Normal Form and Exact Feedback Linearisation of Nonlinear Stochastic Systems: the Ideal Case

Alberto Mellone and Giordano Scarciotti

***Abstract*— This paper introduces the concepts of stochastic relative degree, normal form and exact feedback linearisation for single-input single-output nonlinear stochastic systems. The systems are defined by stochastic differential equations in which both the drift and the diffusion terms are nonlinear functions of the states and the control input. First, we define new differential operators and the concept of stochastic relative degree. Then we introduce a suitable coordinate change and we show that the dynamics of the transformed state has a simplified structure, which we name normal form. Finally, we show that a condition on the stochastic relative degree of the system is sufficient for it to be rendered linear via a coordinate change and a nonlinear feedback. We provide an analytical example to illustrate the theory.**

I. INTRODUCTION

The theory of normal forms is a fundamental topic in the analysis and control of nonlinear deterministic systems. Finding the normal form of a nonlinear system consists in determining a suitable local change of coordinates such that the transformed system is described by "simpler" differential equations. The description of the system in the new form makes it easier to draw properties of the system, as well as to design, where possible, observers and feedback control laws yielding exact linearisation, asymptotic stabilisation, asymptotic output tracking and disturbance decoupling [1]. The first work explicitly addressing the convenience of coordinate transformation to describe a nonlinear system in a simpler way is [2], where, in particular, the coordinate change was used to solve the static state feedback non-interacting control problem. Further works dealing with similar topics are [3], [4], [5]. The exact linearisation of single-input single-output systems was firstly addressed and solved in [6], while the extension to multi-input systems was provided in [7]. The procedure for designing the linearising transformation can be found in [8] and [9], while the existence of global transformations was addressed in [10].

In this paper, we seek a normal form for a class of nonlinear stochastic dynamical systems. Stochastic systems are common tools in the modelling of uncertain processes, as uncertainties arising from approximate models can be represented by stochastic differential equations [11]. Examples of applications of stochastic systems theory can be found in the optimal stopping problem, production planning, finance, technology diffusion and research funding [11], [12].

Alberto Mellone and Giordano Scarciotti are with the Department of Electrical and Electronic Engineering, Imperial College London, London, SW7 2AZ, UK. {a.mellone18,g.scarciotti}@imperial.ac.uk

While the concept of relative degree and the definition of normal forms for control purposes for stochastic systems has not been introduced yet, a normal form was proposed in [13] and, subsequently, in [14]. In these works the multiplicative ergodic theorem and Stratonovich calculus are employed to derive a normal form theory for a class of stochastic differential equations characterised by a pure diffusion term. The transformation yielding the normal form requires anticipating the noise over a short time scale. Further applications of coordinate changes applied to stochastic differential equations can be found, *e.g.*, in [15], where symmetries for stochastic differential equations are introduced, or in [16], where a normal form allows separating between fast and slow stochastic dynamics.

In this paper we address a general class of nonlinear stochastic single-input single-output systems. Namely, both the drift and the diffusion terms of the stochastic differential equation are nonlinear functions of the state and the input, while the output is a nonlinear function of the state. The focus of the paper is on defining a coordinate change for the stochastic system such that the dynamics of the transformed system is described by stochastic differential equations that are simpler, as well as more meaningful from the perspective of control system design. The normal form we introduce follows from the notion of stochastic relative degree, for which we propose a definition. Moreover, we give a sufficient condition for the system to be locally linearised by employing a nonlinear state feedback. It is shown that the measurement of the white noise is needed to perform exact linearisation (and this is somewhat consistent with the anticipating nature of other stochastic changes of coordinates such as [13]). Obviously, measuring the white noise is an unrealistic hypothesis in practical applications. Nevertheless, in this paper we develop the mathematical theory in the *ideal* case, *i.e.* the white noise is measurable, as a preliminary step for a future analysis where the white noise is estimated *a posteriori*. We then plan to reformulate the results on exact feedback linearisation in future publications, adapting the approximation procedure presented in [17], [18] and [19].

The rest of the paper is organised as follows. In Section II we recall some preliminary notions on stochastic systems and differential operators. In Section III we define new differential operators that are used in the remainder of the paper and define the concept of stochastic relative degree. Moreover, we introduce a coordinate transformation that brings the system into a suitable normal form. In Section IV we formulate the problem of exact linearisation via state feedback for stochastic systems and show that a static nonlinear state feedback solves the problem, *i.e.* we can locally render the system linear and controllable under a suitable coordinate change. Finally, Section V contains some concluding remarks.

Notation. The symbol $\mathbb{Z}$ denotes the set of integer numbers, while $\mathbb{R}$ and $\mathbb{C}$ denote the fields of real and complex numbers, respectively; by adding the subscript "< 0" ("≥ 0", "0") to any symbol indicating a set of numbers, we denote that subset of numbers with negative (non-negative, zero) real part. The symbol ∂_x^n is used as a shorthand for the operator $\partial^n/\partial x^n$, while $\alpha^{(n)}$ indicates the n-th time derivative of α. $(\nabla, \mathfrak{A}, \mathfrak{P})$ is a probability space given by the set ∇, the σ-algebra $\mathfrak{A}$ defined on ∇ and the probability measure $\mathfrak{P}$ on the measurable space $(\nabla, \mathfrak{A})$. A *stochastic process* with state space $\mathbb{R}^n$ is a family $\{x_t,\, t \in \mathbb{R}\}$ of $\mathbb{R}^n$-valued random variables, *i.e.* for every fixed $t \in \mathbb{R}$, $x_t(\cdot)$ is an $\mathbb{R}^n$-valued random variable and, for every fixed $w \in \nabla$, $x_\cdot(w)$ is an $\mathbb{R}^n$-valued function of time [20, Section 1.8]. For ease of notation, we often indicate a stochastic process $\{x_t,\, t \in \mathbb{R}\}$ simply with x_t (this is common in the literature, see *e.g.* [20]). With a slight abuse of notation, any subscript different from the symbol "t" indicates the corresponding component of the vector x_t, *e.g.* x_i is the i-th component of the vector x_t.

II. Preliminaries

In this section we shortly recall the theory of generalised stochastic processes and define differential operators that will be used in the remainder of the paper.

Let $C_0^\infty(\mathbb{R})$ be the space of all infinitely differentiable functions on $\mathbb{R}$ with compact support [21, Definition 1.2.1]. The following definition characterises the notion of *distribution*, or, equivalently, of *generalised function*.

Definition 1. [22, Definition 3.1] Let X be an open subset of $\mathbb{R}$. A *distribution on* X is a linear form ψ on $C_0^\infty(\mathbb{R})$ that is also continuous in the sense that

$$\lim_{j\to\infty} \psi(\varphi_j) = \psi(\varphi) \qquad \text{as} \qquad \lim_{j\to\infty} \varphi_j = \varphi \quad \text{in} \quad C_0^\infty(\mathbb{R}).$$

If f is a continuous and differentiable function, then by partial integration we get

$$\int_{\mathbb{R}} \dot{f}\varphi\, dt = -\int_{\mathbb{R}} f\dot{\varphi}\, dt, \quad \forall \varphi \in C_0^\infty(\mathbb{R}).$$

By analogy, for any distribution ψ, its distributional derivative $\dot{\psi}$ is defined as the distribution that satisfies

$$\dot{\psi}(\varphi) = -\psi(\dot{\varphi}), \quad \forall \varphi \in C_0^\infty(\mathbb{R}),$$

see [21, Definition 3.1.1]. Note also that generalised functions have derivatives of all order, which are generalised functions as well.

Definition 2. [20, Section 3.2] A generalised stochastic process is a random generalised function in the sense that a random variable $\psi(\varphi)$ is assigned to every $\varphi \in C_0^\infty$, where ψ is, with probability 1, a generalised function.

We now look at the Brownian motion as a generalised stochastic process. Therefore, its distributional derivative is always defined [20, Section 3.2]. In particular, the generalised stochastic process given by such a derivative has zero mean value and covariance function given by the generalised function $\delta(t-s)$, $t, s \in \mathbb{R}$, *i.e.* the Dirac delta. Consequently, the derivative of the generalised Brownian motion is the generalised *white noise* [20, Section 3.2]. In the remainder, with a slight abuse of notation, we refer to generalised Brownian motion and generalised white noise omitting the attribute "generalised" and we denote them by simply $\mathcal{W}_t$ and ξ_t, respectively, with $\xi_t = \dot{\mathcal{W}}_t$. It should be emphasised that the just mentioned time derivative is meant in the sense of distributions, and not as the limit of the difference quotient as the increment tends to zero, which instead applies to differentiable functions in the classical sense.

Consider the nonlinear single-input, single-output stochastic system expressed in the shorthand integral notation

$$\begin{aligned} dx_t &= (f(x_t) + g(x_t)u)dt + (l(x_t) + m(x_t)u)d\,\mathcal{W}_t, \\ y_t &= h(x_t), \end{aligned} \tag{1}$$

with $x_t \in \mathbb{R}^n$, $u \in \mathbb{R}$, $y_t \in \mathbb{R}$ and $f : \mathbb{R}^n \to \mathbb{R}^n$, $g : \mathbb{R}^n \to \mathbb{R}^n$, $l : \mathbb{R}^n \to \mathbb{R}^n$, $m : \mathbb{R}^n \to \mathbb{R}^n$, $h : \mathbb{R}^n \to \mathbb{R}$ *smooth* functions, *i.e.* they admit continuous partial derivatives of any order. We assume that, for a fixed initial condition $x_{t=0}$, the solution of (1) is unique. Note that, in the light of the previous discussion, system (1) can be rewritten in the following differential notation

$$\dot{x}_t = f(x_t) + g(x_t)u + (l(x_t) + m(x_t)u)\xi_t, \quad y_t = h(x_t). \tag{2}$$

Note that when ξ_t is (generalised) white noise, as in this case, then the differential equation (2) is equivalent to the integral equation (1) if the latter is interpreted in Itô's sense [20, Section 10.3]. Given the equivalence of the two representations in the framework of generalised stochastic processes, in the remainder of the paper equations (1) and (2) are used interchangeably, as convenient, to refer to the same nonlinear stochastic system.

Recall that the derivative of h along the vector field f, which is called Lie derivative and is indicated with the symbol $\mathcal{L}_f\, h$, is defined as

$$\mathcal{L}_f\, h(x) = \partial_x[h]\ f(x) = \sum_{i=1}^{n} \frac{\partial h}{\partial x_i} f_i(x).$$

We indicate the derivative of h first along the vector field f and then along the vector field g as $\mathcal{L}_g\,\mathcal{L}_f\, h(x) = \partial_x[\mathcal{L}_f\, h]\ g(x)$. We use the recursive relation $\mathcal{L}_f^k\, h(x) = \partial_x[\mathcal{L}_f^{k-1}\, h]\ f(x)$, with $\mathcal{L}_f^0\, h(x) = h(x)$, to indicate the k-th differentiation of h along f.

III. Coordinate Transformation and Normal Form

In this section we introduce the concept of stochastic relative degree and show that a suitable coordinate transformation brings the system into a simpler form, which is convenient for analysis and stabilisation.

We first introduce three new operators. The first one, which indicates the second derivative of h along the vector fields f

and g, is defined as

$$ {}^{g}\mathcal{G}_{f}h(x) = g(x)^{\top}\partial_x^2[h]\ f(x) = \sum_{j=1}^{n} g_j(x) \sum_{i=1}^{n} \frac{\partial^2 h}{\partial x_j \partial x_i} f_i(x). $$

Similarly to the Lie derivative, we use the notation ${}^{b}\mathcal{G}_{a}\,{}^{g}\mathcal{G}_{f}h(x) = b(x)^{\top}\partial_x^2[{}^{g}\mathcal{G}_{f}h]\ a(x)$, and ${}^{g}\mathcal{G}_{f}^{k}h(x) = g(x)^{\top}\partial_x^2[{}^{g}\mathcal{G}_{f}^{k-1}h]\ f(x)$, to indicate the reiterated operations. The operator ${}^{l}\mathcal{S}_{f}h$ is employed to define the stochastic Lie derivative of h along the drift vector field f and diffusion vector field l as

$$ {}^{l}\mathcal{S}_{f}h(\xi_t, x) = \mathcal{L}_f\, h(x) + \mathcal{L}_l\, h(x)\xi_t + \frac{1}{2}\,{}^{l}\mathcal{G}_{l}h(x). $$

Similarly to the deterministic Lie derivative, if ${}^{l}\mathcal{S}_{f}h(\xi_t, x) = {}^{l}\mathcal{S}_{f}h(x)$ is a deterministic expression, *i.e.* the white noise does not appear explicitly, we use the notation ${}^{l}\mathcal{S}_{f}^{2}h(\xi_t, x) = {}^{l}\mathcal{S}_{f}\,{}^{l}\mathcal{S}_{f}h(\xi_t, x)$ and, iteratively, if ${}^{l}\mathcal{S}_{f}^{k-1}h(\xi_t, x) = {}^{l}\mathcal{S}_{f}^{k-1}h(x)$ is deterministic, ${}^{l}\mathcal{S}_{f}^{k}h(\xi_t, x) = {}^{l}\mathcal{S}_{f}\,{}^{l}\mathcal{S}_{f}^{k-1}h(\xi_t, x)$, with ${}^{l}\mathcal{S}_{f}^{0}h(x) = h(x)$ by definition. Finally, we define the third operator

$$ {}^{m}_{g}\mathcal{A}_{l}h(\xi_t, x) = \mathcal{L}_g\, h(x) + \mathcal{L}_m\, h(x)\xi_t + {}^{m}\mathcal{G}_{l}h(x). $$

By using Itô's formula, it is easy to see that the first derivative of the output of system (2) is given by

$$ y_t^{(1)} = {}^{l}\mathcal{S}_{f}h(\xi_t, x_t) + {}^{m}_{g}\mathcal{A}_{l}h(\xi_t, x_t)u + \frac{1}{2}\,{}^{m}\mathcal{G}_{m}h(x_t)u^2. \tag{3} $$

We now define the concept of stochastic relative degree and then point out the rationale of such a definition.

Definition 3. (Stochastic Relative Degree) Assume that there exists $\bar{r}$ such that

$$ \mathcal{L}_l\,{}^{l}\mathcal{S}_{f}^{k}h(\xi_t, x) = 0, \qquad \forall\, k \in \{0, ..., \bar{r}-2\}, \tag{4} $$

and for all x in a neighbourhood of $\bar{x}$. System (2) is said to have *stochastic relative degree* r at a point $\bar{x}$ if $\bar{r} = r$ and

1) $\mathcal{L}_g\,{}^{l}\mathcal{S}_{f}^{k}h(\xi_t, x) + {}^{m}\mathcal{G}_{l}\,{}^{l}\mathcal{S}_{f}^{k}h(\xi_t, x) = 0$ and $\mathcal{L}_m\,{}^{l}\mathcal{S}_{f}^{k}h(\xi_t, x) = 0$ and ${}^{m}\mathcal{G}_{m}\,{}^{l}\mathcal{S}_{f}^{k}h(\xi_t, x) = 0$ for all x in a neighborhood of $\bar{x}$ and all $k \in \{0, ..., r-2\}$.
2) $\mathcal{L}_g\,{}^{l}\mathcal{S}_{f}^{r-1}h(\xi_t, \bar{x}) + {}^{m}\mathcal{G}_{l}\,{}^{l}\mathcal{S}_{f}^{r-1}h(\xi_t, \bar{x}) \neq 0$ or $\mathcal{L}_m\,{}^{l}\mathcal{S}_{f}^{r-1}h(\xi_t, \bar{x}) \neq 0$ or ${}^{m}\mathcal{G}_{m}\,{}^{l}\mathcal{S}_{f}^{r-1}h(\xi_t, \bar{x}) \neq 0$.

Before discussing the implications of the definition just given, we provide two lemmas and a standing assumption which are useful in clarifying the meaning of Definition 3.

Lemma 1. Let $x \in U \subset \mathbb{R}^n$ and $k \in \{0, ..., r-2\}$. Then ${}^{m}_{g}\mathcal{A}_{l}\,{}^{l}\mathcal{S}_{f}^{k}h(\xi_t, x) = 0$ for all $\xi_t \in \mathbb{R}$ if and only if $\mathcal{L}_g\,{}^{l}\mathcal{S}_{f}^{k}h(\xi_t, x) + {}^{m}\mathcal{G}_{l}\,{}^{l}\mathcal{S}_{f}^{k}h(\xi_t, x) = 0$ and $\mathcal{L}_m\,{}^{l}\mathcal{S}_{f}^{k}h(\xi_t, x) = 0$.

Lemma 2. Let $\bar{x} \in \mathbb{R}^n$. Then ${}^{m}_{g}\mathcal{A}_{l}\,{}^{l}\mathcal{S}_{f}^{r-1}h(\xi_t, \bar{x}) \neq 0$ almost surely if and only if $\mathcal{L}_g\,{}^{l}\mathcal{S}_{f}^{r-1}h(\xi_t, \bar{x}) + {}^{m}\mathcal{G}_{l}\,{}^{l}\mathcal{S}_{f}^{r-1}h(\xi_t, \bar{x}) \neq 0$ or $\mathcal{L}_m\,{}^{l}\mathcal{S}_{f}^{r-1}h(\xi_t, \bar{x}) \neq 0$.

To understand the meaning of the stochastic relative degree let $x_{t=\bar{t}} = \bar{x}$ be the state of system (2) at time $\bar{t}$ and assume, for instance, that system (2) has stochastic relative degree $r > 2$ at $\bar{x}$. Then, by assumption (4), Lemma 1 and Definition 3, expression (3) simplifies to

$$ y_t^{(1)} = {}^{l}\mathcal{S}_{f}h(\xi_t, x_t) = {}^{l}\mathcal{S}_{f}h(x_t) = \mathcal{L}_f\, h(x) + \frac{1}{2}\,{}^{l}\mathcal{G}_{l}h(x), $$

which does not depend on the white noise or on the control input. Now, computing the second derivative and using assumption (4), Lemma 1 and Definition 3 yields

$$ \begin{aligned} y_t^{(2)} &= {}^{l}\mathcal{S}_{f}^{2}h(\xi_t, x_t) + {}^{m}_{g}\mathcal{A}_{l}\,{}^{l}\mathcal{S}_{f}h(\xi_t, x_t)u + \frac{1}{2}\,{}^{m}\mathcal{G}_{m}\,{}^{l}\mathcal{S}_{f}h(x_t)u^2 \\ &= {}^{l}\mathcal{S}_{f}^{2}h(x_t) = \mathcal{L}_f\,{}^{l}\mathcal{S}_{f}h(x_t) + \frac{1}{2}\,{}^{l}\mathcal{G}_{l}\,{}^{l}\mathcal{S}_{f}h(x_t), \end{aligned} $$

which, again, does not directly depend on the white noise or on the control input. Iterating this procedure yields that, for all $k < r$ and t in a neighbourhood of $\bar{t}$,

$$ y_t^{(k)} = {}^{l}\mathcal{S}_{f}^{k}h(\xi_t, x_t) = {}^{l}\mathcal{S}_{f}^{k}h(x_t), $$

where $y_t^{(k)}$ is a deterministic function, *i.e.* ξ_t does not explicitly appears in $y_t^{(k)}$. In the remainder we will omit the dependency of the operators $\mathcal{S}$ and $\mathcal{A}$ on the white noise ξ_t whenever this does not appear explicitly because of assumption (4). Finally,

$$ \begin{aligned} y_{t=\bar{t}}^{(r)} = {}^{l}\mathcal{S}_{f}^{r}h(\xi_t, \bar{x}) + {}^{m}_{g}\mathcal{A}_{l}\,{}^{l}\mathcal{S}_{f}^{r-1}h(\xi_t, \bar{x})u(\bar{t}) + \\ \frac{1}{2}\,{}^{m}\mathcal{G}_{m}\,{}^{l}\mathcal{S}_{f}^{r-1}h(\bar{x})u(\bar{t})^2, \end{aligned} \tag{5} $$

where by Lemma 2 $u(\bar{t})$ explicitly appears in the expression of $y_{\bar{t}}^{(r)}$. Then, analogously to the deterministic case, the stochastic relative degree is equal to the order of the derivative of the output at time $\bar{t}$ in which the input $u(\bar{t})$ explicitly appears. Two observations are in order: first, while the white noise does not appear in all the derivatives up to order $r-1$ because of (4), it may or may not appear in the r-th derivative; second, differently from the deterministic case, the control u appears linearly and quadratically in (5).

Remark 1. If $\mathcal{L}_l\,{}^{l}\mathcal{S}_{f}^{k}h(x) \neq 0$ for a $k < r-1$, the differentiation of y_t up to the r-th time would require us to introduce successive derivatives of the white noise. We exclude this possibility with the standing assumption (4). The reasons for avoiding this are twofold. The first one is of theoretical nature, as at this stage we are not able to provide a complete theory on the stochastic differential $d\xi_t$ (and successive ones) appearing by applying iteratively Itô's lemma. The second is of numerical nature, as at this stage we are not able to implement the derivative(s) of the white noise.

Remark 2. There might be points where a stochastic relative degree cannot be defined, in analogy with deterministic systems (see, *e.g.*, [1, Section 4.1]). Nevertheless, the set of points where a stochastic relative degree can be defined is open and dense in $\mathbb{R}^n$.

We are now interested in finding a diffeomorphism $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ that locally (*i.e.* in a neighbourhood $\bar{U}$ of $\bar{x} \in U \subset \mathbb{R}^n$) transforms system (2) in such a way that its dynamics is somewhat "simpler".

Remark 3. Since the Jacobian of a diffeomorphism Φ is invertible by definition, the dynamics of a deterministic system in the transformed state is always well-defined. In fact, recall that, if $\dot{x} = f(x,u)$ and $z = \Phi(x)$, then $\dot{z} = (\partial_x[\Phi]\ f(x,u))_{x=\Phi^{-1}(z)} = \tilde{f}(z,u)$ and $\tilde{f}$ is non-zero for any non-zero f. For stochastic systems, the same holds almost surely. To see this, suppose $\dot{x}_t = f(x_t,u) + l(x_t,u)\xi_t$ and let $z_t = \Phi(x_t)$. Then, applying Itô's lemma,

$$\dot{z}_t = (\partial_x[\Phi](f(x_t,u) + l(x_t,u)\xi_t))_{x_t=\Phi^{-1}(z_t)} + \frac{1}{2}\begin{bmatrix} {}^l\mathcal{G}_l\Phi_1(x_t) & \dots & {}^l\mathcal{G}_l\Phi_n(x_t)\end{bmatrix}^\top_{x_t=\Phi^{-1}(z_t)}.$$

Then, the case $l \equiv 0$ is equivalent to the deterministic case. If $l \neq 0$, then the first term on the right-hand side is non-zero and, given the randomness induced by the white noise, almost surely different from the second term. Therefore, a change of coordinates, defined by a diffeomorphism, is sufficient to ensure that the dynamics of the system, expressed in the transformed state applying the Itô chain rule, is well-defined almost surely.

We now make the following assumption on the stochastic Lie derivatives of $y_t = h(x_t)$ along the drift vector f and the diffusion vector l.

Assumption 1. Let r be the stochastic relative degree of system (2) at $\bar{x}$. Then the row vectors

$$\partial_x[h]_{x=\bar{x}},\ \partial_x[{}^l\mathcal{S}_f h]_{x=\bar{x}},\ \dots,\ \partial_x[{}^l\mathcal{S}_f^{r-1}h]_{x=\bar{x}},$$

are linearly independent.

Observe that if Assumption 1 holds, then necessarily $r \leq n$.

Remark 4. For deterministic nonlinear systems Assumption 1, *i.e.* the linear independence of the gradients of the first $r-1$ successive derivatives of the output at $\bar{x}$, is a fact that can be proved, see *e.g*, [1, Lemma 4.1.1]. The proof of the stochastic counterpart is a topic under investigation at this stage. Nevertheless, making this assumption is sufficient to develop the theory presented in the remainder of the paper. At present, no counter-example has been found for which this property is not satisfied.

Proposition 1. Suppose that system (2) has stochastic relative degree r at $\bar{x}$ and let Assumption 1 hold. Set

$$\phi_1(x)=h(x),\ \ \phi_2(x)={}^l\mathcal{S}_f h(x),\ \ \dots,\ \ \phi_r(x)={}^l\mathcal{S}_f^{r-1}h(x).$$

If $r < n$, then there exist smooth functions $\phi_{r+1}(x), ..., \phi_n(x)$, with $\phi_j \in \mathbb{R}$ for all $j \in \{r+1,..,n\}$, such that the Jacobian of the mapping

$$\Phi(x) = \begin{bmatrix}\phi_1(x) & \phi_2(x) & \dots & \phi_n(x)\end{bmatrix}^\top$$

is invertible at $\bar{x}$ almost surely, thus defining a coordinate transformation in a neighbourhood of $\bar{x}$. Then the state-space representation of system (2) in the transformed state $z_t = \Phi(x_t)$ is

$$\begin{aligned}
\dot{z}_i &= z_{i+1}, & i &= 1,...,r-1\\
\dot{z}_r &= c(\xi_t,z_t) + b(\xi_t,z_t)u + a(z_t)u^2, & &\\
\dot{z}_j &= p_j(\xi_t,z_t) + q_j(\xi_t,z_t)u + s_j(z_t)u^2, & j &= r+1,...,n,
\end{aligned}$$

where

$$\begin{aligned}
c(\xi_t,z_t) &= {}^l\mathcal{S}_f^r h(\xi_t,\Phi^{-1}(z_t)),\\
b(\xi_t,z_t) &= {}^m_g\mathcal{A}_l\,{}^l\mathcal{S}_f^{r-1}h(\xi_t,\Phi^{-1}(z_t)),\\
a(z_t) &= \frac{1}{2}\,{}^m\mathcal{G}_m\,{}^l\mathcal{S}_f^{r-1}h(\Phi^{-1}(z_t)),\\
p_j(\xi_t,z_t) &= {}^l\mathcal{S}_f\phi_j(\xi_t,\Phi^{-1}(z_t)),\\
q_j(\xi_t,z_t) &= {}^m_g\mathcal{A}_l\phi_j(\xi_t,\Phi^{-1}(z_t)),\\
s_j(z_t) &= \frac{1}{2}\,{}^m\mathcal{G}_m\phi_j(\Phi^{-1}(z_t)),
\end{aligned}$$

with the output $y_t = z_1$. According to the definitions of the operators $\mathcal{S}$ and $\mathcal{A}$ given in Section III, the dependency of c, b, p_j and q_j on the white noise ξ_t is linear.

Note that it might be possible to find smooth functions $\phi_{r+1},...,\phi_n$ such that the dynamics of the last $n-r$ transformed coordinates is independent of the input u, *i.e.* $q_j(\cdot,z_t) \equiv 0$, $s_j(\cdot,z_t) \equiv 0$, for all $j \in \{r+1,...,n\}$, in a neighbourhood of $\Phi(\bar{x})$. This observation motivates the next definition.

Definition 4. (Stochastic Normal Form) Let x_t be the unique solution of (2) and $z_t = \Phi(x_t)$ be a local diffeomorphism in a subset U of $\mathbb{R}^n$ such that

$$\begin{aligned}
\dot{z}_i &= z_{i+1}, & i &= 1,...,r-1,\\
\dot{z}_r &= c(\xi_t,z_t) + b(\xi_t,z_t)u + a(z_t)u^2, & &\\
\dot{z}_j &= p_j(\xi_t,z_t), & j &= r+1,...,n,\\
y_t &= z_1. & &
\end{aligned} \tag{6}$$

System (6) is said to be the *stochastic normal form* of system (2).

Obviously, if the stochastic relative degree at $\bar{x}$ is equal to the order of the system, then the system admits a stochastic normal form in a neighbourhood U of $\bar{x}$.

Remark 5. While for deterministic systems it can be proved (see, *e.g.*, [1, Proposition 4.1.3]) that functions $\phi_{r+1},...,\phi_n$ always exist such that a normal form exists when $r < n$, in the stochastic case the validity of an analogous result is currently under investigation.

Example 1. Consider the nonlinear stochastic system

$$\dot{x}_t = \begin{bmatrix}\sin(x_3)^2(x_1^2 - x_2^2 + x_3)\\ \cos(x_3)^2(x_1^2 - x_2^2 + x_3)\\ f_3(x)\end{bmatrix} + \begin{bmatrix}\cos(x_2)\\ -\cos(x_2)\\ -2\cos(x_2)(x_1+x_2)\end{bmatrix}u + \mu\begin{bmatrix}x_1x_3\\ -x_1x_3\\ -2x_1x_3(x_1+x_2)\end{bmatrix}\xi_t + \begin{bmatrix}x_3\\ -x_3\\ -2x_3(x_1+x_2)\end{bmatrix}u\xi_t,$$

with $\mu \in \mathbb{R}$ and

$$f_3(x) = x_1^\beta + (1-\beta)x_2 + \\ -2(x_1 \sin(x_3)^2 - x_2 \cos(x_3)^2)(x_1^2 - x_2^2 + x_3),$$

$\beta \in \mathbb{Z}_{>0}$. We study the system in a neighbourhood $U \subseteq (-\pi/2, \pi/2)^3$ of $\bar{x} = 0$. Let the output be $y_t = h(x_t) = x_1 + x_2$. Set $z_1 = h(x_t)$, then it is straightforward to compute its derivative

$$\dot{z}_1 = y_t^{(1)} = \dot{x}_1 + \dot{x}_2 = x_1^2 - x_2^2 + x_3.$$

Since neither ξ_t nor u appears explicitly in the previous expression, we conclude that assumption (4) is satisfied so far and the stochastic relative degree of the system, if defined, is larger than 1 at $\bar{x} = 0$. Setting $z_2 = \dot{z}_1$, we proceed to compute the second derivative of the output. Using Itô's lemma we have

$$dz_2 = \begin{bmatrix} 2x_1 & -2x_2 & 1 \end{bmatrix} dx_t + \frac{1}{2} dx_t^\top \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} dx_t,$$

which yields

$$\dot{z}_2 = y_t^{(2)} = x_1^\beta + (1-\beta)x_2.$$

The standing assumption (4) is still satisfied. Moreover, the stochastic relative degree of the system, if defined, is 3 at $\bar{x} = 0$, hence we set $z_3 = \dot{z}_2$. Using Itô's lemma on z_3 we obtain

$$dz_3 = \begin{bmatrix} \beta x_1^{\beta-1} & 1-\beta & 0 \end{bmatrix} dx_t + \\ \frac{1}{2} dx_t^\top \begin{bmatrix} \beta(\beta-1)x_1^{\beta-2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} dx_t.$$

Setting

$$\begin{aligned} \tilde{c}(\xi_t, x_t) &= \beta x_1^{\beta-1}(\sin(x_3)^2(x_1^2 - x_2^2 + x_3) + x_1 x_3) + \\ &\quad (1-\beta)(\cos(x_3)^2(x_1^2 - x_2^2 + x_3)) + \\ &\quad \mu x_1 x_3(\beta x_1^{\beta-1} - (1-\beta))\xi_t + \\ &\quad \tfrac{1}{2}\beta(\beta-1)x^{\beta-2}(\mu x_1 x_3)^2, \\ \tilde{b}(\xi_t, x_t) &= (\beta x_1^{\beta-1} - (1-\beta))(\cos(x_2) + x_3\xi_t) + \\ &\quad \beta(\beta-1)x^{\beta-2}x_1 x_3^2, \\ \tilde{a}(x_t) &= \tfrac{1}{2}\beta(\beta-1)x^{\beta-2}x_3^2, \end{aligned}$$

we finally get

$$\dot{z}_3 = y_t^{(3)} = \tilde{c}(\xi_t, x_t) + \tilde{b}(\xi_t, x_t)u + \tilde{a}(x_t)u^2.$$

It is easily checked that, for all $\beta \in \mathbb{Z}_{>0}$, $\tilde{b}$ is non-zero at $\bar{x} = 0$, therefore the relative degree of the system at the origin is 3. Observe that the function

$$z_t = \Phi(x_t) = \begin{bmatrix} x_1 + x_2 \\ x_1^2 - x_2^2 + x_3 \\ x_1^\beta + (1-\beta)x_2 \end{bmatrix}$$

has Jacobian

$$\partial_x \Phi(x_t) = \begin{bmatrix} 1 & 1 & 0 \\ 2x_1 & -2x_2 & 1 \\ \beta x_1^{\beta-1} & 1-\beta & 0 \end{bmatrix},$$

which is nonsingular at $\bar{x} = 0$, hence it is a local diffeomorphism in U, for all $\beta \in \mathbb{Z}_{>0}$. Substituting $x_t = \Phi^{-1}(z_t)$ in the expression of $\dot{z}_3$, the system expressed in the new coordinates is given by

$$\begin{aligned} \dot{z}_1 &= z_2, \\ \dot{z}_2 &= z_3, \\ \dot{z}_3 &= c(\xi_t, z_t) + b(\xi_t, z_t)u + a(z_t)u^2, \\ y_t &= z_1, \end{aligned}$$

with $c(\xi_t, z_t) = \tilde{c}(\xi_t, \Phi^{-1}(z_t))$, $b(\xi_t, z_t) = \tilde{b}(\xi_t, \Phi^{-1}(z_t))$ and $a(z_t) = \tilde{a}(\Phi^{-1}(z_t))$. The system is in stochastic normal form, since the relative degree coincides with the order of the system.

IV. Exact Linearisation via State Feedback

In this section we give a sufficient condition for the feedback linearisation of stochastic nonlinear systems of the form (2) when the state x_t, as well as the white noise ξ_t, are available for measure. This last assumption is clearly unrealistic, as far as practical applications are concerned. Nevertheless, we hereby develop the theory in an ideal framework where the noise is available for feedback; the obtained results will be preliminary to forthcoming works where the noise will be approximated *a posteriori* via a procedure similar to the one proposed in [17] and [18]. In the remainder of the section we show how the change of coordinates yielding the normal form (6) can be employed to design a static feedback control law which renders the transformed system linear around a set point $\bar{x}$.

First, we formulate the problem we aim to solve.

Problem 1. (State-Space Exact Linearisation Problem) Consider the nonlinear stochastic system without output

$$\dot{x}_t = f(x_t) + g(x_t)u + (l(x_t) + m(x_t)u)\xi_t. \tag{7}$$

Given a point $\bar{x}$, the *state-space exact linearisation problem* consists in finding a neighbourhood U of $\bar{x}$, a feedback law $u_t = k(\xi_t, x_t, v)$, with $v \in \mathbb{R}$, defined on U and a stochastic coordinates transformation $z_t = \Phi(x_t)$ defined on U such that the closed-loop system

$$\dot{x}_t = f(x_t) + g(x_t)k(\xi_t, x_t, v) + (l(x_t) + m(x_t)k(\xi_t, x_t, v))\xi_t,$$

in the coordinates $z_t = \Phi(x_t)$, is linear, deterministic and controllable. ■

In other words, the transformed state-space model that we seek in U has the form $\dot{z} = Az + Bv$ with the matrix $\begin{bmatrix} B & AB & \dots & A^{n-1}B \end{bmatrix}$ with rank n. The solution to this problem is provided by the next proposition.

Proposition 2. If there exists a real valued function $h(x_t)$ such that system (7) with the output $y_t = h(x_t)$ has stochastic relative degree n at $\bar{x}$, then Problem 1 is solvable. In addition, if ${}^m\mathcal{G}_m {}^l\mathcal{S}_f^{n-1} h(x) \equiv 0$, then the control law

$$u_t = \tilde{k}(\xi_t, z_t, v) = \frac{1}{b(\xi_t, z_t)}(-c(\xi_t, z_t) + v) \tag{8}$$

is well-defined almost surely in a neighborhood of $\bar{x}$ and solves Problem 1.

Remark 6. In general b and c are depend on ξ_t, thus implying that the state feedback also requires the knowledge of the exact value of the white noise for all t. Clearly, the feedback linearising control (8) can be implemented in real applications only when the noise does not appear directly in the expression of $\dot{z}_n$. Alternatively, approximations could be provided by *a-posteriori* estimations of the white noise, adapting the results presented in [17], [18].

V. Conclusion and Further Research Directions

In this paper we have introduced the concept of stochastic relative degree and we have used this to define a normal form for nonlinear stochastic systems and to solve the exact feedback linearisation problem. We have pointed out that the computation of the linearising control requires, in general, the knowledge of the white noise. This, however, is unrealistic.

This paper is a first step towards the study and solution of several control problems for stochastic nonlinear system. Future work includes, *e.g.*, the study of the zero dynamics and of the problem of output tracking. On a different note, in order to develop a practically sound theory, it is essential to design state feedbacks which do not require perfect knowledge of the white noise. For instance, the works [17], [18] and [19] can be used as a reference to build an *a-posteriori* estimate of the noise affecting the system, which can then be used to design approximate feedback linearising controllers. These topics are currently under investigation and will appear in forthcoming works.

References

[1] A. Isidori, *Nonlinear Control Systems*, ser. Communications and Control Engineering. Springer-Verlag London, 1995.

[2] A. Isidori, A. Krener, C. Gori-Giorgi, and S. Monaco, "Nonlinear decoupling via feedback: A differential geometric approach," *IEEE Transactions on Automatic Control*, vol. 26, no. 2, pp. 331–345, April 1981.

[3] M. Zeitz, "Controllability canonical (phase-variable) form for non-linear time-variable systems," *International Journal of Control*, vol. 37, no. 6, pp. 1449–1457, 1983.

[4] D. Bestle and M. Zeitz, "Canonical form observer design for non-linear time-variable systems," *International Journal of Control*, vol. 38, no. 2, pp. 419–431, 1983.

[5] A. J. Krener, "Normal forms for linear and nonlinear systems," *Contemporary Mathematics*, vol. 68, pp. 157–189, 1987.

[6] R. Brockett, "Feedback invariants for nonlinear systems," *IFAC Proceedings Volumes*, vol. 11, no. 1, pp. 1115–1120, 1978.

[7] B. Jakubczyk and W. Respondek, "On linearization of control systems," *Bulletin de l'Académie Polonaise des Sciences. Serie des sciences mathematiques*, vol. 28, pp. 517–522, 1980.

[8] R. Su, "On the linear equivalents of nonlinear systems," *Systems & Control Letters*, vol. 2, no. 1, pp. 48–52, 1982.

[9] L. Hunt, , and G. Meyer, "Global transformations of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 28, no. 1, pp. 24–31, January 1983.

[10] W. Dayawansa, W. Boothby, and D. Elliott, "Global state and feedback equivalence of nonlinear systems," *Systems & Control Letters*, vol. 6, no. 4, pp. 229–234, 1985.

[11] B. Øksendal, *Stochastic Differential Equations (Sixth Edition)*. Springer-Verlag, 2003.

[12] J. Yong and X. Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, ser. Stochastic Modelling and Applied Probability. Springer New York, 1999.

[13] L. Arnold and P. Imkeller, "Normal forms for stochastic differential equations," *Probability Theory and Related Fields*, vol. 110, no. 4, pp. 559–588, May 1998.

[14] L. Arnold, *Random Dynamical Systems*, ser. Springer Monographs in Mathematics. Springer-Verlag Berlin Heidelberg, 2003.

[15] G. Gaeta and N. Rodríguez Quintero, "Lie-point symmetries and stochastic differential equations," *Journal of Physics A: Mathematical and General*, vol. 32, no. 48, pp. 8485–8505, 1999.

[16] A. Roberts, "Normal form transforms separate slow and fast modes in stochastic dynamical systems," *Physica A: Statistical Mechanics and its Applications*, vol. 387, no. 1, pp. 12–38, 2008.

[17] A. Mellone and G. Scarciotti, "ε-Approximate Output Regulation of Linear Stochastic Systems: a Hybrid Approach," in *2019 European Control Conference (ECC)*, June 2019, pp. 287–292.

[18] ——, "Error-Feedback Output Regulation of Linear Stochastic Systems: a Hybrid Nonlinear Approach," in *Joint Conference 8th IFAC Symposium on Mechatronic Systems (MECHATRONICS 2019), and 11th IFAC Symposium on Nonlinear Control Systems (NOLCOS 2019)*, September 2019, pp. 907–912.

[19] ——, "Output Regulation of Linear Stochastic Systems," *IEEE Transactions on Automatic Control, under review*, 2020.

[20] L. Arnold, *Stochastic Differential Equations*, ser. A Wiley-Interscience publication. Wiley, 1974.

[21] L. Hörmander, *The Analysis of Linear Partial Differential Operators I*, ser. Classics in Mathematics. Springer-Verlag Berlin Heidelberg, 1983.

[22] J. J. Duistermaat and J. A. C. Kolk, *Distributions*, ser. Cornerstones. Birkäuser Basel, 1995.