j,$$

$$X_n^+ := \sum_{j=0}^m \int_{0}^1 \Phi_n(t)^{-1} p_{j,n}(\varphi_1(t)\kappa_1, \ldots, \varphi_{n-1}(t,\kappa_1,\ldots,\kappa_{n-2})\kappa_{n-1}) \circ dW_t^j.$$

The sequences in (3.29) and (3.30) converge $\mathbb{P}$*-a. s. absolutely and geometrically, and the limits* κ_n^s *and* κ_n^u *as well as* $\kappa_n = \kappa_n^s \oplus \kappa_n^u$ *are tempered random vectors.*

Proof. Since Theorem 3.12 holds, and since $(\kappa_1, \ldots, \kappa_{n-1})$ is a tempered vector, Theorem 3.14 applies and entails that X_n^- and X_n^+ are tempered. Hence by Lemma 3.16 and Theorem 3.15, the series in equations (3.29) and (3.30) have the convergence properties claimed, and the limits $\kappa_n^{s,u}$ are tempered. Hence finally κ_n is tempered.

Invariance and uniqueness of κ_n are proved as in Theorem 3.17. □

3.19. Remark. (i) By means of the substitution rule (3.28), κ_1 can also be written as

$$\kappa_1^s = \sum_{j=0}^m \int_{-\infty}^0 \Phi_1(t)^{-1}(\pi_1^s(\theta_t\cdot)b_j) \circ dW_t^j,$$

$$\kappa_1^u = -\sum_{j=0}^m \int_{0}^\infty \Phi_1(t)^{-1}(\pi_1^u(\theta_t\cdot)b_j) \circ dW_t^j,$$

where the integrals exist as the $\mathbb{P}$-a. s. limits of the non–adapted Stratonovich integrals $\int_{-T}^0$ and $\int_0^T$ for $T \to \infty$. Similarly for κ_n.

(ii) Again by the substitution rule (3.28), the cocycle (3.17) evaluated at $\kappa := (\kappa_1, \ldots, \kappa_n)$ is equal to the solution of the first n equations with non–adapted initial value κ, and is the unique stationary solution of these equations, $\varphi(t)\kappa = \kappa \circ \theta_t$, $t \in \mathbb{R}$. ■

This completes the proof of Theorem 3.2. We close with an example.

3.20. Example (The one–dimensional case). For $d = 1$, $H_{n,1} \cong \mathbb{R}^1$ (choose the basis $x^n e_1$) for all $n \geq 2$. The linearization of the SDE $dx = \sum_{j=0}^m f_j(x) \circ dW^j$ is $dv = \sum_{j=0}^m A_j v \circ dW^j$ whose explicit solution is

$$\Phi(t,\omega) = \exp(A_0 t + \sum_{j=1}^m A_j W_t^j(\omega)).$$

Hence $\lambda = A_0$ is the Lyapunov exponent. Further, $\mathrm{ad}_n A_j = -(n-1)A_j$, and $\Sigma(\Phi_n) = \{-(n-1)\lambda\}$. We have nonresonance for all n if and only if $\lambda \neq 0$. The unique stationary solution of the nth cohomological SDE

$$dh_n = \sum_{j=0}^{m} ((\mathrm{ad}_n A_j) h_n + k_{j,n}(\theta_t \cdot)) \circ dW_t^j$$

is

$$h_n = \begin{cases} \sum_{j=0}^m \int_{-\infty}^0 \exp((n-1)\sum_{l=0}^m A_l W_t^l) k_{j,n}(\theta_t \cdot) \circ dW_t^j, & \lambda > 0, \\ -\sum_{j=0}^m \int_0^\infty \exp((n-1)\sum_{l=0}^m A_l W_t^l) k_{j,n}(\theta_t \theta_t \cdot) \circ dW_t^j, & \lambda < 0. \end{cases}$$

■

4 The resonant and small noise case

4.1 Resonant case

If $0 \in \Sigma(\Phi_n)$ we cannot guarantee anymore that the cohomological equation (2.11) of order n has a solution for any right-hand side . Assume for simplicity that Φ has simple spectrum (all Lyapunov exponents Λ_i are different, equivalently all E_i are one–dimensional). Then by Theorem 9 of Arnold and Imkeller [5], the linear SDE (2.1) can be diagonalized by means of a random (anticipative) coordinate transformation $w_t = P(\theta_t\omega)v_t$, so that the linear cocycle

$$\Psi(t,\omega) = P(\theta_t\omega)\Phi(t,\omega)P(\omega)^{-1} = \mathrm{diag}\,(\psi_1(t,\omega),\ldots,\psi_d(t,\omega))$$

has the same spectrum as Φ, is diagonal, and has generator

$$dw_t = \sum_{j=0}^{m} \mathrm{diag}\,(Q_j(u_1(\theta_t\cdot)),\ldots,Q_j(u_d(\theta_t\cdot)))w_t \circ dW_t^j, \tag{4.1}$$

where $Q_j(u) = \langle A_j u, u\rangle$, and $F = (u_1,\ldots,u_d)$ is a basis of random eigenvectors such that $u_i \in E_i$. The anticipative Stratonovich SDE (4.1) makes sense. Further, all $\mathrm{ad}_n A_j$ become diagonal matrices in this basis (see Remark 1.1), so that the random cohomological equation (2.11) decouples into D scalar SDE. Recall that for $n \geq 2$, $\Sigma(\Phi_n) = \{\Lambda_i - (\Lambda,\tau) : \Lambda_i \in \Sigma(\Phi), |\tau| = n\}$.

If (i,τ) is such that $\Lambda_i - (\Lambda,\tau) \neq 0$ we put $dG_n^{i,\tau} = 0$, and solve the corresponding affine cohomological equation as usual. If for some (i,τ), $\Lambda_i - (\Lambda,\tau) = 0$ (resonance of order n) we make the **convention** to choose $h_n^{i,\tau} = 0$, hence $dG_n^{i,\tau} = dK_n^{i,\tau}$, i.e., we will not try to simplify $G_n^{i,\tau}$ (which will only be possible in rare exceptions anyway). We call the result of this procedure after N steps **normal form of order** N. The formal handling of the hierarchical system of anticipative SDE can be justified as in the nonresonant case which we refrain from making explicit here.

4.2 Small noise: a case study

The engineering and physics literature on stochastic normal forms has worked exclusively in a center/stable situation, and with a smallness parameter multiplying the noise terms, thus obtaining a stochastic normal form as a small perturbation of the deterministic one.

We will now connect our general approach developed in the previous sections (which is independent of any smallness assumptions) with the existing physics and engineering literature by presenting a stochastic analogue of a very successful procedure proposed by Elphick et al. [13] for simultaneously obtaining the normal form, eliminating the stable variables from the center equation, and determining the center manifold. This was done for random differential equations by Arnold and Xu [9].

We are now in a position to carry this procedure over to the SDE case. We will, however, not repeat details and just emphasize that all nonresonant (anticipative Stratonovich) cohomological equations in the SDE analogues of Theorems 2.1 and 3.1 of [9] are solved as in the nonresonant case of section 3, and, following our convention made in subsection 4.1, for all resonant equations we make the trivial choice $h_n = 0$ for the random transformation. This mathematically rigorous procedure finally justifies earlier important work on normal forms for SDE quoted at the beginning of this article.

We now discuss the following application: In stochastic bifurcation theory one needs to study an SDE

$$dx_t = f_0(x_t, \alpha)dt + \sigma \sum_{j=1}^{m} f_j(x_t, \alpha) \circ dW_t^j \tag{4.2}$$

in $\mathbb{R}^d$, where the smooth vector fields $f_j(x, \alpha)$ smoothly depend on a parameter $\alpha \in \mathbb{R}^m$, and $\sigma \in \mathbb{R}$ is a (small) intensity parameter. We assume that $f_j(0, \alpha) = 0$ for all $\alpha \in \mathbb{R}^m$, so that the linearization of (4.2) is

$$dv_t = A_0(\alpha)v_t dt + \sigma \sum_{j=1}^{m} A_j(\alpha)v_t \circ dW_t^j, \quad A_j(\alpha) := D_x f_j(x, \alpha)|_{x=0}, \; 0 \le j \le m.$$

Assume further that for $(\alpha, \sigma) = (0, 0)$ the deterministic linear equation

$$\dot{v} = A_0 v, \quad A_0 = D_x f_0(x, 0)|_{x=0}$$

is in a center/stable situation (no eigenvalues with positive real part) and has been brought into Jordan canonical form by a (deterministic!) coordinate transformation, so that $\mathbb{R}^d = \mathbb{R}^{d_c} \times \mathbb{R}^{d_s}$ is the invariant splitting for $A_0 = \begin{pmatrix} A_0^c & 0 \\ 0 & A_0^s \end{pmatrix}$ into a center and stable part.

We treat $(\alpha, \sigma) \in \mathbb{R}^{m+1}$ as a small parameter and seek the normal form and center manifold reduction for small (x, α, σ) as a polynomial in $(x^c)^p (x^s)^q \alpha^l \sigma^r$.

All cohomological operators (hence all linear cocycles) are now deterministic and of the form

$$dL_n(h_n) = dh_n - (\mathrm{ad}\,_n A_0) h_n dt.$$

We add the trivial equations $d\alpha = 0$ and $d\sigma = 0$. The result is given by the (rather obvious) white noise analogue of Theorem 3.1 of [9] and is (omitting remainder terms) as follows: The center variable satisfies

$$dx_t^c = A_0^c x_t^c dt + \sum_{j=0}^{m} g_j^c(\theta_t \cdot, x_t^c, \alpha, \sigma) \circ dW_t^j, \quad d\alpha_t = 0, \quad d\sigma_t = 0,$$

where g_j^c is a random polynomial of order N in x^c and M in (α, σ), and the (approximate) stochastic center manifold is given by its graph

$$\mathbb{R}^{d_c} \times \mathbb{R}^{m+1} \ni (x^c, \alpha, \sigma) \mapsto m_c(\cdot, x^c, \alpha, \sigma) \in \mathbb{R}^{d_s},$$

where m_c is a random polynomial of order N in x^c and M in (α, σ).

The prototypical **Duffing–van der Pol oscillator**

$$\ddot{y} = \alpha y + \beta \dot{y} - y^3 - y^2 \dot{y} \tag{4.3}$$

under the influence of parametric and additive noise has been the subject of numerous investigations (see Arnold, Sri Namachchivaya and Schenk–Hoppé [7], Schenk–Hoppé [17],[18],[19], and the references therein). For $\alpha < 0$ fixed and β the bifurcation parameter, the system (4.3) exhibits a Hopf bifurcation for $\beta = 0$. For $\beta < 0$ fixed and α the bifurcation parameter, it undergoes a pitchfork bifurcation at $\alpha = 0$.

To reduce complexity, we will treat a particular case of parametric noise: Let the parameter α be replaced by $\alpha + \sigma \dot{W}$, where $\dot{W}$ stands, as usual, for white noise, and σ is a strength parameter. With $x = \binom{y}{\dot{y}}$, the perturbed version of (4.3) is

$$dx_t = \left(\begin{pmatrix} 0 & 1 \\ \alpha & \beta \end{pmatrix} x_t + \begin{pmatrix} 0 \\ -x_{1,t}^3 - x_{1,t}^2 x_{2,t} \end{pmatrix} \right) dt + \sigma \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} x_t \circ dW_t. \tag{4.4}$$

Here W is a two–sided one–dimensional Wiener process. The linearization of (4.4) at $x = 0$ is

$$dv_t = \begin{pmatrix} 0 & 1 \\ \alpha & \beta \end{pmatrix} v_t dt + \sigma \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} v_t \circ dW_t. \tag{4.5}$$

The pitchfork scenario under small random perturbations

Put for simplicity $\beta = -1$. Then the eigenvalues of the linear part of the deterministic equation (4.4) ($\sigma = 0$) are $(-1 \pm \sqrt{1 + 4\alpha})/2$. We treat (α, σ) as a small two-dimensional parameter and will determine the simultaneous stochastic normal form and center manifold reduction of (4.4) for small x, α and σ.

The (formidable) calculations are formally the same as in the real noise case (just replace in [9] $\xi(t)$ by $\dot{W}$, with all the obvious consequences). We introduce the auxiliary processes U, V, X, and Y which are the unique stationary solutions of the scalar SDE

$$dU_t = -U_t dt + dW_t, \quad dV_t = V_t dt + dW_t,$$

and

$$dX_t = -X_t dt + dU_t, \quad dY_t = -Y_t dt + U_t \circ dW_t.$$

Note that V anticipates the future of W. The resulting center SDE (up to terms of order 3 in x^c and 2 in (α, σ)) is

$$\begin{aligned} dx_t^c &= (\alpha - \alpha\sigma U_t - \alpha^2) x_t^c dt + (\sigma - \sigma^2 U_t - \alpha\sigma) x_t^c \circ dW_t \\ &\quad + (-1 + 2\sigma U_t - \sigma^2(3U_t V_t + 4Y_t + 2U_t + U_t^2) \\ &\quad + 3\alpha - \alpha\sigma(15U_t + 7X_t + 4V_t) - 18\alpha^2)(x_t^c)^3 dt \\ &\quad + (\sigma - \sigma^2(4U_t + 3X_t + V_t) - 10\alpha\sigma)(x_t^c)^3 \circ dW_t. \end{aligned} \tag{4.6}$$

Note that this is an anticipative SDE since V is anticipative.

The approximate center manifold is the random function $(x^c, \alpha, \sigma) \mapsto m_c(\cdot, x^c, \alpha, \sigma) = x^s$ given by

$$\begin{aligned} m_c(\cdot, x^c, \alpha, \sigma) &= (-\sqrt{2}\sigma U_0 + 2\sqrt{2}\sigma^2 Z_0 - \sqrt{2}\alpha \\ &\quad + 2\sqrt{2}\alpha\sigma(U_0 + X_0) + 2\sqrt{2}\alpha^2) x^c \\ &\quad + (\sqrt{2} - \sqrt{2}\sigma(4U_0 + 3X_0) \\ &\quad + \sigma^2 H_{3002}^s - 7\sqrt{2}\alpha + \alpha\sigma H_{3011}^s - 47\sqrt{2}\alpha^2)(x^c)^3, \end{aligned}$$

where H_{3002}^s and H_{3011}^s are certain anticipative random variables which are very complicated functions of lower order terms.

The computational effort for these results is enormous and could only be accomplished by using the computer algebra program MAPLE. There are 106 cohomological equations to be solved to determine the coefficients. For the real noise case, these cohomological equations fill 18 pages and can be found in an appendix to a reprint version of [9].

The scalar SDE (4.6) can now be utilized for stochastic bifurcation theory of the two–dimensional SDE (4.4), similarly as in Arnold and Boxler [4].

For example, linearizing the SDE (4.6) at $x_c = 0$ gives the scalar SDE

$$dv_t = (\alpha - \alpha\sigma U_t - \alpha^2) v_t dt + (\sigma - \sigma^2 U_t - \alpha\sigma) v_t \circ dW_t$$

which can be explicitly solved to give its Lyapunov exponent

$$\lambda_c(\alpha, \sigma) = \alpha - \alpha^2 - \sigma^2 \lim_{t\to\infty} \frac{1}{t} \int_0^t U_s \circ dW_s = \alpha - \alpha^2 - \frac{\sigma^2}{2},$$

where we have used

$$\int_0^t U_s \circ dW_s = \int_0^t U_s\, dW_s + \frac{t}{2}$$

and $\int_0^t U_s\, dW_s/t \to 0$ $\mathbb{P}$-a. s. This is in full agreement with an asymptotic formula for the top Lyapunov exponent $\lambda_1(\alpha, \sigma)$ of the two–dimensional linear SDE (4.5) (for $\beta = -1$) by Pardoux and Wihstutz [16] (Theorem 5.3) which gives

$$\lambda_1(\alpha, \sigma) = \lambda_c(\alpha, \sigma) + O(\sigma^4).$$

In particular $\lambda_1(0, \sigma) = -\frac{\sigma^2}{2} + O(\sigma^4)$, hence the disturbed system is still stable at the deterministic critical value $\alpha = 0$, and the stochastic pitchfork bifurcation (bifurcation of two new invariant measures from the trivial reference measure δ_0) is delayed to the parameter value

$$\alpha_c = \frac{\sigma^2}{1 + \sqrt{1 - 2\sigma^2}} + \ldots = \frac{\sigma^2}{2} + \ldots.$$

at which $\lambda_1(\alpha_c, \sigma) = 0$.

The Hopf scenario under small random perturbations

We now assume that $\alpha < 0$ is fixed and β is the bifurcation parameter which varies in the interval $|\beta| < \sqrt{-4\alpha}$ so that the frequency

$$\omega_d := \sqrt{-\alpha - \beta^2/4}$$

is well–defined. We follow [7] and obtain the normal form for small x, β and σ. The linearized SDE (4.5) is for $\sigma = 0$

$$\dot{v} = \begin{pmatrix} 0 & 1 \\ \alpha & \beta \end{pmatrix} v,$$

with eigenvalues $\frac{\beta}{2} \pm i\omega_d$ which are purely imaginary for $\beta = 0$. We have dimension $d = d_c = 2$ for the center space, while the stable component is not present. All cohomological equations are hence resonant (in the random sense, i. e., all Lyapunov exponents vanish). Our recipe for their solution is as follows: If in a cohomological equation the right–hand side is random, we choose $h_n = 0$ (hence do not simplify g_n), while if the right–hand side is nonrandom we search for a nonrandom h_n for which g_n is "as simple as possible" by means of deterministic normal form theory.

The truncated stochastic normal form for $N = 3$ and $M = 1$ in polar coordinates (r, φ), $x_1 = r\cos\varphi$, $x_2 = r\sin\varphi$, is as follows (again after formidable computational efforts which we omit):

$$
\begin{aligned}
dr_t &= \left(\frac{\beta}{2}r_t - \frac{1}{2}r_t^3\right) dt + \\
&\quad \frac{r_t}{8\omega_d^2}\left(-2r_t^2 + \left(\frac{5r_t^2}{\omega_d} - 4\omega_d\right)\sin 2\varphi_t - \frac{2r_t^2}{\omega_d}\sin 4\varphi_t \right.\\
&\quad \left. + r_t^2 \cos 2\varphi_t\right) \sigma \circ dW_t,
\end{aligned}
$$

$$
\begin{aligned}
d\varphi_t &= \left(\omega_d + \frac{6+3\beta}{4\omega_d} r_t^2\right) dt + \\
&\quad \frac{1}{2\omega_d}\left(-1 + \frac{3r_t^2}{\omega_d^2} + \left(-1 + \frac{5r_t^2}{2\omega_d^2}\right)\cos 2\varphi_t - \frac{r_t^2}{2\omega_d^2}\cos 4\varphi_t \right.\\
&\quad \left. - \frac{3r_t^2}{2\omega_d}\sin 2\varphi_t\right) \sigma \circ dW_t.
\end{aligned}
$$

For $\sigma = 0$ we recover the deterministic truncated normal form, from which the Hopf bifurcation at $\beta = 0$ can be read off. For $\sigma \neq 0$, the corresponding Pardoux–Wihstutz formula ([16], Theorem 4.1) for the Lyapunov exponents of (4.5) yields for fixed $\alpha < 0$

$$\lambda_{1,2}(\beta, \sigma) = \frac{\beta}{2} \pm \frac{\sigma^2}{8\omega_d^2} + O(\sigma^4).$$

At $\beta = 0$, $\lambda_2(0, \sigma) < 0 < \lambda_1(0, \sigma)$, so that the trivial solution has already lost its stability. Stochastic Hopf bifurcation consists of the premature bifurcation of a first invariant measure from the trivial one at $\beta_1 < 0$ where $\lambda_1(\beta_1, \sigma) = 0$, and then of a second measure at $\beta_2 > 0$ where $\lambda_2(\beta_2, \sigma) = 0$.

References

[1] D. V. Anosov and V. I. Arnold. *Dynamical systems I.* Springer, Berlin Heidelberg New York, 1988.

[2] L. Arnold. *Random dynamical systems.* Preliminary version, 1994.

[3] L. Arnold. Six lectures on random dynamical systems. In R. Johnson, editor, *Dynamical Systems (CIME Summer School 1994)*, volume 1609 of *Lecture Notes in Mathematics*, pages 1–43. Springer, Berlin Heidelberg New York, 1995.

[4] L. Arnold and P. Boxler. Stochastic bifurcation: instructive examples in dimension one. In M. Pinsky and V. Wihstutz, editors, *Diffusion processes*

and related problems in analysis, volume II: Stochastic flows, volume 27 of *Progress in Probability*, pages 241–255. Birkhäuser, Boston Basel Stuttgart, 1992.

[5] L. Arnold and P. Imkeller. Stratonovich calculus with spatial parameters and anticipative problems in multiplicative ergodic theory. *Stochastic Processes and their Applications*, 62:19–54, 1996.

[6] L. Arnold and M. Scheutzow. Perfect cocycles through stochastic differential equations. *Probab. Th. Rel. Fields*, 101:65–88, 1995.

[7] L. Arnold, N. Sri Namachchivaya, and K. R. Schenk–Hoppé. Toward an understanding of stochastic Hopf bifurcation: a case study. *International Journal of Bifurcation and Chaos*, 1996.

[8] L. Arnold and Xu Kedai. Normal forms for random diffeomorphisms. *J. Dynamics and Differential Equations*, 4:445–483, 1992.

[9] L. Arnold and Xu Kedai. Simultaneous normal form and center manifold reduction for random differential equations. In C. Perelló, C. Simó, and J. Solá-Morales, editors, *EQUADIFF–91*, volume 1, pages 68–80. World Scientific, Singapore, 1993.

[10] L. Arnold and Xu Kedai. Normal forms for random differential equations. *Journal of Differential Equations*, 116:484–503, 1995.

[11] M. Barlow and M. Yor. Semi–martingale inequalities via the Garsia–Rodemich–Rumsey lemma and applications to local times. *Journal of Functional Analysis*, 49:198–229, 1982.

[12] P. H. Coullet, C. Elphick, and E. Tirapegui. Normal form of a Hopf bifurcation with noise. *Pysics Letters*, 111A:277–282, 1985.

[13] C. Elphick, E. Tirapegui, M. E. Brachet, P. H. Coullet, and G. Iooss. A simple global characterization for normal forms of singular vector fields. *Physica D*, 29:95–127, 1987.

[14] A. Katok and B. Hasselblatt. *Introduction to the modern theory of dynamical systems.* Cambridge University Press, 1995.

[15] C. Nicolis and G. Nicolis. Normal form analysis of stochastically forced dynamical systems. *Dynam. Stabil. Syst.*, 1:249–253, 1986.

[16] E. Pardoux and V. Wihstutz. Lyapunov exponent and rotation number of two–dimensional linear stochastic systems with small diffusion. *SIAM J. Appl. Math.*, 48:442–457, 1988.

[17] K. R. Schenk–Hoppé. Bifurcation scenarios of the noisy Duffing–van der Pol oscillator. *Nonlinear Dynamics*, 1996.

[18] K. R. Schenk–Hoppé. Deterministic and stochastic Duffing–van der Pol oscillators are non–explosive. *ZAMP*, 1996.

[19] K. R. Schenk–Hoppé. *The stochastic Duffing–van der Pol equation.* PhD thesis, Institut für Dynamische Systeme, Universität Bremen, 1996.

[20] G. Schöner and H. Haken. The slaving principle for Stratonovich stochastic differential equations. *Z. Phys. B*, 63:493–504, 1986.

[21] G. Schöner and H. Haken. A systematic elimination procedure for Ito stochastic differential equations and the adiabatic approximation. *Z. Phys. B*, 68:89–103, 1987.

[22] N. Sri Namachchivaya and Y. K. Lin. Methods of stochastic normal forms. *Int. J. Nonlin. Mech.*, 26:931–943, 1991.

[23] A. Vanderbauwhede. Center manifolds, normal forms and elementary bifurcations. In U. Kirchgraber and H. O. Walter, editors, *Dynamics Reported*, volume 2, pages 89–169. Teubner and Wiley, New York, 1989.