Winter 2009

Interconnections of Nonlinear Systems Driven by L_2-ITÔ Stochastic Processes

Luis A. Duffaut Espinosa
Old Dominion University

INTERCONNECTIONS OF NONLINEAR SYSTEMS DRIVEN BY L_2-ITÔ STOCHASTIC PROCESSES

by

Luis A. Duffaut Espinosa
B.Sc. Physics 2003, Universidad Nacional de Ingeniería, Lima, Perú
M.Sc. Mathematics 2007, Pontificia Universidad Católica del Perú, Lima, Perú

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirement for the Degree of

DOCTOR OF PHILOSOPHY

ELECTRICAL AND COMPUTER ENGINEERING

OLD DOMINION UNIVERSITY
December 2009

Approved by:

W. Steven Gray (Director)

N. Rao Chaganty (Member)

Oscar R. González (Member)

Dimitrie C. Popescu (Member)

ABSTRACT

INTERCONNECTIONS OF NONLINEAR SYSTEMS DRIVEN BY L_2-ITÔ STOCHASTIC PROCESSES

Luis A. Duffaut Espinosa
Old Dominion University, 2009
Director: Dr. W. Steven Gray

Fliess operators have been an object of study in connection with nonlinear systems acting on deterministic inputs since the early 1970's. They describe a broad class of nonlinear input-output maps using a type of functional series expansion, but in most applications, a system's inputs have noise components. In such circumstances, new mathematical machinery is needed to properly describe the input-output map via the Chen-Fliess algebraic formalism. In this dissertation, a class of L_2-Itô stochastic processes is introduced specifically for this purpose. Then, an extension of the Fliess operator theory is presented and sufficient conditions are given under which these operators are convergent in the mean-square sense. Next, three types of system interconnections are considered in this context: the parallel, product and cascade connections. This is done by first introducing the notion of a formal Fliess operator driven by a formal stochastic process. Then the generating series induced by each interconnection is derived. Finally, sufficient conditions are given under which the generating series of each composite system is convergent. This allows one to determine when an interconnection of Fliess operators driven by a class of L_2-Itô stochastic processes is well-defined.

To my wife, Lorena

ACKNOWLEDGMENTS

First, I would like to express my deepest gratitude to my advisor, Dr. W. Steven Gray, for his guidance, constant support and warm encouragement during the years I pursued my Ph.D. degree. His persistence and extremely organized research methodology provides a role model for me in my future career.

I am also indebted to Dr. Oscar R. González for his wisdom, effort and quick grasp of the big picture in each of our formal and informal discussions. They have been of great help in my development as a researcher.

Next I want to thank my other committee members, Dr. N. Rao Chaganty (Department of Mathematics and Statistics, Old Dominion University) and Dr. Dimitrie C. Popescu for their invaluable time and comments devoted to improving this dissertation.

I can also not forget the faculty and staff of the Department of Electrical and Computer Engineering at Old Dominion University for all their generous assistance and support.

Mark Twain said, "*Good friends, good books and a sleepy conscience: this is the ideal life.*" I would also like to acknowledge all my colleagues in the Systems Research Lab (SRL) who have stayed and worked with me in good and bad circumstances for so many years.

I also wish to recognize the generous support from the NASA Langley Research Center under contracts NNX07AD52A, NNL04AA03A and NCC-1-03026.

Finally, I would like to thank my family: my wife, Lorena; my mother, Ana Maria; my father, Gabriel; my sister, Dennisse; my brothers, Gabriel and Gonzalo;

my nephew, Gabriel (38); my nieces, Fabiana (39) and Marcela (41); my mother in law, Edda; and my wife's children, Lorenita and Gianfranco for their everlasting love, support and encouragement.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SYMBOLS

$\mathbb{N}$	The set of natural numbers
$\mathbb{Q}$	The set of rational numbers
$\mathbb{R}$	The set of real numbers
$\mathbb{R}^+$	The set of positive real numbers
$\mathcal{T}$	Time interval $[0,T]$, $[0,\infty]$ or $[0,\tau_R]$
Π	Partition of interval $\mathcal{T}$
$\mathcal{C}(\mathcal{T})$	Set of all continuous functions over $\mathcal{T}$
$\mathcal{C}_b^\infty$	Set of all smooth and bounded functions
$\mathbb{1}_A$	Indicator function for the set A
$\#A$	Cardinality of the set A
$\mathcal{B}(\mathbb{R})$	Borel σ-algebra of $\mathbb{R}^n$
$(\Omega,\mathcal{F})$	Measurable space
$(\Omega,\mathcal{F},P)$	Probability space
$(X^*,\mathscr{C},\emptyset)$	Free monoid formed with the alphabet X
$P(A)$	Probability of the event A
$L_2(P\otimes\lambda)$	Space of all square integrable stochastic processes
$L^2(\Omega,\mathcal{F},P)$	Space of all square integrable random variables
$\mathscr{I}$	Set of all L_2-Itô processes
$\lVert\cdot\rVert_p$	Norm for random variables in $L_p^m(\Omega,\mathcal{F},P)$
$\lVert\cdot\rVert_{L_p}$	Norm for stochastic processes in $L_p^m(\Omega\times\mathcal{T},\mathcal{P},P\otimes\lambda)$
$\mathcal{N}$	Set of measure zero in Ω
$\mathcal{F}$	σ-algebra of Ω
$\mathcal{F}_X$	σ-algebra induced by the random variable X
$\mathbf{F}$	Filtration of $(\Omega,\mathcal{F})$
$\mathbf{F}_X$	Natural filtration of stochastic process X
$\mathbf{E}[\cdot]$	Expectation
$\mathbf{E}[\cdot\mid\mathcal{F}]$	Conditional expectation with respect to $\mathcal{F}$
$X^{-1}(A)$	Inverse image of A with respect to the random variable X
$X^\tau(t,\omega)$	Truncated process X at time τ
X_ω	Path defined by a stochastic process for a fixed $\omega\in\Omega$
X_t	Random variable defined by a stochastic process at time t

W	Wiener process over $(\Omega, \mathcal{F}, P)$
τ_R	First time a stochastic process reaches boundary R
$\tau \wedge \sigma$	Minimum of τ and σ
$\mathbb{R}\langle X \rangle$	Set of all polynomials generated with the alphabet X
$\mathbb{R}\langle\langle X \rangle\rangle$	Set of all formal power series generated with the alphabet X
$\mathbb{R}^\ell\langle\langle X \rangle\rangle$	ℓ-dimensional vector-valued formal power series generated by X
$\mathscr{C}$	Catenation product
$\circ$	Composition product
$\odot$	Hadamard product
$\sqcup\!\sqcup$	Shuffle product
ξ^{-1}	Left-shift operator associated with the word ξ
E_η	Iterated integral associated with η
$J(\eta)$	Position of all letters from alphabet Y in η
F_c	Fliess operator associated with the formal power series c
X^*	Set of all words formed under the alphabet X
X^k	Set of words of length k formed under the alphabet X
X^kY^n	Set of words formed with k letters from X and n letters from Y
$\boldsymbol{L}$	Characteristic series associated with the language L
$B(t)$	Square of the L_2 process norm of the diffusion input $b(t)$
$U(t)$	L_1 process norm of a deterministic drift input
$\bar{U}(t)$	L_1 process norm of a drift input
$\tilde{U}(t)$	Square of the L_2 process norm of a drift input
$V(t)$	Square of the L_2 process norm of a diffusion input
$\bar{V}^2(t)$	Square of the L_4 process norm of a diffusion input
δ_{ij}	Kronecker delta function
$LV(f)_{[0,T]}$	Linear variation of the function f over $[0, T]$
$\langle f, g \rangle_{[0,T]}$	Quadratic covariation of the functions f and g over $[0, T]$
$\langle f \rangle_{[0,T]}$	Quadratic variation of the function f over $[0, T]$

CHAPTER I

INTRODUCTION

The main goal of this dissertation is to describe the interconnection of nonlinear input-output systems that are driven by stochastic inputs. For this purpose, a mathematical representation of an input-output system is introduced using the theory of formal power series. It is presented as a stochastic extension of the known theory for Fliess operators [17, 20]. At the same time, one must ensure that there exists some compatibility between the inputs and outputs of these systems in order to permit their interconnection. The main contributions of this dissertation are sufficient conditions for the well definedness of nonlinear input-output systems accepting L_2-Itô stochastic processes as inputs and a description of the generating series for the parallel, product and cascade interconnections.

This chapter is organized as follows. Section I.1 provides the background and motivation for the dissertation. Section I.2 states the primary research problems addressed herein. Finally, Section I.3 outlines the basic structure of this dissertation.

I.1 BACKGROUND AND MOTIVATION

Functional series expansions of nonlinear input-output operators have been utilized since the early 1900's in engineering, mathematics and physics. Among the more representative approaches are those of M. Fliess [16–21], V. Volterra [52,60] and

This dissertation follows the style of the *IEEE Transactions on Automatic Control* for placement of the figure titles and the format of the bibliography.

N. Wiener [52, 64]. In the early 1970's, the area of nonlinear control systems started to use noncommutative algebra. The theory of formal languages and automata employed in computer science and linguistics found common ground with nonlinear realization and control theory through the work of Marcel-Paul Schützenberger, who introduced powerful new ideas like rationality into the context of noncommutative formal power series [53]. In the 1980's, M. Fliess introduced formal power series together with the path integrals of K. T. Chen to provide an algebraic description of functional expansions known as Chen-Fliess series [7,17]. From a deterministic point of view, operators constructed utilizing this formalism describe a large class of nonlinear input-output systems. For example, any Volterra operator with analytic kernel functions can be described using a Chen-Fliess series. Specifically, let $X = \{x_0, x_1, \ldots, x_m\}$ be an alphabet and X^* be the free monoid comprised of all words over X (including the empty word $\emptyset$) under the catenation product. A formal power series in X is any mapping of the form $X^* \to \mathbb{R}^\ell$, and the set of all such mappings will be denoted by $\mathbb{R}^\ell\langle\langle X\rangle\rangle$. For each $c \in \mathbb{R}^\ell\langle\langle X\rangle\rangle$, one can formally associate an m-input, ℓ-output operator F_c in the following manner. Let $p \geq 1$ and $a < b$ be given. For a measurable function $u : [a, b] \to \mathbb{R}^m$, define $\|u\|_{L_p} = \max\{\|u_i\|_{L_p} : \; 1 \leq i \leq m\}$, where $\|u_i\|_{L_p}$ is the usual L_p-norm for a measurable real-valued function, u_i, defined on $[a, b]$. Let $L_p^m[a, b]$ denote the set of all measurable functions defined on $[a, b]$ having a finite $\|\cdot\|_{L_p}$-norm and $B_p^m(R)[a, b] := \{u \in L_p^m[a, b] : \|u\|_{L_p} \leq R\}$. With $t_0, T \in \mathbb{R}$ fixed and $T > 0$, define recursively for each $\eta \in X^*$ the mapping $E_\eta : L_1^m[t_0, t_0 + T] \to C[t_0, t_0 + T]$ by $E_\emptyset = 1$, and

$$E_{x_i\eta'}[u](t) = \int_{t_0}^{t} u_i(\tau) E_{\eta'}[u](\tau)\, d\tau, \tag{I.1.1}$$

where $x_i \in X$, $\eta' \in X^*$ and $u_0 = 1$. Also, without loss of generality, it is assumed that $t_0 = 0$. The input-output operator corresponding to c is then

$$F_c[u](t) = \sum_{\eta \in X^*} (c, \eta)\, E_\eta[u](t), \tag{I.1.2}$$

which is called a *Fliess operator*, and all (c, η) is called a coefficient of c. Properties of Fliess operators have been widely studied. For example, their continuity, local convergence, global convergence, differentiability, analyticity and realizability have been characterized [17, 19, 20, 29, 31, 61–63]. In the classical literature where these operators first appeared, it was normally assumed that there exist real numbers $K, M > 0$ such that

$$|(c, \eta)| \leq KM^{|\eta|}|\eta|!,$$

for all $\eta \in X^*$, where $|z| = \max\{|z_1|, |z_2|, \ldots, |z_\ell|\}$ when $z \in \mathbb{R}^\ell$, and $|\eta|$ denotes the number of symbols in η [17, 19, 20, 57]. This growth condition on the coefficients of c ensures that there exist positive real numbers R and T such that for all piecewise continuous u with $\|u\|_{L_\infty} \leq R$ the series (I.1.2) converges uniformly and absolutely on $[t_0, t_0 + T]$. Such a power series c is said to be *locally convergent*. More recently, Gray and Wang showed in [29] that

$$|E_\eta[u](t)| \leq \prod_{i=0}^{m} \frac{U_i^{\alpha_i}(t)}{\alpha_i!}, \tag{I.1.3}$$

where for each x_i, $U_i(t) = \int_{t_0}^{t} |u_i(s)|\, ds$, and $\alpha_i = |\eta|_{x_i}$ is the number of times the letter x_i appears in η. This bound can be used to show that $F_c[u]$ converges absolutely in $[t_0, \infty)$ for $u \in L_{p,e}[t_0, \infty)$ when c satisfies the growth condition

$$|(c, \eta)| \leq KM^{|\eta|}.$$

Series satisfying this condition are said to be *globally convergent.* Moreover, if c is locally convergent then (I.1.3) can be used to show that $F_c[u]$ constitutes a well-defined operator from $B_p^m(R)[t_0,\, t_0+T]$ into $B_q^\ell(S)[t_0,\, t_0+T]$ for sufficiently small $R, S, T > 0$, where the numbers $p, q \in [1, \infty]$ are conjugate exponents, i.e., $1/p + 1/q = 1$ with $(1, \infty)$ being a conjugate pair by convention. This also allows one to characterize the well-posedness of interconnections of Fliess operators defined on L_p spaces [28].

A convenient property of any Fliess operator is that its input-output behavior is completely determined by its generating series, independent of whether a state space representation is available. When a state space repalization exists for the system, a coordinate frame has been intrinsically assigned, and, therefore, the input-output system may be localized to a coordinate neighborhood on the state manifold. This is not a necessary setting for the analysis of Fliess operators. The behavior (free of coordinate frames) of an input-output system restricted to a ball in an L_p-space can be studied extensively using purely combinatoric/algebraic tools. It is also worth mentioning that these concepts are intimately related to differential geometric methods commonly used in nonlinear control theory [35].

In most applications, a system's inputs usually have *noise* components. In such circumstances, additional mathematical machinery is needed to properly describe an input-output map in the sense of Fliess. Several authors have formulated approaches under which stochastic processes are admissible inputs. One example is the series expansion of the solution of a stochastic differential equation, where iterated Itô and Stratonovich integrals play a central role [23, 39]. H. Sussmann gave a detailed description of the situation using Lie series and showed that a particularly suitable

mathematical formulation involves the use of Stratonovich integrals because they obey the rules of ordinary differential calculus [56, 59]. This is also supported by the transfer principle developed by P. Malliavin, where geometric constructions involving manifold-valued curves can be extended to manifold-valued processes by replacing ordinary calculus with Stratonovich stochastic calculus [12, 44]. On the other hand, Itô integrals are useful for computing estimates of process moments [1, 45, 50, 51]. Several approaches for systems driven by Wiener process inputs have been presented in [1, 16, 18, 21, 23, 42]. It is easily verified, however, that the corresponding output process of a nonlinear input-output system is, in general, not a Wiener process. Hence, these approaches are not well suited for modeling interconnected systems. In this dissertation, a broader class of stochastic processes called *L_2-Itô processes* are considered as possible input processes [8, 39]. It is argued that this input class is more appropriate for practical applications. Then, stochastic versions of (I.1.1) and (I.1.2) will be defined in Chapter IV using Lebesgue and Stratonovich integrals.

As a motivating example, consider an autonomous system modeled by the stochastic differential equation

$$dz(t) = \bar{f}(z(t))\, dt + g(z(t))\, dW(t), \tag{I.1.4}$$

where W is a standard Wiener process. Equation (I.1.4) in integral form is written as

$$z(t) = z(0) + \int_0^t \bar{f}(z(s))\, ds + \int_0^t g(z(s))\, dW(s), \tag{I.1.5}$$

where $\bar{f}(z)$ and $g(z)$ are suitably defined functions [39]. By Itô's differentiation rule,

(I.1.5) can be written in Stratonovich form as

$$z(t) = z(0) + \int_0^t \underbrace{\left(\bar{f}(z(s)) + \frac{g(z(s))}{2}\frac{\partial}{\partial z}g(z(s))\right)}_{f(z(s))} ds + \int_0^t g(z(s))\, dW(s). \qquad \text{(I.1.6)}$$

For a $\mathcal{C}^2$ function, F, the Stratonovich differential chain rule in integral form is

$$F(z(t)) = F(z(0)) + \int_0^t \left(f(z(s))\frac{\partial}{\partial z}F(z(s))\right) ds + \oint_0^t g(z(s))\frac{\partial}{\partial z}F(z(s))\, dW(s). \qquad \text{(I.1.7)}$$

From these equations, one can identify the Lie differentiation operators $L_f = f(z)\frac{\partial}{\partial z}$ and $L_g = g(z)\frac{\partial}{\partial z}$ so that (I.1.7) becomes

$$F(z(t)) = F(z(0)) + \int_0^t L_f F(z(s))\, ds + \oint_0^t L_g F(z(s))\, dW(s).$$

Now, let $F(z)$ in (I.1.7) be replaced by either f or g from (I.1.6) and substitute $f(z(t))$ and $g(z(t))$ back into (I.1.6). This yields

$$\begin{aligned} z(t) &= z(0) + f(z(0))\int_0^t ds + g(z(0))\oint_0^t dW(s) \\ &\quad + \int_0^t\int_0^s L_f f(z(r))\, dr\, ds + \int_0^t\oint_0^s L_g f(z(r))\, dW(r)\, ds \\ &\quad + \oint_0^t\int_0^s L_f g(z(r))\, dr\, dW(s) + \oint_0^t\oint_0^s L_g g(z(r))\, dW(r)\, dW(s) \\ &= z(0) + f(z(0))\int_0^t ds + g(z(0))\oint_0^t dW(s) + R_1(z(t)), \end{aligned}$$

where $R_1(z(t))$ contains all the integrals whose integrands do not depend on $z(0)$. Continuing in this way produces the usual Peano-Baker formula [52, p. 95]. Since the integrals involved here have a similar nature as the iterated integral (I.1.1), define alphabets $X = \{x_0\}$, $Y = \{y_0\}$, $XY = X \cup Y$ and the iterated Lie derivatives

$L_{g_{x_0\eta}} = L_{g_\eta} L_{g_{x_0}}$ and $L_{g_{y_0\eta}} = L_{g_\eta} L_{g_{y_0}}$, where $g_{x_0} = f$, $g_{y_0} = g$, and $\eta \in XY^*$. Then, slightly abusing the notation in (I.1.5), one can write

$$\begin{aligned} z(t) &= z(0) + L_{g_{x_0}} I(z(0)) E_{x_0}[0](t) + L_{g_{y_0}} I(z(0)) E_{y_0}[0](t) \\ &\quad + L_{g_{x_0 x_0}} I(z(0)) E_{x_0 x_0}[0](t) + L_{g_{y_0 x_0}} I(z(0)) E_{y_0 x_0}[0](t) \\ &\quad + L_{g_{x_0 y_0}} I(z(0)) E_{x_0 y_0}[0](t) + L_{g_{y_0 y_0}} I(z(0)) E_{y_0 y_0}[0](t) + R_2(z(t)), \end{aligned}$$

where I denotes the identity map and once again $R_2(z(t))$ contains all the integrals whose integrands do not depend on $z(0)$. This produces the general series solution of the stochastic differential equation (I.1.4)

$$z(t) = \sum_{\eta \in XY^*} L_{g_\eta} I(z(0))\ E_\eta[0](t). \tag{I.1.8}$$

Here, $(f, g, I, z(0))$ realizes the operator F_c driven by noise when $(c, \eta) = L_{g_\eta} I(z(0))$, $\forall \eta \in XY^*$. This example is analogous to the one presented in [17,35] for deterministic inputs. It also suggests an extension of the Fliess operator theory to the case where the inputs are stochastic processes. Such a generalization also has an underlying noncommutative algebraic structure. M. Fliess observed from the work of R. Ree [49] that the multiplication of two deterministic iterated integrals can be described using the shuffle product. Thus, the set of generating series for Fliess operators forms a commutative $\mathbb{R}$-algebra [17]. This *shuffle algebra* plays a key role in the theory of systems interconnections [13,25,28]. In the stochastic setting, an analogous algebraic structure needs to be proposed. Fortunately, Stratonovich iterated integrals induce a shuffle algebra, which is identical to that in the deterministic case since they obey the same rules of ordinary integral calculus [23].

Now, continuing with the previous example, suppose in equation (I.1.5) that $f(z) = 0$ and $g(z) = z$. A simple inductive procedure shows that the series coefficients

in (I.1.8) are

$$(c,\eta) = L_{g_\eta} I(z(0)) = \begin{cases} 1 & : \quad \forall \eta \in Y^* \\ 0 & : \quad \text{otherwise.} \end{cases} \tag{I.1.9}$$

Therefore, using integration by parts,

$$z(t) = \sum_{k=0}^{\infty} \oint_0^t \cdots \oint_0^{t_2} dW(t_1) \cdots dW(t_k) = \sum_{k=0}^{\infty} \frac{W^k(t)}{k!} = e^{W(t)}.$$

This shows that the series c can be associated with the well-defined random variable $e^{W(t)}$ for any fixed $t > 0$. On the other hand, if $f(z) = 0$ and $g(z) = z^2$, it can be shown that

$$(c,\eta) = L_{g_\eta} I(z(0)) = \begin{cases} |\eta|! & : \quad \forall \eta \in Y^* \\ 0 & : \quad \text{otherwise.} \end{cases} \tag{I.1.10}$$

By the same argument

$$z(t) = \sum_{k=0}^{\infty} k! \oint_0^t \cdots \oint_0^{t_2} dW(t_1) \cdots dW(t_k) = \sum_{k=0}^{\infty} W^k(t) = \frac{1}{1 - W(t)},$$

only if $t < \tau = \inf\{t > 0 : |W(t)| < 1\}$. Thus, $z(t)$ is a well-defined random variable only when $t < \tau$, where τ is also a random variable. One of the objectives of this dissertation is to find conditions under which the generating series of a Fliess operator can be related to a well-defined random variable in the sense that it is the mean square limit of an infinite summation of random variables. Observe that the coefficients in (I.1.9) and (I.1.10) are upper bounded by

$$|(c,\eta)| \leq KM^{|\eta|}, \; \forall \eta \in Y^*,$$

and

$$|(c,\eta)| \leq KM^{|\eta|}\, |\eta|!, \; \forall \eta \in Y^*,$$

respectively, when $K = 1$ and $M = 1$. In the deterministic case, it is known that the former bound leads to the *global convergence* of a Fliess operator, while the latter

bound leads to *local convergence* [29]. In the stochastic case, another type of series is important in addition to the mentioned above. Consider the series $c \in \mathbb{R}\langle\langle X \rangle\rangle$ such that for every word in X^* the image under c is independent of the order of the letters in the word, i.e., $(c, x_{i_n} x_{i_{n-1}} \cdots x_{i_1}) = (c, x_{i_{\sigma(n)}} x_{i_{\sigma(n-1)}} \cdots x_{i_{\sigma(1)}})$, where σ denotes any permutation of $\{1, \ldots, n-1, n\}$. This type of series is called *exchangeable*. It was introduced by Fliess in [17]. Exchangeability introduces a degree of commutativity to a generating series. Fliess operators associated with exchangeable generating series can be written more compactly and conveniently. Specifically, the properties of globally convergent series will be exploited to obtain the global convergence of Fliess operators driven by L_2-Itô processes, while the properties of locally convergent and exchangeable series will be used to obtain their local convergence.

Interconnections of dynamical systems are found everywhere in applications. In Figure 1, the most elementary configurations: the parallel, product, cascade and feedback connections are presented. Characterizing the nature of these interconnections allows engineers to understand how to design system controllers so that a collection of subsystems can work together to achieve a common task. For linear systems, a complete treatment of their interconnection theory can be found in [36]. On the other hand, interconnections involving nonlinear systems are not well understood. For example, it is known that the set of linear systems is closed under parallel (addition), cascade and feedback interconnections. That is, the composite system is another linear system. However, when nonlinear systems of a given class are interconnected, the composite system may fall into a different class of systems.

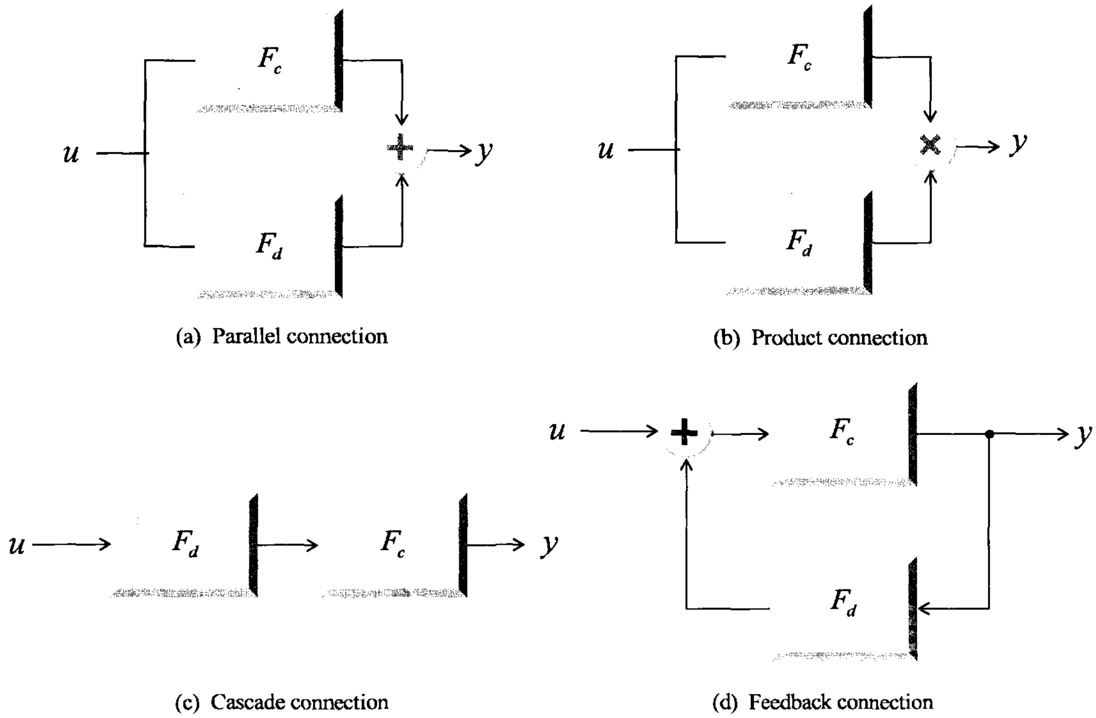

Fig. 1: Elementary system interconnections.

Consider as an example two bilinear state space systems

$$\begin{aligned} \dot{z}_i(t) &= A_i z_i(t) + \sum_{j=1}^{m} N_{j,i} z_i(t) u_{j,i}(t), \quad z_i(0) = z_{i,0} \\ y_i(t) &= C_i z_i(t), \end{aligned}$$

where $i = 1, 2$ and $z_i(t) \in \mathbb{R}^n$; $u_{j,i}(t) \in \mathbb{R}$; $y_i(t) \in \mathbb{R}^\ell$; and A_i, $N_{j,i}$ and C_i are matrices of appropriate dimensions. It can be easily verified that if they are interconnected in a cascade fashion, that is, if $m = \ell$ and one feeds the outputs of one system into the inputs of the other, then one possible state space realization for the input-output

mapping $u_1 \mapsto y_2$ is

$$\dot{z}_1(t) = A_1 z_1(t) + \sum_{j=1}^{m} N_{j,1} z_1(t) u_{j,1}(t), \; z_1(0) = z_{1,0} \tag{I.1.11}$$

$$\dot{z}_2(t) = A_2 z_2(t) + \sum_{j=1}^{m} N_{j,2} z_2(t) (C_1 z_1(t))_j, \; z_2(0) = z_{2,0} \tag{I.1.12}$$

$$y_2(t) = C_2 z_2(t), \tag{I.1.13}$$

which is an affine-input nonlinear system, i.e., a system of the form

$$\begin{aligned} \dot{z}(t) &= f(z(t)) + \sum_{j=1}^{m} g_j(z(t)) u_j(t), \quad z(0) = z_0 \\ y(t) &= h(z(t)), \end{aligned}$$

where f and g_j are vector fields defined in terms of local coordinates on a state space manifold, and h is the output function [35]. In this example, in particular, the g_j's have quadratic polynomial components.

From a Fliess operator point of view, the four elementary interconnections shown in Figure 1 were first studied by Ferfera in [13]. Ferfera's work described the generating series of the four interconnections. Since his work only focused on interconnection of Fliess operators with rational generating series, the convergence analysis was straightforward. On the other hand, Wang in [63] showed that the product connection of two Fliess operators having arbitrary locally convergent generating series is another Fliess operator associated to a locally convergent generating series. Later, Gray and Li in [28] analyzed the cascade and feedback connections of Fliess operators. They showed in particular that the set of local convergent Fliess operators is closed under the cascade connection. In summary, let F_c and F_d be two Fliess operators with generating series series $c, d \in \mathbb{R}\langle\langle X \rangle\rangle$, respectively. Then, the parallel,

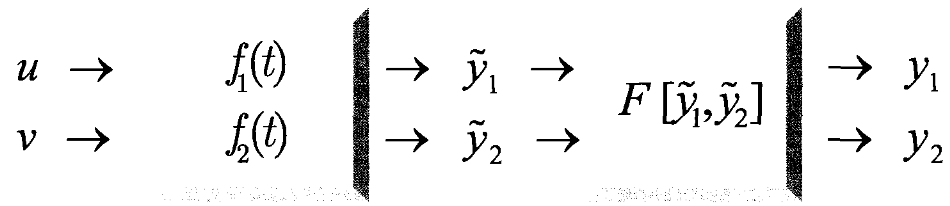

Fig. 2: Cascade of input-output maps.

product and cascade connection of F_c and F_d are described by

$$F_c[u] + F_d[u] = F_{c+d}[u] \tag{I.1.14}$$

$$F_c[u] \cdot F_d[u] = F_{c \sqcup\!\sqcup\, d}[u] \tag{I.1.15}$$

$$F_c[F_d[u]] = F_{c \circ d}[u], \tag{I.1.16}$$

where u is a deterministic input, and $+$, $\sqcup\!\sqcup$ and $\circ$ are addition, the shuffle product and the composition product, respectively [28]. No similar treatment, however, is available for Fliess operators driven by stochastic inputs. To illustrate the problems encountered in the interconnection of systems driven by stochastic processes, consider the cascade connection of an arbitrary input-output map $F[\tilde{y}_1, \tilde{y}_2]$ and the input-output map defined componentwise by

$$\begin{aligned} \tilde{y}_1(t) &= f_1(u) = \int_0^t u_1(s) \fint_0^s u_2(r)\, dW(r)\, ds \\ \tilde{y}_2(t) &= f_2(u) = \fint_0^t u_2(s) \int_0^s u_1(r)\, dr\, dW(s), \end{aligned}$$

where u_1 and u_2 are suitable L_2-bounded stochastic processes (see Figure 2). When dealing with stochastic processes, a desired property for system inputs is that they

are independent processes. In this setting, consider u_1 and u_2 to be mutually independent. However, the intermediate signals $\tilde{y}_1$ and $\tilde{y}_2$ may be dependent processes since f_1 and f_2 depend at the same time upon u_1 and u_2. Thus, if the input-output system F is only defined for independent inputs then it cannot be driven by $\tilde{y}_1$ and $\tilde{y}_2$ since they are dependent of each other. From this point of view, the cascade connection presented in Figure 2 is not well-posed because the inputs and outputs are not compatible with each other in the sense that the outputs must preserve the properties, such as independence, established for the inputs. In order to provide a characterization similar to (I.1.14), (I.1.15) and (I.1.16) for the stochastic case, appropriate extensions of the series operations $+$, $\sqcup\!\sqcup$ and $\circ$ need to be defined to properly describe each interconnection. Then, the following questions can be formulated: what conditions need to be imposed to obtain a well-defined stochastic process at the output of the interconnected system? Can each interconnecion of Fliess operators be represented by a Fliess operator? If so, what is the nature of the generating series of the composite Fliess operator given that the component generating series are either globally convergent or locally convergent?

I.2 PROBLEM STATEMENT

The main objectives of this dissertation are to:

i. Define a class of L_2-Itô stochastic processes that are admissible as inputs to a Fliess operator.

ii. Define a Fliess operator over this input class and provide sufficient conditions

under which the operator converges to produce an output process that is well-defined over a nonzero interval of time.

iii. Characterize the corresponding set of outputs, giving their main properties and describing in what sense there is compatibility between the input class and the output class.

iv. Describe the generating series for parallel, product and cascade interconnections of two Fliess operators for formal input processes and over the class of L_2-Itô input processes.

v. Provide sufficient conditions under which convergence in some sense is preserved for such interconnected systems.

I.3 DISSERTATION OUTLINE

This dissertation is organized as follows. In Chapter II, the probabilistic framework is presented. After the preliminaries, three basic topics are considered: stochastic processes, the Itô and Stratonovich integrals and stopping times.

In Chapter III, a basic introduction to formal power series is presented. The classes of rational and recognizable series are described in detail, and the equivalence between these two classes is set forth. Then a variety of formal power series products are defined. It is shown under certain assumptions that some of these products preserve rationality. Next, Ferfera's condition for the rationality of the composition product is presented as well as an extension of this condition in terms of the Hankel rank.

In Chapter IV, Fliess operators with stochastic input processes are considered. First, the relevant space of stochastic inputs is described. Then a formula for transforming Stratonovich iterated integrals into a sum of Itô integrals is developed, and L_2-bounds for Stratonovich and Itô iterated integrals are determined. This yields sufficient conditions for both local and global convergence of Fliess operators over the class of stochastic inputs considered. Next, the shuffle algebra formed by the set of generating series of Fliess operators with stochastic inputs is presented and various examples of its utility are given. Finally, the relationship between Fliess operators driven by stochastic inputs and Chen series is described.

In Chapter V, the interconnection of systems with stochastic inputs is addressed. In the first part of the chapter, a short summary of the deterministic case is presented as a point of reference. In the next section, the notion of a formal Fliess operator is introduced in order to study the interconnection problem independent of convergence issues. This is followed by the characterization of the non-recursive interconnections using the addition, shuffle product and composition product. Then, the mean square convergence of the parallel and product connections is studied in for the global case, the local case and for exchangeable series. Finally, conditions for preserving rationality under the composition product are given. These latter conditions then play a central role in ensuring that the cascade of two systems driven by stochastic processes is convergent in the mean square sense.

In Chapter VI, the main conclusions are summarized, and future research topics are given.

CHAPTER II

PROBABILISTIC FRAMEWORK

This chapter introduces all the necessary stochastic machinery for the remainder of this dissertation. First, the basic probabilistic setup is given, followed by the definition of a stochastic process. The intent is not to give a complete development of stochastic processes, but only the essential concepts, such as almost sure continuity, Martingales, stopping times, etc. Then, two basic stochastic integrals are presented, the Itô integral and the Stratonovich integral, together with their main properties. Finally, a section on stopping times is presented in order to facilitate the study of convergence over stochastic intervals of time undertaken in Chapter IV. It is worth mentioning that this chapter will serve as one of two pillars on which the new Fliess operator theory will rely. The majority of the concepts presented in this chapter have been taken from [2, 8, 46–48, 54, 55].

II.1 PRELIMINARIES

Definition II.1.1. Let Ω be a non empty set and $\mathcal{F}$ a σ-algebra on Ω. A **measurable space** is defined by the pair $(\Omega, \mathcal{F})$.

The smallest σ-algebra containing all the open sets in $\mathbb{R}$ is denoted by $\mathcal{B}(\mathbb{R})$ and is called the Borel σ-algebra. The elements of $\mathcal{B}(\mathbb{R})$ are called *Borelians*.

Definition II.1.2. Let $(\Omega, \mathcal{F})$ be a measurable space. A **probability measure** over $(\Omega, \mathcal{F})$ is a function $P : \mathcal{F} \rightarrow \mathbb{R}^+ \cup \{\infty\}$ such that

i. $P(A) \geq 0$ for all $A \in \mathcal{F}$.

ii. $P(\Omega) = 1$ and $P(\varnothing) = 0$.

iii. For any disjoint sequence $A_1, A_2 \ldots \in \mathcal{F}$,

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

The triple $(\Omega, \mathcal{F}, P)$ defines a **probability space**.

Definition II.1.3. A **random variable**, X, defined on $(\Omega, \mathcal{F})$ is a real-valued function $X : \Omega \to \mathbb{R}$ such that

$$X^{-1}(A) = \{\omega \in \Omega : X(\omega) \in A\} \in \mathcal{F}$$

for any $A \in \mathcal{B}(\mathbb{R})$.

A random variable X defined on $(\Omega, \mathcal{F})$ is called $\mathcal{G}$-measurable, where $\mathcal{G}$ is a σ-algebra, if

$$X^{-1}(A) = \{\omega \in \Omega \ : \ X(\omega) \in A\} \in \mathcal{G}$$

for any set $A \in \mathcal{B}(\mathbb{R})$. Furthermore, a random variable X induces a σ-algebra, $\mathcal{F}_X$, defined as

$$\mathcal{F}_X = X^{-1}(\mathcal{B}(\mathbb{R})) \triangleq \{X^{-1}(B) \ : \ B \in \mathcal{B}(\mathbb{R})\},$$

provided that X satisfies the following properties:

i. $X^{-1}(\varnothing) = \varnothing \in \mathcal{F}_X$,

ii. $(X^{-1}(B))^c = X^{-1}((B)^c) \in \mathcal{F}_X$, where $(B)^c = \mathbb{R} \backslash B \in \mathcal{B}(\mathbb{R})$,

iii. $\bigcup_{i=1}^{\infty} X^{-1}(B_i) = X^{-1}\left(\bigcup_{i=1}^{\infty} B_i\right)$ and $\bigcup_{i=1}^{\infty} B_i \in \mathcal{B}(\mathbb{R})$.

It is also clear that $\mathcal{F}_X$ is the smallest σ-algebra which contains all the sets $X^{-1}(B)$, where $B \in \mathcal{B}(\mathbb{R})$.

There are several different senses in which random variables can be considered to be equivalent.

Definition II.1.4. Let X and Y be two random variables on $(\Omega, \mathcal{F}, P)$. Then

i. X and Y are **equal** if $X(\omega) = Y(\omega)$, $\forall\, \omega \in \Omega$;

ii. X and Y are **almost surely equal** if $P(\omega \in \Omega, X(\omega) = Y(\omega)) = 1$.

iii. X and Y are **equal in distribution** if $P(X \in F) = P(Y \in F)$ for $F \in \mathcal{F}$.

II.2 STOCHASTIC PROCESSES

Definition II.2.1. Let $\mathcal{T}$ be a set of indexes. A **stochastic process** is a parametrized collection of random variables, $X = \{X_t\}_{t\in\mathcal{T}}$, defined on the probability space $(\Omega, \mathcal{F}, P)$, where each element of the collection takes values in $\mathbb{R}^n$, i.e.,

$$\begin{aligned} X \;:\; & \mathcal{T} \times \Omega \to \mathbb{R}^n \\ & (t, \omega) \mapsto X(t, \omega). \end{aligned}$$

Generally, $\mathcal{T}$ can be any set, for instance, $[0, \infty)$, $[a, b]$ with $a < b < \infty$ or a non negative set of integers. In this dissertation, $\mathcal{T} = [0, \infty)$ or $[0, T]$, $T > 0$. Note that for a fixed $t \in \mathcal{T}$, $X_t : \Omega \to \mathbb{R}^n$ is an n-dimensional random variable. On the other hand, for a fixed $\omega \in \Omega$, $X_\omega : \mathcal{T} \to \mathbb{R}^n$ is a vector-valued function over $\mathcal{T}$. This latter function of time is known as a *path* of the stochastic process X. For

simplicity of notation, the argument ω will normally be suppressed, i.e., $X_t(\omega) = X_t$ and $X_\omega(t) = X(t)$.

Given the temporal nature of a stochastic process, one can consider in some sense a past, a present and a future. Therefore, the concept of a filtration is used to represent changes on the set of events that can be measured due to an increase or decrease of information over time.

Definition II.2.2. A **filtration**, $\mathbf{F}$, is a non-decreasing family of σ-algebras on a measurable space $(\Omega, \mathcal{F})$, i.e., a family $\{\mathcal{F}_t\}_{t\in\mathcal{T}}$ which satisfies $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for $0 \leq s < t$.

A common way to construct a σ-algebras $\mathcal{F}_t$ is $\mathcal{F}_t = \sigma\left(\bigcup_{i\leq t} \mathcal{F}_i\right)$. Moreover, given a process $X = \{X_t\}_{t\in\mathcal{T}}$ defined on $(\Omega, \mathcal{F}, P)$, the filtration $\mathbf{F}_X$ formed with the σ-algebras generated by the random variables X_t, $t \in \mathcal{T}$, is known as the *natural filtration of* X, i.e.,

$$\mathbf{F}_X \triangleq \left\{\sigma\left(\bigcup_{s\leq t} \mathcal{F}_{X_s}\right) \ : \ t \in \mathcal{T}\right\},$$

where $\mathcal{F}_{X_s}$ is the σ-algebra generated by the stochastic process at time s.

Definition II.2.3. A stochastic process $X = \{X_t\}_{t\in\mathcal{T}}$ is called **adapted** to the filtration $\mathbf{F}$ if for each $t \in \mathcal{T}$, X_t is a $\mathcal{F}_t$-measurable random variable.

Naturally, every stochastic process is adapted to its natural filtration. Notions of equivalence between two stochastic processes are given next.

Definition II.2.4. Let $X = \{X_t\}_{t\in\mathcal{T}}$ and $Y = \{Y_t\}_{t\in\mathcal{T}}$ be two stochastic processes on $(\Omega, \mathcal{F}, P)$. Then

i. X and Y are **equivalent** if $X(t,\omega) = Y(t,\omega)$, $\forall\, t \in \mathcal{T}$ and $\forall\, \omega \in \Omega$;

ii. X and Y are **stochastically equivalent** if $P(X_t = Y_t) = 1$, $\forall t \in \mathcal{T}$;

iii. X and Y are **wide sense stochastically equivalent** if for all $k = 1, 2 \dots$ and for all $\{t_1, t_2, \dots, t_k\} \subset \mathcal{T}$ the following is satisfied:

$$P(X_{t_1} \in F_1, X_{t_2} \in F_2, \dots, X_{t_k} \in F_k) = P(Y_{t_1} \in F_1, Y_{t_2} \in F_2, \dots, Y_{t_k} \in F_k),$$

for $F_i \in \mathcal{F}_{t_i}$, $i = 1, 2, \dots, k$;

iv. X and Y are **indistinguishable** if almost all paths are equal, i.e.,

$$P\left(\omega \in \Omega, X_t = Y_t,\ \forall t \in \mathcal{T}\right) = 1.$$

For any of the equivalences defined above, if there exist a stochastic process Y equivalent to $X = \{X_t\}_{t \in \mathcal{T}}$, then Y is called a *version of* X.

Definition II.2.5. A stochastic process $X = \{X_t\}_{t \in \mathcal{T}}$ is said to be **almost surely continuous** if $X_\omega : \mathcal{T} \to \mathbb{R}^n$ is continuous for all $\omega \in \Omega \backslash \mathcal{N}$, where $\mathcal{N}$ satisfies $P\{\mathcal{N}\} = 0$.

An interesting fact about equivalency is that, from a stochastic point of view, two processes can be equivalent while at the same time their paths can have different properties. For example, a discontinuous process can have a version which is almost sure continuous. Thus, one could utilize the continuous version rather than the original version with no probabilistic consequences.

Theorem II.2.1. *(Kolmogorov Continuity Theorem) Let X be a stochastic process. If there exists positive real numbers α, β and γ such that*

$$\mathbf{E}[|X(t) - X(s)|^\alpha] \le \gamma \left|t - s\right|^{1+\beta}, \quad 0 \le s < t \le T < \infty,$$

then there exists a continuous version of the process X.

Definition II.2.6. A stochastic process X is said to be **càdlàg** if (almost surely) it has sample paths which are right continuous with left limits. Similarly, a stochastic process X is said to be **càglàd** if (almost surely) it has sample paths which are left continuous with right limits.

Another property that is preserved through equivalence of stochastic processes is adaptability.

Proposition II.2.1. *Let X and Y be two equivalent processes and $\mathbf{F}$ be a filtration of $(\Omega, \mathcal{F}, P)$. If X is adapted to $\mathbf{F}$, then Y is adapted to $\mathbf{F}$ as well.*

Definition II.2.7. A process X defined on $(\Omega, \mathcal{F}, P)$ is called a **Martingale** with respect to the filtration $\mathbf{F}$ if

i. $\mathbf{E}[X] < \infty$;

ii. X is adapted to $\mathbf{F}$;

iii. $\mathbf{E}[X_t|\mathcal{F}_s] = X_s$ a.s., where s, $t \in \mathcal{T}$ and $0 \leq s \leq t$.

Condition (i) is a technical requirement. Condition (ii) means that one can measure the actual value of X at each instant $t \in \mathcal{T}$. The last condition ensures that the expected value remains constant throughout time.

II.2.1 The Wiener process

In 1828 the Scottish botanist Robert Brown observed that when pollen particles are suspended over a liquid they make irregular and erratic movements even though

no external force is being applied [5]. This phenomenon was called *Brownian motion.* The fact that there was not a physical or mathematical explanation for this phenomenon created some concern within the scientific comunity. The first accurate explanation of Brownian motion was advanced by Desaulx in 1877: "In my way of thinking the phenomenon is a result of thermal molecular motion in the liquid environment (of the particles)." This is indeed the case. A suspended particle is constantly and randomly bombarded from all sides by the molecules of the liquid. If the particle is very small, the collisions it experiences on one side will sometimes be stronger than those from the opposite side, causing it to jump. These small random jumps are what make up Brownian motion. It was not until 1905, when Einstein brought an accepted explanation of Brownian motion to the physics community, that an indirect confirmation of the existence of molecules was presented [11]. This type of movement was later described mathematically by Norbert Wiener. Therefore, it is called a *Wiener process.*

Definition II.2.8. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathbf{F} = \{\mathcal{F}_t\}_{t\in\mathcal{T}}$ a filtration of that space. An n-dimensional **Wiener process** is a stochastic process $W = \{W_t\}_{t\in\mathcal{T}}$, $\mathcal{T} = [0, \infty)$ satisfying the following properties:

i. W is adapted to the filtration $\mathbf{F}$, and for $0 \leq s < t$, the increment $W_t - W_s$ is independent of $\mathcal{F}_s$.

ii. For $s < t$, the increment $W_t - W_s \sim N(0, (t-s)I_{n\times n})$, where $I_{n\times n}$ is the $n \times n$ identity matrix.

iii. The processe starts at 0 almost surely, i.e., $P\{W_0 = 0\} = 1$.

iv. The sample paths of $W = \{W_t\}_{t\in\mathcal{T}}$ are continuous in the sense of Definition II.2.5.

Condition (*iii*) is often relaxed by noting that if $X = \{X_t\}_{t\in\mathcal{T}}$ satisfies (*i*), (*ii*) and (*iv*), then $Y = X - X_0 = \{X(t) - X(0)\}_{t\in\mathcal{T}}$ is taken to be a Wiener process. In addition to the description of Brownian motion as a physical phenomenon and a Wiener process as a mathematical abstraction, some authors (for example, Nielsen [46]) refer to a *Brownian process.* The difference between a Brownian process and the Wiener process is that the latter is related to a particular filtration while the former is not. In other words, if the filtration is changed then the Brownian process remains the same, but the Wiener process may vary.

Given that Wiener processes play a central role in subsequent sections, some of their key properties are presented next.

Theorem II.2.2. *Let W be a Wiener process. Then $W_t \sim N(0,t)$, $\mathbf{E}\left[W_t W_s\right] = \min\{t,s\}$ and $\mathbf{E}\left[|W_t - W_s|^2\right] = |t-s|$.*

Theorem II.2.3. *Let W be a Wiener process with natural filtration $\mathbf{F}$, then W is a Martingale respect to $\mathbf{F}$.*

An important question about the Wiener process is whether or not most of the paths of a Wiener process are differentiable. This will be important when one tries to construct Stieltjes type integrals using a Wiener process as an integrator function. If a Wiener process were differentiable then it would be easy to define $\frac{dW}{dt}$. Unfortunately, the next theorem shows that the opposite is true.

Theorem II.2.4. [2] *Let W be a Wiener process defined in the probability space*

$(\Omega, \mathcal{F}, P)$. *Then*

$$P\left(\omega \in \Omega : W_\omega(t) \text{ is not differentiable at any } t \in \mathcal{T}\right) = 1.$$

There is an easy heuristic argument behind this conclusion. First, recall that in Theorem II.2.2 it was stated that $W(t)$ is a Gaussian random variable with zero mean and with variance t. Taking the ratio

$$\frac{\Delta W}{\Delta t} \triangleq \frac{W(t+h) - W(t)}{(t+h) - t},$$

it is easily observed that the numerator has variance h, and the denominator is the constant h. Thus, $\frac{\Delta W}{\Delta t}$ has variance $1/h$, which tends to infinity as h goes to 0. In other words, a path of $W(t)$ on smaller and smaller scales becomes more and more erratic, and the slope ultimately diverges.

II.3 ITÔ AND STRATONOVICH INTEGRALS

The main objective of this section is to describe integration of a stochastic process with respect to a Wiener process, i.e., to define the integral

$$\int_0^t X(s)\, dW(s), \; t > 0.$$

This type of integration serves as a way to create a process with an increment whose variance changes over time. Processes with this chararacteristic appear in many engineering and finance applications [46, 47].

II.3.1 Linear and quadratic variation

It is known that if a function has a bounded linear variation, then it can act as the integrator of a Stieltjes integral. Informally speaking, the linear variation measures

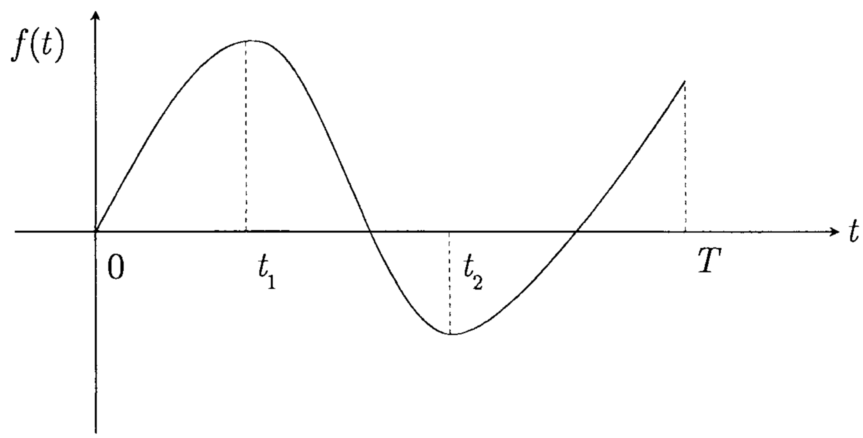

Fig. 3: Partition $\Pi = \{0, t_1, t_2, T\}$ of $f(t)$ on $[0, T]$.

the total amount of up's and down's made by a function over an interval of time. For example, the linear variation of a differentiable function $f(t)$, as shown in Figure 3, for the partition $\Pi = \{0, t_1, t_2, T\}$ is

$$
\begin{aligned}
LV(f)_{[0,T]} &\triangleq [f(t_1) - f(0)] - [f(t_2) - f(t_1)] + [f(T) - f(t_2)] \\
&= \int_0^{t_1} f'(t)\, dt + \int_{t_1}^{t_2} (-f'(t))\, dt + \int_{t_2}^{T} f'(t)\, dt \\
&= \int_0^{T} |f'(t)|\, dt.
\end{aligned}
$$

Definition II.3.1. Let $\Pi = \{t_0 = 0, t_1, \ldots, t_n = T\}$ be a partition of $[0, T]$, $T < \infty$. Then the **linear variation** of a (not necessarily differentiable) function f is defined as

$$
LV(f)_{[0,T]} \triangleq \lim_{\|\Pi\| \to 0} \sum_{k=0}^{n-1} |f(t_{k+1}) - f(t_k)|,
$$

where $\|\Pi\| = \max_{k=0,\ldots,n-1} (t_{k+1} - t_k)$ is the measure of the partition Π.

Example II.3.1. Suppose that f is differentiable. Using the mean value theorem, there exists a time $t_k^* \in (t_k, t_{k+1}]$ such that

$$f(t_{k+1}) - f(t_k) = f'(t_k^*)(t_{k+1} - t_k).$$

Thus,

$$\sum_{k=0}^{n-1} |f(t_{k+1}) - f(t_k)| = \sum_{k=0}^{n-1} |f'(t_k^*)| (t_{k+1} - t_k),$$

and therefore

$$LV(f)_{[0,T]} = \lim_{\|\Pi\| \to 0} \sum_{k=0}^{n-1} |f'(t_k^*)| (t_{k+1} - t_k) = \int_0^T |f'(t)| \, dt.$$

□

There exist functions whose linear variation is unbounded (see Figure 4) [40]. In this situation, the quadratic variation of the function can still be finite.

Definition II.3.2. The **quadratic variation** of a function $f : [0,T] \to \mathbb{R}$ is defined as

$$\langle f \rangle_{[0,T]} = \lim_{\|\Pi\| \to 0} \sum_{k=0}^{n-1} |f(t_{k+1}) - f(t_k)|^2 .$$

The **quadratic covariation** of two functions f and g over $[0,T]$ is defined as

$$\langle f, g \rangle_{[0,T]} = \lim_{\|\Pi\| \to 0} \sum_{k=0}^{n-1} (f(t_{k+1}) - f(t_k))(g(t_{k+1}) - g(t_k)) . \tag{II.3.1}$$

Two important properties of the quadratic covariation as a function of time are that it has bounded variation and satisfies

$$\langle f, g \rangle_{[0,T]} = \frac{1}{4} \left(\langle f + g \rangle_{[0,T]} - \langle f - g \rangle_{[0,T]} \right) .$$

This is called the *polarization identity*.

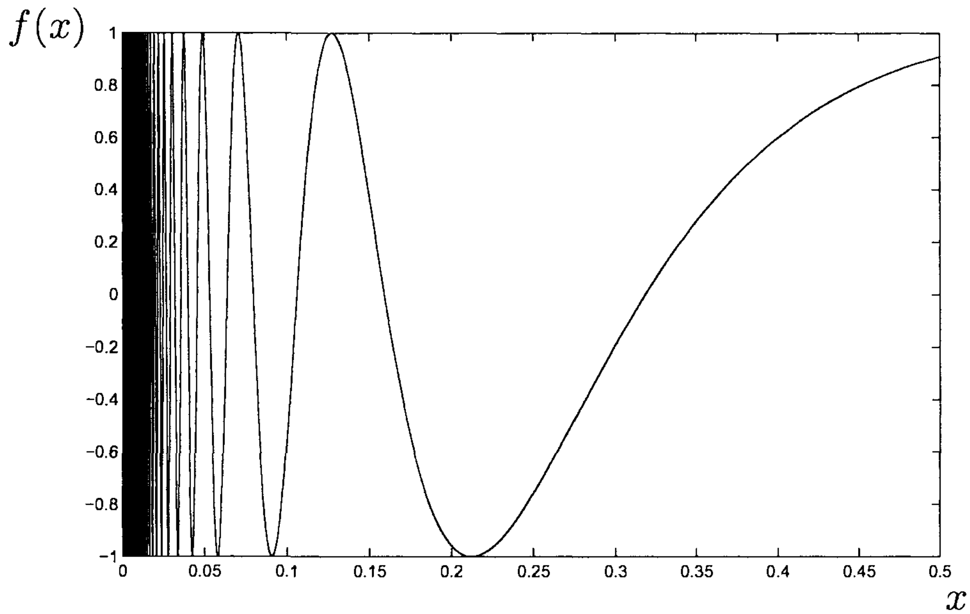

Fig. 4: Unbounded linear variation function $f(x) = \sin(1/x)$.

Example II.3.2. Suppose that f has a continuous derivative. Then

$$\begin{aligned}\sum_{k=0}^{n-1} |f(t_{k+1}) - f(t_k)|^2 &= \sum_{k=0}^{n-1} (f')^2(t_k^*)(t_{k+1} - t_k)^2 \\ &\leq \left(\sum_{k=0}^{n-1} (f')^2(t_k^*)(t_{k+1} - t_k)\right) \|\Pi\| .\end{aligned}$$

Taking limits,

$$\begin{aligned}\langle f \rangle_{[0,T]} &\leq \lim_{\|\Pi\| \to 0} \left[\left(\sum_{k=0}^{n-1} (f')^2(t_k^*)(t_{k+1} - t_k))\right) \|\Pi\|\right] \\ &= \lim_{\|\Pi\| \to 0} \|\Pi\| \lim_{\|\Pi\| \to 0} \left[\sum_{k=0}^{n-1} (f')^2(t_k^*)(t_{k+1} - t_k))\right] \\ &= \lim_{\|\Pi\| \to 0} \|\Pi\| \int_0^T |f'(t)|^2 \, dt \\ &= 0.\end{aligned}$$

The last step above used the fact that $f'(x)$ is continuous, which ensures that $\int_0^T |f'(t)|^2 \, dt$ is finite. $\square$

In ordinary calculus, integrable functions are exactly equivalent to those functions whose linear variation is bounded. A disadvantage of Wiener processes is that they are nowhere differentiable; as a consequence, the mean value theorem does not apply. However, if their linear variation is finite, then a Stieltjes integral can be constructed.

Theorem II.3.1. *Let W be a Wiener process. The quadratic variation of W over $[0,T]$ is*

$$\langle W\rangle_{[0,T]} = T.$$

Example II.3.3. Suppose that the Wiener process W is a function of bounded variation on $[0,T]$. It then follows that

$$\sum_{k=0}^{n-1} |W(t_{k+1}) - W(t_k))|^2 \leq \max_{0\leq k\leq n-1} |W(t_{k+1}) - W(t_k))| \left|\sum_{k=0}^{n-1} |W(t_{k+1}) - W(t_k))|\right|.$$

Since W is a.s. continuous on $[0,T]$, it is necessarily a.s. uniformly continuous on $[0,T]$. Therefore,

$$\max_{0\leq k\leq n-1} |W(t_{k+1}) - W(t_k))| \to 0 \text{ as } \|\Pi\| \to 0,$$

from which it follows that

$$\lim_{\|\Pi\|\to 0} \sum_{k=0}^{n-1} |W(t_{k+1}) - W(t_k))|^2 \to 0 \text{ a.s.}$$

This contradicts Theorem II.3.1, and thus $LV(W)_{[0,t]} = \infty$. □

This example shows that a Stieltjes type integral cannot be constructed if a Wiener process acts as an integrator. However, using the fact that a Wiener process has finite quadratic variation, an approximation procedure may be used to define a stochastic integral in terms of a Wiener process. This procedure is a density argument

similar to that used for defining Lebesgue integrals, except that all the limits involved are taken in the mean-square sense.

Before constructing an integral for the Wiener process, the necessary setting is introduced. Consider the probability space $(\Omega, \mathcal{F}, P)$, the Wiener process W, the natural filtration $\mathbf{F}$ of W and the process X adapted to $\mathbf{F}$, which is $\mathcal{F}_t \times \mathcal{B}([0,t])$-measurable for any fixed $t > 0$.

Definition II.3.3. The $\mathbb{R}$-vector space formed by all square integrable random variables is denoted by $L^2(\Omega, \mathcal{F}, P)$, or, in abbreviated form, $L^2(P)$, where square integrability of X means

$$\mathbf{E}\left[|X|^2\right] < \infty.$$

Theorem II.3.2. *The space $L^2(P)$ with norm $\|\cdot\|_{L^2(P)} = \sqrt{\mathbf{E}\left[|\cdot|^2\right]}$ is a complete normed space, i.e., a Banach space.*

In a similar manner, consider the set of all square integrable stochastic processes defined on the probability space $(\Omega \times \mathcal{T}, \mathcal{P}, P \otimes \lambda)$ under the measure $P \otimes \lambda$. $\mathcal{P}$ is known as the *predictable* σ-algebra, i.e., the σ-algebra generated by sets of the form $A_s \times (s,t]$ and $A_0 \times \{0\}$, where $\{0 < s < t \leq T\}$, $A_s \in \mathcal{F}_s$, and $A_0 \in \mathcal{F}_0$.

Definition II.3.4. The $\mathbb{R}$-vector space formed by all square integrable stochastic processes is denoted by $L_2(\Omega \times [0,T], \mathcal{P}, P \otimes \lambda)$, or, in abbreviated form, $L_2(P \otimes \lambda)$, where square integrability of the stochastic process X means

$$\mathbf{E}\left[\int_0^T |X_t|^2 \, dt\right] < \infty.$$

Theorem II.3.3. *The space $L_2(P \otimes \lambda)$ with the norm $\|\cdot\|_{L_2(P \otimes \lambda)} = \sqrt{\mathbf{E}\left[\int_0^T |\cdot|^2 \, dt\right]}$ is a Banach space.*

The construction of the integral for a Wiener process, referred to as an *Itô integral*, is outlined in three steps.

Step 1: The Itô integral for simple processes

Definition II.3.5. Let $\Pi = \{t_0, t_1, \dots, t_{k-1}, t_k\}$ be a partition of $[0, T]$, $T < \infty$. A **simple process** is a stochastic process X that can be expressed as

$$X = \sum_{j=0}^{k-1} X^j \mathbb{1}_{[t_j, t_{j+1})},$$

where X^j is a square integrable random variable, $\mathcal{F}_{t_j}$-measurable, and $\mathbb{1}_{[t_j, t_{j+1})}$ is the indicator function for the interval $[t_j, t_{j+1})$.

Note that a simple process is adapted, square integrable and has càdlàg (right-continuous with left limits everywhere) paths. It is not difficult to show that the set of simple processes forms a real vector space, denoted here by $\mathcal{H}_0^2$.

Definition II.3.6. The **Itô integral** $\mathbf{I} : \mathcal{H}_0^2 \to L^2(P)$ is defined as

$$\mathbf{I}(X) = \int_0^T X(s)\, dW(s) \triangleq \sum_{j=0}^{k-1} X(t_j)(W(t_{j+1}) - W(t_j)).$$

Theorem II.3.4. *(Basic properties) Let $X, Y \in \mathcal{H}_0^2$ and k be a real number. Then*

i. (Linearity): $\int_0^T (X(s) + kY(s))\, dW(s) = \int_0^T X(s)\, dW(s) + k \int_0^T Y(s)\, dW(s)$

ii. (Zero Expectation): $\mathbf{E}\left[\int_0^T X(s)\, dW(s)\right] = 0$

iii. (Isometry):
$$\mathbf{E}\left[\left(\int_0^T X(s)\, dW(s)\right)\left(\int_0^T Y(s)\, dW(s)\right)\right] = \mathbf{E}\left[\int_0^T X(s)Y(s)\, ds\right].$$

Note that if $X = Y$ a.s., then this identity reduces to

$$\mathbf{E}\left[\left(\int_0^T X(s)\,dW(s)\right)^2\right] = \mathbf{E}\left[\int_0^T X^2(s)\,ds\right]. \tag{II.3.2}$$

In terms of the norms defined previously, identity (II.3.2) is equivalent to

$$\|I(X)\|_{L^2(P)} = \|X\|_{L_2(P\otimes\lambda)}\,.$$

Step 2: Extension via approximations

Definition II.3.7. Define $\mathcal{H}^2$ as the subset of $L_2(P \otimes \lambda)$ consisting of all measurable and adapted stochastic processes that satisfy

$$\mathbf{E}\left[\int_0^T |X(s)|^2\,ds\right] < \infty.$$

Clearly $\mathcal{H}_0^2 \subset \mathcal{H}^2 \subset L_2(P \otimes \lambda)$.

Theorem II.3.5. *$\mathcal{H}_0^2$ is dense in $\mathcal{H}^2$, i.e., for any $X \in \mathcal{H}^2$ there exists a family of processes in $\mathcal{H}_0^2$, $\{X_n\}_{n>0}$, such that $\lim_{n\to\infty} \|X_n - X\|_{L_2(P\otimes\lambda)} = 0$.*

Theorem II.3.6. *Let $X \in \mathcal{H}^2$ and $\{X_n\}_{n>0}$ be a sequence of processes in $\mathcal{H}_0^2$ that converges to X, i.e., $\lim_{n\to\infty} \|X_n - X\|_{L_2(P\otimes\lambda)} \to 0$. Then, the sequence of integrals $\mathbf{I}(X_n) = \int_0^T X_n(s)\,dW(s)$, $n > 0$, is a Cauchy sequence in the complete space $L^2(P)$, and the limit is independent of the sequence of approximating simple processes.*

Definition II.3.8. Let W be a Wiener process and $X \in \mathcal{H}^2$. The **Itô integral** of X is defined as

$$\mathbf{I}(X) = \int_0^T X(s)\,dW(s) \triangleq \lim_{n\to\infty} \int_0^T X_n(s)\,dW(s),$$

where the limit is taken in the $L^2(P)$ sense, and $\{X_n\}_{n>0}$ is a sequence in $\mathcal{H}_0^2$ such that

$$\lim_{n\to\infty} \mathbf{E}\left[\int_0^T (X(s) - X_n(s))^2 ds\right] = 0.$$

It is important to mention that an additional extension of the Itô integral concept is possible. This extension allows one to integrate processes that are adapted, measurable, and satisfy

$$P\left(\omega \in \Omega : \int_0^T X^2(s)ds < \infty, \ \ \forall t \geq 0\right) = 1.$$

The set of these processes is denoted by $\mathcal{L}_{LOC}^2(P)$. This generalization will not be needed in this dissertation. Note that the Itô integral

$$\begin{aligned} \mathbf{I} \ &: \ \mathcal{H}^2 \to L^2(P) \\ X &\mapsto \mathbf{I}(X) = \int_0^t X(s)dW(s) \end{aligned}$$

has been defined as an isometric linear extension of the Itô integral on $\mathcal{H}_0^2$. Thus, the properties of the Itô integral for simple processes are also valid for the integral of processes in $\mathcal{H}^2$.

Theorem II.3.7. *The integral operator* $\mathbf{I} : \mathcal{H}^2 \to L^2(P)$ *satisfies:*

i. (Linearity): $\displaystyle\int_0^T (X(s) + kY(s))\, dW(s) = \int_0^T X(s)\, dW(s) + k\int_0^T Y(s)\, dW(s)$

ii. (Zero Expectation): $\displaystyle\mathbf{E}\left[\int_0^T X(s)\, dW(s)\right] = 0$

iii. (Isometry): $\displaystyle\mathbf{E}\left[\left(\int_0^T X(s)\, dW(s)\right)^2\right] = \mathbf{E}\left[\int_0^T X^2(s)\, ds\right].$

Here $X, Y \in \mathcal{H}^2$ *and* $k \in \mathbb{R}$.

It is clear that if $X \in \mathcal{H}^2$ and $\mathbb{1}_{\Omega \times A}$ is the indicator function for the measurable set $\Omega \times A$, where $A \in \mathcal{B}(\mathcal{T})$, then the product $X\mathbb{1}_{\Omega \times A} \in \mathcal{H}^2$. This allows one to define the Itô integral over subsets of $\mathcal{T}$.

Definition II.3.9. Let $A \subset \mathcal{B}(\mathcal{T})$. The **Itô integral** of $X \in \mathcal{H}^2$ over A is

$$\int_A X(s)\, dW(s) \triangleq \int_{\mathcal{T}} X(s)\mathbb{1}_{\Omega \times A}\, dW(s).$$

Step 3: The Itô integral as a stochastic process

The possibility of integrating over any subset of $\mathcal{B}(\mathcal{T})$ allows one to introduce the notion of a stochastic process generated by Itô integrals. Indeed, for each $t \in \mathcal{T}$, the Itô integral $\mathbf{I}(X)_t \triangleq \int_0^t X(s)\, dW(s) = \int_0^T X(s)\mathbb{1}_{\Omega \times [0,t]}\, dW(s)$ is $\mathcal{F}_t$-measurable. Naturally, $\mathbf{I}(X)_t$ is now parametrized by $t \in \mathcal{T}$. Thus, this family constitutes a natural choice to define the process $\{\mathbf{I}(X)_t\}_{t \in \mathcal{T}}$ induced by X. However, the Itô integral $\int_0^t X(s)\, dW(s)$ is a random variable on $L^2(P)$, so $\mathbf{I}(X)_t$ can be specified arbitrarily for a set $A_t \in \mathcal{F}$ with $P(A_t) = 0$. In other words, $\mathbf{I}(X)_t$ is ambiguous on a null set A_t. If there are countable null sets A_t, then there is no problem since the measure of a countable union of measure zero sets is again zero. But $[0, T]$ is an uncountable set, and the union of all A_t over $t \in [0, T]$ might well be all of Ω. Then, the construction could be ambiguous for some $\omega \in \Omega$. The next result solves the problem.

Theorem II.3.8. *If $X \in \mathcal{H}^2$, then there exists a continuous process $X' : \Omega \times \mathcal{T} \to \mathbb{R}$ such that*

$$P\left(\omega \in \Omega : X'_t(\omega) = \int_0^t X(s)\, dW_s\right) = 1.$$

Proof: Given that $\mathcal{H}_0^2$ is dense in $\mathcal{H}^2$, one can take a sequence $\{X_n\}_{n>0} \in \mathcal{H}_0^2$ approximating X. For each $t \in \mathcal{T}$, $X_n \mathbb{1}_{[0,t]} \in \mathcal{H}_0^2$, and, therefore, the integral $\int_0^t X_n(s)\, dW(s) = \int_{\mathcal{T}} X_n(s) \mathbb{1}_{[0,t]}\, dW(s) \in L^2(P)$ is well defined on Ω. For each n, the process X_n does not have the ambiguity described in the previous paragraph. Hence, it is possible to define the sequence $\{X'_{t,n}\}_{n>0}$ as

$$X'_{t,n} = \int_0^t X_n(s)\, dW(s).$$

More explicitly, for $t_i < t \leq t_{i+1}$,

$$X'_{t,n} = \sum_{i=0}^{j-1} X_n(t_i)(W(t_{i+1}) - W(t_i)).$$

Now, from the Wiener process properties, each $X'_{t,n}$ is a $\mathcal{F}_t$-adapted continuous Martingale. Applying the isometry property and Doob's maximal inequality [8] for $n \geq m$

$$\begin{aligned} P\left(\sup_{\mathcal{T}} \left\|X'_{t,n} - X'_{t,m}\right\| \geq \epsilon\right) &\leq \frac{1}{\epsilon^2} \mathbf{E}\left[\left\|X'_{t,n} - X'_{t,m}\right\|^2\right] \\ &\leq \frac{1}{\epsilon^2} \left\|X_n - X_m\right\|_{L_2(P\otimes\lambda)}. \end{aligned}$$

Since $\{X_n\}_{n>0}$ is convergent in $\mathcal{H}^2$, it is a Cauchy sequence. Thus, one can consider an increasing subsequence $\{n_k\}_{k>0}$ such that $\left\|X_{n_{k+1}} - X_{n_k}\right\|_{L_2(P\otimes\lambda)} < 2^{-3k}$, for a sufficiently large n_k. If $\epsilon = 2^{-k}$ then

$$P\left(\sup_{\mathcal{T}} \left\|X'_{t,n_{k+1}} - X'_{t,n_k}\right\| \geq 2^{-k}\right) \leq 2^{-k},\ k = 1, 2, \ldots$$

Note that $\sum_{k=1}^{\infty} 2^{-k} < \infty$, and by the Borel-Cantelli Lemma, there exist a set $\Omega_0 \subseteq \Omega$ with $P(\omega \in \Omega_0) = 1$ and a random variable Y such that $Y(\omega) < \infty\ \forall\ \omega \in \Omega_0$. If $\omega \in \Omega_0$, then $\sup_{\mathcal{T}} \left|X'_{t,n_{k+1}}(\omega) - X'_{t,n_k}(\omega)\right| \geq 2^{-k}$ for a finite number of k's. It then follows that

$$\sup_{\mathcal{T}} \left|X'_{t,n_{k+1}}(\omega) - X'_{t,n_k}(\omega)\right| < 2^{-k}\ \forall\ k > Y(\omega).$$

This indicates that $\{X'_{t,n_k}(\omega)\}$, for any $\omega \in \Omega_0$, is a Cauchy sequence under the supremum norm in $\mathcal{C}(\mathcal{T})$. Since this space is complete, the sequence $\{X_{t,n_k}(\omega)\}$ is convergent for each $\omega \in \Omega_0$ to a continuous function, namely $X'_t(\omega)$.

Concerning the stochastic equivalence claim, note that $X_{n_k}\mathbb{1}_{[0,t]} \rightarrow X\mathbb{1}_{[0,t]} \in L_2(P \otimes \lambda)$ as $k \rightarrow \infty$. By the isometry property, $\int_{\mathcal{T}} X_{n_k}(s)\mathbb{1}_{[0,t]}\, dW(s) \rightarrow \int_{\mathcal{T}} X(s)\mathbb{1}_{[0,t]}\, dW(s) \in L^2(P)$. Finally, by the uniqueness of the limit,

$$P\left(X'_t = \int_0^t X(s)dW_s\right) = 1,$$

which concludes the proof. ∎

The process X' will be considered the integral process induced by the process X. Some properties of this integral process are presented next.

Theorem II.3.9. *Let $X \in \mathcal{H}^2$. The integral process*

$$X'(t) = \int_0^t X(s)\, dW(s)$$

is adapted to $\mathbf{F}$*, which is the natural filtration of the Wiener process W.*

Hereafter the filtration under consideration will always be the one generated by the Wiener process W.

Theorem II.3.10. *The quadratic variation of $X' = \mathbf{I}(X)$ over $[0,t]$ is*

$$\langle \mathbf{I}(X) \rangle_{[0,t]} = \int_0^t X^2(s)\, ds.$$

Theorem II.3.11. *(Martingale property) Let $X \in \mathcal{H}^2$. The integral process induced by X is a Martingale process with respect to the filtration* $\mathbf{F}$*, i.e.*

$$\mathbf{E}(\mathbf{I}(X)_t | \mathcal{F}_s) = \mathbf{I}(X)_s \text{ almost surely, where } s,t \in \mathcal{T} \text{ and } 0 \leq s \leq t.$$

It is also interesting to consider the converse claim, i.e., under what circumstances can a Martingale be written as an Itô integral?. The next theorem presents one of the most important results in Martingale theory and explains why Martingales play a central role in Itô stochastic calculus.

Theorem II.3.12. *(The Martingale representation theorem) Suppose Y is a Martingale with respect to the filtration $\mathbf{F}$ and that $Y_t \in L^2(P)$ for all $t \geq 0$. Then there exists a unique adapted stochastic process $X \in \mathcal{H}^2$ such that a.s.*

$$Y(t) = \mathbf{E}[Y(0)] + \int_0^t X(s)\, dW(s), \quad \text{for all} \quad t \geq 0.$$

Corollary II.3.1. *Let Y be a Martingale with respect to the filtration $\mathbf{F}$. If $\langle Y \rangle_{[0,t]} = 0$ for all $t \in [0,T]$, then a.s. $Y = 0$.*

Proof: From Theorem II.3.12, one can express the process Y as

$$Y(t) = \mathbf{E}[Y_0] + \int_0^t X(s)\, dW(s).$$

Calculating the quadratic variation of Y over $[0,t]$, it then follows that

$$\begin{aligned} \langle Y \rangle_{[0,t]} &= \underbrace{\langle \mathbf{E}[Y_0] \rangle_{[0,t]}}_{0} + \left\langle \int_0^t X(s)\, dW(s) \right\rangle_{[0,t]} \\ &= \int_0^t X^2(s)\, ds. \end{aligned}$$

Observe that this is a non-decreasing function. It is also known that for Lebesgue integrals, if $\int_0^t |f(s)|\, ds = 0$, then a.s. $f = 0$. By hypothesis,

$$\langle Y \rangle_{[0,t]} = \int_0^t X^2(s)\, ds = 0,$$

which implies that a.s. $X = 0$. Therefore, Y can only be the zero Martingale. ∎

II.3.2 The Stratonovich integral

The Itô integral was first defined for simple processes as

$$\mathbf{I}(X)_t = \sum_{k=0}^{n-1} X(t_k^*)\left(W(t_{k+1}) - W(t_k)\right),$$

where there is an intrinsically defined partition of the interval $[0,t]$, i.e., $0 = t_1 < t_2 < \cdots < t_{n-1} < t_n = t$, and $t_k^* = t_k$ for each subinterval $[t_k, t_{k+1})$. Then by a density argument the definition was extended for processes in $\mathcal{H}^2$. Suppose now that the simple processes are slightly modified, i.e., rather than considering $t_k^* = t_k$ one considers $t_k^* = \lambda t_k + (1-\lambda)t_{k+1}$ with $\lambda \in [0,1]$. Then, applying the same density argument, the resulting integral for $\lambda \neq 0$ is not equivalent to the Itô integral.

Example II.3.4. Calculate $\int_0^t W(s)\,dW(s)$ for $\lambda = 0$ and $\frac{1}{2}$.

i. $\lambda = 0$ (Itô integral)

First, calculate the partial sums

$$\begin{aligned} S_n &= \sum_{k=0}^{n-1} W(t_k)\left(W(t_{k+1}) - W(t_k)\right) \\ &= \frac{1}{2}\sum_{k=0}^{n-1}\left(W^2(t_{k+1}) - W^2(t_k)\right) - \frac{1}{2}\sum_{k=0}^{n-1}\left(W(t_{k+1}) - W(t_k)\right)^2 \\ &= \frac{1}{2}\left(W^2(T) - W^2(0)\right) - \frac{1}{2}\sum_{k=0}^{n-1}\left(W(t_{k+1}) - W(t_k)\right)^2. \end{aligned}$$

Then, the mean-square limit of S_n converges to

$$\int_0^t W(s)\,dW(s) = \frac{1}{2}W^2(t) - \frac{1}{2}t.$$

ii. $\lambda = \dfrac{1}{2}$

Similarly,

$$S_n = \sum_{k=0}^{n-1} W\left(\frac{t_{k+1}+t_k}{2}\right)(W(t_{k+1}) - W(t_k)).$$

By the continuity of the Wiener process, S_n can be replaced by

$$\begin{aligned} S'_n &= \sum_{k=0}^{n-1}\left(\frac{W(t_{k+1})+W(t_k)}{2}\right)(W(t_{k+1}) - W(t_k)) \\ &= \frac{1}{2}\sum_{k=0}^{n-1}\left(W^2(t_{k+1}) - W^2(t_k)\right) \\ &= \frac{1}{2}W^2(t). \end{aligned}$$

Thus, the mean-square limit of S'_n gives

$$\int_0^t W(s)\,dW(s) = \frac{1}{2}W^2(t).$$

In general, for any $\lambda \in [0,1]$

$$\int_0^t W(s)\,dW(s) = \frac{1}{2}W^2(t) + \left(\lambda - \frac{1}{2}\right)t.$$

□

The advantage of using $\lambda = 0$ is that the resulting integral is adapted, and the induced integral process is a Martingale process. This is not true for any other value of λ. If $\lambda = \frac{1}{2}$, the extra term obtained from Itô integration disappears; therefore, it resembles Stieltjes integration. This is known as the *Stratonovich integral.* This integral can also be defined in terms of the Itô integral.

Definition II.3.10. The **Stratonovich integral** of $X \in \mathcal{H}^2$ is defined as

$$\mathbf{S}(X) = \oint_0^t X(s)\,dW(s) \triangleq \int_0^t X(s)\,dW(s) + \frac{1}{2}\langle X, W\rangle_{[0,t]}, \qquad \text{(II.3.3)}$$

where $\langle X, W\rangle_{[0,t]}$ is the quadratic covariation defined in (II.3.1).

Observe that from the Itô integral and quadratic covariation definitions, the definition of the Stratonovich integral is equivalent to the following limit

$$\mathbf{S}(X) = \lim_{n\to\infty} \sum_{k=0}^{n-1} \frac{(X(t_{k+1}) + X(t_k))}{2} \left(W(t_{k+1}) - W(t_k)\right), \tag{II.3.4}$$

where the limit is taken in the $L_2(P)$ sense. If in addition the regularity of the integrand and the continuity of W are utilized, then (II.3.4) is also equivalen to

$$\mathbf{S}(X) = \lim_{n\to\infty} \sum_{k=0}^{n-1} X\left(\frac{t_{k+1}+t_k}{2}\right) \left(W(t_{k+1}) - W(t_k)\right),$$

which is similar to the Itô integral definition with the difference that now the limit is taken on the processes defined using the middle point, $\frac{t_{k+1}+t_k}{2}$, of the time interval $[t_k, t_{k+1}]$ rather than its left extreme t_k.

Corollary II.3.2. *Consider $Y \in \mathcal{H}^2$ such that $\langle Y, W\rangle_{[0,t]} < \infty$ and a Lebesgue integrable function X. If*

$$\oint_0^t Y(s)\, dW(s) = \int_0^t X(s)\, ds,$$

then a.s. $Y = X = 0$.

Proof: Since $\int_0^t X(s)\, ds$ has bounded variation, then $\left\langle \int_0^\cdot X(s)\, ds\right\rangle_{0,t} = 0$. Using Definition II.3.10

$$\left\langle \oint_0^\cdot Y(s)\, dW(s)\right\rangle_{[0,t]} = \left\langle \int_0^\cdot Y(s)\, dW(s)\right\rangle_{[0,t]} + \underbrace{\left\langle \frac{1}{2}\langle Y, W\rangle_{[0,\cdot]}\right\rangle_{[0,t]}}_{0} = 0.$$

Now, by Corollary II.3.1, $Y = 0$ a.s. This implies that $\int_0^\cdot X(s)\, ds = 0$ a.s. Thus, $Y = X = 0$ a.s. as claimed. ■

At first glance, the Itô integral seems enough to represent a variety of stochastic processes. But, the extra terms appearing when the Itô integral is computed cannot

be expressed as such. For example,

$$\int_0^t W(s)\,dW(s) = \frac{1}{2}W^2(t) - \frac{1}{2}t,$$

or equivalently,

$$W^2(t) = t + \int_0^t 2W(s)\,dW(s) = \int_0^t 1\,ds + \int_0^t 2W(s)\,dW(s), \tag{II.3.5}$$

where t has no representation as an Itô integral, and W^2 is not a Martingale. Although Theorem II.3.12 cannot be applied to the process W^2, it can be represented by a sum of a Lebesgue and an Itô integral. In general, a large number of interesting and important phenomena can be described by combining Lebesgue integrals and Itô integrals.

Definition II.3.11. An **Itô process** is a stochastic process on $(\Omega, \mathcal{F}, P)$ that can be written as

$$X(t) \triangleq X(0) + \int_0^t a(s)\,ds + \int_0^t b(s)\,dW(s),$$

where a is Lebesgue integrable, $b \in \mathcal{H}^2$, $X(0)$ is a real number and $t \in \mathcal{T}$. These processes can be represented in differential form as

$$dX(t) = a(t)\,dt + b(t)\,dW(t).$$

In addition, if $a \in \mathcal{H}^2$, the process X will be called an $\boldsymbol{L_2}$**-Itô process**, and the set of all L_2-Itô processes will be denoted by $\mathscr{I}$.

Note that at $t = 0$ there is no information accumulated for X, i.e., the trivial σ-algebra $\mathcal{F}_0 = \{\emptyset, \Omega\}$ defines the past information of X. This is the reason why $X(0)$ is assumed to be a real number. Thus, all the paths of X start at the specified

value $X(0)$ (see [54]). Moreover, the Lebesgue integral component of an Itô process satisfies the following theorem.

Theorem II.3.13. [46] *If $a \in L_1(\Omega \times [0,T], \mathcal{P}, P \otimes \lambda)$ then the process*

$$X(t) = \int_0^t a(s)ds$$

is continuous and adapted.

Observe that from Theorems II.3.8 and II.3.13 one can conclude that every Itô process has a.s. continuous paths and is adapted.

Example II.3.5. Let X be an L_2-Itô process with Lebesgue integrand a and Itô integrand b. The Stratonovich integral of X is

$$\begin{aligned}
\oint_0^t X(s)\, dW(s) &= \int_0^t X(s)\, dW(s) + \frac{1}{2}\langle X, W\rangle_{[0,t]} \\
&= \int_0^t X(s)\, dW(s) + \int_0^t \frac{1}{2} b(s)\, ds.
\end{aligned}$$

Thus, the Stratonovich integral of an L_2-Itô process is always well-defined and is an L_2-Itô process since $b \in \mathcal{H}^2$. □

In stochastic calculus, Itô's formula for functions of Itô processes is the analogue of the chain rule in ordinary calculus.

Theorem II.3.14. *(Itô's formula) Let X be the Itô process $dX(t) = a(t)\, dt + b(t)\, dW(t)$ and consider $g(x,t) \in \mathcal{C}^2([0,\infty) \times \mathbb{R})$. Then, $Y(t) = g(X(t),t)$ is an Itô process and satisfies*

$$dY(t) = \frac{\partial g}{\partial t}(X(t),t)\, dt + \frac{\partial g}{\partial x}(X(t),t)\, dX(t) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(X(t),t)(dX(t))^2.$$

In general, since $dtdt = dtdW(t) = dW(t)dt = 0$ and $(dW(t))^2 = dt$, it follows that

$$\begin{aligned} dY(t) &= \left(\frac{\partial g}{\partial t}(X(t),t) + \frac{\partial g}{\partial x}(X(t),t)a(t) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(X(t),t)b^2(t)\right) dt \\ &\quad + \frac{\partial g}{\partial x}(X(t),t)b(t)\, dW(t). \end{aligned}$$

II.4 STOPPING TIMES

In this section, the book by Protter [48] has been used as the main reference.

Definition II.4.1. Let $\mathbf{F} = \{\mathcal{F}_t\}_{t\in\mathcal{T}}$ be a filtration with $\mathcal{T} = [0,\infty)$. A random variable $\tau : \mathbb{R} \to [0,\infty)$ is a **stopping time** with respect to $\mathbf{F}$ if the event $\{\tau \le t\} \in \mathcal{F}_t$ for every $0 \le t < \infty$.

The left and right limits of a filtration can be defined for any time, t, as

$$\mathcal{F}_{t+} = \bigcap_{s>t} \mathcal{F}_s, \quad \mathcal{F}_{t-} = \sigma\left(\bigcup_{s<t} \mathcal{F}_s\right).$$

For $\mathcal{T} = [0,T]$, if $t = T$, it is convenient to set $\mathcal{F}_{t+} = \mathcal{F}_T$ or, if $t = 0$, $\mathcal{F}_{t-} = \mathcal{F}_0$. It is easy to verify that $\mathcal{F}_s \subseteq \mathcal{F}_{s+} \subseteq \mathcal{F}_{t-} \subseteq \mathcal{F}_t$ for all times $s < t$. Furthermore, $\{\mathcal{F}_{t+}\}_{t\ge 0}$ and $\{\mathcal{F}_{t-}\}_{t\ge 0}$ are themselves filtrations. A filtration is said to be right-continuous if $\mathcal{F}_t = \mathcal{F}_{t+}$ for every t, so, in particular, $\{\mathcal{F}_{t+}\}_{t\ge 0}$ is always the smallest right-continuous filtration larger than $\{\mathcal{F}_t\}_{t\ge 0}$. Hereafter, all filtrations under consideration will be right-continuous. One important consequence of the right continuity of a filtration is given in the following theorem.

Theorem II.4.1. *The event $\{\tau < t\} \in \mathcal{F}_t$, $0 \le t \le \infty$, if and only if τ is a stopping time.*

Proof: Since $\{\tau \leq t\} = \bigcap_{t+\epsilon>s>t} \{\tau < s\}$, for any $\epsilon > 0$, one has $\{\tau \leq t\} \in \bigcap_{s>t} \mathcal{F}_s = \mathcal{F}_t$. So τ is a stopping time. Conversely, assume that τ is a stopping time. Then, it satisfies $\{\tau < t\} = \bigcup_{t>\epsilon>0} \{\tau \leq t-\epsilon\}$, and $\{\tau \leq t-\epsilon\} \in \mathcal{F}_{t-\epsilon}$. Hence $\{\tau < t\} \in \mathcal{F}_t$. ∎

Definition II.4.2. Let X be a stochastic process and let D be a Borel set in $\mathbb{R}$. Define $\tau(\omega) = \inf\{t > 0 : X_t \in D\}$. Then τ is called the **hitting time** of D for X.

Theorem II.4.2. *Let X be an adapted càdlàg stochastic process, and let D be either an open or closed set. Then the hitting time, τ, of D is a stopping time.*

Proof: Let D be an open set. By Theorem II.4.1, it suffices to show that $\{\tau < t\} \in \mathcal{F}_t$, $0 \leq t < \infty$. But

$$\{\tau < t\} = \bigcup_{s \in \mathbb{Q} \cap [0,t)} \{X_s \in D\},$$

because D is open and X has right continuous paths. Given $\{X_s \in D\} = X_s^{-1}(D) \in \mathcal{F}_s$, the result follows. Next, assume that D is a closed set. Define $A_n = \{x : d(x,D) < 1/n\}$, where $d(x,D)$ denotes the distance from x to D. Then A_n is an open set, and the event

$$\{\tau \leq t\} = \bigcap_{n=1}^{\infty} \bigcup_{s \in \mathbb{Q} \cap [0,t)} \underbrace{\{X_s \in A_n\}}_{\mathcal{F}_s \subseteq \mathcal{F}_t} \in \mathcal{F}_t.$$

Thus, τ is a stopping time. ∎

Example II.4.1. Consider a continuous stochastic process X. A special case of a hitting time is $\tau = \inf\{t > 0 : X_t = R\}$, $R \in \mathbb{R}$. It is usually called the *first passage time* for the barrier R. □

Example II.4.2. The random variable

$$\kappa \triangleq \sup\{t \geq 0 : X_t \in D\},$$

is the last time that X_t hits D. In general, it is not a stopping time. The heuristic reason is that the event $\{\kappa \leq t\}$ depends upon the entire future of the process and, thus, would not in general be $\mathcal{F}_t$-measurable. □

Definition II.4.3. Let $(\Omega, \mathcal{F})$ be a measurable space. A **stopped σ-algebra**, $\mathcal{F}_\tau$, is the set of events $A \in \mathcal{F}$ for which

$$A \cap \{\tau \leq t\} \in \mathcal{F}_t, \ t \in \mathbb{R}.$$

Theorem II.4.3. *Let F be a filtration. If τ is a stopping time, then τ is $\mathcal{F}_\tau$-mesurable.*

Proof: One needs to prove that $\{\tau \leq s\} \in \mathcal{F}_\tau$ for all s. Let t be arbitrary. Since τ is a stopping time then

$$\{\tau \leq s\} \cap \{\tau \leq t\} = \{\tau \leq s \wedge t\} \in \mathcal{F}_{s \wedge t} \subseteq \mathcal{F}_t.$$

Thus $\{\tau \leq s\} \in \mathcal{F}_\tau$ from the definition of $\mathcal{F}_\tau$. Hence, τ is $\mathcal{F}_t$ measurable. ■

Lemma II.4.1. *Let τ and σ be stopping times. The random variable $\tau \wedge \sigma$ is also a stopping time.*

Proof: Observe that $\{\tau \wedge \sigma \leq t\} = \{\tau \leq t\} \bigcup \{\sigma \leq t\} \in \mathcal{F}_t$. Then $\tau \wedge \sigma$ is a stopping time. ■

Theorem II.4.4. *Let τ and σ be stopping times. If $\tau(\omega) \leq \sigma(\omega)$ for all $\omega \in \Omega$, then $\mathcal{F}_\tau \subset \mathcal{F}_\sigma$.*

Proof: Observe that $A\bigcap\{\tau \leq t\} = (A\bigcap\{\tau \leq t\})\bigcap\{\sigma \leq t\} \in \mathcal{F}_t$. This implies that $\mathcal{F}_\tau \subset \mathcal{F}_\sigma$. ∎

Definition II.4.4. Let X be a stochastic process on $\mathcal{T}$, and let τ be a stopping time. A **stopped** or **truncated process** is any process of the form

$$X^\tau(t,\omega) \triangleq X(\tau(\omega)\wedge t,\omega) = X(t,\omega)\mathbb{1}_{\{t<\tau\}} + X(\tau(\omega),\omega)\mathbb{1}_{\{t\geq\tau\}},$$

where the random variable $X_\tau(\omega) \triangleq X(\tau(\omega),\omega)$ is called a **stopped** random variable.

Observe that if the stopped process $X^\tau(t,\omega)$ is restricted to the *stochastic interval*

$$[0,\tau] \triangleq \{(t,\omega) \in [0,\infty)\times\omega : 0 \leq t \leq \tau(\omega)\},$$

then $X^\tau(t,\omega) = X(t,\omega)\mathbb{1}_{[0,\tau(\omega)]}$. Usually, a path of this process is denoted simply by $X(t\wedge\tau)$.

Theorem II.4.5. *Let $X \in \mathcal{H}^2$ and τ be a stopping time. Define the stopped random variable $Y_\tau = \int_0^\tau X(s)\,dW(s) = \int_0^t X(s)\mathbb{1}_{[0,\tau]}(s)\,dW(s)$. Then*

$$Y^\tau(t) = Y(t\wedge\tau) = \int_0^{t\wedge\tau} X(s)\,dW(s) = \int_0^t X(s)\mathbb{1}_{[0,\tau]}(s)\,dW(s).$$

The proof of this theorem is not trivial. A nice treatment can be found in [50, 55]. A direct application of the previous theorem gives a similar result for Stratonovich integrals.

Corollary II.4.1. *Let $X(t) = \displaystyle\oint_0^t v(s)\,dW(s)$, where v is an L_2-Itô process. Then, the stopped random process $Y^\tau = \displaystyle\oint_0^{t\wedge\tau} X(s)\,dW(s)$ satisfies*

$$Y^\tau(t) = \oint_0^t X(s)\mathbb{1}_{[0,\tau]}(s)\,dW(s).$$

Proof: Since X can be writen as

$$X(t) = \int_0^t v(s)\, dW(s) + \int_0^t \frac{b(s)}{2}\, ds,$$

it follows, by direct application of Theorem II.4.5, that

$$\begin{aligned}
X(t \wedge \tau_R) &= X(t)\mathbb{1}_{[0,\tau_R]}(t) \\
&= \int_0^{t\wedge\tau_R} v(s)\, dW(s) + \int_0^{t\wedge\tau_R} \frac{b(s)}{2}\, ds \\
&= \int_0^t v(s)\mathbb{1}_{[0,\tau_R]}(s)\, dW(s) + \int_0^t \frac{b(s)}{2}\mathbb{1}_{[0,\tau_R]}(s)\, ds \\
&= \int_0^t v(s)\mathbb{1}_{[0,\tau_R]}(s)\, dW(s) \\
&\quad +\frac{1}{2}\left\langle \int_0^{\cdot} a(s)\mathbb{1}_{[0,\tau_R]}(s)\, ds + \int_0^{\cdot} b(s)\mathbb{1}_{[0,\tau_R]}(s)\, dW(s), W \right\rangle_{[0,t]} \\
&= \int_0^t v(s)\mathbb{1}_{[0,\tau_R]}(s)\, dW(s) + \frac{1}{2}\left\langle v\mathbb{1}_{[0,\tau_R]}, W\right\rangle_{[0,t]} \\
&= \oint_0^t v(s)\mathbb{1}_{[0,\tau_R]}(s)\, dW(s).
\end{aligned}$$

This completes the proof. ∎

The main result of this section is described in the following theorem. It will be used later in Chapter IV for the convergence of Fliess operators over stochastic time intervals.

Theorem II.4.6. *Let* $X(t) = \oint_0^t v(s)\, dW(s)$*, where* v *is an* L_2*-Itô process. Then:*

i. There exists a strictly positive stopping time $\tau_R = \inf\{t \in \mathcal{T} : |X(t)| = R\}$ *for any finite real* $R > 0$*.*

ii. $X^{\tau_R}(t,\omega)$ *restricted to* $[0, \tau_R]$ *is a well-defined* L_2*-bounded, a.s. continuous and adapted* L_2*-Itô process.*

Proof: For part (i), since v is an L_2-Itô process, it can be written as

$$v(t) = \int_0^t a(s)\,ds + \int_0^t b(s)\,dW(s).$$

Therefore, τ_R is a well-defined stopping time since (by Theorems II.3.8 and II.3.13)

$$X(t) = \int_0^t \frac{b(s)}{2}\,ds + \int_0^t v(s)\,dW(s),$$

and the absolute value function is continuous. Regarding part ii, it is well known that since X is adapted and a.s. continuous, the stopped process is also adapted and a.s. continuous [48]. Now evaluate $\int_0^T \mathbf{E}\left[X^2(s)\right] ds$. By Itô's formula,

$$X^2(t) = 2\int_0^t X(s)\bar{b}(s)\,ds + 2\int_0^t X(s)v(s)\,dW(s) + \int_0^t v^2(s)\,ds,$$

where $\bar{b}(s) = b(s)/2$. Since $2k_1k_2 \leq k_1^2 + k_2^2$, for all $k_1, k_2 \in \mathbb{R}$, it follows from Theorem II.3.7 that

$$\mathbf{E}[X^2(t)] \leq \mathbf{E}\left[\int_0^t \left(X^2(s) + \bar{b}^2(s)\right)\,ds + \int_0^t v^2(s)\,ds\right].$$

The L_2 bound for X restricted to $[0, \tau_R]$ can be calculated from the previous expression as

$$\mathbf{E}[X^2(t \wedge \tau_R)] \leq \mathbf{E}\left[\int_0^t \left(X^2(s \wedge \tau_R) + \bar{b}^2(s \wedge \tau_R)\right)\,ds + \int_0^t v^2(s \wedge \tau_R)\,ds\right].$$

Given that $\bar{b}, v \in L_2(\Omega \times [0,T], \mathcal{P}, P \otimes \lambda)$, define a real number $M \geq \mathbf{E}\left[\int_0^t \bar{b}^2(s)\,ds\right] + \mathbf{E}\left[\int_0^t v^2(s)\,ds\right]$. Then by Fubini's theorem

$$\mathbf{E}[X^2(t \wedge \tau_R)] \leq M + \int_0^t \mathbf{E}\left[X^2(s \wedge \tau_R)\right]\,ds. \tag{II.4.1}$$

If this inequality is used in the right-hand-side of (II.4.1) then

$$\mathbf{E}[X^2(t \wedge \tau_R)] \leq M + Mt + \int_0^t \int_0^s \mathbf{E}\left[X^2(r \wedge \tau_R)\right]\,drds.$$

Repeating this procedure infinitely many times and knowing that $\mathbf{E}\left[X^2(r \wedge \tau_R)\right] \leq R^2$, it follows that

$$\mathbf{E}[X^2(t \wedge \tau_R)] \leq \lim_{p \to \infty} M \sum_{n=0}^{p} \frac{t^n}{n!} + R^2 \frac{t^p}{p!} = Me^t$$

for a fixed $t \in [0, T]$. This implies that $\int_0^T \mathbf{E}\left[X^2(s)\right] ds < \infty$, and thus $X^{\tau_R}(t, \omega) \in L_2(\Omega \times [0, T], \mathcal{P}, P \otimes \lambda)$. This completes the proof. ■

CHAPTER III

FORMAL POWER SERIES

This chapter presents some elements from the fundamental theory of formal power series. The treatment relies heavily on [4, 24]. Formal power series appear naturally in the context of language theory; therefore, the terminology of this subject will be adopted. The definition of formal languages and formal power series are introduced first. Then, various products of formal power series are defined along with their basic properties. Some important classes of formal power series are also addressed, in particular, rational and recognizable series. Then, the relationship between these classes of series is presented in Schützenberger's Theorem. This naturally leads to the question of which products of formal power series are closed over these classes. Here, a new and simpler proof of Ferfera's sufficient condition for the rationality of the composition product of two formal power series is presented. Finally, the main properties and algebraic tools that will be used in Chapter V to define the interconnection of systems are described.

III.1 FORMAL LANGUAGES

An *alphabet* is a non-empty set of symbols, $X = \{x_0, x_1, \ldots, x_m\}$. Each element of X is called a *letter*, and any string of symbols in X, $\eta = x_{i_k} \cdots x_{i_1}$ is called a *word*. The length of η, $|\eta|$, is the number of symbols in η, and $|\eta|_{x_i}$ is the number of times the letter x_i appears in η. The set of all words of length k is denoted by X^k. The

set of all words, including the empty word, $\emptyset$, is denoted by X^*. A *language* is any subset of X^*. The set X^* forms a monoid under the catenation product,

$$\mathscr{C} \;:\; X^* \times X^* \to X^*$$

$$(\eta, \xi) \mapsto \eta\xi.$$

Clearly, for any $\eta, \xi, \nu \in X^*$, $\mathscr{C}$ is associative: $(\eta\xi)\nu = \eta(\xi\nu)$. The *empty word* $\emptyset$ is the identity element for $\mathscr{C}$, $\emptyset\eta = \eta\emptyset = \eta, \;\; \forall \eta \in X^*$. The triple $(X^*, \mathscr{C}, \emptyset)$ is in fact a *free monoid* of X.

Definition III.1.1. Let $(M, \Box, e)$ and $(M', \Box', e')$ be two arbitrary monoids. A mapping $\varrho : M \to M'$ is called a **morphism** if

$$\varrho(\eta \Box \xi) = \varrho(\eta) \Box' \varrho(\xi), \;\; \forall \eta, \xi \in M, \tag{III.1.1}$$

where $\varrho(e) = e'$. If, in addition, ϱ is bijective, then it is called an **isomorphism**.

Note that any morphism ϱ on X^* is completely determined by its action on X. In other words, for any word $\eta = x_{i_k} \cdots x_{i_1} \in X^*$,

$$\varrho(x_{i_k} \cdots x_{i_1}) = \varrho(x_{i_k}) \cdots \varrho(x_{i_1}).$$

If ϱ is injective, i.e., $\varrho(\eta) = \varrho(\xi)$ implies $\eta = \xi, \;\; \forall \eta, \xi \in X^*$, then ϱ is called a *code*.

III.1.1 Formal power series

Given the alphabet $X = \{x_0, x_1, \ldots, x_m\}$ and finite $\ell \in \mathbb{N}$, a formal power series is any function $c : X^* \to \mathbb{R}^\ell$. The value of c for a specific word $\eta \in X^*$ is denoted by (c, η). In particular, $(c, \emptyset)$ is referred to as the *constant term*. Typically, c is written as the formal sum

$$c = \sum_{\eta \in X^*} (c, \eta)\eta.$$

The collection of all series generated with the alphabet X is denoted by $\mathbb{R}^\ell\langle\langle X\rangle\rangle$. If $c, d \in \mathbb{R}^\ell\langle\langle X\rangle\rangle$, their *sum* is given by

$$c + d = \sum_{\eta \in X^*} (c + d, \eta)\eta = \sum_{\eta \in X^*} \left[(c, \eta) + (d, \eta)\right]\eta,$$

and their *scalar product* is defined by

$$rc = \sum_{\eta \in X^*} (rc, \eta)\eta = \sum_{\eta \in X^*} r(c, \eta)\eta$$

for $r \in \mathbb{R}$. The set of formal power series with these two operations forms an $\mathbb{R}$-vector space. The *Cauchy product* is defined by

$$cd = \sum_{\eta \in X^*} (cd, \eta)\eta = \sum_{\eta \in X^*} \sum_{\xi\zeta = \eta} (c, \xi)(d, \zeta)\eta,$$

where the interior summation is obviously finite. The $\mathbb{R}$-vector space $\mathbb{R}\langle\langle X\rangle\rangle$ endowed with the Cauchy product forms an $\mathbb{R}$-algebra. For $c \in \mathbb{R}^\ell\langle\langle X\rangle\rangle$, the language

$$\mathrm{supp}(c) = \{\eta \in X^* : (c, \eta) \neq 0\}$$

is called the *support* of c. The subset of $\mathbb{R}^\ell\langle\langle X\rangle\rangle$ consisting of all series with finite support is denoted by $\mathbb{R}^\ell\langle X\rangle$. Its elements are called *polynomials*. c is said to be *proper* if $\emptyset \notin \mathrm{supp}(c)$. $\mathbb{R}^\ell\langle\langle X\rangle\rangle$ forms a complete ultrametric space under the mapping

$$\mathrm{dist} : \mathbb{R}^\ell\langle\langle X\rangle\rangle \times \mathbb{R}^\ell\langle\langle X\rangle\rangle \to \mathbb{R}^+ \cup \{0\}$$

$$(c, d) \mapsto \mathrm{dist}(c, d) = \sigma^{\mathrm{ord}(c-d)},$$

where $0 < \sigma < 1$ and $\mathrm{ord}(c) \triangleq \inf\{|\eta| : (c, \eta) \neq 0\}$ if $c \neq 0$, otherwise $\mathrm{ord}(c) = \infty$. dist satisfies the inequality

$$\mathrm{dist}(c, d) \leq \max\{\mathrm{dist}(c, e), \mathrm{dist}(e, d)\}, \quad \forall c, d, e \in \mathbb{R}\langle\langle X\rangle\rangle,$$

which can be viewed as a stronger version of the usual triangle inequality.

Let $\{c_i\}_{i\in I}$ be a family of series, where I is a set of indexes in $\mathbb{N}$. This family is called *summable* if there exists a formal series c such that for all $\epsilon > 0$, there is a finite subset I' of I such that for all finite subsets J of I that contain I'

$$\text{dist}\left(\sum_{j\in J} c_j, c\right) \leq \epsilon.$$

The series c is called the sum of the family $\{c_i\}_{i\in I}$, and it is unique.

A family $\{c_i\}_{i\in I}$ is said to be *locally finite* if for every word $\eta \in X^*$ there exists only a finite number of indexes $i \in I$ such that $(c_i, \eta) \neq 0$. It is easy to show that every locally finite familly is summable, and

$$(c, \eta) = \sum_{i\in I}(c_i, \eta), \ \ \eta \in X^*,$$

where this summation is finite. It is not true in general that a summable family is always locally finite. It is useful to define the *left-shift operator*, ξ^{-1}, as follows:

$$\begin{aligned} \xi^{-1} \ : \ & X^* \to \mathbb{R}\langle X\rangle \\ & \eta \mapsto \begin{cases} \eta' & : \ \eta = \xi\eta' \\ 0 & : \ \text{otherwise.} \end{cases} \end{aligned}$$

Note from the definition that η can be mapped to the polynomial 0, and that clearly there exists a bijection between the non-empty words in X^* and the monic polynomials in $\mathbb{R}\langle X\rangle$. For any series $c \in \mathbb{R}\langle\langle X\rangle\rangle$, the definition can be extended as

$$\begin{aligned} \xi^{-1}(c) &= \sum_{\eta\in X^*}(c,\eta)\xi^{-1}(\eta) \\ &= \sum_{\eta\in X^*}(c,\xi\eta)\eta. \end{aligned}$$

It is simple to verify that $\xi^{-1}(\cdot)$ is a linear operator on the $\mathbb{R}$-vector space $\mathbb{R}\langle\langle X\rangle\rangle$. That is,

$$\xi^{-1}(r_1c_1 + r_2c_2) = r_1\xi^{-1}(c_1) + r_2\xi^{-1}(c_2)$$

for all $r_i \in \mathbb{R}$ and $c_i \in \mathbb{R}^{\ell}\langle\langle X\rangle\rangle$.

Lemma III.1.1. *For any $x \in X$, $\xi, \nu \in X^*$ and series $c, d \in \mathbb{R}\langle\langle X\rangle\rangle$ the following properties hold:*

i. $(\xi\nu)^{-1}(c) = \nu^{-1}(\xi^{-1}(c))$

ii. $x^{-1}(cd) = x^{-1}(c)d + (c, \emptyset)x^{-1}(d)$.

III.1.2 Formal power series products

Definition III.1.2. Let $c, d \in \mathbb{R}\langle\langle X\rangle\rangle$. The **Hadamard product** of c and d is defined as

$$c \odot d \triangleq \sum_{\nu \in X^*} (c, \nu)(d, \nu)\nu.$$

Here, $\mathrm{supp}(c \odot d) = \mathrm{supp}(c) \cap \mathrm{supp}(d)$.

Example III.1.1. Let $X = \{x_0, x_1, \ldots, x_m\}$ be an alphabet and define for a fixed x_i and k the language $L = \{\eta \in X^*, |\eta|_{x_i} = k\}$. The characteristic series generated by L is defined as $\boldsymbol{L} = \sum_{\eta \in L} \eta$. If $c \in \mathbb{R}\langle\langle X\rangle\rangle$ is arbitrary, the restriction of c to the language L can be written using the Hadamard product as

$$c_L \triangleq c|_L = c \odot \boldsymbol{L}.$$

□

Definition III.1.3. The **shuffle** of two words $\eta, \xi \in X^*$ is defined to be the language

$$\begin{aligned} S_{\eta,\xi} &= \{\nu \in X^* \, : \nu = \eta_1\xi_1\eta_2\xi_2\cdots\eta_n\xi_n, \ \eta_i, \xi_i \in X^*, \\ &\qquad \eta = \eta_1\eta_2\cdots\eta_n, \ \xi = \xi_1\xi_2\cdots\xi_n, \ n \geq 0\}. \end{aligned}$$

In particular, $S_{\eta,\emptyset} = \{\eta\}$ and $S_{\emptyset,\xi} = \{\xi\}$.

Definition III.1.4. Let $\eta = x_j\eta'$ and $\xi = x_k\xi'$, where $\eta', \xi' \in X^*$ and $x_j, x_k \in X$. The **shuffle product** of η and ξ is recursively defined as

$$\eta \sqcup\!\sqcup \xi = x_j[\eta' \sqcup\!\sqcup \xi] + x_k[\eta \sqcup\!\sqcup \xi'],$$

where $\emptyset \sqcup\!\sqcup \emptyset = \emptyset$ and $\xi \sqcup\!\sqcup \emptyset = \emptyset \sqcup\!\sqcup \xi = \xi$.

It is easily verified that $\eta \sqcup\!\sqcup \xi$ yields a polynomial involving words of only length $|\eta| + |\xi|$ and $\mathrm{supp}\{\eta \sqcup\!\sqcup \xi\} = S_{\eta,\xi}$.

Example III.1.2. Suppose $X = \{x_0, x_1, x_2, x_3\}$. Then

$$\begin{aligned} x_0x_1 \sqcup\!\sqcup x_2x_3 &= x_0[x_1 \sqcup\!\sqcup x_2x_3] + x_2[x_0x_1 \sqcup\!\sqcup x_3] \\ &= x_0(x_1[\emptyset \sqcup\!\sqcup x_2x_3] + x_2[x_1 \sqcup\!\sqcup x_3]) + x_2(x_0[x_1 \sqcup\!\sqcup x_3] + x_3[x_0x_1 \sqcup\!\sqcup \emptyset]) \\ &= x_0x_1x_2x_3 + x_0x_2x_1x_3 + x_0x_2x_3x_1 + x_2x_0x_1x_3 + x_2x_0x_3x_1 + x_2x_3x_0x_1, \end{aligned}$$

$$\begin{aligned} x_0x_1 \sqcup\!\sqcup x_1x_3 &= x_0[x_1 \sqcup\!\sqcup x_1x_3] + x_1[x_0x_1 \sqcup\!\sqcup x_3] \\ &= x_0(x_1[\emptyset \sqcup\!\sqcup x_1x_3] + x_1[x_1 \sqcup\!\sqcup x_3]) + x_1(x_0[x_1 \sqcup\!\sqcup x_3] + x_3[x_0x_1 \sqcup\!\sqcup \emptyset]) \\ &= 2x_0x_1x_1x_3 + x_0x_1x_3x_1 + x_1x_0x_1x_3 + x_1x_0x_3x_1 + x_1x_3x_0x_1, \end{aligned}$$

and

$$S_{x_0x_1,x_2x_3} = \{x_0x_1x_2x_3, x_0x_2x_1x_3, x_0x_2x_3x_1, x_2x_0x_1x_3, x_2x_0x_3x_1, x_2x_3x_0x_1\},$$

$$S_{x_0x_1,x_1x_3} = \{x_0x_1x_1x_3, x_0x_2x_3x_1, x_2x_0x_1x_3, x_2x_0x_3x_1, x_2x_3x_0x_1\}.$$

Observe that each element of the shuffle language involves a combination of the letters in each given word which preserves their relative order, i.e., x_0 is always to the left of x_1 and x_2 is always to the left of x_3. □

Consider the next example where the behavior of the product of deterministic iterated integrals is presented.

Example III.1.3. Let u be an m-dimensional piecewise continuous, real-valued function defined over the finite interval $[t_0, t_1]$. Recall that for $\eta \in X^*$ the iterated integral E_η was defined recursively as

$$E_\eta[u](t) = E_{x_j\eta'}[u](t) = \int_{t_0}^{t} u_j(\tau)E_{\eta'}[u](\tau)d\tau$$

with $E_\emptyset[u](t) \equiv 1$ for all $t \in [t_0, t_1]$, and t_0 was assumed to be 0. This iterated integral can be extended for any polynomial $p \in \mathbb{R}\langle X\rangle$ in the following manner,

$$E_p[u](t) = \sum_{\eta \in \text{supp}(p)} (p, \eta)E_\eta[u](t).$$

Let $\mathcal{E}(\mathbb{R}\langle X\rangle)$ be the set of all such iterated integrals. Now observe for each $\eta, \xi \in X^*$ that from the integration by parts formula

$$\begin{aligned} E_\eta[u](t)E_\xi[u](t) &= E_{x_j\eta'}[u](t)E_{x_k\xi'}[u](t) \\ &= \int_0^t u_j(\tau)\left[E_{\eta'}[u](\tau)E_\xi[u](\tau)\right]d\tau \end{aligned}$$

$$
\begin{aligned}
&+\int_0^t u_k(\tau)\left[E_\eta[u](\tau)E_{\xi'}[u](\tau)\right]d\tau \\
=\;& E_{x_j}\left[E_{\eta'}[u](\tau)E_\xi[u](\tau)\right](t) \\
&+E_{x_k}(\tau)\left[E_\eta[u](\tau)E_{\xi'}[u](\tau)\right](t). \\
=\;& E_{x_j(\eta' \sqcup\!\sqcup\, \xi)+x_k(\eta \sqcup\!\sqcup\, \xi')}[u](t).
\end{aligned}
$$

As a consequence, any product of two integrals is again an iterated integral in $\mathcal{E}(\mathbb{R}\langle X\rangle)$, and it can be expresed in terms of the shuffle product. $\square$

The shuffle product definition is linearly extended directly to series $c, d \in \mathbb{R}\langle\langle X\rangle\rangle$ as

$$
c \sqcup\!\sqcup\, d = \sum_{\eta,\xi\in X^*} [(c,\eta)(d,\xi)]\eta \sqcup\!\sqcup\, \xi. \tag{III.1.2}
$$

For a specific word $\nu \in X^*$, the coefficient

$$
(\eta \sqcup\!\sqcup\, \xi, \nu) = 0 \;\text{ if }\; |\eta| + |\xi| \neq |\nu|\,.
$$

Therefore, the summation in (III.1.2) is well-defined since the family of polynomials $\{\eta \sqcup\!\sqcup\, \xi : \forall \eta, \xi \in X^*\}$ is locally finite. An equivalent expression for this product is

$$
c \sqcup\!\sqcup\, d = \sum_{\nu\in X^*} (c \sqcup\!\sqcup\, d, \nu)\nu,
$$

where

$$
(c \sqcup\!\sqcup\, d, \nu) = \sum_{\eta,\xi\in X^*} (c,\eta)(d,\xi)(\eta \sqcup\!\sqcup\, \xi, \nu).
$$

In the case where $c, d \in \mathbb{R}^\ell\langle\langle X\rangle\rangle$, the shuffle product is defined componentwise, i.e., the i-th component of $c \sqcup\!\sqcup\, d$ is $(c \sqcup\!\sqcup\, d, \nu)_i = (c_i \sqcup\!\sqcup\, d_i, \nu)$, where $\nu \in X^*$ and $1 \leq i \leq \ell$. Moreover, it is easy to show that the shuffle product is commutative, associative and distributive with respect to the addition. As a consequence, $\mathbb{R}\langle X\rangle$ and $\mathbb{R}\langle\langle X\rangle\rangle$ are $\mathbb{R}$-algebras with multiplicative element 1.

Theorem III.1.1. *Let $\eta, \xi \in X^*$ such that $|\eta| = n_1$ and $|\xi| = n_2$. Then* $\mathrm{supp}\{\eta \sqcup\!\sqcup\, \xi\}$ *has at most $\binom{n_1+n_2}{n_1}$ unique words.*

Proof: The proof is by induction over $n = n_1 + n_2$. For $n = 0$ and $n = 1$ the result is trivial. If $n_1 = 0$ and $n_2 \in \mathbb{N}$, then $S_{\emptyset,\xi} = \{\xi\}$, and therefore, $\#(S_{\emptyset,\xi}) = \binom{0+n_2}{0} = 1$, where $\#$ denotes the cardinality of $S_{\emptyset,\xi}$. Now assume the hypothesis holds up to some fixed $n \geq 0$, i.e., $\#(S_{\eta,\xi}) \leq \binom{n}{n_1}$. Let $\eta = x_{j_{n_1+1}} \cdots x_{j_1}$ and $\xi = x_{i_{n_2}} \cdots x_{i_1}$. Then by the shuffle product definition

$$\eta \sqcup\!\sqcup\, \xi = x_{j_{n_1+1}} \underbrace{(x_{j_{n_1}} \cdots x_{j_1} \sqcup\!\sqcup\, \xi)}_{\leq \binom{n}{n_1} \text{ words}} + x_{i_{n_2}} \underbrace{(\eta \sqcup\!\sqcup\, x_{i_{n_2-1}} \cdots x_{i_1})}_{\leq \binom{n}{n_1+1} \text{ words}}.$$

Since $\binom{n}{n_1} + \binom{n}{n_1+1} = \binom{n+1}{n_1+1}$, it follows that $\#(S_{\eta,\xi}) \leq \binom{n+1}{n_1+1}$, and the theorem is proved. ■

Example III.1.4. From Example III.1.2,

$$\#(S_{x_0x_1,x_2x_3}) = 6 = \binom{4}{2} \quad \text{and} \quad \#(S_{x_0x_1,x_1x_3}) = 5 < \binom{4}{2} = 6.$$

□

Theorem III.1.2. [24] *Let be X an alphabet. For any $x \in X$:*

i. $x^{\sqcup\!\sqcup\, k} \triangleq \underbrace{x \sqcup\!\sqcup\, x \sqcup\!\sqcup \cdots \sqcup\!\sqcup\, x}_{k \text{ times}} = k!\, x^k$

ii. $x^k \sqcup\!\sqcup\, x^{n-k} = \binom{n}{k} x^n, \quad 0 \leq k \leq n$

iii. $(x^j)^{\sqcup\!\sqcup\, k} = \dfrac{(jk)!}{(j!)^k}\, x^{jk}, \quad j, k \geq 1$

iv. $\left(\sum_{j=0}^{\infty} x^j\right)^{\sqcup\!\sqcup\, k} = \sum_{j=0}^{\infty} k^j x^j, \quad k \geq 1.$

Theorem III.1.3. *The left-shift operator acts as a derivation with respect to the shuffle product. That is,* $\forall c, d \in \mathbb{R}\langle\langle X\rangle\rangle$ *and* $x_k \in X$

$$x_k^{-1}(c \sqcup\!\sqcup d) = x_k^{-1}(c) \sqcup\!\sqcup d + c \sqcup\!\sqcup x_k^{-1}(d).$$

Proof: Without loss of generality, consider the nonempty words $\eta = x_j\eta', \xi = x_i\xi' \in X^*$. From Definition III.1.4

$$\begin{aligned}
x_k^{-1}(\eta \sqcup\!\sqcup \xi) &= x_k^{-1}(x_j(\eta' \sqcup\!\sqcup \xi) + x_i(\eta \sqcup\!\sqcup \xi')) \\
&= x_k^{-1}x_j(\eta' \sqcup\!\sqcup \xi) + x_k^{-1}x_i(\eta \sqcup\!\sqcup \xi') \\
&= \begin{cases} 0 & : \ j \neq k, i \neq k \\ \eta' \sqcup\!\sqcup \xi & : \ j = k \neq i \\ \eta \sqcup\!\sqcup \xi' & : \ i = k \neq j \\ \eta' \sqcup\!\sqcup \xi + \eta \sqcup\!\sqcup \xi' & : \ i = j = k \end{cases} \\
&= x_k^{-1}(\eta) \sqcup\!\sqcup \xi + \eta \sqcup\!\sqcup x_k^{-1}(\xi).
\end{aligned}$$

Next, let $c, d \in \mathbb{R}\langle\langle X\rangle\rangle$. By the linearity of the left-shift operator and the previous identity

$$\begin{aligned}
x_k^{-1}(c \sqcup\!\sqcup d) &= x_k^{-1}\left(\sum_{\eta,\xi \in X^*} (c,\eta)(d,\xi)\eta \sqcup\!\sqcup \xi\right) \\
&= \sum_{\eta,\xi \in X^*} (c,\eta)(d,\xi)x_k^{-1}(\eta \sqcup\!\sqcup \xi) \\
&= \sum_{\eta,\xi \in X^*} (c,\eta)(d,\xi)\left[x_k^{-1}(\eta) \sqcup\!\sqcup \xi + \eta \sqcup\!\sqcup x_k^{-1}(\xi)\right] \\
&= \sum_{\eta,\xi \in X^*} (c,\eta)(d,\xi)x_k^{-1}(\eta) \sqcup\!\sqcup \xi + \sum_{\eta,\xi \in X^*} (c,\eta)(d,\xi)\eta \sqcup\!\sqcup x_k^{-1}(\xi) \\
&= \sum_{\eta,\xi \in X^*} (x_k^{-1}(c),\eta)\eta \sqcup\!\sqcup (d,\xi)\xi + \sum_{\eta,\xi \in X^*} (c,\eta)\eta \sqcup\!\sqcup (x_k^{-1}(d),\xi)\xi \\
&= x_k^{-1}(c) \sqcup\!\sqcup d + c \sqcup\!\sqcup x_k^{-1}(d).
\end{aligned}$$

■

Definition III.1.5. [13, 14] Given $\eta = x_0^{n_k} x_{i_k} x_0^{n_{k-1}} x_{i_{k-1}} \cdots x_0^{n_1} x_{i_1} x_0^{n_0} \in X^*$ and the series $d \in \mathbb{R}^m\langle\langle X \rangle\rangle$, the **composition** of η with d is defined recursively as

$$\eta \circ d = \begin{cases} \eta & : \quad |\eta|_{x_i} = 0,\ \forall\, i \neq 0 \\ x_0^{k+1}[d_i \sqcup\!\sqcup (\eta' \circ d)] & : \quad \eta = x_0^k x_i \eta',\ k \in \mathbb{N}, \\ & \qquad i \neq 0,\ \eta' \in X^*, \end{cases}$$

where $d_i : \xi \mapsto (d, \xi)_i$, and $(d, \xi)_i$ is the i-th component of (d, ξ). Furthermore, the **composition** of a series $c \in \mathbb{R}^\ell\langle\langle X \rangle\rangle$ with d is

$$c \circ d = \sum_{\eta \in X^*} (c, \eta) \eta \circ d.$$

Theorem III.1.4. [24, 28] *Let $d \in \mathbb{R}^m\langle\langle X \rangle\rangle$ then the family of series $\{\eta \circ d \ : \ \eta \in X^*\}$ is locally finite and therefore summable.*

Proof: Given that any word in X^* can be written as

$$\eta = x_0^{n_k} x_{i_k} x_0^{n_{k-1}} x_{i_{k-1}} \cdots x_0^{n_1} x_{i_1} x_0^{n_0},$$

it follows that

$$\mathrm{ord}(\eta \circ d) = n_0 + k + \sum_{j=1}^{k} n_j + \mathrm{ord}(d_{i_j}) = |\eta| + \sum_{j=1}^{|\eta| - |\eta|_{x_0}} \mathrm{ord}(d_{i_j}).$$

Thus, for any $\xi \in X^*$,

$$\begin{aligned} I_d(\xi) &\triangleq \{\eta \in X^* \ : \ (\eta \circ d) \neq 0\} \\ &\subset \{\eta \in X^* \ : \ \mathrm{ord}(\eta \circ d) \leq |\xi|\} \\ &= \left\{ \eta \in X^* \ : \ |\eta| + \sum_{j=1}^{|\eta| - |\eta|_{x_0}} \mathrm{ord}(d_{i_j}) \leq |\xi| \right\}. \end{aligned}$$

One can see that the latter set is finite, ensuring $I_d(\xi)$ is finite, which in turn implies summability. ■

It is easy to verify that the composition product is linear with respect to its first argument. That is, for all $c, d, e \in \mathbb{R}^m\langle\langle X\rangle\rangle$, and $\alpha, \beta \in \mathbb{R}$,

$$(\alpha c + \beta d) \circ e = \alpha(c \circ e) + \beta(d \circ e).$$

However, in general, $c \circ (\alpha d + \beta e) \neq \alpha(c \circ d) + \beta(c \circ e)$. An exception is the case when a series is linear. A series $c \in \mathbb{R}^\ell\langle\langle X\rangle\rangle$ is called *linear* if

$$\mathrm{supp}(c) \subseteq \{\eta \in X^* \ : \ \eta = x_0^{n_1} x_i x_0^{n_0},\ i \in \{1, 2, \ldots, n\},\ n_0, n_1 \geq 0\}.$$

Given a word $\eta = x_0^{n_1} x_i x_0^{n_0}$ and using the bilinearity of the shuffle product, observe that

$$\begin{aligned}
\eta \circ (\alpha d + \beta e) &= x_0^{n_1+1}(\alpha d + \beta e)_i \sqcup\!\sqcup\, x_0^{n_0} \\
&= \alpha x_0^{n_1+1}(d_i \sqcup\!\sqcup\, x_0^{n_0}) + \beta x_0^{n_1+1}(e_i \sqcup\!\sqcup\, x_0^{n_0}) \\
&= \alpha(\eta \circ d) + \beta(\eta \circ e).
\end{aligned}$$

Hence,

$$\begin{aligned}
c \circ (\alpha d + \beta e) &= \sum_{\eta \in X^*} \alpha(c, \eta)\eta \circ (\alpha d + \beta e) \\
&= \sum_{\eta \in X^*} (\alpha(c, \eta)\eta \circ d + \beta(c, \eta)\eta \circ e) \\
&= \alpha(c \circ d) + \beta(c \circ e).
\end{aligned}$$

One can verify that the composition product is associative and distributive from the right with respect to the shuffle product. Unfortunately, it lacks an identity element and is not commutative. Consequently, $(\mathbb{R}^\ell\langle\langle X\rangle\rangle, \circ)$ and $(\mathbb{R}^\ell\langle X\rangle, \circ)$ are only semigroups. Some additional properties are given next.

Lemma III.1.2. [28] *The following properties hold for the composition product (Here $\mathbb{1}$ is a column vector of m ones):*

i. $0 \circ d = 0$, $\forall d \in \mathbb{R}\langle\langle X \rangle\rangle$.

ii. $c \circ 0 = c_0 \triangleq \sum_{n \geq 0} (c, x_0^n) x_0^n$. *(Thereby, $c \circ 0 = 0$ if and only if $c_0 = 0$).*

iii. $c_0 \circ d = c_0$, $\forall d \in \mathbb{R}^m\langle\langle X \rangle\rangle$. *(In particular, $1 \circ d = 1$.)*

iv. $c \circ \mathbb{1} = c_{\mathbb{1}} \triangleq \sum_{\eta \in X^*} (c, \eta) x_0^{|\eta|}$. *(Thus, $c \circ \mathbb{1} = c$ if and only if $c_0 = c$.)*

III.2 RATIONAL AND RECOGNIZABLE SERIES

In this section, the notion of rationality is introduced in terms of four rational operations on $\mathbb{R}\langle\langle X \rangle\rangle$: addition, scalar multiplication, catenation and inversion. Then, conditions for the existence of a linear representation of a series are presented. This defines the concept of recognizable series. Finally, it is determined whether or not certain products of formal power series preserve rationality. When rationality is not preserved in general, sufficient conditions for preserving rationality are provided.

III.2.1 Rational series

A series $c \in \mathbb{R}\langle\langle X \rangle\rangle$ is called **invertible** if there exists a series $c^{-1} \in \mathbb{R}\langle\langle X \rangle\rangle$ such that $cc^{-1} = c^{-1}c = 1$. If c is not proper, one can always write

$$c = (c, \emptyset)(1 - c'),$$

where $c' \in \mathbb{R}\langle\langle X \rangle\rangle$ is proper. In which case,

$$c^{-1} = \frac{1}{(c, \emptyset)}(1 - c)^{-1} = \frac{1}{(c, \emptyset)}(c')^*,$$

where

$$(c')^* \triangleq \sum_{i=0}^{\infty} (c')^i \ .$$

It can be easily shown that c is invertible if and only if c is not proper. Now, let S be any subalgebra of the catenation $\mathbb{R}$-algebra on $\mathbb{R}\langle\langle X \rangle\rangle$. S is called *rationally closed* when every invertible series $c \in S$ has $c^{-1} \in S$. The *rational closure* of any subset $E \subset \mathbb{R}\langle\langle X \rangle\rangle$ is the smallest rationally closed subalgebra of $\mathbb{R}\langle\langle X \rangle\rangle$ such that it contains E.

Definition III.2.1. A series $c \in \mathbb{R}\langle\langle X \rangle\rangle$ is **rational** if it belongs to the rational closure of $\mathbb{R}\langle X \rangle$.

Consequently, all rational series can be computed by a finite number of sums, scalar multiplication, catenations and inversions. Any operation preserving rationality is referred to as a *rational operation.*

III.2.2 Recognizable series

Definition III.2.2. A **linear representation** of $c \in \mathbb{R}\langle\langle X \rangle\rangle$ is any triple (μ, γ, λ), such that

$$(c, \eta) = \lambda \mu(\eta) \gamma, \ \ \forall \eta \in X^*,$$

where γ and $\lambda^T \in \mathbb{R}^{n \times 1}$, and $\mu : X^* \to \mathbb{R}^{n \times n}$ is a monoid morphism. The integer n is called the dimension of the linear representation.

Definition III.2.3. [4] A series is called **recognizable** if it has a linear representation.

The next concept will be used to provide a necessary and sufficient condition for recognizability.

Definition III.2.4. [4, 24] A subset $V \subset \mathbb{R}\langle\langle X\rangle\rangle$ is called **stable** if $\xi^{-1}(c) \in V$ for all $c \in V$ and all $\xi \in X^*$.

Theorem III.2.1. [4, 24] *A series $c \in \mathbb{R}\langle\langle X\rangle\rangle$ is recognizable if and only if there exists a stable vector subspace of $\mathbb{R}\langle\langle X\rangle\rangle$ that contains c.*

Example III.2.1. Let $p \in \mathbb{R}\langle X\rangle$, where $X = \{x_0, \ldots, x_m\}$. Let V_p be the vector subspace of $\mathbb{R}\langle X\rangle$ such that $\hat{p} \in V_p$ if $\deg(\hat{p}) \leq \deg(p)$ for all $\hat{p} \in \mathbb{R}\langle X\rangle$. Clearly $p \in V_p$ and $\dim(V_p) = \sum_{i=0}^{deg(p)}(m+1)^i$. Furthermore, V_p is stable since $\xi^{-1}(\hat{p}) \leq \deg(p)$ for any $\xi \in X^*$ and any $\hat{p} \in V_p$. Hence, all polynomials are recognizable. □

The following theorem establishes the connection between rational series and recognizability.

Theorem III.2.2. (Schützenberger, 1961) [4, 24, 53] *A series is rational if and only if it is recognizable.*

From Theorem III.2.2 a bound for the growth of the coefficients of a rational series can be obtained.

Corollary III.2.1. *If $c \in \mathbb{R}\langle\langle X\rangle\rangle$ is rational, then there exist $K, M > 0$ such that*

$$|(c, \eta)| \leq KM^{|\eta|}, \quad \forall\, \eta \in X^*.$$

Proof: In light of the previous theorem, it is known that if c is rational, then there exists a linear representation (λ, μ, γ) for the coefficients of c such that

$$(c, \eta) = \lambda\mu(\eta)\gamma, \quad \forall \eta \in X^*,$$

where $\gamma, \lambda^T \in \mathbb{R}^{n\times 1}$ and $\mu : X^* \to \mathbb{R}^{n\times n}$ is a monoid homomorphism. Define $\#X = m + 1$, $K = \|\lambda\|\,\|\gamma\|$ and $M = \max_{0\leq i\leq m}\{\|\mu(x_i)\|\}$, where $\|\cdot\|$ denotes the

matrix norm. From the morphism property, $\mu(\eta) = \mu(x_{i_k})\mu(x_{i_{k-1}})\cdots\mu(x_{i_1})$ when $\eta = x_{i_k}x_{i_{k-1}}\cdots x_{i_1}$, $x_{i_j} \in X$, $1 \leq j \leq k$. Therefore, the Cauchy-Schwartz inequality gives

$$|(c,\eta)| = \|\lambda\mu(\eta)\gamma\| \leq \|\lambda\|\,\|\gamma\|\,\|\mu(\eta)\| \leq KM^{|\eta|}, \quad \forall \eta \in X^*.$$

■

Corollary III.2.1 implies that all rational series are globally convergent series.

III.2.3 Products of formal power series as rational operations

Rationality is important because it implies the existence of a bilinear state space realization for the corresponding Fliess operator [17, 35]. It is therefore desirable to know when the product of two rational series is again rational.

Theorem III.2.3. [17] *If c and d are rational series in $\mathbb{R}\langle\langle X\rangle\rangle$, then $c \odot d$ is also a rational series in $\mathbb{R}\langle\langle X\rangle\rangle$.*

Proof: From Theorem III.2.1, let V_c and V_d be stable finite dimensional real vector subspaces of $\mathbb{R}\langle\langle X\rangle\rangle$ such that they contain c and d, respectively. Let $\{\bar{c}_i\}_{i=1}^{n_c}$ and $\{\bar{d}_i\}_{i=1}^{n_d}$ be the corresponding bases for V_c and V_d. Define $V_{c\odot d} \subset \mathbb{R}\langle\langle X\rangle\rangle$ as

$$V_{c\odot d} = \operatorname{span}\{\bar{c}_i \odot \bar{d}_j \;:\; i = 1,\ldots,n_c \text{ and } j = 1,\ldots,n_d\}.$$

Clearly, $V_{c\odot d} \subset \mathbb{R}\langle\langle X\rangle\rangle$ is finite dimensional. If one writes

$$c = \sum_{i=1}^{n_c} \alpha_i \bar{c}_i, \quad d = \sum_{j=1}^{n_d} \beta_j \bar{d}_j,$$

then it follows that

$$c \odot d = \sum_{i,j=1}^{n_c,n_d} \alpha_i\beta_j\, \bar{c}_i \odot \bar{d}_j \in V_{c\odot d}.$$

It only remains to be shown that $V_{c\odot d}$ is stable. Observe, by the definition of the Hadamard product, that for any $x \in X$

$$x^{-1}(\bar{c}_i \odot \bar{d}_j) = x^{-1}(\bar{c}_i) \odot x^{-1}(\bar{d}_j).$$

Since V_c and V_d are stable it follows that $x^{-1}(\bar{c}_i \odot \bar{d}_j) \in V_{c\odot d}$, and therefore, $V_{c\odot d}$ is stable. In light of Theorems III.2.1 and III.2.2, the series $c \odot d$ is recognizable, and, consequently, rational. ■

Theorem III.2.4. [17, 24] *If c and d are rational series in $\mathbb{R}\langle\langle X\rangle\rangle$, then $c \sqcup\!\sqcup d$ is also a rational series in $\mathbb{R}\langle\langle X\rangle\rangle$.*

Proof: A proof similar to the one above will work, but here define the following vector subspace of $\mathbb{R}\langle\langle X\rangle\rangle$

$$V_{c \sqcup\!\sqcup d} = \text{span}\{\bar{c}_i \sqcup\!\sqcup \bar{d}_j \ : \ i = 1, \ldots, n_c \text{ and } j = 1, \ldots, n_d\}.$$

It is easy to see that $V_{c \sqcup\!\sqcup d}$ is finite dimensional. If

$$c = \sum_{i=1}^{n_c} \alpha_i \bar{c}_i, \quad d = \sum_{j=1}^{n_d} \beta_j \bar{d}_j,$$

then

$$c \sqcup\!\sqcup d = \sum_{i,j=1}^{n_c, n_d} \alpha_i \beta_j \ \bar{c}_i \sqcup\!\sqcup \bar{d}_j \in V_{c \sqcup\!\sqcup d}.$$

To see that $V_{c \sqcup\!\sqcup d}$ is stable, observe from Theorem III.1.3 that for every letter $x \in X$

$$x^{-1}(\bar{c}_i \sqcup\!\sqcup \bar{d}_j) = x^{-1}(\bar{c}_i) \sqcup\!\sqcup \bar{d}_j + \bar{c}_i \sqcup\!\sqcup x^{-1}(\bar{d}_j), \ \ \forall i, j.$$

Therefore, since V_c and V_d are stable, $x^{-1}(\bar{c}_i \sqcup\!\sqcup \bar{d}_j) \in V_{c \sqcup\!\sqcup d}$, and $V_{c \sqcup\!\sqcup d}$ is stable. Again by Theorems III.2.1 and III.2.2, $c \sqcup\!\sqcup d$ is recognizable, and, consequently, rational. ■

In contrast, the composition product does not behave as nicely as the previous

products. The following example, given by Ferfera [13], shows that the composition product is not a rational operation.

Example III.2.2. [26, 27] Suppose $X = \{x_0, x_1\}$ and consider the rational series $c = (1 - x_1)^{-1} = x_1^*$. The claim is that c composed with itself is not rational. The main goal is to show that

$$(c \circ c, x_0^{k_0} x_1^{k_1}) = (k_0)^{k_1}, \quad k_0 \geq 0, \quad k_1 \geq 0,$$

or equivalently,

$$(x_1^{-k_1} x_0^{-k_0}(c \circ c), \emptyset) = (k_0)^{k_1}. \tag{III.2.1}$$

The claim is trivial when $k_0 = k_1 = 0$ provided that $0^0 := 1$. If $k_0 = 1$ and $k_1 = 0$, observe that

$$x_0^{-1}(c \circ c) = \underbrace{x_0^{-1}(c)}_{0} \circ c + c \,⧢\, (\underbrace{x_1^{-1}(c)}_{c} \circ c) = c \,⧢\, (c \circ c).$$

The intermediate claim then is that

$$x_0^{-k_0}(c \circ c) = c^{⧢\, k_0} \,⧢\, (c \circ c), \quad k_0 \geq 1,$$

where the shuffle power of c is defined as

$$c^{⧢\, k} = \underbrace{c \,⧢\, c \,⧢\, \cdots \,⧢\, c}_{k \text{ times}}, \; k > 1$$

and $c^{⧢\, 0} = 1$. If the identity above holds up to some fixed $k_0 \geq 1$ then

$$\begin{aligned}
x_0^{-k_0-1}(c \circ c) &= x_0^{-1}(c^{⧢\, k_0} \,⧢\, (c \circ c)) \\
&= x_0^{-1}(c^{⧢\, k_0}) \,⧢\, (c \circ c) + c^{⧢\, k_0} \,⧢\, x_0^{-1}(c \circ c) \\
&= \left[k_0 c^{⧢\, (k_0-1)} \,⧢\, \underbrace{x_0^{-1}(c)}_{0} \right] ⧢\, (c \circ c) + c^{⧢\, k_0} \,⧢\, (c \,⧢\, (c \circ c)) \\
&= c^{⧢\, (k_0+1)} \,⧢\, (c \circ c).
\end{aligned}$$

Hence, the intermediate identity in question holds for $k_0 \geq 0$. Observe that

$$\begin{aligned}
x_1^{-1}x_0^{-k_0}(c \circ c) &= x_1^{-1}(c^{\sqcup\!\sqcup\, k_0} \sqcup\!\sqcup\, (c \circ c)) \\
&= x_1^{-1}(c^{\sqcup\!\sqcup\, k_0}) \sqcup\!\sqcup\, (c \circ c) + c^{\sqcup\!\sqcup\, k_0} \sqcup\!\sqcup\, \underbrace{x_1^{-1}(c \circ c)}_{0} \\
&= k_0 c^{\sqcup\!\sqcup\, (k_0-1)} \sqcup\!\sqcup\, \underbrace{x_1^{-1}(c)}_{c} \sqcup\!\sqcup\, (c \circ c) \\
&= k_0 c^{\sqcup\!\sqcup\, k_0} \sqcup\!\sqcup\, (c \circ c).
\end{aligned}$$

The next proposition is that

$$x_1^{-k_1}x_0^{-k_0}(c \circ c) = (k_0)^{k_1} c^{\sqcup\!\sqcup\, k_0} \sqcup\!\sqcup\, (c \circ c).$$

If this is the case up to some fixed $k_1 \geq 1$ then

$$\begin{aligned}
x_1^{-k_1-1}x_0^{-k_0}(c \circ c) &= x_1^{-1}((k_0)^{k_1} c^{\sqcup\!\sqcup\, k_0} \sqcup\!\sqcup\, (c \circ c)) \\
&= (k_0)^{k_1} \left[x_1^{-1}(c^{\sqcup\!\sqcup\, k_0}) \sqcup\!\sqcup\, (c \circ c) + c^{\sqcup\!\sqcup\, k_0} \sqcup\!\sqcup\, \underbrace{x_1^{-1}(c \circ c)}_{0} \right] \\
&= (k_0)^{k_1} \left[k_0 c^{\sqcup\!\sqcup\, k_0} \sqcup\!\sqcup\, (c \circ c) \right] \\
&= (k_0)^{k_1+1} c^{\sqcup\!\sqcup\, k_0} \sqcup\!\sqcup\, (c \circ c).
\end{aligned}$$

Hence, the proposition holds for all $k_1, k_0 \geq 0$. To validate (III.2.1), simply compare the constant coefficients in the above identity:

$$\begin{aligned}
(x_1^{-k_1}x_0^{-k_0}(c \circ c), \emptyset) &= ((k_0)^{k_1} c^{\sqcup\!\sqcup\, k_0} \sqcup\!\sqcup\, (c \circ c), \emptyset) \\
(c \circ c, x_0^{k_0}x_1^{k_1}) &= (k_0)^{k_1}.
\end{aligned}$$

Setting $k_0 = k_1$ reduces the expression to

$$(c \circ c, x_0^k x_1^k) = k^k, \quad k \geq 0.$$

The key observation is that these coefficients are growing faster than any sequence of coefficients from a rational series can possibly grow, namely, at a rate $KM^{|\eta|}$ for some $K, M > 0$ (Corollary III.2.1). Hence, the series $c \circ c$ *cannot* be rational. $\square$

Although the composition product does not preserve rationality in general, there exist a special condition under which rationality is preserved. Below, such condition is given.

Definition III.2.5. [13,14] A series $c \in \mathbb{R}\langle\langle X\rangle\rangle$ is **limited relative to x_i** if there exists an integer $\mathcal{N}_i \geq 0$ such that

$$\sup_{\eta \in \mathrm{supp}(c)} |\eta|_{x_i} = \mathcal{N}_i < \infty.$$

If c is limited relative to x_i for every $i = 1, \ldots, m$ then c is said to be **input-limited**. In such cases, let $\mathcal{N}_c := \max_i \mathcal{N}_i$. A series $c \in \mathbb{R}^\ell\langle\langle X\rangle\rangle$ is **input-limited** if each component series, c_j, is input-limited for $j = 1, \ldots, \ell$. In this case, $\mathcal{N}_c := \max_j \mathcal{N}_{c_j}$.

Theorem III.2.5. [13,14] *Let $c \in \mathbb{R}^\ell\langle\langle X\rangle\rangle$ and $d \in \mathbb{R}^m\langle\langle X\rangle\rangle$ be two rational series. If c is input-limited then the series $c \circ d$ is rational.*

The proof presented here relies on the following lemma.

Lemma III.2.1. [26,27] *Let $c \in \mathbb{R}^\ell\langle\langle X\rangle\rangle$ be a rational series with a linear representation (μ, γ, λ). Let $N_i \triangleq \mu(x_i) \in \mathbb{R}^{n\times n}$, $i = 0, 1, \ldots, m$. Then for any $d \in \mathbb{R}^m\langle\langle X\rangle\rangle$ it follows that*

$$c \circ d = \sum_{\eta \in \hat{X}} \lambda D_\eta((N_0 x_0)^*)\gamma,$$

where $\hat{X} \triangleq \{x_1, x_2, \ldots, x_m\}$, and the set of operators $\{D_\eta : \eta \in \hat{X}^*\}$ is the monoid under composition uniquely specified by

$$D_{x_i} : \mathbb{R}^{n\times n}\langle\langle X\rangle\rangle \to \mathbb{R}^{n\times n}\langle\langle X\rangle\rangle : E \mapsto x_0(N_0 x_0)^* N_i(d_i \sqcup\!\sqcup E)$$

with $D_\emptyset$ equivalent to the identity map.

Proof: Without lost of generality, assume $\ell = 1$. Directly from the definition of the composition product observe that

$$\begin{aligned} c \circ d &= \sum_{k\geq 0} \sum_{i_1,\ldots,i_k=1}^{m} \sum_{n_0,\ldots,n_k\geq 0} \lambda N_0^{n_k} N_{i_k} N_0^{n_{k-1}} N_{i_{k-1}} \cdots N_0^{n_1} N_{i_1} N_0^{n_0} \gamma \cdot \\ &\qquad x_0^{n_k} x_{i_k} x_0^{n_{k-1}} x_{i_{k-1}} \cdots x_0^{n_1} x_{i_1} x_0^{n_0} \circ d \\ &= \sum_{k\geq 0} \sum_{i_1,\ldots,i_k=1}^{m} \sum_{n_0,\ldots,n_k\geq 0} \lambda N_0^{n_k} N_{i_k} N_0^{n_{k-1}} N_{i_{k-1}} \cdots N_0^{n_1} N_{i_1} N_0^{n_0} \gamma \cdot \\ &\qquad x_0^{n_k+1} \left[d_{i_k} ⧢ \left[x_0^{n_{k-1}+1} \left[d_{i_{k-1}} ⧢ \cdots x_0^{n_1+1} \left[d_{i_1} ⧢ x_0^{n_0} \right] \cdots \right] \right] \right]. \end{aligned}$$

From the bilinearity and continuity of the shuffle product (in the ultrametric sense), it follows that

$$\begin{aligned} c \circ d = \sum_{k\geq 0} \sum_{i_1,\ldots,i_k=1}^{m} \lambda x_0 \left(\sum_{n_k\geq 0} (N_0 x_0)^{n_k} \right) N_{i_k} \left[d_{i_k} ⧢ \left[x_0 \left(\sum_{n_{k-1}\geq 0} (N_0 x_0)^{n_{k-1}} \right) N_{i_{k-1}} \cdot \right.\right. \\ \left.\left. \left[d_{i_{k-1}} ⧢ \cdots x_0 \left(\sum_{n_1\geq 0} (N_0 x_0)^{n_1} \right) N_{i_1} \left[d_{i_1} ⧢ \left(\sum_{n_0\geq 0} (N_0 x_0)^{n_0} \right) \right] \cdots \right] \right] \right] \gamma \\ = \sum_{k\geq 0} \sum_{i_1,\ldots,i_k=1}^{m} \lambda x_0 (N_0 x_0)^* N_{i_k} \left[d_{i_k} ⧢ \left[x_0 (N_0 x_0)^* N_{i_{k-1}} \left[d_{i_{k-1}} ⧢ \cdots \right.\right.\right. \\ \left.\left.\left. x_0 (N_0 x_0)^* N_{i_1} \left[d_{i_1} ⧢ (N_0 x_0)^* \right] \cdots \right]\right]\right] \gamma. \end{aligned}$$

Finally, applying the definition of D_η,

$$\begin{aligned} c \circ d &= \sum_{k\geq 0} \sum_{x_{i_k}\cdots x_{i_1} \in \hat{X}^k} \lambda D_{x_{i_k}} D_{x_{i_{k-1}}} \cdots D_{x_{i_1}} ((N_0 x_0)^*) \gamma \\ &= \sum_{\eta \in \hat{X}^*} \lambda D_\eta ((N_0 x_0)^*) \gamma, \end{aligned}$$

and the lemma is proved. ■

Proof of Theorem III.2.5*:* Since c is input-limited, it follows from Lemma III.2.1 that

$$c \circ d = \sum_{k=0}^{\mathcal{N}_c} \sum_{\eta \in \hat{X}^k} \lambda D_\eta ((N_0 x_0)^*) \gamma.$$

Clearly, each operator D_η is mapping a rational series to another rational series as it involves only a finite number of rational operations (including the shuffle product). Therefore, for any integer $k \geq 0$ the formal power series

$$\sum_{\eta \in \hat{X}^k} \lambda D_\eta((N_0 x_0)^*)\gamma$$

is again rational since the summation is finite. Thus, $c \circ d$ must be rational. ∎

Example III.2.3. [26] Reconsider the series $c \circ c$, where $c = x_1^*$ as in Example III.2.2. The nested inductive argument used there can be directly extended to establish the identity

$$\begin{aligned}(c \circ c, x_0^{k_0} x_1^{k_1} \cdots x_0^{k_{l-1}} x_1^{k_l}) &= (k_0)^{k_1}(k_0 + k_2)^{k_3} \cdots \\ & \quad (k_0 + k_2 + \cdots + k_{l-1})^{k_l} \end{aligned} \tag{III.2.2}$$

for all $l \geq 0$ and $k_i \geq 0$, $i = 0, 1, \ldots, l$. In which case,

$$\begin{aligned}&(c \circ c, x_0^{n_0} x_1 x_0^{n_1} x_1 \cdots x_0^{n_{j-1}} x_1 x_0^{n_j}) \\ &\quad = n_0(n_0 + n_1) \cdots (n_0 + n_1 + \cdots + n_{j-1}) \end{aligned} \tag{III.2.3}$$

$$(c \circ c, x_1^{m_0} x_0 x_1^{m_1} \cdots x_0 x_1^{m_k}) = 0^{m_0} 1^{m_1} 2^{m_2} \cdots k^{m_k} \tag{III.2.4}$$

for all $j \geq 0$ and $n_i \geq 0$, $i = 0, 1, \ldots, j$; and all $k \geq 0$ and $m_i \geq 0$, $i = 0, 1, \ldots, k$. Using identity (III.2.4), observe that

$$\begin{aligned} c \circ c &= \sum_{m_0 \geq 0} (c \circ c, x_1^{m_0})\, x_1^{m_0} + \sum_{k \geq 1} \sum_{m_0, \ldots, m_k \geq 0} (c \circ c, x_1^{m_0} x_0 x_1^{m_1} \cdots x_0 x_1^{m_k}) \cdot \\ & \quad x_1^{m_0} x_0 x_1^{m_1} \cdots x_0 x_1^{m_k} \\ &= 1 + \sum_{k \geq 1} \sum_{m_1, \ldots, m_k \geq 0} 1^{m_1} 2^{m_2} \cdots k^{m_k}\, x_0 x_1^{m_1} x_0 x_1^{m_2} \cdots x_0 x_1^{m_k} \end{aligned}$$

$$
\begin{aligned}
&= 1+\sum_{k\geq 1} x_0\left(\sum_{m_1\geq 0} x_1^{m_1}\right) x_0\left(\sum_{m_1\geq 0}(2x_1)^{m_2}\right) x_0\left(\sum_{m_k\geq 0}(kx_1)^{m_k}\right)\\
&= 1+\sum_{k\geq 1} x_0x_1^*x_0(2x_1)^*\cdots x_0(kx_1)^*. \qquad \text{(III.2.5)}
\end{aligned}
$$

Alternatively, observe that x_1^* has a linear representation with $N_0=0$, $N_1=1$ and $\lambda=\gamma=1$. Thus, $D_{x_1}: e\longrightarrow x_0(x_1^* \sqcup\!\sqcup\, e)$, and from Lemma III.2.1

$$
\begin{aligned}
c\circ c &= \sum_{\eta\in\hat{X}^*}\lambda D_\eta((N_0x_0)^*)\gamma\\
&= \sum_{k\geq 0} D_{x_1^k}(1)\\
&= 1+\sum_{k\geq 1} x_0(x_1^*\sqcup\!\sqcup\,(x_0(x_1^*\sqcup\!\sqcup\,(\cdots x_0(x_1^*\sqcup\!\sqcup\,1)\cdots))))\\
&= 1+\sum_{k\geq 1} x_0x_1^*x_0(2x_1)^*\cdots x_0(kx_1)^*,
\end{aligned}
$$

which is consistent with (III.2.5). Clearly, if the first argument in $c\circ c$ is truncated, then the resulting series composition produces a rational series as expected from Theorem III.2.5. □

Example III.2.4. Let $c=x_1^*=(1-x_1)^{-1}$, which is rational but obviously *not* input-limited, and $d=1$. Trivially, $c\circ d=(1-x_0)^{-1}$. Thus, having c input-limited is a sufficient but not necessary condition for the composition product to preserve rationality. On the other hand, if one sets $d=x_1$ then it can be verified that $(c\circ d, x_0^kx_1^k)=k!$, $k\geq 1$ [13]. In which case, requiring d to be input-limited instead of c is not a sufficient condition for preserving rationality. □

Another sufficient condition for the rationality of the composition product can be described in terms of the Hankel rank of a formal power series.

Definition III.2.6. [15] For any $c\in\mathbb{R}^\ell\langle\langle X\rangle\rangle$, the $\mathbb{R}$-linear mapping $\mathcal{H}_c:\mathbb{R}\langle X\rangle\longrightarrow$

$\mathbb{R}^{\ell}\langle\langle X\rangle\rangle$ on the vector space $\mathbb{R}\langle X\rangle$ uniquely specified by

$$(\mathcal{H}_c(\eta), \xi) = (c, \xi\eta), \quad \forall \xi, \eta \in X^*$$

is called the **Hankel mapping** of c.

$\mathcal{H}_c$ has a matrix representation whose (ξ, η) component is given by $(\mathcal{H}_c)_{\xi,\eta} = (c, \xi\eta)$ for all $\xi, \eta \in X^*$. Its range space, $\mathcal{H}_c(\mathbb{R}\langle X\rangle)$, is an $\mathbb{R}$-vector subspace of $\mathbb{R}^{\ell}\langle\langle X\rangle\rangle$, which is not necessarily finite dimensional. Consider the following definition and theorem.

Definition III.2.7. [15] The **Hankel rank** of $c \in \mathbb{R}^{\ell}\langle\langle X\rangle\rangle$ is $\rho_H(c) = \dim(\mathcal{H}_c(\mathbb{R}\langle X\rangle))$.

Theorem III.2.6. [15] *A series $c \in \mathbb{R}^{\ell}\langle\langle X\rangle\rangle$ is rational if and only if its Hankel rank is finite.*

Note that for an arbitrary input-limited rational series $c \in \mathbb{R}^{\ell}\langle\langle X\rangle\rangle$, there exists a natural number $\mathcal{N}_c$ such that its support is

$$\mathrm{supp}(c) = \left\{ \eta \in X^* : \sum_{i=1}^{m} |\eta|_{x_i} \leq \mathcal{N}_c \right\}.$$

It is easy to verify that the composition of any input-limited rational series $c \in \mathbb{R}^{\ell}\langle\langle X\rangle\rangle$ with a rational series $d \in \mathbb{R}^{m}\langle\langle X\rangle\rangle$ can be written as

$$c \circ d = \sum_{k=0}^{N_c} c_k \circ d,$$

where $c_k = c \odot \boldsymbol{L_k}$, and $\boldsymbol{L_k}$ is the characteristic series of the language $L_k = \{\eta \in X^* : \sum_{i=1}^{m} |\eta|_{x_i} = k\}$. From the definition of the Hadamard product $\mathrm{supp}(c_k) = \mathrm{supp}(c) \cap L_k$. Moreover, since $L_k = x_0^*(x_{i_k}x_0^*)\cdots(x_{i_1}x_0^*)$ for $i_j \in \{1, \ldots, m\}$,

$j = 1, 2, \ldots, m$, then L_k is a rational language, and its characteristic series is rational. By Theorem III.2.3, the series c_k is also a rational series. Clearly, the partial sum $\tilde{c}_r = \sum_{k=0}^{r} c_k$ is also input-limited. Thus, it follows from Theorem III.2.5 that for any rational $c \in \mathbb{R}^{\ell}\langle\langle X\rangle\rangle$ and $d \in \mathbb{R}^{m}\langle\langle X\rangle\rangle$, each series in the sequence $\{\tilde{c}_r \circ d\}_{r\geq 0}$ is rational, or equivalently, in light of Theorem III.2.6, $\{\rho_H(\tilde{c}_r \circ d)\}_{r\geq 0}$ is a well-defined sequence of nonnegative integers. In this context, consider the following theorem.

Theorem III.2.7. [9] *Let $c \in \mathbb{R}^{\ell}\langle\langle X\rangle\rangle$ and $d \in \mathbb{R}^{m}\langle\langle X\rangle\rangle$ be two rational series. If the sequence $\{\rho_H(\tilde{c}_r \circ d)\}_{r\geq 0}$ has a limit then $c \circ d$ is rational.*

Proof: The claim follows directly from Theorem III.2.6 and the following (easily verified) property concerning infinite matrices. Let $\{M_r\}_{r\geq 0}$ be a sequence of doubly infinite matrices with real coefficients. Assume that $\lim_{r\to\infty} M_r = M$ (componentwise in the usual topology). If each M_r has finite rank (meaning that $\rho_H(M_r) \triangleq \dim(M_r(\mathbb{R}\langle X\rangle))$ is finite), it is not necessarily the case that M has finite rank (such examples abound). But if the sequence $\{\rho_H(M_r)\}_{r\geq 0}$ has a limit, then it does follow that $\rho_H(M) \leq \lim_{r\to\infty} \rho_H(M_r)$, that is, M must have finite rank. ■

Example III.2.5. Suppose $X = \{x_0, x_1, x_2\}$, and let $c = 1 + (1 - x_1)^{-1} - (1 - x_2)^{-1}$ and $d = [1\ 1]$. Clearly in this case c is not input-limited since $c_0 = 1$ and $c_j = x_1^j - x_2^j$ for all $j \geq 1$. Now observe that $c_0 \circ d = 1$ and $c_j \circ d = 0$ for all $j \geq 1$. Thus, $\tilde{c}_r \circ d = 1$ for all $r \geq 0$, which in turn implies that $\rho_H(\tilde{c}_r \circ d) = 1$ for all $r \geq 0$. From Theorem III.2.7 it then follows that $c \circ d$ must be rational. It is trivial to check that indeed $c \circ d = 1$. □

If c is input-limited, then obviously $\lim_{r\to\infty} \rho_H(\tilde{c}_r \circ d) = \rho_H(\tilde{c}_r \circ d)|_{r=\mathcal{N}_c} = \rho_H(c \circ d)$. Conversely, if $\lim_{r\to\infty} \rho_H(\tilde{c}_r \circ d)$ exists, then there must exist an integer $r^* \geq 0$ such

that $\rho_H(\tilde{c}_r \circ d) = \rho_H(c \circ d)$ for all $r \geq r^*$ (which, of course, does not imply that $\tilde{c}_r \circ d = c \circ d$). In general, however, it is possible for $\mathcal{H}_{\tilde{c}_r \circ d} \rightarrow \mathcal{H}_{c \circ d}$ as $r \rightarrow \infty$, while at the same time the integer sequence $\rho_H(\tilde{c}_r \circ d)$ diverges. In which case, Theorem III.2.7 would not apply.

Example III.2.6. Reconsider Example III.2.4, where $c = (1 - x_1)^{-1}$, $d = 1$, and $c \circ d = (1 - x_0)^{-1}$. As noted earlier, c is not input-limited, but rationality is still preserved, specifically $\rho_H(c \circ d) = 1$. Now observe that $c_j = x_1^j$, $c_j \circ d = x_0^j$ for $j \geq 0$, and thus, $\tilde{c}_r \circ d = \sum_{j=0}^{r} x_0^j$. In which case, $\rho_H(\tilde{c}_r \circ d) = r$ for all $r \geq 0$. This example clearly falls outside the realm of Theorem III.2.5 and its generalization, Theorem III.2.7, even though rationality is in fact preserved. $\square$

It is worth noting in the previous example that if each Hankel matrix $\mathcal{H}_{\tilde{c}_r \circ d}$ is truncated to an $r \times (r + 1)$ matrix, then the resulting matrix always has rank equivalent to that of $\mathcal{H}_{c \circ d}$ for every $r > 0$. This is reminiscent of classical Hankel matrix analysis done for the partial (linear) realization problem [37]. In fact, when $c \circ d$ is not rational, $\tilde{c}_r \circ d$ can be viewed as a rational approximation of $c \circ d$, i.e., a type of partial bilinear realization problem or a noncommutative Padé approximation (e.g., see [32,33]). Therefore, it seems unlikely that any *finite* test for rationality can be devised by considering only the ranks of truncated Hankel matrices.

CHAPTER IV

FLIESS OPERATORS

In this chapter, Fliess operators are introduced in such a way that dissertation problems (i), (ii) and (iii) are solved. As noted in Chapter I, a series $c \in \mathbb{R}^{\ell}\langle\langle X \rangle\rangle$ can be formally associated with the m-input, ℓ-output Fliess operator

$$F_c[u](t) = \sum_{\eta \in X^*} (c, \eta)\, E_\eta[u](t),$$

when the inputs are measurable functions. In order to develop a stochastic version of a Fliess operator, a class of suitable stochastic input processes needs to be introduced. Then, stochastic integrals will be used to provide an extension of the definition of Fliess operators. To develop the corresponding convergence results, it is necessary to write a Stratonovich iterated integral as a sum of Itô integrals, so that the convenient properties of the Itô integral can be used. The calculus of upper bounds for stochastic iterated integrals in L_2 will be used to prove that a Fliess operator with stochastic inputs converges in the mean square sense when its corresponding series is globally convergent. In addition, it will be shown that the notion of a stopping time can be used to significantly expand the set of input-output systems that can be studied by this formalism. The chapter is concluded by characterizing the nature of the output process and some of the underlying algebraic structure concerning the new class of Fliess operators. Specifically, it is shown that the given stochastic formalism for Fliess operators can be directly related to Chen series, which have a well known relationship to Fliess operators driven by deterministic inputs.

IV.1 FLIESS OPERATORS WITH STOCHASTIC INPUTS

IV.1.1 The $\mathcal{UV}^m[0,T]$ space

Consider a one-dimensional Wiener process, W, defined over a probability space $(\Omega, \mathcal{F}, P)$. For a predictable function $u : \Omega \times [t_0, t_0 + T] \to \mathbb{R}^m$, let $\|u\|_p = \max\{\|u_i\|_{L_p} : 1 \leq i \leq m\}$, where $\|\cdot\|_{L_p}$ is the usual norm on $L_p(\Omega \times [t_0, t_0 + T], \mathcal{P}, P \otimes \lambda)$, the set of all predictable functions defined on $[t_0, t_0 + T]$ having finite $\|\cdot\|_{L_p}$-norm with $\mathcal{P}$ the predictable algebra and λ the Lebesgue measure.

Definition IV.1.1. Let $0 \leq t_0 < T$. Consider the set of all m-dimensional stochastic processes over $[t_0, t_0 + T]$, denoted by $\widetilde{\mathcal{UV}}^m[t_0, t_0 + T]$, which can be written as

$$w(t) = \int_{t_0}^{t} u(s)\, ds + \oint_{t_0}^{t} v(s)\, dW(s) \tag{IV.1.1}$$

for some $u, v \in \mathscr{I}$. The latter are called the **drift** and **diffusion** inputs, respectively. Moreover, the subset $\mathcal{UV}^m[t_0, t_0+T] \subset \widetilde{\mathcal{UV}}^m[t_0, t_0+T]$ will refer to all processes where:

a) Each integrand consists of m components such that $\mathbf{E}[u_i(t)] < \infty$, $\mathbf{E}[v_i(t)] < \infty$, $t \in [t_0, t_0 + T]$.

b) The integrands u and v are such that

$$\|u\|_{L_2}, \|v\|_{L_2}, \|v\|_{L_4} \leq R \in \mathbb{R}^+.$$

c) The random variables $u_i(t_1)$, $u_i(t_2)$, $v_i(t_1)$ and $v_i(t_2)$ are mutually independent for $1 \leq i \leq m$ and $t_1 \neq t_2$.

Observe that since $u, v \in \mathscr{I}$, by Definition II.3.10 any $w \in \widetilde{\mathcal{UV}}^m[t_0, t_0 + T]$ is also an L_2-Itô process because the Stratonovich integral of v can be expressed as the Itô

process

$$\oint_{t_0}^{t} v(s)\, dW(s) = \int_{t_0}^{t} v(s)\, dW(s) + \frac{1}{2}\int_{t_0}^{t} b(s)\, ds.$$

To describe an iterated integral over $\mathcal{UV}^m[t_0, t_0+T]$, consider the following alphabets: $X = \{x_0, x_1, \ldots, x_m\}$, $Y = \{y_0, y_1, \ldots, y_m\}$ and $XY = X \cup Y$. An arbitrary element of $X \cup Y$ will be denoted by q. For each $\eta \in XY^*$, define recursively the mapping $E_\eta : L_2^m(\Omega \times [t_0, t_0+T], \mathcal{P}, P \otimes \lambda) \to \mathcal{C}_{a.s.}[t_0, t_0+T]$ by first setting $E_\emptyset = 1$ and then letting

$$E_{x_i\eta'}[w](t) = \int_{t_0}^{t-} u_i(s) E_{\eta'}[w](s)\, ds,\ x_i \in X \tag{IV.1.2}$$

$$E_{y_i\eta'}[w](t) = \oint_{t_0}^{t-} v_i(s) E_{\eta'}[w](s)\, dW(s),\ y_i \in Y, \tag{IV.1.3}$$

where $\eta' \in XY^*$, $u_0 = v_0 = 1$, and the notation $t-$ indicates that the integration is over $[t_0, t)$. The notation $t-$ will be suppressed in subsequent sections. Also, without loss of generality, it is assumed hereafter that $t_0 = 0$.

The following terminology will be used throughout the upcoming sections. Let $\mathbb{N}^{m+1}$ be the set of all vectors with components in $\mathbb{N} = \{0, 1, \ldots\}$. Define the language $X^kY^n \triangleq \{\eta \in XY^*, |\eta|_X = k, |\eta|_Y = n\}$ formed by words having k letters in X and n letters in Y. For a fixed word $\eta \in X^kY^n$, define the vectors $\boldsymbol{\alpha} = (\alpha_m, \cdots, \alpha_0) \in \mathbb{N}^{m+1}$ and $\boldsymbol{\beta} = (\beta_m, \cdots, \beta_0) \in \mathbb{N}^{m+1}$, where $\alpha_i = |\eta|_{x_i}$, $\beta_i = |\eta|_{y_i}$, $k = \sum_{i=0}^m \alpha_i$ and $n = \sum_{i=0}^m \beta_i$. The summations over all possible α_i's that sum to k and all possible β_i's that sum to n are denoted, respectively, by $\sum_{\|\alpha\|=k}$ and $\sum_{\|\beta\|=n}$. Since one is interested in working with arbitrary letters in the alphabet XY, hereafter q_i^l will denote an element of XY, where $q_i^l = x_i$ if $l = 1$ and $q_i^l = y_i$ if $l = 2$. The symbol $w_{q_i^l}$ will denote either a drift or a diffusion input, and dq_i^l will denote either Lebesgue or

Stratonovich integration according to the value of l.

IV.1.2 An Itô-Stratonovich identity

Now given that $E_\eta[w](t)$ uses Stratonovich integrals in its definition, a special approach has to be employed to calculate $\|E_\eta[w](t)\|_2$. It is known that Stratonovich integrals lack certain important properties such as isometry [39], but if a Stratonovich integral is written in terms of Itô integrals, then all the properties associated with Itô integrals are available. There exist several formulas for writing iterated Stratonovich integrals as sums of iterated Itô integrals [34,41]. If $\eta \in X^kY^n$, an analogous formula for $E_\eta[w]$ can be obtained by a sucessive application of Definition II.3.10. To develop this identity, first define $J(\eta) = (j_n, \ldots, j_1) \in \mathbb{N}^n$ to be those places in η where all the letters belonging to Y are located. For example, if $\eta = x_{i_5}y_{i_4}x_{i_3}y_{i_2}y_{i_1}$ then $J(\eta) = (j_3, j_2, j_1) = (4, 2, 1)$.

Theorem IV.1.1. *Let $\eta \in X^kY^n$ and $w \in \mathcal{UV}^m[0,T]$ be arbitrary. Then*

$$E_\eta[w](t) = \sum_{r_1=0, r_2=0}^{n, \lfloor \frac{n}{2} \rfloor} \frac{1}{2^{r_1}2^{r_2}} \sum_{\substack{\boldsymbol{s_{r_1}} \in A_{nr_1}^{\bar{\boldsymbol{s}}_{\boldsymbol{r_2}}} \\ \bar{\boldsymbol{s}}_{\boldsymbol{r_2}} \in \bar{A}_{nr_2}}} \mathbf{I}_\eta^{\substack{\boldsymbol{s_{r_1}} \\ \bar{\boldsymbol{s}}_{\boldsymbol{r_2}}}}[w](t) \ , \tag{IV.1.4}$$

where

$$\bar{A}_{nr_2} = \Big\{ \bar{\boldsymbol{s}}_{\boldsymbol{r_2}} = (\bar{s}_{r_2}, \ldots, \bar{s}_1) \in \mathbb{N}^{r_2} : \bar{s}_{l_2} + 1 < \bar{s}_{l_2+1},\ l_{\bar{s}_{l_2}+1} = 2,\ 1 \le l_2 \le r_2 - 1,$$
$$\bar{s}_{l_2} \in J(\eta) \Big\}$$

for $1 \le r_2 \le \lfloor \frac{n}{2} \rfloor$, $\bar{A}_{n0} = \emptyset$,

$$A_{nr_1}^{\bar{\boldsymbol{s}}_{\boldsymbol{r_2}}} = \Big\{ \boldsymbol{s_{r_1}} = (s_{r_1}, \ldots, s_1) \in \mathbb{N}^{r_1} : s_{l_1} < s_{l_1+1},\ 1 \le l_1 \le r_1 - 1,\ s_{l_1} \ne \bar{s}_{l_2} \text{ or } \bar{s}_{l_2} + 1,$$
$$\bar{s}_{l_2} \in \bar{\boldsymbol{s}}_{\boldsymbol{r_2}},\ s_{l_1} \in J(\eta) \Big\}$$

for $1 \leq r_1 \leq n$, $A_{n0}^{\bar{s}_{r_2}} = \emptyset$, *and* $\lfloor \cdot \rfloor$ *is the floor function. In addition, if* $\eta = q_{i_{k+n}}^{l_{k+n}} \eta' \in X^k Y^n$ *then*

$$\mathbf{I}_\eta[w](t) \triangleq \int_0^t w_{q_{i_{k+n}}^{l_{k+n}}}(s) \mathbf{I}_{\eta'}[w](t_{k+n})\, dq_{i_{k+n}}^{l_{k+n}}(s) \tag{IV.1.5}$$

$$\mathbf{I}_\eta^{\substack{s_{r_1} \\ \bar{s}_{r_2}}}[w](t) \triangleq \mathbf{I}_\eta[w](t) \Bigg|_{\substack{\int v_{i_{\bar{s}_{l_2}+1}} \int v_{i_{\bar{s}_{l_2}}} dW(t') dW(t) \longrightarrow \int v_{i_{\bar{s}_{l_2}+1}} v_{i_{\bar{s}_{l_2}}} dt \\ \int v_{i_{s_{l_1}}} dW(t) \longrightarrow \int b_{i_{s_{l_1}}} dt}} \tag{IV.1.6}$$

with $b_{i_{s_l}} \in L_2(\Omega \times [0,T], \mathcal{P}, P \otimes \lambda)$, $1 \leq l_1 \leq r_1$, $1 \leq l_2 \leq r_2$, *and* $i_{s_{l_1}}, i_{\bar{s}_{l_2}} \in \{0, \ldots, m\}$ *are the indices of the* s_{l_1}*-th and* $\bar{s}_{l_2}$*-th elements of* $J(\eta)$.

Proof: Before starting the proof, it is important to observe that $\bar{A}_{nr_2}$ and $A_{nr_1}^{\bar{s}_{r_2}}$ only affect the stochastic integrals in $E_\eta[w]$, i.e., they are related exclusively to letters in Y. The proof is by induction on the length of the word $\eta \in X^k Y^n$. The $k+n=0$ case is trivial. For $k+n=1$, if $\eta = x_{i_1} \in X$ then it is clear that

$$E_{x_{i_1}}[w](t) = \int_0^t u_{i_1}(s)\, ds$$

since $r_1 = r_2 = 0$. On the other hand, if $\eta = y_{i_1} \in Y$ then from (II.3.3) it follows that

$$\begin{aligned} E_{y_{i_1}}[w](t) &= \int_0^t v_{i_1}(s)\, dW(s) + \frac{1}{2} \int_0^t b_{i_1}(s)\, ds \\ &= \mathbf{I}_{y_{i_1}}[w](t) + \frac{1}{2} \mathbf{I}_{y_{i_1}}^{\substack{(1) \\ \emptyset}}[w](t), \end{aligned} \tag{IV.1.7}$$

since $s_1 = 1$ ($A_{11} = \{(1)\}$) and $r_2 = 0$. The $k+n=2$ case is helpful for understanding the general case. Here $\eta \in \{x_{i_2}x_{i_1}, x_{i_2}y_{i_1}, y_{i_2}x_{i_1}, y_{i_2}y_{i_1}\}$. The $\eta = x_{i_2}x_{i_1}$ case is trivial since there are no letters in Y involved. If $\eta = x_{i_2}y_{i_1}$ or $y_{i_2}x_{i_1}$, then from formula (II.3.3)

$$E_{x_{i_2}y_{i_1}}[w](t) = \mathbf{I}_{x_{i_2}y_{i_1}}[w](t) + \frac{1}{2} \mathbf{I}_{x_{i_2}y_{i_1}}^{\substack{(1) \\ \emptyset}}[w](t),$$

and

$$E_{y_{i_2}x_{i_1}}[w](t) = \mathbf{I}_{y_{i_2}x_{i_1}}[w](t) + \frac{1}{2}\mathbf{I}_{y_{i_2}x_{i_1}}^{\overset{(2)}{\emptyset}}[w](t),$$

respectively. When $\eta = y_{i_2}y_{i_1}$, the following identity needs to be proved

$$E_\eta[w](t) = \mathbf{I}_\eta[w](t) + \frac{1}{2}\mathbf{I}_\eta^{\overset{(1)}{\emptyset}}[w](t) + \frac{1}{2}\mathbf{I}_\eta^{\overset{(2)}{\emptyset}}[w](t) + \frac{1}{2}\mathbf{I}_\eta^{\overset{\emptyset}{(1)}}[w](t) + \frac{1}{4}\mathbf{I}_\eta^{\overset{(2,1)}{\emptyset}}[w](t). \quad \text{(IV.1.8)}$$

Suppose

$$\begin{aligned} f_1(t) &= v_{i_2}(t)\int_0^t v_{i_1}(s)\, dW(s) \\ f_2(t) &= v_{i_2}(t)\int_0^t b_{i_1}(s)\, ds. \end{aligned}$$

From the Itô derivation rule

$$d(V_1V_2)(s) = V_1(s)\, dV_2(s) + V_2(s)\, dV_1(s) + B_1(s)B_2(s)\, ds, \quad \text{(IV.1.9)}$$

where $dV_i(s) = A_i(s)\, ds + B_i(s)\, dW(s)$, one can determine which terms in df_1 and df_2 generate a quadratic covariation different from zero. For example,

$$\begin{aligned} f_1(t) &= \left(\int_0^t a_{i_2}(s)\, ds + \int_0^t b_{i_2}(s)\, dW(s)\right)\left(\int_0^t v_{i_1}(s)\, dW(s)\right) \quad \text{(IV.1.10)} \\ &= \int_0^t b_{i_2}(s)\int_0^s v_{i_1}(r)\, dW(r)dW(s) + \int_0^t v_{i_2}(s)v_{i_1}(s)\, dW(s) + \Phi(dt). \end{aligned}$$

Here $\Phi(dt)$ denotes a generic term that together with the Wiener process W generates quadratic covariations equal to zero. For f_2, given that the right factor of f_2 is a Lebesgue integral, only v_{i_2} contributes to a quadratic covariation different from zero. Hence,

$$f_2(t) = \int_0^t b_{i_2}(s)\int_0^s b_{i_1}(r)\, drdW(s) + \Phi(dt). \quad \text{(IV.1.11)}$$

Continuing with the proof of (IV.1.8), substitute (IV.1.7) into $E_{y_{i_2}y_{i_1}}[w](t)$ and employ (IV.1.10) and (IV.1.11). Then

$$
\begin{aligned}
&E_{y_{i_2}y_{i_1}}[w](t) \\
&= \oint_0^t v_{i_2}(t_2)\left(\int_0^{t_1} v_{i_1}(t_1)\,dW(t_1) + \frac{1}{2}\int_0^{t_2} b_{i_1}(t_1)\,dt_1\right) dW(t_2) \\
&= \int_0^t v_{i_2}(t_2)\left(\int_0^{t_1} v_{i_1}(t_1)\,dW(t_1) + \frac{1}{2}\int_0^{t_2} b_{i_1}(t_1)\,dt_1\right) dW(t_2) \\
&\quad + \frac{1}{2}\left\langle v_{i_2}(\cdot)\int_0^{\cdot} v_{i_1}(t_1)\,dW(t_1) + \frac{1}{2}v_{i_2}(\cdot)\int_0^{\cdot} b_{i_1}(t_1)\,dt_1, W\right\rangle_{[0,t]} \\
&= \int_0^t v_{i_2}(t_2)\int_0^{t_1} v_{i_1}(t_1)\,dW(t_1)dW(t_2) + \frac{1}{2}\int_0^t v_{i_2}(t_2)\int_0^{t_2} b_{i_1}(t_1)\,dt_1 dW(t_2) \\
&\quad + \frac{1}{2}\int_0^t b_{i_2}(t_2)\int_0^{t_2} v_{i_1}(t_1)\,dW(t_1)dt_2 + \frac{1}{2}\int_0^t v_{i_2}(t_1)v_{i_1}(t_1)\,dt_1 \\
&\quad + \frac{1}{4}\int_0^t b_{i_2}(t_2)\int_0^{t_2} b_{i_1}(t_1)\,dt_1 dt_2 \\
&= \mathbf{I}_\eta[w](t) + \frac{1}{2}\mathbf{I}_\eta^{\overset{(1)}{\emptyset}}[w](t) + \frac{1}{2}\mathbf{I}_\eta^{\overset{(2)}{\emptyset}}[w](t) + \frac{1}{2}\mathbf{I}_\eta^{\overset{\emptyset}{(1)}}[w](t) + \frac{1}{4}\mathbf{I}_\eta^{\overset{(2,1)}{\emptyset}}[w](t).
\end{aligned}
$$

Now assume (IV.1.4) holds up to $n+k$ and let $\eta = q_{i_{k+n+1}}^{l_{k+n+1}}\eta' \in XY^{k+n+1}$. If $l_{k+n+1} = 1$ then

$$
\begin{aligned}
&E_\eta[w](t) \\
&= \int_0^t u_{i_{k+n+1}}(t_{k+n+1})\left(\sum_{r_1=0,r_2=0}^{n,\lfloor\frac{n}{2}\rfloor} \frac{1}{2^{r_1}2^{r_2}} \sum_{\substack{s_{r_1}\in A_{nr_1}^{\bar{s}_{r_2}} \\ \bar{s}_{r_2}\in \bar{A}_{nr_2}}} \mathbf{I}_{\eta'}^{\overset{s_{r_1}}{\bar{s}_{r_2}}}[w](t_{k+n+1})\right) dt_{k+n+1} \\
&= \sum_{r_1=0,r_2=0}^{n,\lfloor\frac{n}{2}\rfloor} \frac{1}{2^{r_1}2^{r_2}} \sum_{\substack{s_{r_1}\in A_{nr_1}^{\bar{s}_{r_2}} \\ \bar{s}_{r_2}\in \bar{A}_{nr_2}}} \mathbf{I}_{\eta}^{\overset{s_{r_1}}{\bar{s}_{r_2}}}[w](t).
\end{aligned}
\tag{IV.1.12}
$$

If $l_{k+n+1} = 2$ then

$$
E_\eta[w](t) \;=\; \oint_0^t v_{i_{k+n+1}}(t_{k+n+1})\cdot
$$

$$\left(\sum_{r_1=0, r_2=0}^{n, \lfloor \frac{n}{2} \rfloor} \frac{1}{2^{r_1} 2^{r_2}} \sum_{\substack{\boldsymbol{s}_{r_1} \in A_{nr_1}^{\bar{\boldsymbol{s}}_{r_2}} \\ \bar{\boldsymbol{s}}_{r_2} \in \bar{A}_{nr_2}}} \mathbf{I}_{\eta'}^{\substack{\boldsymbol{s}_{r_1} \\ \bar{\boldsymbol{s}}_{r_2}}}[w](t_{k+n+1}) \right) dW(t_{k+n+1}). \qquad \text{(IV.1.13)}$$

Since any $s_{l_1} \neq \bar{s}_{l_2}, \bar{s}_{l_2}+1$ for $0 \leq l_2 \leq \lfloor \frac{n}{2} \rfloor$, using (II.3.3) and (IV.1.9),

$$\left\langle v_{i_{k+n+1}}(\cdot) \left(\mathbf{I}_{\eta'}^{\substack{\boldsymbol{s}_{r_1} \\ \bar{\boldsymbol{s}}_{r_2}}}[w](\cdot) \right), W \right\rangle_{[0,t]}$$
$$= \begin{cases} \mathbf{I}_{\eta}^{\substack{\boldsymbol{s}'_{r_1} \\ \bar{\boldsymbol{s}}_{r_2}}}[w](t) + \mathbf{I}_{\eta}^{\substack{\boldsymbol{s}_{r_1} \\ \bar{\boldsymbol{s}}'_{r_2}}}[w](t) & : \ s_{r_1} < k+n, \ \bar{s}_{r_2} < k+n-1 \\ & \quad \text{and } q_{i_{n+k}}^{l_{n+k}} \in Y, \\ \mathbf{I}_{\eta}^{\substack{\boldsymbol{s}'_{r_1} \\ \bar{\boldsymbol{s}}_{r_2}}}[w](t) & : \ s_{r_1} = k+n \ \text{ or } \ \bar{s}_{r_2} = k+n-1 \\ & \quad \text{or } q_{i_{n+k}}^{l_{n+k}} \in X, \end{cases} \qquad \text{(IV.1.14)}$$

where $\boldsymbol{s}'_{r_1} = (k+n+1, s_{r_1}, \ldots, s_1)$ and $\bar{\boldsymbol{s}}'_{r_2} = (k+n, \bar{s}_{r_2}, \ldots, \bar{s}_1)$. Substituting (IV.1.14) into (IV.1.13) and regrouping gives

$$\begin{aligned} E_\eta[w](t) &= \sum_{r_1=0, r_2=0}^{n, \lfloor \frac{n}{2} \rfloor} \frac{1}{2^{r_1} 2^{r_2}} \sum_{\substack{\boldsymbol{s}_{r_1} \in A_{nr_1}^{\bar{\boldsymbol{s}}_{r_2}} \\ \bar{\boldsymbol{s}}_{r_2} \in \bar{A}_{nr_2}}} \oint_0^t v_{i_{k+n+1}}(t_{k+n+1}) \left(\mathbf{I}_{\eta'}^{\substack{\boldsymbol{s}_{r_1} \\ \bar{\boldsymbol{s}}_{r_2}}}[w](t_{k+n+1}) \right) dW(t_{k+n+1}) \\ &= \sum_{r_1=0, r_2=0}^{n, \lfloor \frac{n}{2} \rfloor} \frac{1}{2^{r_1} 2^{r_2}} \sum_{\substack{\boldsymbol{s}_{r_1} \in A_{nr_1}^{\bar{\boldsymbol{s}}_{r_2}} \\ \bar{\boldsymbol{s}}_{r_2} \in \bar{A}_{nr_2}}} \left(\mathbf{I}_{\eta}^{\substack{\boldsymbol{s}_{r_1} \\ \bar{\boldsymbol{s}}_{r_2}}}[w](t) + \frac{1}{2} \left\langle v_{i_{k+n+1}}(\cdot) \left(\mathbf{I}_{\eta'}^{\substack{\boldsymbol{s}_{r_1} \\ \bar{\boldsymbol{s}}_{r_2}}}[w](\cdot) \right), W \right\rangle_{[0,t]} \right). \end{aligned}$$

For n even,

$$E_\eta[w](t) = \sum_{r_1=0, r_2=0}^{n+1, \lfloor \frac{n}{2} \rfloor} \frac{1}{2^{r_1} 2^{r_2}} \sum_{\substack{\boldsymbol{s}_{r_1} \in A_{(n+1)r_1}^{\bar{\boldsymbol{s}}_{r_2}} \\ \bar{\boldsymbol{s}}_{r_2} \in \bar{A}_{(n+1)r_2}}} \mathbf{I}_{\eta}^{\substack{\boldsymbol{s}_{r_1} \\ \bar{\boldsymbol{s}}_{r_2}}}[w](t), \qquad \text{(IV.1.15)}$$

and for $n = n'+1$ odd,

$$E_\eta[w](t) = \sum_{r_1=0, r_2=0}^{n'+1, \lfloor \frac{n}{2} \rfloor + 1} \frac{1}{2^{r_1} 2^{r_2}} \sum_{\substack{\boldsymbol{s}_{r_1} \in A_{(n'+1)r_1}^{\bar{\boldsymbol{s}}_{r_2}} \\ \bar{\boldsymbol{s}}_{r_2} \in \bar{A}_{(n'+1)r_2}}} \mathbf{I}_{\eta}^{\substack{\boldsymbol{s}_{r_1} \\ \bar{\boldsymbol{s}}_{r_2}}}[w](t). \qquad \text{(IV.1.16)}$$

Together (IV.1.15) and (IV.1.16) become

$$E_\eta[w](t) = \sum_{r_1=0,r_2=0}^{n+1,\left\lfloor\frac{n+1}{2}\right\rfloor} \frac{1}{2^{r_1}2^{r_2}} \sum_{\substack{\boldsymbol{s}_{r_1}\in A^{\bar{\boldsymbol{s}}_{r_2}}_{(n+1)r_1} \\ \bar{\boldsymbol{s}}_{r_2}\in\bar{A}_{(n+1)r_2}}} \mathbf{I}_\eta^{\substack{\boldsymbol{s}_{r_1} \\ \bar{\boldsymbol{s}}_{r_2}}}[w](t). \tag{IV.1.17}$$

Finally, from equations (IV.1.13) and (IV.1.17) the induction is completed. ∎

IV.1.3 Upper bounds for stochastic iterated Itô integrals

The next two theorems present L_2 upper bounds for the iterated Itô integrals (IV.1.5) and (IV.1.6).

Theorem IV.1.2. *Let $\eta \in X^kY^n$ and $w \in \mathcal{UV}^m[0,T]$ be arbitrary. An L_2 upper bound for the iterated Itô integral* (IV.1.5) *at a fixed $t \in [0,T]$ is*

$$\|\mathbf{I}_\eta[w](t)\|_2^2 \leq t^k \prod_{i=0}^{m} \frac{\tilde{U}_i^{\alpha_i}(t)}{\alpha_i!}\frac{V_i^{\beta_i}(t)}{\beta_i!}, \tag{IV.1.18}$$

where $\tilde{U}_i(t) = \int_0^t \mathbf{E}\left[u_i^2(s)\right] ds$ and $V_i(t) = \int_0^t \mathbf{E}\left[v_i^2(s)\right] ds$.

Proof: The inequality is proved by induction over the total number of $k+n$ integrals. For $k+n=0$, the claim follows trivially. If $k+n=1$, then there are two cases. First, if $\eta = x_{i_1}$ then by Hölder's inequality

$$\left\|\mathbf{I}_{x_{i_1}}[w](t)\right\|_2^2 \leq t \int_0^t \mathbf{E}\left[u_{i_1}^2(t_1)\right] dt_1 = t\,\tilde{U}_{i_1}(t).$$

The second case is when $\eta = y_{i_1}$. By the isometry property

$$\left\|\mathbf{I}_{y_{i_1}}[w](t)\right\|_2^2 = \int_0^t \mathbf{E}\left[v_{i_1}^2(t_1)\right] dt_1 = V_{i_1}(t).$$

Now calculate the bound for $n+k+1$ integrals assuming that (IV.1.18) holds up to some fixed $n+k \geq 0$. Set $\eta = q_{i_{k+n+1}}^{l_{k+n+1}}\eta'$ with $\eta' \in X^kY^n$. If $l_{k+n+1} = 1$, then the

independence property assumed in Definition IV.1.1 gives

$$
\begin{aligned}
&\|\mathbf{I}_\eta[w](t)\|_2^2 \\
&\quad = \mathbf{E}\left[\left(\int_0^t u_{i_{k+n+1}}(t_{k+n+1})\mathbf{I}_{\eta'}[w](t_{k+n+1})\,dt_{k+n+1}\right)^2\right] \\
&\quad \leq t\int_0^t \mathbf{E}\left[u^2_{i_{k+n+1}}(t_{k+n+1})\left(\mathbf{I}_{\eta'}[w](t_{k+n+1})\right)^2\right]\,dt_{k+n+1} \\
&\quad = t\int_0^t \mathbf{E}\left[u^2_{i_{k+n+1}}(t_{k+n+1})\right]\mathbf{E}\left[\left(\mathbf{I}_{\eta'}[w](t_{k+n+1})\right)^2\right]\,dt_{k+n+1} \\
&\quad = t\int_0^t \mathbf{E}\left[u^2_{i_{k+n+1}}(t_{k+n+1})\right]\left[t^k_{k+n+1}\prod_{i=0}^m \frac{\tilde{U}_i^{\bar{\alpha}_i}(t_{k+n+1})}{\bar{\alpha}_i!}\frac{V_i^{\beta_i}(t_{k+n+1})}{\beta_i!}\right]\,dt_{k+n+1} \\
&\quad \leq t^{k+1}\prod_{\substack{i=0\\ i\neq i_{k+n+1}}}^m \frac{\tilde{U}_i^{\bar{\alpha}_i}(t)}{\bar{\alpha}_i!}\prod_{i=0}^m\frac{V_i^{\beta_i}(t)}{\beta_i!}\underbrace{\int_0^t \mathbf{E}\left[u^2_{i_{k+n+1}}(t_{k+n+1})\right]\frac{\tilde{U}_{i_{k+n+1}}^{\bar{\alpha}_{i_{k+n+1}}}(t_{k+n+1})}{\bar{\alpha}_{i_{k+n+1}}!}dt_{k+n+1}}_{=\frac{\tilde{U}_i^{\bar{\alpha}_{i_{k+n+1}}+1}(t)}{(\bar{\alpha}_{i_{k+n+1}}+1)!}} \\
&\quad \leq t^{k+1}\prod_{i=0}^m \frac{\tilde{U}_i^{\bar{\alpha}_i}(t)}{\bar{\alpha}_i!}\frac{V_i^{\beta_i}(t)}{(\beta_i)!},
\end{aligned}
$$

where $\sum_{i=0}^m \bar{\alpha}_i = k$. Letting $\alpha_i = \bar{\alpha}_i + \delta_{i(i_{k+n+1})}$ (here δ_{ij} denotes the Kronecker delta function), then

$$\|\mathbf{I}_\eta[w](t)\|_2^2 \leq t^{k+1}\prod_{i=0}^m \frac{\tilde{U}_i^{\alpha_i}(t)}{\alpha_i!}\frac{V_i^{\beta_i}(t)}{\beta_i!},$$

where $\sum_{i=0}^m \alpha_i = k+1$. Now, when $l_{k+n+1} = 2$, the isometry property gives

$$
\begin{aligned}
&\|\mathbf{I}_\eta[w](t)\|_2^2 \\
&\quad = \mathbf{E}\left[\left(\int_0^t v_{i_{k+n+1}}(t_{k+n+1})\mathbf{I}_{\eta'}[w](t_{k+n+1})\,dW(t_{k+n+1})\right)^2\right] \\
&\quad = \int_0^t \mathbf{E}\left[v^2_{i_{k+n+1}}(t_{k+n+1})\left(\mathbf{I}_{\eta'}[w](t_{k+n+1})\right)^2\right]\,dt_{k+n+1}
\end{aligned}
$$

$$
\begin{aligned}
&= \int_0^t \mathbf{E}\left[v_{i_{k+n+1}}^2(t_{k+n+1})\right] \mathbf{E}\left[(\mathbf{I}_{\eta'}[w](t_{k+n+1}))^2\right]\, dt_{k+n+1} \\
&= \int_0^t \mathbf{E}\left[v_{i_{k+n+1}}^2(t_{k+n+1})\right] \left[t_{k+n+1}^k \prod_{i=0}^m \frac{\tilde{U}_i^{\alpha_i}(t_{k+n+1})}{\alpha_i!} \prod_{i=0}^m \frac{V_i^{\bar{\beta}_i}(t_{k+n+1})}{\bar{\beta}_i!}\right] dt_{k+n+1} \\
&\le t^k \prod_{i=0}^m \frac{\tilde{U}_i^{\alpha_i}(t)}{\alpha_i!} \prod_{\substack{i=0 \\ i\neq i_{k+n+1}}}^m \frac{V_i^{\bar{\beta}_i}(t)}{\bar{\beta}_i!} \underbrace{\int_0^t \mathbf{E}\left[v_{i_{k+n+1}}^2(t_{k+n+1})\right] \frac{V_{i_{k+n+1}}^{\bar{\beta}_{i_{k+n+1}}}(t_{k+n+1})}{\bar{\beta}_{i_{k+n+1}}!} dt_{k+n+1}}_{= \frac{V_i^{\bar{\beta}_{i_{k+n+1}}+1}(t)}{(\bar{\beta}_{i_{k+n+1}}+1)!}},
\end{aligned}
$$

where $\sum_{i=0}^m \bar{\beta}_i = n$. Similarly, let $\beta_i = \bar{\beta}_i + \delta_{i(i_{k+n+1})}$, then

$$
||\mathbf{I}_\eta[w](t)||_2^2 \le t^k \prod_{i=0}^m \frac{\tilde{U}_i^{\alpha_i}(t)}{\alpha_i!} \frac{V_i^{\beta_i}(t)}{\beta_i!},
$$

where $\sum\limits_{i=0}^m \beta_i = n+1$. Hence, the proposition is proved for all $k+n \ge 0$. ■

Theorem IV.1.3. *Let $\eta \in X^kY^n$ and $w \in \mathcal{UV}^m[0,T]$ be arbitrary. An L_2 upper bound for the iterated Itô integral* (IV.1.6) *at a fixed $t \in [0,T]$ is*

$$
\left\| \mathbf{I}_\eta^{\substack{\boldsymbol{s}_{r_1} \\ \bar{\boldsymbol{s}}_{r_2}}}[w](t) \right\|_2^2 \le 2^{r_2} t^{k+r_1+r_2} \prod_{i=0}^m \frac{\tilde{U}_i^{\alpha_i}(t)}{\alpha_i!} \frac{V_i^{\bar{\beta}_i}(t)}{\bar{\beta}_i!} \frac{\bar{V}_i^{\bar{\gamma}_i}(t)}{\sqrt{\bar{\gamma}_i!}} \frac{B_i^{\gamma_i}(t)}{\gamma_i!}, \tag{IV.1.19}
$$

where $\bar{\gamma}_i = \sum_{l_2=1}^{r_2}(\delta_{ii_{\bar{s}_{l_2}}} + \delta_{i(i_{\bar{s}_{l_2}}+1)})$, $\bar{s}_{l_2} \in \bar{\boldsymbol{s}}_{\boldsymbol{r_2}}$; $\gamma_i = \sum_{l_1=1}^{r_1} \delta_{ii_{s_{l_1}}}$, $s_{l_1} \in \boldsymbol{s}_{\boldsymbol{r_1}}$; $\bar{\beta}_i = \beta_i - \bar{\gamma}_i - \gamma_i$; $\tilde{U}_i(t) = \int_0^t \mathbf{E}[u_i^2(s)]\, ds$, $V_i(t) = \int_0^t \mathbf{E}[v_i^2(s)]\, ds$, $\bar{V}_i^2(t) = \int_0^t \mathbf{E}[v_i^4(s)]\, ds$ and $B_i(t) = \int_0^t \mathbf{E}[b_i^2(s)]\, ds$.

Proof: This inequality is proved by induction over $r = r_1 + r_2$. If $r_1 = 0$ and $r_2 = 0$ then inequality (IV.1.19) reduces directly to inequality (IV.1.18). Now assume that (IV.1.19) holds up to $r-1 \ge 0$. To illustrate the inductive step, consider the following calculations. By Hölder's inequality

$$
\begin{aligned}
\left\| \int_0^t b_{i_1}(t_1) dt_1 \right\|_2^2 &\le t \int_0^t \mathbf{E}\left[b_{i_1}^2(t_1)\right] dt_1 \\
&\le t\, B_{i_1}(t).
\end{aligned}
$$

To calculate the norm squared of $\int_0^t v_{i_2}(t_1)v_{i_1}(t_1)dt_1$, there are two cases.

If $i_1 \neq i_2$ then

$$\begin{aligned}\left\|\int_0^t v_{i_2}(t_1)v_{i_1}(t_1)dt_1\right\|_2^2 &\leq t\int_0^t \mathbf{E}\left[v_{i_2}^2(t_1)v_{i_1}^2(t_1)\right]dt_1\\ &\leq t\left(\int_0^t \mathbf{E}\left[v_{i_2}^4(t_1)\right]dt_1\right)^{\frac{1}{2}}\left(\int_0^t \mathbf{E}\left[v_{i_1}^4(t_1)\right]dt_1\right)^{\frac{1}{2}}\\ &= t\ \bar{V}_{i_1}(t)\bar{V}_{i_2}(t) \leq 2t\ \bar{V}_{i_1}(t)\bar{V}_{i_2}(t).\end{aligned}$$

If $i_1 = i_2$ then

$$\left\|\int_0^t v_{i_1}^2(t_1)dt_1\right\|_2^2 \leq t\int_0^t \mathbf{E}\left[v_{i_1}^4(t_1)\right]dt_1 \leq 2t\ \frac{\left(\bar{V}_{i_1}(t)\right)^2}{\sqrt{2!}}.$$

Keeping in mind the inequalities above, one can proceed with the induction. Without loss of generality, set $\eta = q_{i_{k+n}}^{l_{k+n}}\cdots q_{i_{s_r+1}}^{l_{s_r+1}}y_{i_{s_r}}\eta' \in X^kY^n$, $\eta' \in X^{k-k_1}Y^{s_r-1}$ with $s_r = s_{r_1}$ or $\bar{s}_{r_2}+1$, and $k_1 \leq k$. For $s_{r_1} > \bar{s}_{r_2}$, applying the isometry property $n - s_{r_1}$ times and Hölder's inequality k_1 times gives

$$\begin{aligned}&\left\|\mathbf{I}_{\eta}{}^{\substack{s_{r_1}\\ \bar{s}_{r_2}}}[w](t)\right\|_2^2\\ &= \mathbf{E}\Bigg[\Bigg(\int_0^t w_{q_{i_{k+n}}^{l_{k+n}}}(t_{k+n})\cdots\int_0^{t_{s_{r_1}+2}} w_{q_{i_{s_{r_1}+1}}^{l_{s_{r_1}+1}}}(t_{s_{r_1}+1})\int_0^{t_{s_{r_1}+1}} b_{i_{s_{r_1}}}(t_{s_{r_1}})\mathbf{I}_{\eta'}{}^{\substack{s_{r_1}-1\\ \bar{s}_{r_2}}}[w](t_{s_{r_1}})dt_{s_{r_1}}\cdot\\ &\qquad dq_{i_{s_{r_1}+1}}^{l_{s_{r_1}+1}}(t_{s_{r_1}+1})\cdots dq_{i_{k+n}}^{l_{k+n}}(t_{k+n})\Bigg)^2\Bigg]\\ &\leq t^{k_1}\int_0^t \mathbf{E}\left[w^2_{q_{i_{k+n}}^{l_{k+n}}}(t_{k+n})\right]\cdots\int_0^{t_{s_{r_1}+2}}\mathbf{E}\left[w^2_{q_{i_{s_{r_1}+1}}^{l_{s_{r_1}+1}}}(t_{s_{r_1}+1})\right]\mathbf{E}\Bigg[\Bigg(\int_0^{t_{s_{r_1}+1}} b_{i_{s_{r_1}}}(t_{s_{r_1}})\cdot\\ &\qquad \mathbf{I}_{\eta'}{}^{\substack{s_{r_1}-1\\ \bar{s}_{r_2}}}[w](t_{s_{r_1}})\,dt_{s_{r_1}}\Bigg)^2\Bigg]dt_{s_{r_1}+1}\cdots dt_{k+n}.\end{aligned}$$

Using Hölder's inequality once more and applying the inductive step with $r-1 = r_1+r_2-1$,

$$
\begin{aligned}
&\left\| \mathbf{I}_{\eta}^{\substack{s_{r_1} \\ \bar{s}_{r_2}}}[w](t) \right\|_2^2 \\
&\leq t^{k_1} \int_0^t \mathbf{E}\left[w^2_{\substack{l_{k+n} \\ q_{i_{k+n}}}}(t_{k+n}) \right] \cdots \int_0^{t_{s_{r_1}+2}} \mathbf{E}\left[w^2_{\substack{l_{s_{r_1}+1} \\ q_{i_{s_{r_1}+1}}}}(t_{s_{r_1}+1}) \right] t_{s_{r_1}+1} \int_0^{t_{s_{r_1}+1}} \mathbf{E}\left[b^2_{i_{s_{r_1}}}(t_{s_{r_1}}) \right] \cdot \\
&\qquad \left\| \mathbf{I}_{\eta'}^{\substack{s_{r_1-1} \\ \bar{s}_{r_2}}}[w](t_{s_r}) \right\|_2^2 \, dt_{s_{r_1}} dt_{s_{r_1}+1} \cdots dt_{k+n} \\
&\leq t^{k_1} \int_0^t \mathbf{E}\left[w^2_{\substack{l_{k+n} \\ q_{i_{k+n}}}}(t_{k+n}) \right] \cdots \int_0^{t_{s_{r_1}+2}} \mathbf{E}\left[w^2_{\substack{l_{s_{r_1}+1} \\ q_{i_{s_{r_1}+1}}}}(t_{s_{r_1}+1}) \right] t_{s_{r_1}+1} \int_0^{t_{s_{r_1}+1}} \mathbf{E}\left[b^2_{i_{s_{r_1}}}(t_{s_{r_1}}) \right] \cdot \\
&\qquad 2^{r_2} t_{s_r}^{k-k_1+r_1+r_2-1} \prod_{i=0}^{m} \frac{\tilde{U}^{\bar{\alpha}_i}(t_{s_{r_1}})}{\bar{\alpha}_i!} \frac{V_i^{\bar{\beta}}(t_{s_{r_1}})}{\bar{\beta}_i!} \frac{\bar{V}_i^{\bar{\gamma}_i}(t_{s_{r_1}})}{\sqrt{\bar{\gamma}_i!}} \frac{B_i^{\gamma_i}(t_{s_{r_1}})}{\gamma_i!} \, dt_{s_{r_1}} dt_{s_{r_1}+1} \cdots dt_{k+n} \\
&\leq 2^{r_2} t^{k+r_1+r_2} \int_0^t \mathbf{E}\left[w^2_{\substack{l_{k+n} \\ q_{i_{k+n}}}}(t_{k+n}) \right] \cdots \int_0^{t_{s_{r_1}+2}} \mathbf{E}\left[w^2_{\substack{l_{s_{r_1}+1} \\ q_{i_{s_{r_1}+1}}}}(t_{s_{r_1}+1}) \right] \cdot \\
&\qquad \prod_{i=0}^{m} \frac{\tilde{U}^{\bar{\alpha}_i}(t_{s_{r_1}})}{\bar{\alpha}_i!} \frac{V_i^{\bar{\beta}}(t_{s_{r_1}+1})}{\bar{\beta}_i!} \frac{\bar{V}_i^{\bar{\gamma}_i}(t_{s_{r_1}+1})}{\sqrt{\bar{\gamma}_i!}} \prod_{\substack{l=0 \\ l \neq i_{s_{r_1}}}}^{m} \frac{B_l^{\gamma_l}(t_{s_{r_1}+1})}{\gamma_l!} \cdot \\
&\qquad \int_0^{t_{s_{r_1}+1}} \mathbf{E}\left[b^2_{i_{s_{r_1}}}(t_{s_{r_1}}) \right] \frac{B_{i_{s_{r_1}}}^{\gamma_{i_{s_{r_1}}}}(t_{s_{r_1}})}{\gamma_{i_{s_{r_1}}}!} \, dt_{s_{r_1}} dt_{s_{r_1}+1} \cdots dt_{k+n} \\
&\leq 2^{r_2} t^{k+r_1+r_2} \int_0^t \mathbf{E}\left[w^2_{\substack{l_{k+n} \\ q_{i_{k+n}}}}(t_{k+n}) \right] \cdots \int_0^{t_{s_{r_1}+2}} \mathbf{E}\left[w^2_{q_{i_{s_{r_1}+1}}}(t_{s_{r_1}+1}) \right] \cdot \\
&\qquad \prod_{i=0}^{m} \frac{\tilde{U}^{\bar{\alpha}_i}(t_{s_{r_1}})}{\bar{\alpha}_i!} \frac{V_i^{\bar{\beta}}(t_{s_{r_1}+1})}{\bar{\beta}_i!} \frac{\bar{V}_i^{\bar{\gamma}_i}(t_{s_{r_1}+1})}{\sqrt{\bar{\gamma}_i!}} \frac{B_l^{\gamma_i'}(t_{s_{r_1}+1})}{\gamma_i'!} \, dt_{s_{r_1}+1} \cdots dt_{k+n},
\end{aligned}
$$

where $\gamma_i' = \gamma_i + \delta_{ii_{s_{r_1}}}$. Observe that $\sum_{i=0}^{m} \gamma_i' = r_1$. The remaining $(n+k_1-s_{r_1})$ integrals are evaluated exactly as in the proof of Theorem IV.1.2. Thus, the $\bar{\alpha}_i$'s and $\bar{\beta}_i$'s increase rather than the γ_i's, and

$$
\left\| \mathbf{I}_{\eta}^{\substack{s_{r_1} \\ \bar{s}_{r_2}}}[w](t) \right\|_2^2 \leq 2^{r_2} t^{k+r_1+r_2} \prod_{i=0}^{m} \frac{\tilde{U}^{\alpha_i}(t)}{\alpha_i!} \frac{V_i^{\bar{\bar{\beta}}}(t)}{\bar{\bar{\beta}}_i!} \frac{\bar{V}_i^{\bar{\gamma}_i}(t)}{\sqrt{\bar{\gamma}_i!}} \frac{B_l^{\gamma_i'}(t)}{\gamma_i'!},
$$

where $\sum_{i=0}^{m}\alpha_i = k$, $\alpha_i = \bar{\alpha}_i + \left| q_{i_{k+n}}^{l_{k+n}} \cdots q_{i_{s_r+1}}^{l_{s_{r_1}+1}} \right|_{x_i}$, $\sum_{i=0}^{m}\left(\bar{\bar{\beta}}_i + \bar{\gamma}_i + \gamma_i'\right) = n$, $\bar{\bar{\beta}}_i = \bar{\beta}_i + \left| q_{i_{k+n}}^{l_{k+n}} \cdots q_{i_{s_r+1}}^{l_{s_{r_1}+1}} \right|_{y_i}$, and $\beta_i = \bar{\bar{\beta}}_i + \bar{\gamma}_i + \gamma_i'$. Similarly for $\bar{s}_{r_2} > s_{r_1}$, one instead applies the isometry property $n - (\bar{s}_{r_2} - 1)$ times and Hölder's inequality k_1 times. There are two situations. The first is when $\bar{s}_{r_2} + 1 \neq \bar{s}_{r_2}$. It then follows by Hölder's inequality that

$$
\begin{aligned}
&\left\| \mathbf{I}_{\eta}^{\substack{s_{r_1}\\ \bar{s}_{r_2}}}[w](t) \right\|_2^2 \\
&= \mathbf{E}\left[\left(\int_0^t w_{q_{i_{k+n}}^{l_{k+n}}}(t_{k+n}) \cdots \int_0^{t_{\bar{s}_{r_2}+3}} w_{q_{i_{\bar{s}_{r_2}+2}}^{l_{\bar{s}_{r_2}+2}}}(t_{\bar{s}_{r_2}+1}) \int_0^{t_{\bar{s}_{r_2}+1}} v_{i_{\bar{s}_{r_2}}+1}(t_{\bar{s}_{r_2}}) v_{i_{\bar{s}_{r_2}}}(t_{\bar{s}_{r_2}}) \cdot \right.\right. \\
&\qquad \left.\left. \mathbf{I}_{\eta'}^{\substack{s_{r_1}\\ \bar{s}_{r_2}-1}}[w](t_{\bar{s}_{r_2}})\, dt_{\bar{s}_{r_2}} dq_{i_{\bar{s}_{r_2}+1}}^{l_{\bar{s}_{r_2}+1}}(t_{\bar{s}_{r_2}+1}) \cdots dq_{i_{k+n}}^{l_{k+n}}(t_{k+n}) \right)^2 \right] \\
&\leq t^{k_1} \int_0^t \mathbf{E}\left[w^2_{q_{i_{k+n}}^{l_{k+n}}}(t_{k+n}) \right] \cdots \int_0^{t_{\bar{s}_{r_2}+3}} \mathbf{E}\left[w^2_{q_{i_{\bar{s}_{r_2}+2}}^{l_{\bar{s}_{r_2}+2}}}(t_{\bar{s}_{r_2}+1}) \right] t_{\bar{s}_{r_2}+1} \cdot \\
&\qquad \int_0^{t_{\bar{s}_{r_2}+1}} \mathbf{E}\left[v^2_{i_{\bar{s}_{r_2}}+1}(t_{\bar{s}_{r_2}}) v^2_{i_{\bar{s}_{r_2}}}(t_{\bar{s}_{r_2}}) \right] \left\| \mathbf{I}_{\eta'}^{\substack{s_{r_1}\\ \bar{s}_{r_2}-1}}[w](t_{s_r}) \right\|_2^2 dt_{\bar{s}_{r_2}} dt_{\bar{s}_{r_2}+1} \cdots dt_{k+n} \\
&\leq t^{k_1} \int_0^t \mathbf{E}\left[w^2_{q_{i_{k+n}}^{l_{k+n}}}(t_{k+n}) \right] \cdots \int_0^{t_{\bar{s}_{r_2}+3}} \mathbf{E}\left[w^2_{q_{i_{\bar{s}_{r_2}+2}}^{l_{\bar{s}_{r_2}+2}}}(t_{\bar{s}_{r_2}+1}) \right] t_{\bar{s}_{r_2}+1} \cdot \\
&\qquad \int_0^{t_{\bar{s}_{r_2}+1}} \mathbf{E}\left[v^2_{i_{\bar{s}_{r_2}}+1}(t_{\bar{s}_{r_2}}) v^2_{i_{\bar{s}_{r_2}}}(t_{\bar{s}_{r_2}}) \right] 2^{r_2-1} t_{\bar{s}_{r_2}}^{k-k_1+r_1+r_2-1} \cdot \\
&\qquad \prod_{i=0}^{m} \frac{\tilde{U}_i^{\bar{\alpha}_i}(t_{\bar{s}_{r_2}})}{\bar{\alpha}_i!} \frac{V_i^{\bar{\beta}_i}(t_{\bar{s}_{r_2}})}{\bar{\beta}_i!} \frac{\bar{V}_i^{\bar{\gamma}_i'}(t_{\bar{s}_{r_2}})}{\sqrt{\bar{\gamma}_i'!}} \frac{B_i^{\gamma_i}(t_{\bar{s}_{r_2}})}{\gamma_i!} dt_{\bar{s}_{r_2}} dt_{\bar{s}_{r_2}+1} \cdots dt_{k+n} \\
&\leq 2^{r_2-1} t^{k+r_1+r_2} \int_0^t \mathbf{E}\left[w^2_{q_{i_{k+n}}^{l_{k+n}}}(t_{k+n}) \right] \cdots \int_0^{t_{\bar{s}_{r_2}+3}} \mathbf{E}\left[w^2_{q_{i_{\bar{s}_{r_2}+2}}^{l_{\bar{s}_{r_2}+2}}}(t_{\bar{s}_{r_2}+1}) \right] \cdot \\
&\qquad \prod_{i=0}^{m} \frac{\tilde{U}_i^{\bar{\alpha}_i}(t_{\bar{s}_{r_2}+1})}{\bar{\alpha}_i!} \frac{V_i^{\bar{\beta}_i}(t_{\bar{s}_{r_2}+1})}{\bar{\beta}_i!} \frac{B_i^{\gamma_i}(t_{\bar{s}_{r_2}+1})}{\gamma_i!} \prod_{\substack{l=0\\ l\neq i_{\bar{s}_{r_2}}, i_{\bar{s}_{r_2}+1}}}^{m} \frac{\bar{V}_l^{\bar{\gamma}_l'}(t_{\bar{s}_{r_2}+1})}{\sqrt{\bar{\gamma}_l'!}} \cdot
\end{aligned}
$$

$$\left(\int_0^{t_{\bar{s}_{r_2}+1}} \mathbf{E}\left[v^2_{i_{\bar{s}_{r_2}+1}}(t_{\bar{s}_{r_2}})\right] \frac{\bar{V}_{i_{\bar{s}_{r_2}+1}}^{2\bar{\gamma}'_{i_{\bar{s}_{r_2}+1}}}(t_{\bar{s}_{r_2}})}{\bar{\gamma}'_{i_{\bar{s}_{r_2}}}!} dt_{\bar{s}_{r_2}}\right)^{\frac{1}{2}} \cdot$$

$$\left(\int_0^{t_{\bar{s}_{r_2}+1}} \mathbf{E}\left[v^2_{i_{\bar{s}_{r_2}}}(t_{\bar{s}_{r_2}})\right] \frac{\bar{V}_{i_{\bar{s}_{r_2}}}^{2\bar{\gamma}'_{i_{\bar{s}_{r_2}}}}(t_{\bar{s}_{r_2}})}{\bar{\gamma}'_{i_{\bar{s}_{r_2}}}!} dt_{\bar{s}_{r_2}}\right)^{\frac{1}{2}} dt_{\bar{s}_{r_2}+1} \cdots dt_{k+n}$$

$$\leq 2^{r_2-1} t^{k+r_1+r_2} \int_0^t \mathbf{E}\left[w^2_{q_{i_{k+n}}^{l_{k+n}}}(t_{k+n})\right] \cdots \int_0^{t_{\bar{s}_{r_2}+3}} \mathbf{E}\left[w^2_{q_{i_{\bar{s}_{r_2}+2}}^{l_{\bar{s}_{r_2}+2}}}(t_{\bar{s}_{r_2}+1})\right] \cdot$$

$$\prod_{i=0}^{m} \frac{\tilde{U}_i^{\bar{\alpha}_i}(t_{\bar{s}_{r_2}+1})}{\bar{\alpha}_i!} \frac{V_i^{\bar{\beta}_i}(t_{\bar{s}_{r_2}+1})}{\bar{\beta}_i!} \frac{B_i^{\gamma_i}(t_{\bar{s}_{r_2}+1})}{\gamma_i!} \cdot$$

$$\prod_{\substack{l=0 \\ l\neq i_{\bar{s}_{r_2}}, i_{\bar{s}_{r_2}+1}}}^{m} \frac{\bar{V}_l^{\bar{\gamma}_l}(t_{\bar{s}_{r_2}+1})}{\sqrt{\bar{\gamma}_l!}} \frac{\bar{V}_{i_{\bar{s}_{r_2}+1}}^{\bar{\gamma}'_{i_{\bar{s}_{r_2}+1}}+1}(t_{\bar{s}_{r_2}+1})}{\sqrt{(\bar{\gamma}'_{i_{\bar{s}_{r_2}+1}}+1)!}} \frac{\bar{V}_{i_{\bar{s}_{r_2}}}^{\bar{\gamma}'_{i_{\bar{s}_{r_2}}}+1}(t_{\bar{s}_{r_2}+1})}{\sqrt{(\bar{\gamma}'_{i_{\bar{s}_{r_2}}}+1)!}} dt_{\bar{s}_{r_2}+1} \cdots dt_{k+n}$$

$$\leq 2^{r_2-1} t^{k+r_1+r_2} \int_0^t \mathbf{E}\left[w^2_{q_{i_{k+n}}^{l_{k+n}}}(t_{k+n})\right] \cdots \int_0^{t_{\bar{s}_{r_2}+3}} \mathbf{E}\left[w^2_{q_{i_{\bar{s}_{r_2}+2}}^{l_{\bar{s}_{r_2}+2}}}(t_{\bar{s}_{r_2}+1})\right] \cdot$$

$$\prod_{i=0}^{m} \frac{\tilde{U}_i^{\bar{\alpha}_i}(t_{\bar{s}_{r_2}+1})}{\bar{\alpha}_i!} \frac{V_i^{\bar{\beta}_i}(t_{\bar{s}_{r_2}+1})}{\bar{\beta}_i!} \frac{\bar{V}_i^{\bar{\gamma}_i}(t_{\bar{s}_{r_2}+1})}{\sqrt{\bar{\gamma}_i!}} \frac{B_i^{\gamma_i}(t_{\bar{s}_{r_2}+1})}{\gamma_i!} dt_{\bar{s}_{r_2}+1} \cdots dt_n,$$

where $\bar{\gamma}_i = \bar{\gamma}'_i + \delta_{i\bar{s}_{r_2}} + \delta_{i(\bar{s}_{r_2}+1)}$ and $\sum_{i=0}^m \bar{\gamma}_i = 2r_2$. Thus, the remaining $(n + k_1 - (s_{r_2}+1))$ integrals are calculated as in the case when $s_{r_1 > \bar{s}_{r_2}}$, and again the $\bar{\alpha}_i$'s and $\bar{\beta}_i$'s increase instead of the $\bar{\gamma}'_i$'s. Therefore,

$$\left\|\mathbf{I}_{\eta}^{\substack{s_{r_1} \\ \bar{s}_{r_2}}}[w](t)\right\|_2^2 \leq 2\left(2^{r_2-1} t^{k+r_1+r_2} \prod_{i=0}^{m} \frac{\tilde{U}^{\alpha_i}(t)}{\alpha_i!} \frac{V_i^{\bar{\bar{\beta}}}(t)}{\bar{\bar{\beta}}_i!} \frac{\bar{V}_i^{\bar{\gamma}_i}(t)}{\sqrt{\bar{\gamma}_i!}} \frac{B_l^{\gamma_i}(t)}{\gamma_i!}\right),$$

where $\sum_{i=0}^m \alpha_i = k$, $\alpha_i = \bar{\alpha}_i + \left|q_{i_{k+n}}^{l_{k+n}} \cdots q_{i_{s_{r_2}+2}}^{l_{s_{r_2}+2}}\right|_{x_i}$, $\sum_{i=0}^m \left(\bar{\bar{\beta}}_i + \bar{\gamma}_i + \gamma_i\right) = n$, $\bar{\bar{\beta}}_i = \bar{\beta}_i + \left|q_{i_{k+n}}^{l_{k+n}} \cdots q_{i_{s_{r_2}+2}}^{l_{s_{r_2}+2}}\right|_{y_i}$, and $\beta_i = \bar{\bar{\beta}}_i + \bar{\gamma}_i + \gamma_i$. In the second situation, $i_{\bar{s}_{r_2}} + 1 = i_{\bar{s}_{r_2}}$. Since $(\bar{\gamma}'_i + 2) \geq \sqrt{\bar{\gamma}'_i + 1}\sqrt{\bar{\gamma}'_i + 2}$ it follows that

$$\left\|\mathbf{I}_{\eta}^{\substack{s_{r_1} \\ \bar{s}_{r_2}}}[w](t)\right\|_2^2$$

$$\leq 2^{r_2-1} t^{k+r_1+r_2} \int_0^t \mathbf{E}\left[w^2_{q_{i_{k+n}}^{l_{k+n}}}(t_{k+n})\right] \cdots \int_0^{t_{\bar{s}_{r_2}+3}} \mathbf{E}\left[w^2_{q_{i_{\bar{s}_{r_2}+2}}^{l_{\bar{s}_{r_2}+2}}}(t_{\bar{s}_{r_2}+1})\right] \cdot$$

$$\prod_{i=0}^{m} \frac{\tilde{U}_i^{\bar{\alpha}_i}(t_{\bar{s}_{r_2}+1})}{\bar{\alpha}_i!} \frac{V_i^{\bar{\beta}_i}(t_{\bar{s}_{r_2}+1})}{\bar{\beta}_i!} \frac{B_i^{\gamma_i}(t_{\bar{s}_{r_2}+1})}{\gamma_i!} \prod_{\substack{l=0\\ l\neq i_{\bar{s}_{r_2}}, i_{\bar{s}_{r_2}+1}}}^{m} \frac{\bar{V}_l^{\bar{\gamma}'_l}(t_{\bar{s}_{r_2}+1})}{\sqrt{\bar{\gamma}'_l!}} \cdot$$

$$\int_0^{t_{\bar{s}_{r_2}+1}} \mathbf{E}\left[v^4_{i_{\bar{s}_{r_2}+1}}(t_{\bar{s}_{r_2}})\right] \frac{\bar{V}_{i_{\bar{s}_{r_2}+1}}^{\bar{\gamma}'_{i_{\bar{s}_{r_2}+1}}}(t_{\bar{s}_{r_2}})}{\sqrt{\bar{\gamma}'_{i_{\bar{s}_{r_2}+1}}!}} dt_{\bar{s}_{r_2}} dt_{\bar{s}_{r_2}+1} \cdots dt_{k+n}$$

$$\leq 2^{r_2-1} t^{k+r_1+r_2} \int_0^t \mathbf{E}\left[w^2_{q_{i_{k+n}}^{l_{k+n}}}(t_{k+n})\right] \cdots \int_0^{t_{\bar{s}_{r_2}+3}} \mathbf{E}\left[w^2_{q_{i_{\bar{s}_{r_2}+2}}^{l_{\bar{s}_{r_2}+2}}}(t_{\bar{s}_{r_2}+1})\right] \cdot$$

$$\prod_{i=0}^{m} \frac{\tilde{U}_i^{\bar{\alpha}_i}(t_{\bar{s}_{r_2}+1})}{\bar{\alpha}_i!} \frac{V_i^{\bar{\beta}_i}(t_{\bar{s}_{r_2}+1})}{\bar{\beta}_i!} \frac{B_i^{\gamma_i}(t_{\bar{s}_{r_2}+1})}{\gamma_i!} \prod_{\substack{l=0\\ l\neq i_{\bar{s}_{r_2}}, i_{\bar{s}_{r_2}+1}}}^{m} \frac{\bar{V}_l^{\bar{\gamma}'_l}(t_{\bar{s}_{r_2}+1})}{\sqrt{\bar{\gamma}'_l!}} \cdot$$

$$\int_0^{t_{\bar{s}_{r_2}+1}} \mathbf{E}\left[v^4_{i_{\bar{s}_{r_2}}}(t_{\bar{s}_{r_2}})\right] \frac{\bar{V}_{i_{\bar{s}_{r_2}}}^{\bar{\gamma}'_{i_{\bar{s}_{r_2}}}}(t_{\bar{s}_{r_2}})}{\sqrt{\bar{\gamma}'_{i_{\bar{s}_{r_2}}}!}} dt_{\bar{s}_{r_2}}\, dt_{\bar{s}_{r_2}+1} \cdots dt_{k+n}$$

$$\leq 2^{r_2} t^{k+r_1+r_2} \int_0^t \mathbf{E}\left[w^2_{q_{i_{k+n}}^{l_{k+n}}}(t_{k+n})\right] \cdots \int_0^{t_{\bar{s}_{r_2}+3}} \mathbf{E}\left[w^2_{q_{i_{\bar{s}_{r_2}+2}}^{l_{\bar{s}_{r_2}+2}}}(t_{\bar{s}_{r_2}+1})\right] \cdot$$

$$\prod_{i=0}^{m} \frac{\tilde{U}_i^{\bar{\alpha}_i}(t_{\bar{s}_{r_2}+1})}{\bar{\alpha}_i!} \frac{V_i^{\bar{\beta}_i}(t_{\bar{s}_{r_2}+1})}{\bar{\beta}_i!} \frac{B_i^{\gamma_i}(t_{\bar{s}_{r_2}+1})}{\gamma_i!} \prod_{\substack{l=0\\ l\neq i_{\bar{s}_{r_2}}, i_{\bar{s}_{r_2}+1}}}^{m} \frac{\bar{V}_l^{\bar{\gamma}'_l}(t_{\bar{s}_{r_2}+1})}{\sqrt{\bar{\gamma}'_l!}} \frac{\bar{V}_{i_{\bar{s}_{r_2}}}^{\bar{\gamma}'_{i_{\bar{s}_{r_2}}}+2}(t_{\bar{s}_{r_2}})}{\sqrt{\bar{\gamma}'_{i_{\bar{s}_{r_2}}}!(\bar{\gamma}'_{i_{\bar{s}_{r_2}}}+2)}} \cdot$$

$$dt_{\bar{s}_{r_2}+1} \cdots dt_{k+n}$$

$$\leq 2^{r_2} t^{k+r_1+r_2} \int_0^t \mathbf{E}\left[w^2_{q_{i_{k+n}}^{l_{k+n}}}(t_{k+n})\right] \cdots \int_0^{t_{\bar{s}_{r_2}+3}} \mathbf{E}\left[w^2_{q_{i_{\bar{s}_{r_2}+2}}^{l_{\bar{s}_{r_2}+2}}}(t_{\bar{s}_{r_2}+1})\right] \cdot$$

$$\prod_{i=0}^{m} \frac{\tilde{U}_i^{\bar{\alpha}_i}(t_{\bar{s}_{r_2}+1})}{\bar{\alpha}_i!} \frac{V_i^{\bar{\beta}_i}(t_{\bar{s}_{r_2}+1})}{\bar{\beta}_i!} \frac{\bar{V}_i^{\bar{\gamma}_i}(t_{\bar{s}_{r_2}+1})}{\sqrt{\bar{\gamma}_i!}} \frac{B_i^{\gamma_i}(t_{\bar{s}_{r_2}+1})}{\gamma_i!} dt_{\bar{s}_{r_2}+1} \cdots dt_n,$$

where $\bar{\gamma}_i = \bar{\gamma}'_i + 2\delta_{i i_{\bar{s}_{r_2}}}$ and $\sum_{i=0}^m \bar{\gamma}'_i = 2r_2$. The remaining $n + k_1 - (\bar{s}_{r_2} - 1)$ integrals are evaluated as in the previous steps. Thus,

$$\left\| \mathbf{I}_{\eta}^{\substack{s_{r_1}\\ \bar{s}_{r_2}}}[w](t) \right\|_2^2 \leq 2^{r_2} t^{k+r_1+r_2} \prod_{i=0}^{m} \frac{\tilde{U}^{\alpha_i}(t)}{\alpha_i!} \frac{V_i^{\bar{\bar{\beta}}}(t)}{\bar{\bar{\beta}}_i!} \frac{\bar{V}_i^{\bar{\gamma}_i}(t)}{\sqrt{\bar{\gamma}_i!}} \frac{B_l^{\gamma_i}(t)}{\gamma_i!},$$

where $\sum_{i=0}^m \alpha_i = k$, $\alpha_i = \bar{\alpha}_i + \left| q_{i_{k+n}}^{l_{k+n}} \cdots q_{i_{s_{r_2}+2}}^{l_{s_{r_2}+2}} \right|_{x_i}$, $\sum_{i=0}^m \left(\bar{\bar{\beta}}_i + \bar{\gamma}_i + \gamma_i\right) = n$, $\bar{\bar{\beta}}_i = \bar{\beta}_i + \left| q_{i_{k+n}}^{l_{k+n}} \cdots q_{i_{s_{r_2}+2}}^{l_{s_{r_2}+2}} \right|_{y_i}$, and $\beta_i = \bar{\bar{\beta}}_i + \bar{\gamma}_i + \gamma_i$. Since all the possible situations are

covered, the proof is complete. ■

A consequence of the previous two theorems is the L_2 upper bound for the random variable $E_\eta[w](t)$ when $\eta \in X^kY^n$.

Theorem IV.1.4. *Let $\eta \in X^kY^n$ and $w \in \mathcal{UV}^m$ be arbitrary. Then for a fixed $t \in [0,T]$*

$$\|E_\eta[w](t)\|_2 < \frac{(R\sqrt{t})^k(\sqrt{2}R(\sqrt{t}+2))^{2n}}{(\alpha!)^{\frac{1}{2}}(\beta!)^{\frac{1}{4}}}, \tag{IV.1.20}$$

where $\alpha! \triangleq \alpha_0!\cdots\alpha_m!$, $\beta! \triangleq \beta_0!\cdots\beta_m!$ and $\max\{\|u\|_{L_2}, \|v\|_{L_2}, \|v_0\|_{L_2}, \|v\|_{L_4}\} \le R$.

Proof: A Stratonovich integral can be written in terms of Itô integrals using (IV.1.4). Note that $\#(\bar{A}_{nr_2}) \le \binom{n-r_2}{r_2} < \binom{n}{r_2}$ and $\#(A_{nr_1}^{\bar{s}_{r_2}}) = \binom{n-2r_2}{r_1} \le \binom{n}{r_1}$. Using the triangle inequality, Theorem IV.1.3 and the binomial theorem, observe

$$\begin{aligned}
\|E_\eta[w](t)\|_2 &\le \sum_{r_1=0,r_2=0}^{n,\lfloor\frac{n}{2}\rfloor} \frac{1}{2^{r_1}2^{r_2}} \sum_{\substack{s_{r_1}\in A_{nr_1}^{\bar{s}_{r_2}}\\ \bar{s}_{r_2}\in\bar{A}_{nr_2}}} \left\| \mathbf{I}_\eta^{\substack{s_{r_1}\\ \bar{s}_{r_2}}}[w](t)\right\|_2 \\
&\le \sum_{r_1=0,r_2=0}^{n,\lfloor\frac{n}{2}\rfloor} \frac{t^{\frac{k+r_1+r_2}{2}}}{2^{r_1}2^{\frac{r_2}{2}}} \sum_{\substack{s_{r_1}\in A_{nr_1}^{\bar{s}_{r_2}}\\ \bar{s}_{r_2}\in\bar{A}_{nr_2}}} \sqrt{\prod_{i=0}^{m} \frac{\bar{U}_i^{\alpha_i(t)}}{\alpha_i!}\frac{V_i^{\bar{\beta}_i(t)}}{\bar{\beta}_i!}\frac{\bar{V}_i^{\bar{\gamma}_i}(t)}{\sqrt{\bar{\gamma}_i!}}\frac{B_i^{\gamma_i}(t)}{\gamma_i!}} \\
&\le \left(\frac{R^{k+n}}{(\alpha!)^{\frac{1}{2}}}\right) \sum_{r_1=0,r_2=0}^{n,\lfloor\frac{n}{2}\rfloor} \frac{t^{\frac{k+r_1+r_2}{2}}}{2^{r_1}2^{\frac{r_2}{2}}} \sum_{\substack{s_{r_1}\in A_{nr_1}^{\bar{s}_{r_2}}\\ \bar{s}_{r_2}\in\bar{A}_{nr_2}}} \prod_{i=0}^{m} \frac{1}{(\bar{\beta}_i!)^{\frac{1}{2}}(\bar{\gamma}_i!)^{\frac{1}{4}}(\gamma_i!)^{\frac{1}{2}}} \\
&\le \left(\frac{R^{k+n}t^{\frac{k}{2}}}{(\alpha!)^{\frac{1}{2}}(\beta!)^{\frac{1}{4}}}\right) \sum_{r_1=0,r_2=0}^{n,\lfloor\frac{n}{2}\rfloor} \frac{t^{\frac{r_1+r_2}{2}}}{2^{r_1}2^{\frac{r_2}{2}}} \sum_{\substack{s_{r_1}\in A_{nr_1}^{\bar{s}_{r_2}}\\ \bar{s}_{r_2}\in\bar{A}_{nr_2}}} \prod_{i=0}^{m} \frac{\beta!}{\bar{\beta}_i!\bar{\gamma}_i!\gamma_i!} \\
&\le \left(\frac{R^{k+n}t^{\frac{k}{2}}}{(\alpha!)^{\frac{1}{2}}(\beta!)^{\frac{1}{4}}}\right) 3^n \sum_{r_1=0}^{n} \frac{t^{\frac{r_1}{2}}}{2^{r_1}}\binom{n}{r_1} \sum_{r_2=0}^{n} \frac{t^{\frac{r_2}{2}}}{2^{\frac{r_2}{2}}}\binom{n}{r_2} \\
&\le \frac{(R\sqrt{t})^k(3\sqrt{2}R(\sqrt{t}+2)(\sqrt{t}+\sqrt{2}))^n}{4^n(\alpha!)^{\frac{1}{2}}(\beta!)^{\frac{1}{4}}}
\end{aligned}$$

$$< \frac{(R\sqrt{t})^k(\sqrt{2R}(\sqrt{t}+2))^{2n}}{(\alpha!)^{\frac{1}{2}}(\beta!)^{\frac{1}{4}}}.$$

■

IV.1.4 Fliess operators and their global convergence

In this section Fliess operators are suitably extended in order to accept stochastic processes from $\mathcal{UV}^m[0,T]$ as inputs. Now recall that for any $w \in \mathcal{UV}^m[0,T]$, R is an upperbound for $\|u\|_{L_2}$, $\|v\|_{L_2}$ and $\|v\|_{L_4}$. Thus, the concrete objective of this section is to show the mean square convergence of the stochastic extension of Fliess operator for all $t \in [0,T]$, where $T, R \geq 0$ are arbitrarily large but finite. This type of convergence will be known as *global convergence.*

Definition IV.1.2. A causal m-input, ℓ-output **Fliess operator** F_c, $c \in \mathbb{R}^\ell\langle\langle XY\rangle\rangle$, driven by a stochastic process in $\mathcal{UV}^m[0,T]$ is formally defined as

$$F_c[w](t) = \sum_{\eta \in XY^*} (c,\eta)\, E_\eta[w](t), \tag{IV.1.21}$$

where each E_η is given in (IV.1.2)-(IV.1.3).

The operator F_c lacks real meaning unless its convergence is described in some manner. Since (IV.1.21) involves stochastic integrals, a mean square notion of convergence is assumed, i.e., the infinite series $\sum_{\eta\in XY^*} |(c,\eta)|\, \|E_\eta[w]\|_2$ is finite. The procedure used here is motivated by Riccomagno in [50, 51]. In this regard, consider the following definition.

Definition IV.1.3. For a fixed $t \in [0,T]$, the series $F_c[w](t)$ in (IV.1.21) is said to be a **Cauchy series** if for any $\epsilon > 0$ there exist an $N > 0$ such that

$$\left\| \sum_{j=N_1}^{N_2} \sum_{k=0}^{j} \sum_{\eta \in X^k Y^{j-k}} (c,\eta) E_\eta[w](t) \right\|_2 < \epsilon,$$

when $N_2 > N_1 > N$.

It is well known that $L_2(\Omega, \mathcal{F}, P)$ with its usual norm is a Hilbert space modulo the almost sure equivalence relation in Definition II.1.4. The following theorem ensures that a Fliess operator converges *absolutely* in the mean square sense to produce a well-defined output process when its corresponding series is globally convergent.

Theorem IV.1.5. *Suppose for a series $c \in \mathbb{R}^{\ell}\langle\langle XY\rangle\rangle$ there exists real numbers $K, M > 0$ such that*

$$|(c,\eta)| \leq KM^{|\eta|}, \quad \forall \eta \in XY^*.$$

Then for any stochastic process $w \in \mathcal{UV}^m[0,T]$, $T > 0$, the series (IV.1.21) *converges absolutely in the mean square sense to a well-defined random vector $y(t) = F_c[w](t)$, $t \in [0,T]$.*

Proof: Without loss of generality it is assumed that $\ell = 1$. Pick a $t \in [0,T]$ and any $w \in \mathcal{UV}^m[0,T]$. Let $R = \max\{\|u\|_{L_2}, \|v\|_{L_2}, \|v_0\|_{L_2}, \|v\|_{L_4}\}$. For a word $\eta \in X^kY^n$, recall $k = \sum_{i=0}^m \alpha_i$ is the number of Lebesgue integrals in η, while $n = \sum_{i=0}^m \beta_i$ is the number of stochastic integrals in η. Define

$$a_{k,n}(t) \triangleq \sum_{\eta \in X^kY^n} (c,\eta) E_\eta[w](t).$$

Note that the language $L_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \{\eta \in X^kY^n : |\eta|_{x_i} = \alpha_i, |\eta|_{y_i} = \beta_i, i = 0,\ldots,m\}$ consists of $(k+n)!/(\alpha!\beta!)$ words. Applying Theorem IV.1.4,

$$\begin{aligned}
\|a_{k,n}(t)\|_2 &\leq \sum_{\eta \in X^kY^n} |(c,\eta)|\, \|E_\eta[w](t)\|_2 \\
&\leq KM^{n+k} \sum_{\|\alpha\|=k, \|\beta\|=n} \frac{(R\sqrt{t})^k(\sqrt{2R}(\sqrt{t}+2))^{2n}}{(\alpha!)^{\frac{1}{2}}(\beta!)^{\frac{1}{4}}} \frac{(k+n)!}{\alpha!\beta!}.
\end{aligned}$$

Without loss of generality, it is assumed that $R \geq 1$. If $R' \triangleq 4R(R+4)$, then by the multinomial theorem,

$$
\begin{aligned}
\|a_{k,n}(t)\|_2 &\leq K(MR')^{k+n} \sum_{\|\alpha\|=k, \|\beta\|=n} \frac{(k+n)!}{(\alpha!)^{\frac{3}{2}}(\beta!)^{\frac{5}{4}}} \\
&\leq K(2MR')^{k+n} \sum_{\|\alpha\|=k, \|\beta\|=n} \frac{k!\; n!}{(\alpha!)^{\frac{3}{2}}(\beta!)^{\frac{5}{4}}} \\
&\leq \frac{K(2MR')^{k+n}}{(k!)^{\frac{1}{2}}(n!)^{\frac{1}{4}}} \sum_{\|\alpha\|=k} \frac{(k!)^{\frac{3}{2}}}{(\alpha!)^{\frac{3}{2}}} \sum_{\|\beta\|=n} \frac{(n!)^{\frac{5}{4}}}{(\beta!)^{\frac{5}{4}}} \\
&\leq \frac{K(2MR')^{k+n}}{(k!)^{\frac{1}{2}}(n!)^{\frac{1}{4}}} \left(\sum_{\|\alpha\|=k} \frac{k!}{\alpha!}\right)^2 \left(\sum_{\|\beta\|=n} \frac{n!}{\beta!}\right)^2 \\
&\leq \frac{K(2MR'(m+1)^2)^{k+n}}{(k!)^{\frac{1}{2}}(n!)^{\frac{1}{4}}}.
\end{aligned}
\tag{IV.1.22}
$$

To show that (IV.1.21) is mean square convergent, it is sufficient to show that it is a Cauchy series. Since $|\eta| = |\eta|_X + |\eta|_Y = k + n = j$, it follows immediately from the triangle inequality that

$$
\left\| \sum_{j=N_1}^{N_2} \sum_{k=0}^{j} \sum_{\eta \in X^k Y^{j-k}} (c, \eta) E_\eta[w](t) \right\|_2 \leq \sum_{j=N_1}^{N_2} \sum_{k=0}^{j} \|a_{k,j-k}(t)\|_2 \tag{IV.1.23}
$$

for any $N_2 > N_1 \in \mathbb{N}$. Now for any $\epsilon > 0$ there exist an $N > 0$ such that by (IV.1.22)

$$
\begin{aligned}
&\sum_{j=N_1}^{N_2} \sum_{k=0}^{j} \|a_{k,j-k}(t)\|_2 \\
&\quad \leq \sum_{j=N_1}^{N_2} \sum_{k=0}^{j} \sum_{\eta \in X^k Y^{j-k}} |(c, \eta)| \, \|E_\eta[w](t)\|_2 \\
&\quad \leq K \sum_{j=0}^{\infty} \sum_{k=0}^{j} \frac{(2MR'(m+1)^2)^k (2MR'(m+1)^2)^{j-k}}{(k!)^{\frac{1}{2}}((j-k)!)^{\frac{1}{4}}} \\
&\quad = K \sum_{k=0}^{\infty} \frac{(2MR'(m+1)^2)^k}{(k!)^{\frac{1}{2}}} \sum_{n=0}^{\infty} \frac{(2MR'(m+1)^2)^n}{(n!)^{\frac{1}{4}}} \\
&\quad = K \sum_{k=0}^{\infty} \frac{(M'')^k}{(k!)^{\frac{1}{2}}} \sum_{n=0}^{\infty} \frac{(M'')^n}{(n!)^{\frac{1}{4}}} \\
&\quad < \epsilon
\end{aligned}
$$

for all $N_2 > N_1 > N$, where $M'' \triangleq (2MR'(m+1)^2)$. Note that $\frac{(M'')^k}{(k!)^{\frac{1}{2}}}$ and $\frac{(M'')^n}{(n!)^{\frac{1}{4}}}$ are the k-th and n-th term of an absolutely convergent series, respectively. By the ratio test,

$$\lim_{k\to\infty} \frac{s_{k+1}}{s_k} = \lim_{k\to\infty} \frac{M''}{(k+1)^{\frac{1}{2}}} = 0$$

and

$$\lim_{n\to\infty} \frac{s_{n+1}}{s_n} = \lim_{n\to\infty} \frac{M''}{(n+1)^{\frac{1}{4}}} = 0.$$

Thus the series (IV.1.21) is Cauchy. This implies that $\sum_{\eta\in XY^*} |(c,\eta)| \, \|E_\eta[w]\|_2$ is finite. ■

Example IV.1.1. Consider the following system driven by a Wiener process

$$\begin{aligned} dz(t) &= Mz(t)\, dW(t),\ z(0) = 1 \\ y(t) &= Kz(t). \end{aligned} \tag{IV.1.24}$$

The generating series for (IV.1.24) is $(c, y_0^k) = KM^k$ for $k \geq 0$ and 0 otherwise. Therefore, c satisfies the growth condition in Theorem IV.1.5. The output when $w = 0$ is

$$y(t) = F_c[0](t) = \sum_{k=0}^{\infty} KM^k \oint_0^t \cdots \oint_0^{t_2} dW(t_1) \cdots dW(t_k).$$

Since Stratonovich integrals follow the rules of standard integral calculus,

$$\oint_0^t \frac{W^k(s)}{k!}\, dW(s) = \frac{W^{k+1}(t)}{(k+1)!},\ k \geq 0. \tag{IV.1.25}$$

Hence,

$$y(t) = F_c[0](t) = \sum_{k=0}^{\infty} KM^k \frac{W^k(t)}{k!} = Ke^{MW(t)},\ t \in [0,\infty).$$

□

One application of Fliess operators is that it provides a series solution of the state equation of a nonlinear system [17,21,42,50]. This fact is related to the Borel-Laplace

transform of analytic signals, which was applied by Fliess in [17]. A similar approach was used by Riccomagno in [50] to solve a specific class of nonlinear stochastic differential equations. The following examples illustrate the basic ideas in the stochastic setting utilized here.

Example IV.1.2. Consider the following stochastic linear system

$$dz(t) = -az(t)\,dt + b\dot{w}(t), \quad z(0) = 1, \tag{IV.1.26}$$

where $\dot{w}$ is the formal derivative of $w \in \mathcal{UV}[0,T]$, i.e., if

$$w = \int_0^t u(s)\,ds + \oint_0^t v(s)\,dW(s),$$

then

$$\dot{w} = \frac{d}{dt}w = u(s) + v(s)\bar{w}, \tag{IV.1.27}$$

where $\bar{w}$ stands for white Gaussian noise. Technically speaking, equation (IV.1.26) is only valid in its integral form

$$z(t) - z(0) = -a\int_0^t z(s)\,ds + b\left(\int_0^t u(s)\,ds + \int_0^t v(s)\,dW(s)\right), \quad z(0) = 1.$$

Assume there exists a series $c \in \mathbb{R}\langle\langle XY\rangle\rangle$ such that the solution $z(t) = F_c[w](t)$. Then the solution $z(t)$ can be replaced by the series c, Lebesgue integration by the letter x_0 to the left, Stratonovich integration by the letter y_0 to the left, Lebesgue integration with integrand the i-th component of u by x_i to the left and Stratonovich integration with integrand the i-th component of v by y_i to the left, where $i \neq 0$. Applying these rules, one can write the algebraic equation

$$c - 1 = -ax_0\,c + b\,(x_1 + y_1), \tag{IV.1.28}$$

Solving for c gives

$$c = (1 + ax_0)^{-1}(1 + b(x_1 + y_1)).$$

Given that c is written as the ratio of two polynomials, it resembles the Laplace transform of a linear system. Since c is rational, it follows from Theorem IV.1.5 that the solution to the system has to be globally convergent. The Fliess operator associated with c is

$$\begin{aligned} F_c[w](t) &= F_{(1+ax_0)^{-1}(1+b(x_1+y_1))}[w](t) \\ &= F_{(1+ax_0)^{-1}}[w](t) + F_{b(1+ax_0)^{-1}x_1}[w](t) + F_{b(1+ax_0)^{-1}y_1}[w](t) \\ &= F_{(1+ax_0)^{-1}}[w](t) + \int_0^t bF_{(1+ax_0)^{-1}}[w](s)u(s)\,ds \\ &\quad + \int_0^t bF_{(1+ax_0)^{-1}}[w](s)v(s)\,dW(s). \end{aligned}$$

A simple calculation shows that

$$\begin{aligned} F_{(1+ax_0)^{-1}}[w](t) &= \sum_{k=0}^{\infty} F_{(-ax_0)^k}[w](t) = \sum_{k=0}^{\infty} \frac{(-a)^k}{k!} F_{x_0^{\sqcup\!\sqcup k}}[w](t) \\ &= \sum_{k=0}^{\infty} \frac{(-a)^k}{k!} \left(E_{x_0}[w](t)\right)^k = \sum_{k=0}^{\infty} \frac{(-at)^k}{k!} = e^{-a(t-0)}. \end{aligned}$$

Hence,

$$\begin{aligned} F_c[w](t) &= e^{-at} + b\int_0^t F_{(1+ax_0)^{-1}}[w](s)u(s)\,ds + b\int_0^t F_{(1+ax_0)^{-1}}[w](s)v(s)\,dW(s) \\ &= e^{-at} + b\int_0^t e^{-a(t-s)}u(s)\,ds + b\int_0^t e^{-a(t-s)}v(s)\,dW(s). \end{aligned}$$

This solution is a Volterra series with the following kernels:

$$h_0(t) = e^{-at} \text{ and } h_1(t,s) = be^{-a(t-s)}.$$

$\square$

In stochastic calculus a so-called linear stochastic differential equation has the form of a bilinear system fed with a Wiener process input [39, 47]. Besides, it is known that bilinearity is equivalent to having a rational generating series [4]. For example, consider the stochastic differential equation (I.1.6), where $f(z) = N_0^1 z$ and $g(z) = N_0^2 z$, with $N_0^1, N_0^2 \in \mathbb{R}^{n\times n}$ and $z(t) \in \mathbb{R}^n$. It follows that

$$z(t) = \int_0^t N_0^1 z(s)\, ds + \oint_0^t N_0^2 z(s)\, dW(s). \tag{IV.1.29}$$

If $\dot{w}$ denotes the formal derivative of the Wiener process W, i.e., $W(t) = \int_0^t \dot{w}\, dt$, then (IV.1.29) can be written as a bilinear system driven by the white Gaussian noise $\dot{w}$,

$$\begin{aligned} \dot{z}(t) &= N_0^1 z(t) + N_0^2 z(t)\dot{w}(t), \quad z(0) = \gamma \\ y(t) &= \lambda z(t). \end{aligned}$$

The iterative procedure developed for (I.1.6) gives

$$z(t) = \sum_{\eta \in XY^*} \lambda \mu(\eta)\gamma \, E_\eta[0](t),$$

where μ is defined recursively as $\mu(x_0\eta) = N_0^1\mu(\eta)$ and $\mu(y_0\eta) = N_0^2\mu(\eta)$ for all $\eta \in XY^*$. In Corollary III.2.1, It was shown that there always exists $K, M > 0$ such that $|(c, \eta)| = |\lambda\mu(\eta)\gamma| \leq KM^{|\eta|}$. Therefore, the series associated with $z(t)$ is rational, and thus, globally convergent. The next example shows how to analyze a bilinear system algebraically as in Example IV.1.2.

Example IV.1.3. Consider the following bilinear system

$$dz(t) = -az(t)dt + bz(t)\dot{w}(t), \quad z(0) = 1,$$

where $w \in \mathcal{UV}[0,T]$. In integral form this becomes

$$z(t) - z(0) = -a\int_0^t z(s)ds + b\left(\int_0^t z(t)u(s)\,ds + \int_0^t v(s)\,dW(s)\right), \quad z(0) = 1. \tag{IV.1.30}$$

Assuming there exist a series $c \in \mathbb{R}\langle\langle XY\rangle\rangle$ such that $z(t) = F_c[w](t)$ and using the substitution rules described in the previous example, one obtains

$$c - 1 = -ax_0\,c + b\,(x_1 + y_1)c.$$

Solving for c gives

$$c = (1 + ax_0 - b(x_1 + y_1))^{-1}.$$

Observe that c is a rational series, and therefore, c is also a globally convergent series. From Theorem IV.1.5, it is known that $z(t) = F_c[w](t)$ converges to a well-defined random variable for all $t \in [0,T]$. On the other hand, the operator $F_c[w]$ can now be written as

$$\begin{aligned} F_c[w](t) &= F_{(1+ax_0-b(x_1+y_1))^{-1}}[w](t) \\ &= \sum_{k=0}^{\infty} F_{(-ax_0+b(x_1+y_1))^k}[w](t). \end{aligned}$$

Riccomagno in [50] developed an extension of identity $i.$ in Theorem III.1.2, i.e.,

$$(ax_0 + by_0)^{\sqcup\!\sqcup\, k} = k!(ax_0 + by_0)^k. \tag{IV.1.31}$$

For $i \neq 0$, a simple extension shows that

$$(ax_0 + bx_i + cy_i)^{\sqcup\!\sqcup\, k} = k!(ax_0 + bx_i + cy_i)^k.$$

Then

$$
\begin{aligned}
F_c[w](t) &= \sum_{k=0}^{\infty} \frac{1}{k!} F_{(-ax_0+b(x_1+y_1))^{\sqcup\!\sqcup\, k}}[w](t) \\
&= \sum_{k=0}^{\infty} \frac{1}{k!} \left(F_{(-ax_0+b(x_1+y_1))}[w](t)\right)^k \\
&= \sum_{k=0}^{\infty} \frac{(-at+bw(t))^k}{k!} = e^{-at+bw(t)}.
\end{aligned}
$$

Fliess in [17] also suggested the following iterative scheme.

$$
\begin{aligned}
c_0 &= (1+ax_0)^{-1}, \\
c_k &= b(1+ax_0)^{-1} y_1 c_{k-1},
\end{aligned}
$$

with $c \triangleq \sum_{k=0}^{\infty} c_k$. For simplicity, assume $u(t) = 0$. Then

$$
\begin{aligned}
c_1 &= b(1+ax_0)^{-1} y_1 (1+ax_0)^{-1}, \\
c_2 &= b^2(1+ax_0)^{-1} y_1 (1+ax_0)^{-1} y_1 (1+ax_0)^{-1}, \\
c_3 &= b^3(1+ax_0)^{-1} y_1 (1+ax_0)^{-1} y_1 (1+ax_0)^{-1} y_1 (1+ax_0)^{-1}, \\
&\vdots \\
c_k &= b^k(1+ax_0)^{-1} y_1 (1+ax_0)^{-1} y_1 \cdots y_1 (1+ax_0)^{-1} y_1 (1+ax_0)^{-1}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
F_{c_k}[w](t) &= b^k \oint_0^t e^{-a(t-t_k)} v(t_k) \oint_0^{t_k} e^{-a(t_k-t_{k-1}} v(t_{k-1}) \cdots \\
&\quad \cdots \oint_0^{t_2} e^{-a(t_2-t_1)} e^{-a(t_1)} v(t_{k-1})\, dW(t_1) \cdots dW(t_{k-1}) dW(t_k) \\
&= e^{-at} b^k \oint_0^t v(t_k) \oint_0^{t_k} v(t_{k-1}) \cdots \oint_0^{t_2} v(t_{k-1})\, dW(t_1) \cdots dW(t_{k-1}) dW(t_k) \\
&= e^{-at} \frac{\left(b \oint_0^t v(s)\, dW(s)\right)^k}{k!}.
\end{aligned}
$$

Thus, $F_c[w](t) = \sum_{k=0}^{\infty} F_{c_k}[w](t) = e^{-at+bw(t)}$. $\square$

Example IV.1.4. A *switched linear system* is usually modeled as

$$\dot{z} = A_u z,$$

where $u : \mathbb{R}^+ \longrightarrow \{0,1\}$ is a switching signal and A_0, A_1 are square matrices. This system can be represented as a bilinear system in the following manner:

$$\begin{aligned}\dot{z} &= A_1 z u + A_0 z(1-u) \\ &= A_0 z + (A_1 - A_0) z u \\ &= N_0 z + N_1 z u,\end{aligned}$$

where $N_0 = A_0$ and $N_1 = A_1 - A_0$. If u is stochastic, then an appropriate type of stochastic process modeling the integral process induced by u is a Poisson process. This type of process falls into the class of *jump processes* or *Lévy processes* which are outside the class of processes being considered in this dissertation [48]. It may, however, be possible to extend the notion of Fliess operators for this type of input processes in future work. □

IV.2 THE SHUFFLE ALGEBRA

Given that Stratonovich integration satisfies the integration by parts formula, the $\mathbb{R}$-vector space $\mathbb{R}^\ell\langle\langle XY\rangle\rangle$ together with the shuffle product forms an $\mathbb{R}$-algebra. This algebra is called the *shuffle algebra* of $\mathbb{R}^\ell\langle\langle XY\rangle\rangle$, and it can be seen as a generalization of the shuffle algebra defined on $\mathbb{R}^\ell\langle\langle X\rangle\rangle$. In Chapter V, this algebra will play a central role in the definition of the composition product over XY^*.

Definition IV.2.1. The **addition** of $c, d \in \mathbb{R}^\ell\langle\langle XY\rangle\rangle$ is defined as

$$c + d = \sum_{\eta \in XY^*} ((c,\eta) + (d,\eta))\eta.$$

Definition IV.2.2. The **scalar product** of $r \in \mathbb{R}$ and $c \in \mathbb{R}^\ell\langle\langle XY \rangle\rangle$ is defined as

$$rc = \sum_{\eta \in XY^*} (rc, \eta)\eta = \sum_{\eta \in XY^*} r(c, \eta)\eta.$$

These operations are direct extensions of the operations defined on $\mathbb{R}\langle\langle X \rangle\rangle$.

Definition IV.2.3. The **shuffle product** on XY^* is recursively defined for $\eta = q_{j1}^{l_1}\eta'$, $\xi = q_{j2}^{l_2}\xi'$, $\eta', \xi' \in XY^*$ and $q_{j1}^{l_1}, q_{j2}^{l_2} \in XY$ as

$$\eta \sqcup\!\sqcup \xi = q_{j1}^{l_1}[\eta' \sqcup\!\sqcup \xi] + q_{j2}^{l_2}[\eta \sqcup\!\sqcup \xi'],$$

where $\emptyset \sqcup\!\sqcup \emptyset = \emptyset$ and $\xi \sqcup\!\sqcup \emptyset = \emptyset \sqcup\!\sqcup \xi = \xi$.

The definition is extended to any $c, d \in \mathbb{R}^\ell\langle\langle XY \rangle\rangle$ by

$$c \sqcup\!\sqcup d = \sum_{\eta, \xi \in XY^*} [(c, \eta)(d, \xi)]\eta \sqcup\!\sqcup \xi.$$

It is clear that this extension of the shuffle product to $\mathbb{R}\langle\langle XY \rangle\rangle$ behaves exactly as the one for $\mathbb{R}\langle\langle X \rangle\rangle$. The iterated integral defined in (IV.1.2) and (IV.1.3) can be extended linearly to polynomials as

$$E_p[w](t) = \sum_{\eta \in \mathrm{supp}(p)} (p, \eta) E_\eta[w](t),$$

where $p \in \mathbb{R}\langle XY \rangle$. The set of all such integrals forms a vector space denoted as $\mathcal{E}(\mathbb{R}\langle XY \rangle)$.

Lemma IV.2.1. *Let $w \in \mathcal{UV}^m[0, T]$. Then*

$$E_\eta[w](t) E_\xi[w](t) = E_{\eta \sqcup\!\sqcup \xi}[w](t) \tag{IV.2.1}$$

for $\eta, \xi \in XY^*$. In addition, the set $\mathcal{E}(\mathbb{R}\langle XY \rangle)$ forms an $\mathbb{R}$-algebra with product $\sqcup\!\sqcup$ on XY^* and identity element $E_\emptyset = 1$.

Proof: Consider the iterated integrals $E_\eta[w](t), E_\xi[w](t) \in \mathcal{E}(\mathbb{R}\langle XY\rangle)$ for $\eta = q_{j1}^{l_1}\eta'$, $\xi = q_{j2}^{l_2}\xi'$, $\eta', \xi' \in XY^*$ and $q_{j1}^{l_1}, q_{j2}^{l_2} \in XY$. From the Stratonovich integration by parts rule, their product can be computed as

$$\begin{aligned}
E_\eta[w](t)E_\xi[w](t) &= \int_0^t w_{q_{j1}^{l_1}}(\tau)\left[E_{\eta'}[w](\tau)E_{q_{j2}^{l_2}\xi'}[w](\tau)\right] dq_{j1}^{l_1}(\tau) \\
&\quad + \oint_0^t w_{q_{j2}^{l_2}}(\tau)\left[E_{q_{j1}^{l_1}\eta'}[w](\tau)E_{\xi'}[w](\tau)\right] dq_{j2}^{l_2}(\tau) \\
&= E_{q_{j1}^{l_1}}\left[E_{\eta'}[w](\tau)E_{q_{j2}^{l_2}\xi'}[w](\tau)\right](t) \\
&\quad + E_{q_{j2}^{l_2}}(\tau)\left[E_{q_{j1}^{l_1}\eta'}[w](\tau)E_{\xi'}[w](\tau, t_0)\right](t) \\
&= E_{q_{j1}^{l_1}\left(\eta' \sqcup\!\sqcup\, q_{j2}^{l_2}\xi'\right) + q_{j2}^{l_2}\left(q_{j1}^{l_1}\eta' \sqcup\!\sqcup\, \xi'\right)}[w](t) \\
&= E_{q_{j1}^{l_1}\eta' \sqcup\!\sqcup\, q_{j2}^{l_2}\xi'}[w](t) \\
&= E_{\eta \sqcup\!\sqcup\, \xi}[w](t).
\end{aligned}$$

Now, since $\eta \sqcup\!\sqcup\, \xi \in \mathbb{R}\langle XY\rangle$, the product of two iterated integrals is an element of $\mathcal{E}(\mathbb{R}\langle XY\rangle)$. Hence, $\mathcal{E}(\mathbb{R}\langle XY\rangle)$ forms an $\mathbb{R}$-algebra with product $\sqcup\!\sqcup$ on XY^* and identity element $E_\emptyset = 1$. ■

The next lemma show how to utilize the shuffle product to decompose the characteristic series of an arbitrary alphabet.

Lemma IV.2.2. *Let $Z = \{z_0, \ldots, z_m\}$ be an arbitrary alphabet. The characteristic series of the language Z^* can be written in terms of the shuffle product as*

$$\boldsymbol{Z} \triangleq \sum_{\eta \in Z^*} \eta = \sum_{k=0}^{\infty} \sum_{\|\alpha\|=k} z_0^{\alpha_0} \sqcup\!\sqcup\, z_1^{\alpha_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup\, z_m^{\alpha_m}. \tag{IV.2.2}$$

Proof: Recall that $\sum_{\|\alpha\|=k}$ denotes the summation over all vectors $\boldsymbol{\alpha}$ such that $\alpha_0 + \alpha_1 + \cdots + \alpha_m = k$. For simplicity, consider the case when $Z = \{z_0, z_1\}$. It is sufficient to prove (IV.2.2) for the characteristic series of the language $Z^k = \{\eta \in$

$Z^* \ : |\eta| = k\}$. The proof is done by induction over k. Consider $k = 1$. Clearly,

$$\boldsymbol{Z} = z_0 + q_1 = q_0 \sqcup\!\sqcup \emptyset + \emptyset \sqcup\!\sqcup z_1 = \sum_{\|\alpha\|=1} z_0^{\alpha_0} \sqcup\!\sqcup z_1^{\alpha_1}.$$

For $k = 2$,

$$\boldsymbol{Z^2} = z_0^2 + z_0 z_1 + z_1 z_0 + z_1^2 = z_0^2 \sqcup\!\sqcup \emptyset + z_0 \sqcup\!\sqcup z_1 + \emptyset \sqcup\!\sqcup z_1^2 = \sum_{\|\alpha\|=2} z_0^{\alpha_0} \sqcup\!\sqcup z_1^{\alpha_1}.$$

Assume identity (IV.2.2) holds up to k. For $k+1$ then

$$\begin{aligned}
\boldsymbol{Z^{k+1}} &= \sum_{\eta \in Z^{k+1}} \eta = z_0 \sum_{\eta \in Z^k} \eta + z_1 \sum_{\eta \in Z^k} \eta \\
&= z_0 \sum_{\|\alpha\|=k} z_0^{\alpha_0} \sqcup\!\sqcup z_1^{\alpha_1} + z_1 \sum_{\|\alpha\|=k} z_0^{\alpha_0} \sqcup\!\sqcup z_1^{\alpha_1} \\
&= \sum_{i=0}^{k} z_0(z_0^i \sqcup\!\sqcup z_1^{k-i}) + \sum_{i=0}^{k} z_1(z_0^i \sqcup\!\sqcup z_1^{k-i}) \\
&= \sum_{i=0}^{k-1} z_0(z_0^i \sqcup\!\sqcup z_1^{k-i}) + \sum_{i=1}^{k} z_1(z_0^i \sqcup\!\sqcup z_1^{k-i}) + z_0(z_0^k \sqcup\!\sqcup \emptyset) + z_1(\emptyset \sqcup\!\sqcup z_1^k) \\
&= \sum_{i=0}^{k-1} z_0(z_0^i \sqcup\!\sqcup z_1^{k-i}) + \sum_{i=0}^{k-1} z_1(z_0^{i+1} \sqcup\!\sqcup z_1^{k-i-1}) + (z_0^{k+1} \sqcup\!\sqcup \emptyset) + (\emptyset \sqcup\!\sqcup z_1^{k+1}) \\
&= \sum_{i=0}^{k-1} \left(z_0(z_0^i \sqcup\!\sqcup z_1^{k-i}) + z_1(z_0^{i+1} \sqcup\!\sqcup z_1^{k-i-1})\right) + (z_0^{k+1} \sqcup\!\sqcup \emptyset) + (\emptyset \sqcup\!\sqcup z_1^{k+1}) \\
&= \sum_{i=0}^{k-1} (z_0^{i+1} \sqcup\!\sqcup z_1^{k-i}) + (z_0^{k+1} \sqcup\!\sqcup \emptyset) + (\emptyset \sqcup\!\sqcup z_1^{k+1}) \\
&= \sum_{i=1}^{k} (z_0^i \sqcup\!\sqcup z_1^{k-(i-1)}) + (z_0^{k+1} \sqcup\!\sqcup \emptyset) + (\emptyset \sqcup\!\sqcup z_1^{k+1}) \\
&= \sum_{i=0}^{k+1} (z_0^i \sqcup\!\sqcup z_1^{(k+1)-i}) = \sum_{\|\alpha\|=k+1} z_0^{\alpha_0} \sqcup\!\sqcup z_1^{\alpha_1}.
\end{aligned}$$

Therefore, (IV.2.2) is proved. ■

For any $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^{m+1}$ define the polynomials $p_{\boldsymbol{\alpha}} = x_0^{\alpha_0} \sqcup\!\sqcup \cdots \sqcup\!\sqcup x_m^{\alpha_m}$ and $p_{\boldsymbol{\beta}} = y_0^{\beta_0} \sqcup\!\sqcup \cdots \sqcup\!\sqcup y_m^{\beta_m}$, respectively.

Corollary IV.2.1. *The characteristic series, $\boldsymbol{X^kY^n}$, of the language X^kY^n can be written in terms of the shuffle product as*

$$\begin{aligned}\boldsymbol{X^kY^n} &\triangleq \sum_{\eta\in X^kY^n} \eta \\ &= \sum_{\|\alpha\|=k,\|\beta\|=n} p_{\boldsymbol{\alpha}} \sqcup\!\sqcup\, p_{\boldsymbol{\beta}}.\end{aligned}$$

Proof: In Lemma IV.2.2, let $Z = XY = \{x_0, x_1, \ldots, x_m, y_0, y_1, \ldots, y_m\}$. Then by (IV.2.2)

$$\begin{aligned}\boldsymbol{XY} = \sum_{\eta\in XY^*} \eta &= \sum_{j=0}^{\infty} \sum_{\|\alpha\|+\|\beta\|=j} x_0^{\alpha_0} \sqcup\!\sqcup\, x_1^{\alpha_1} \sqcup\!\sqcup \ldots \sqcup\!\sqcup\, x_m^{\alpha_m} \sqcup\!\sqcup\, y_0^{\beta_0} \sqcup\!\sqcup\, y_1^{\beta_1} \sqcup\!\sqcup \ldots \sqcup\!\sqcup\, y_m^{\beta_m} \\ &= \sum_{j=0}^{\infty} \sum_{\|\alpha\|+\|\beta\|=j} p_{\boldsymbol{\alpha}} \sqcup\!\sqcup\, p_{\boldsymbol{\beta}} \\ &= \sum_{j=0}^{\infty} \sum_{k=0}^{j} \sum_{\|\alpha\|=k,\|\beta\|=j-k} p_{\boldsymbol{\alpha}} \sqcup\!\sqcup\, p_{\boldsymbol{\beta}}.\end{aligned}$$

Since $X^kY^{j-k} = \{\eta \in XY^* : |\eta|_X = k, |\eta|_Y = j-k\}$ and $\boldsymbol{XY} = \sum_{j=0}^{\infty} \sum_{k=0}^{j} \sum_{\eta\in X^kY^{j-k}} \eta$, it follows directly that

$$\boldsymbol{X^kY^{j-k}} = \sum_{k=0}^{j} \sum_{\|\alpha\|=k,\|\beta\|=j-k} p_{\boldsymbol{\alpha}} \sqcup\!\sqcup\, p_{\boldsymbol{\beta}}.$$

■

The next example shows how Lemma IV.2.1 leads to an improvement of the proof for the local convergence of Fliess operators driven by deterministic inputs. This is illustrated using a new grouping of the series (I.1.2) in terms of the shuffle product. This grouping will ultimately provide an improved estimate for the radius of convergence over what appears in the literature [29].

Example IV.2.1. [10] In the deterministic case, i.e., when the alphabet Y is empty, the drift inputs are deterministic and $\max\{\|u\|_{L_1}, T\} \leq R$ on $[0, T]$. One can show,

then, that

$$|F_{p_\alpha}[u](t)| \leq \prod_{i=0}^{m} \frac{U_i^{\alpha_i}(t)}{\alpha_i!} \leq \frac{R^k}{\alpha!}. \tag{IV.2.3}$$

Specifically, it is clear that

$$|F_{p_\alpha}[u](t)| = \left|F_{x_0^{\alpha_0} \sqcup\!\sqcup\, x_1^{\alpha_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup\, x_m^{\alpha_m}}[u](t)\right| = \prod_{i=0}^{m} \left|F_{x_j^{\alpha_j}}[u](t)\right|.$$

By induction, it is next shown that

$$\left|F_{x_j^{\alpha_j}}[u](t)\right| \leq \frac{U_j^{\alpha_j}(t)}{\alpha_j!}. \tag{IV.2.4}$$

When $\alpha_j = 0$ or $\alpha_j = 1$, the claim is trivially true. If (IV.2.4) holds up to some fixed integer $\alpha_j \geq 0$ then

$$\begin{aligned}
\left|F_{x_j^{\alpha_j+1}}[u](t)\right| &\leq \int_0^t |u_j(\tau)| \left|F_{x_j^{\alpha_j}}[u](\tau)\right| \, d\tau \\
&\leq \int_0^t |u_j(\tau)| \frac{U_j^{\alpha_j}(t)}{\alpha_j!} \, d\tau \\
&= \frac{U_j^{\alpha_j+1}(t)}{(\alpha_j+1)!}.
\end{aligned}$$

Furthermore, it is also easy to verify by induction that

$$|E_\eta[u](t)| \leq E_\eta[\bar{u}](t), \quad 0 \leq t \leq T, \tag{IV.2.5}$$

where $\bar{u} \in L_1^m[0,T]$ has components $\bar{u}_j = |u_j|$, $j = 0, 1 \ldots, m$. Suppose (IV.2.5) holds for words up to length k. Then for any $x_j \in X$,

$$\begin{aligned}
\left|E_{x_j\eta}[u](t)\right| &\leq \int_0^t |u_j(\tau)| \, |E_\eta[u](\tau)| \, d\tau \\
&\leq \int_0^t \bar{u}_j(\tau) E_\eta[\bar{u}](\tau) \, d\tau \\
&= E_{x_j\eta}[\bar{u}](t).
\end{aligned}$$

Thus, the claim holds for all $\eta \in X^*$. Now fix $T > 0$. Pick any $u \in L_1^m[0,T]$ and let $R = \max\{\|u\|_{L_1}, T\}$. From Lemma IV.2.1, observe that

$$\begin{aligned}
\sum_{\eta \in X^*} |(c,\eta) E_\eta[u](t)| &\leq \sum_{k=0}^{\infty} \sum_{\eta \in X^k} |(c,\eta)|\, E_\eta[\bar{u}](t) \\
&\leq \sum_{k=0}^{\infty} K M^k k! \sum_{\|\alpha\|=k} F_{p_\alpha}[\bar{u}](t) \\
&\leq \sum_{k=0}^{\infty} K M^k k! \sum_{\|\alpha\|=k} \frac{R^k}{\alpha!} \\
&\leq \sum_{k=0}^{\infty} K (MR)^k \sum_{\|\alpha\|=k} \frac{k!}{\alpha!} \\
&= \sum_{k=0}^{\infty} K (MR(m+1))^k.
\end{aligned}$$

Therefore, the series defining F_c converges absolutely and uniformly on an open ball in $L_1[0,T]$ of radius $R < 1/M(m+1)$. In [29], the more conservative radius of convergence $1/M(m+1)^2$ was proved. $\square$

For fixed $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^{m+1}$, $w \in \mathcal{UV}^m[0,T]$ and $t \geq 0$, define the following sum of iterated integrals

$$\mathbf{S}_{\boldsymbol{\alpha},\boldsymbol{\beta}}[w](t) \triangleq F_{p_{\boldsymbol{\alpha}} \sqcup\!\sqcup\, p_{\boldsymbol{\beta}}}[w](t) = F_{p_{\boldsymbol{\alpha}}}[w](t) F_{p_{\boldsymbol{\beta}}}[w](t). \tag{IV.2.6}$$

The importance of (IV.2.6) comes from the fact that, using the commutativity of the shuffle product and equation (IV.2.1), the Lebesgue integrals and Stratonovich integrals can be completely separated. Thus, an L_2 upper bound for $\mathbf{S}_{\boldsymbol{\alpha},\boldsymbol{\beta}}[w](t)$ can be obtained by calculating individual L_2 upper bounds for the random vectors $F_{p_{\boldsymbol{\alpha}}}[w](t)$ and $F_{p_{\boldsymbol{\beta}}}[w](t)$. Then from the independence assumptions in Definition IV.1.1,

$$\|\mathbf{S}_{\boldsymbol{\alpha},\boldsymbol{\beta}}[w](t)\|_2^2 = \|F_{p_{\boldsymbol{\alpha}}}[w](t)\|_2^2 \left\|F_{p_{\boldsymbol{\beta}}}[w](t)\right\|_2^2. \tag{IV.2.7}$$

The following lemma is needed for calculating a bound for $\|F_{p_{\boldsymbol{\alpha}}}[w](t)\|_2^2$.

Lemma IV.2.3. *Let u be the drift input of $w \in \mathcal{UV}^m[0,T]$. Then for $\boldsymbol{\alpha} = (\alpha_0, \ldots, \alpha_m) \in \mathbb{N}^{m+1}$ and any real numbers $t > s \geq 0$*

$$\mathbf{E}\left[\prod_{i=0}^{m} (U_i(t) - U_i(s))^{\alpha_i}\right] \leq \prod_{i=0}^{m} \bar{U}_i^{\alpha_i}(t), \tag{IV.2.8}$$

where $\bar{U}_i(t) \triangleq \int_0^t \mathbf{E}\left[|u_i(s)|\right] ds$.

Proof: Let $k = \sum_{i=0}^{m} \alpha_i$. The identity is satisfied trivially if $k = 0$. If $k = 1$ then

$$\mathbf{E}\left[U_i(t) - U_i(s)\right] \leq \int_s^t \mathbf{E}\left[|u_i(\tau)|\right] d\tau \leq \bar{U}_i(t).$$

Now suppose (IV.2.8) holds for $k - 1 \geq 0$. Clearly u has independent increments. Using Fubini's theorem, it follows that

$$\begin{aligned}
&\mathbf{E}\left[\prod_{i=0}^{m} (U_i(t) - U_i(s))^{\alpha_i}\right] \\
&\leq \mathbf{E}\left[\left|\int_s^t \sum_{i=0}^{m} \alpha_i u_i(r) \left(U_i(t) - U_i(r)\right)^{\alpha_i - 1} \prod_{\substack{l=0 \\ l \neq i}}^{m} \left(U_l(t) - U_l(r)\right)^{\alpha_l} dr\right|\right] \\
&\leq \int_s^t \sum_{i=0}^{m} \alpha_i \, \mathbf{E}\left[|u_i(r)|\right] \left|\mathbf{E}\left[\left(U_i(t) - U_i(r)\right)^{\alpha_i - 1} \prod_{\substack{l=0 \\ l \neq i}}^{m} \left(U_l(t) - U_l(r)\right)^{\alpha_l}\right]\right| dr \\
&\leq \int_s^t \sum_{i=0}^{m} \alpha_i \, \mathbf{E}\left[|u_i(r)|\right] \left(\bar{U}_i(r)\right)^{\alpha_i - 1} \prod_{\substack{l=0 \\ l \neq i}}^{m} \left(\bar{U}_l(r)\right)^{\alpha_l} dr \\
&= \prod_{i=0}^{m} \left(\bar{U}_i(t)\right)^{\alpha_i}.
\end{aligned}$$

Thus, the inequality in question is proved. ■

Using the above lemma, (IV.2.3), and Definition IV.1.1, the L_2-norm of $F_{p_\alpha}[w](t)$ is

$$\|F_{p_\alpha}[w](t)\|_2^2 \leq \prod_{i=0}^{m} \frac{\bar{U}_i^{2\alpha_i}(t)}{(\alpha_i!)^2} \leq \frac{R^{2k}}{(\alpha!)^2}. \tag{IV.2.9}$$

In Theorem IV.1.5, it was shown that $\sum_{\eta\in XY^*}|(c,\eta)|\,\|E_\eta[w](t)\|_2$ is finite when c is globally convergent. Then, as in the deterministic case, the next example shows how the shuffle product on XY^* can be used to group iterated integrals, and thus, obtain conditional global convergence of a Fliess operator driven by an L_2-Itô process.

Example IV.2.2. [10] Pick a $t\in[0,T]$ and any $w\in\mathcal{UV}^m[0,T]$. Let $R=\max\{\|u\|_{L_1},\|v\|_{L_2},\|v_0\|_{L_2},\|v\|_{L_4}\}$. Define

$$a_{k,n}(t)=KM^{n+k}\sum_{\|\alpha\|=k,\|\beta\|=n}\mathbf{S}_{\alpha,\beta}[w](t).$$

Applying the multinomial theorem, identities (IV.2.7) and (IV.2.9), and Theorem IV.1.4

$$\|a_{k,n}(t)\|_2\leq KM^{n+k}\sum_{\|\alpha\|=k,\|\beta\|=n}\frac{(\sqrt{2R}(\sqrt{t}+2))^{2n}}{(\beta!)^{\frac{1}{4}}}\frac{R^k n!}{\alpha!\beta!}.$$

If it is assumed that $R\geq 1$ and $R'\triangleq 4R(R+4)$, then

$$\begin{aligned}\|a_{k,n}(t)\|_2 &\leq K(MR')^{k+n}\frac{(m+1)^k}{k!(n!)^{\frac{1}{4}}}\sum_{\|\beta\|=n}\frac{(n!)^{\frac{5}{4}}}{(\beta!)^{\frac{5}{4}}}\\ &\leq K(MR')^{k+n}\frac{(m+1)^k}{k!(n!)^{\frac{1}{4}}}\left(\sum_{\|\beta\|=n}\frac{n!}{\beta!}\right)^2\\ &= K(MR')^{k+n}\frac{(m+1)^k}{k!(n!)^{\frac{1}{4}}}(m+1)^{2n}\\ &\leq \frac{K(MR'(m+1)^2)^{k+n}}{k!(n!)^{\frac{1}{4}}}.\end{aligned}$$

Immediately from Lemma IV.2.1, for $N_2>N_1\in\mathbb{N}$, equation (IV.1.23) can be re-derived as

$$\begin{aligned}\left\|\sum_{j=N_1}^{N_2}\sum_{k=0}^{j}\sum_{\eta\in X^kY^{j-k}}(c,\eta)E_\eta[w](t)\right\|_2 &< \sum_{j=N_1}^{N_2}\sum_{k=0}^{j}\|a_{k,j-k}(t)\|_2\\ &\leq K\sum_{j=0}^{\infty}\sum_{k=0}^{j}\frac{(MR'(m+1)^2)^k(MR'(m+1)^2)^{j-k}}{k!((j-k)!)^{\frac{1}{4}}}\end{aligned}$$

$$
\begin{aligned}
&= K\sum_{k=0}^{\infty}\frac{(MR'(m+1)^2)^k}{k!}\sum_{n=0}^{\infty}\frac{(MR'(m+1)^2)^n}{(n!)^{\frac{1}{4}}} \\
&= K'\sum_{n=0}^{\infty}\frac{(MR'(m+1)^2)^n}{(n!)^{\frac{1}{4}}} \\
&= K'\sum_{n=0}^{\infty}\frac{(M'')^n}{(n!)^{\frac{1}{4}}},
\end{aligned}
$$

where $M'' \triangleq (MR'(m+1)^2)$ and $K' \triangleq Ke^{(MR'(m+1)^2)}$. Note that $\frac{(M'')^n}{(n!)^{\frac{1}{4}}}$ is the n-th term of an absolutely convergent series. Hence, one reaches the same conclusion as in Theorem IV.1.5. $\square$

Even though in the example above convergence is achieved, it is only conditional convergence, i.e., the ordered sum $\sum_{j=0}^{\infty}\sum_{k=0}^{j}\left\|\sum_{\eta\in X^kY^{j-k}}(c,\eta)E_\eta[w](t)\right\|_2$ is finite. Thus, this grouping works perfectly in the deterministic case, but in the stochastic situation it is not enough to assured global convergence. However, this grouping will be very useful in the next section when local convergence of Fliess operators is taken in account.

IV.3 LOCAL CONVERGENCE

Although global convergence is a desirable property for the generating series of a Fliess operator, many systems of interest are not of this type. So in this section a stochastic notion of local convergence is introduced using the concept of a stopping time. Then a corresponding sufficient condition for local convergence is developed. But first an example is given to motivate the approach taken.

Example IV.3.1. Consider the system

$$
\begin{aligned}
dz(t) &= Mz^2(t)\,dW(t),\ z(0)=1, \\
y(t) &= Kz(t).
\end{aligned}
\tag{IV.3.1}
$$

The generating series for system (IV.3.1) is $(c, y_0^k) = KM^k k!$, $k \geq 0$, and 0 otherwise. The corresponding output is

$$y(t) = F_c[0](t) = \sum_{k=0}^{\infty} KM^k k! \oint_0^t \cdots \oint_0^{t_2} dW(t_1) \cdots dW(t_k).$$

By formula (IV.1.25),

$$y(t) = F_c[0](t) = \sum_{k=0}^{\infty} KM^k W^k(t).$$

At first glance, y appears not to be convergent since Theorem IV.1.5 does not apply. However, if

$$\tau_R = \inf\{t : |MW(t)| = R\} \tag{IV.3.2}$$

and $R < 1$, then $y(t)$ will be an absolutely convergent series with limit $Kz(t) = \frac{K}{1-MW(t)}$ for any $t \leq \tau_R$. Now, let $K = M = 1$. In [38], the theoretical probability distribution function of τ_R is given by

$$f(\tau_R) = \frac{|R|}{\sqrt{2\pi}} \tau_R^{-\frac{3}{2}} e^{-\frac{R^2}{2\tau_R}}, \tag{IV.3.3}$$

which is the distribution function known as the inverse gamma distribution with parameters $\alpha = \frac{1}{2}$ and $\beta = \frac{R^2}{2}$. In Figure 5, a Monte Carlo simulation of the probability density function of τ_R is presented for $R \approx 1^-$ (approximation to 1 from the left.) $\square$

Two important observations concerning Example IV.3.1 are that the condition given in Theorem IV.1.5 is only a sufficient condition, and the convergence time interval has a random nature, i.e.,

$$[0, \tau] = \{0 \leq t \leq \tau(\omega) : (\tau, \omega) \in [0, \infty) \times \Omega\}.$$

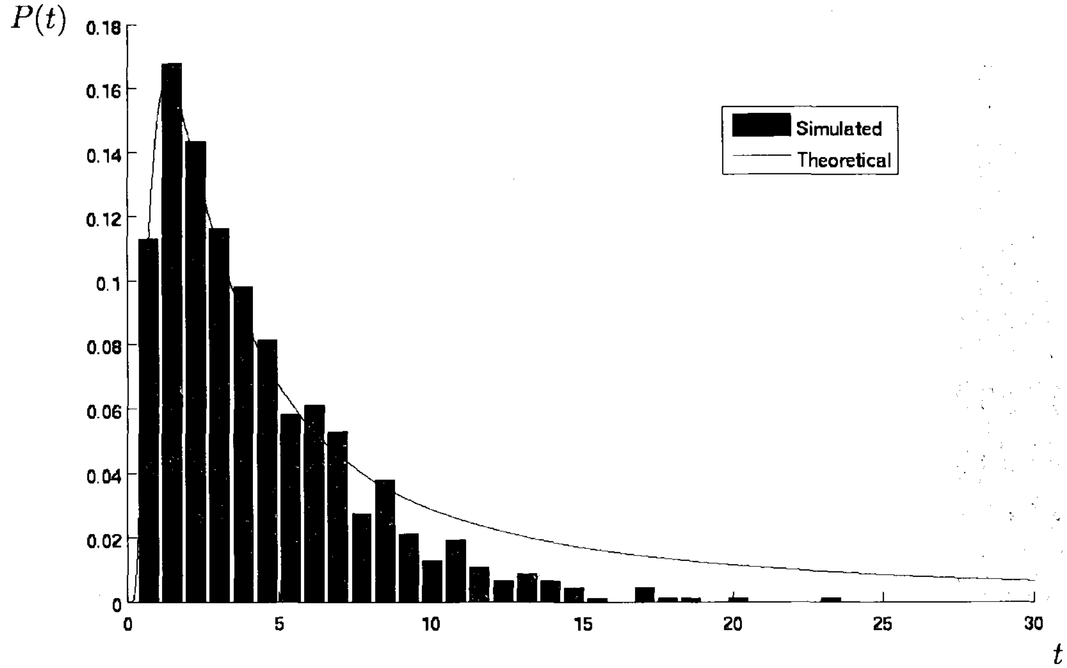

Fig. 5: Estimated probability density function τ_R.

Moreover, the direct solution of (IV.3.1) confirms this since it can be solved by separation of variables, i.e.,

$$\int_0^t \frac{dz(s)}{z^2(s)} = \oint_0^t M\, dW(s) = MW(t).$$

Then,

$$y(t) = Kz(t) = \frac{K}{1 - MW(t)}$$

for any t such that $M_2W(t) < 1$, or for any $t < \tau = \inf\{t \in [0, \infty) : MW(t) < 1\}$. In addition, observe that the stopping time

$$\tau_R \triangleq \min_{i \in \{0,1,\cdots,m\}} \inf \left\{ t \in \mathcal{T} \,:\, \left| \oint_0^t v_i(s)\, dW(s) \right| = R_i \right\} \tag{IV.3.4}$$

will play an important role in the derivation of a the local convergence condition for the case when the generating series of $F_c[w]$ is locally convergent. The next

definition describe the set of admissible inputs for Fliess operators corresponding to locally convergent formal power series.

Definition IV.3.1. Let $X(t) = \oint_0^t v(s)\, dW(s)$, where v is an m-dimensional L_2-Itô process. The set $\mathcal{UV}^m[0, \tau_R]$ is defined as the set of processes $w \in \mathcal{UV}^m[0, T]$ stopped at τ_R defined in (IV.3.4).

The next theorem presents the main result in this section.

Theorem IV.3.1. *Suppose that for a series $c \in \mathbb{R}^\ell\langle\langle XY\rangle\rangle$, there exist real numbers $K > 0$ and $M > 0$ such that*

$$|(c, \eta)| \leq KM^{|\eta|}\, |\eta|!, \quad \forall \eta \in XY^*.$$

Then for any random process $w \in \mathcal{UV}^m[0, \tau_R]$ with τ_R defined as in (IV.3.4), *the series*

$$F_c[w](t) = \sum_{j=0}^{\infty} \sum_{k=0}^{j} \sum_{\eta \in X^k Y^{j-k}} (c, \eta) E_\eta[w](t) \tag{IV.3.5}$$

converges in the mean square sense to the random vector $y(t) = F_c[w](t)$, $t \in [0, \tau_R]$.

Proof: Without loss of generality it is assumed that $\ell = 1$. Pick any $w \in \mathcal{UV}^m[0, \tau_R]$ and a $t \in [0, \tau_R]$. Observe that τ_R^i is the first time the process $X(t) = \oint_0^t v_i(s)\, dW(s)$ hits the barrier $(-R, R)$. Since $X(t)$ is a well-defined Itô process, Theorem II.3.8 and the fact that the absolute value function is a continuous function ensures that one can always choose, without loss of generality, a continuous version of the process X. Then, by Theorem II.4.6, the random variable τ_R^i is a strictly positive stopping time. Thus, the stopped process $X^{\tau_R^i}$ is a well-defined L_2-bounded, a.s. continuous and

adapted L_2-Itô process. Now, using integration by parts and property (IV.2.1)

$$\begin{aligned} F_{p_\beta}[w](t) = F_{y_0^{\beta_0} ⧢ \cdots ⧢ y_m^{\beta_m}}[w](t) &= F_{y_0^{\beta_0}}[w](t) \cdots F_{y_m^{\beta_m}}[w](t) \\ &= \prod_{i=0}^{m} \frac{\left(\oint_0^t v_i(s)\, dW(s) \right)^{\beta_i}}{\beta_i!}. \end{aligned}$$

The L_2-norm for $F_{p_\beta}[w](t)$ truncated at the stopping time τ_R is

$$\begin{aligned} \left\| F_{p_\beta}[w](t \wedge \tau_R) \right\|_2^2 &= \frac{1}{(\beta!)^2} \mathbf{E} \left[\prod_{i=0}^{m} \left(\left| \oint_0^{t \wedge \tau_R} v_i(s)\, dW(s) \right| \right)^{2\beta_i} \right] \\ &\leq \frac{R^{2n}}{(\beta!)^2}. \end{aligned}$$

Define

$$a_{k,n}(t) \triangleq K M^{k+n} (k+n)! \sum_{\|\alpha\|=k, \|\beta\|=n} \mathbf{S}_{\boldsymbol{\alpha},\boldsymbol{\beta}}[w](t).$$

Let $R' = \max\{\|u\|_{L_1}, \|u_0\|_{L_1}, R\}$. Using equations (IV.2.7) and (IV.2.9), and the multinomial theorem, the following bound is obtained

$$\begin{aligned} \|a_{k,n}(t \wedge \tau_R)\|_2 &\leq K M^{k+n} (k+n)! \sum_{\|\alpha\|=k, \|\beta\|=n} \frac{R'^k}{\alpha!} \frac{R'^n}{\beta!} \\ &= K (M R')^{k+n} (k+n)! \frac{(m+1)^k}{k!} \frac{(m+1)^n}{n!} \\ &= K (M R'(m+1))^{k+n} \binom{k+n}{n}. \end{aligned}$$

To show that (IV.3.5) is mean square convergent, it is sufficient to show that it is a Cauchy series. From the triangle inequality,

$$\left\| \sum_{j=N_1}^{N_2} \sum_{k=0}^{j} \sum_{\eta \in X^k Y^{j-k}} (c, \eta) E_\eta[w](t) \right\|_2 \leq \sum_{j=N_1}^{N_2} \sum_{k=0}^{j} \|a_{k,j-k}(t)\|_2$$

for $N_2 > N_1 \in \mathbb{N}$. Now for any $\epsilon > 0$ there exists an $N > 0$ such that

$$\begin{aligned}
\sum_{j=N_1}^{N_2}\sum_{k=0}^{j}\|a_{k,j-k}(t\wedge\tau_R)\|_2 &\leq \sum_{j=N_1}^{N_2}\sum_{k=0}^{j}K(MR'(m+1))^j\frac{j!}{k!(j-k)!} \\
&= \sum_{j=N_1}^{N_2}K(MR'(m+1))^j\sum_{k=0}^{j}\frac{j!}{k!(j-k)!} \\
&= \sum_{j=N_1}^{N_2}K(2MR'(m+1))^j < \epsilon \qquad \text{(IV.3.6)}
\end{aligned}$$

for $2MR'(m+1) < 1$ and $N_2 > N_1 > N$. Note that since $2MR'(m+1) < 1$ then the series on the right hand side of (IV.3.6) is absolutely convergent. Hence, the series $\sum_{j=N_1}^{N_2}\sum_{k=0}^{j}\sum_{\eta\in X^kY^{j-k}}(c,\eta)E_\eta[w](t\wedge\tau_R)$ is Cauchy, and the theorem is proved. ■

Note in (IV.3.5) that there is an implied order of the summation over XY^*. Thus, the current proof for the convergence of F_c is strictly speaking addressing *conditional* convergence.

Definition IV.3.2. [17] Let $\boldsymbol{\alpha},\boldsymbol{\beta} \in \mathbb{N}^{m+1}$ and define the language $\boldsymbol{L}_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \left\{\eta \in XY^*, |\eta|_{x_i} = \alpha_i, |\eta|_{y_i} = \beta_i, i = 0,1,\ldots,m\right\}$. A series $c \in \mathbb{R}^\ell\langle\langle XY\rangle\rangle$ is called **exchangeable** if all the words in $\boldsymbol{L}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ have the same coefficient for all $\boldsymbol{\alpha},\boldsymbol{\beta} \in \mathbb{N}^{m+1}$.

Corollary IV.3.1. *Let $c \in \mathbb{R}^\ell\langle\langle XY\rangle\rangle$ be an exchangeable and locally convergent series. Then for an arbitrary $w \in \mathcal{UV}^m[0,T]$, there exists a stopping time τ_R for $R > 0$ such that the Fliess operator associated with c converges to a well-defined random vector independently of the order in* (IV.3.5).

Proof: Since c is exchangeable, one can group all the iterated integrals associated with words having the same $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, i.e.,

$$F_c[w](t) = \sum_{j=0}^{\infty}\sum_{k=0}^{j}\sum_{\eta\in X^kY^{j-k}}(c,\eta)E_\eta[w](t)$$

$$= \sum_{j=0}^{\infty}\sum_{k=0}^{j}\sum_{\|\alpha\|=k,\|\beta\|=j-k} c_{\boldsymbol{\alpha},\boldsymbol{\beta}} S_{\boldsymbol{\alpha},\boldsymbol{\beta}}[w](t),$$

where for fixed $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^{m+1}$ $c_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ denotes the coefficient for all $\eta \in \boldsymbol{L}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$. Note, then, that formula (IV.2.4) allows to write

$$F_c[w](t) = \sum_{j=0}^{\infty}\sum_{k=0}^{j}\sum_{\|\alpha\|=k,\|\beta\|=j-k} c_{\boldsymbol{\alpha},\boldsymbol{\beta}} \prod_{i=0}^{m+1} \frac{E_{x_i}^{\alpha_i}[w](t)}{\alpha_i!}\frac{E_{y_i}^{\beta_i}[w](t)}{\beta_i!}, \tag{IV.3.7}$$

which is independent of the order indicated in (IV.3.5). Therefore, from Theorem IV.3.1, the infinite series (IV.3.7) is Cauchy, and thus, the series

$$\|F_c[w](t)\|_2 = \sum_{j=0}^{\infty}\sum_{k=0}^{j}\sum_{\|\alpha\|=k,\|\beta\|=j-k} |c_{\boldsymbol{\alpha},\boldsymbol{\beta}}| \prod_{i=0}^{m+1} \left\| \frac{E_{x_i}^{\alpha_i}[w](t)}{\alpha_i!}\frac{E_{y_i}^{\beta_i}[w](t)}{\beta_i!} \right\|_2$$

is finite. This completes the proof. ■

Example IV.3.2. Consider the following nonlinear system

$$\frac{dz(t)}{dt} = z^2\dot{w},$$

where $\dot{w}$ is understood as in (IV.1.27) when $w \in \mathcal{UV}^m[0,T]$. In integral form

$$z(t) = \int_0^t u(s)z^2(s)\,ds + \oint_0^t v(t)z^2(s)\,dW(s). \tag{IV.3.8}$$

Given that the shuffle product represents the product of iterated integrals (see Lemma IV.2.1), the n-th power of $z(t)$ can be substituted with the series $c^{\sqcup\!\sqcup\, n} = \underbrace{c \sqcup\!\sqcup \cdots \sqcup\!\sqcup c}_{n \text{ times}}$ such that (IV.3.8) can be written algebraically as

$$c = (x_1 + y_1)\, c \sqcup\!\sqcup c. \tag{IV.3.9}$$

Thus, c can be calculated as

$$c = \sum_{k=0}^{\infty} c_k,$$

where

$$\begin{aligned} c_0 &= 1 \\ c_k &= (x_1+y_1)\sum_{i=0}^{k-1} c_i \sqcup\!\sqcup\, c_{k-1-i}. \end{aligned}$$

Specifically, the solution of equation (IV.3.9) is obtained calculating all c_k's,

$$\begin{aligned}
c_1 &= (x_1+y_1)\, c_0 \sqcup\!\sqcup\, c_0 = (x_1+y_1) \\
c_2 &= (x_1+y_1)\,(c_0 \sqcup\!\sqcup\, c_1 + c_1 \sqcup\!\sqcup\, c_0) \\
&= (x_1+y_1)\,((x_1+y_1)+(x_1+y_1)) \\
&= 2!(x_1+y_1)^2 \\
c_3 &= (x_1+y_1)\,(c_0 \sqcup\!\sqcup\, c_2 + c_1 \sqcup\!\sqcup\, c_1 + c_2 \sqcup\!\sqcup\, c_0) \\
&= (x_1+y_1)\left(2(x_1+y_1)^2 + (x_1+y_1)^{\sqcup\!\sqcup\, 2} + 2(x_1+y_1)^2\right) = 3(x_1+y_1)^3 \\
&= (x_1+y_1)\left(2(x_1+y_1)^2 + 2(x_1+y_1)^2 + 2(x_1+y_1)^2\right) = 3!(x_1+y_1)^3 \\
c_4 &= (x_1+y_1)\,(c_0 \sqcup\!\sqcup\, c_3 + c_1 \sqcup\!\sqcup\, c_2 + c_2 \sqcup\!\sqcup\, c_1 + c_3 \sqcup\!\sqcup\, c_0) \\
&= (x_1+y_1)\left(3!(x_1+y_1)^3 + 2!(x_1+y_1) \sqcup\!\sqcup\, (x_1+y_1)^2\right. \\
&\quad \left. +2!(x_1+y_1) \sqcup\!\sqcup\, (x_1+y_1)^2 + 3!(x_1+y_1)^3\right) \\
&= (x_1+y_1)\left(3!(x_1+y_1)^3 + 2(x_1+y_1)^{\sqcup\!\sqcup\, 3} + 3!(x_1+y_1)^3\right) \\
&= (x_1+y_1)\left(3!(x_1+y_1)^3 + 2(3!)(x_1+y_1)^3 + 3!(x_1+y_1)^3\right) \\
&= 4!(x_1+y_1)^4 \\
&\vdots
\end{aligned}$$

Inductively, one can show that

$$c_k = k!(x_1+y_1)^k.$$

Thus,

$$c = \sum_{k=0}^{\infty} k!(x_1 + y_1)^k = \sum_{k=0}^{\infty} (x_1 + y_1)^{\sqcup\!\sqcup\, k} \tag{IV.3.10}$$

In this example, the order defined in (IV.3.5) is irrelevant since c is obviously exchangeable. Therfore, $F_c[w]$ is

$$F_c[w](t) = \sum_{k=0}^{\infty} (E_{x_1}[w](t) + E_{y_1}[w](t))^k = \frac{1}{1 - w(t)},$$

where $t \in [0, \tau_1]$ with $\tau_1 = \inf\{t \in [0, \infty), |w(t)| = 1\}$. This algebraic approach gives the same solution as that obtained by solving using variable separation. □

The next example is taken from [20]. It will be used to show how the corresponding generating series of a system can be approximated by globally convergent series.

Example IV.3.3. Consider the circuit in Figure 6. This is formed by four parallel components: a noise current source, an ideal resistor, an ideal capacitor and a nonlinear resistance, where the current is a function of the square of the voltage across it. Applying Kirchoff's current law and assuming $R = C = 1$, the nonlinear differential equation relating the current $\dot{w}$ and the voltage, $z(t)$, across the capacitor is

$$\frac{dz(t)}{dt} + z + z^2 = \dot{w},$$

where $\dot{w}$ is the formal derivative of the input $w \in \mathcal{UV}^m[0, T]$. In integral form

$$z(t) + \int_0^t z(s)\, ds + \int_0^t z^2(s)\, ds = \int_0^t u(s)\, ds + \oint_0^t v(t)\, dW(s).$$

The series that corresponds to this system is definitely not globally convergent, but it is known that at most has a locally convergent growth bound. Using the substitution

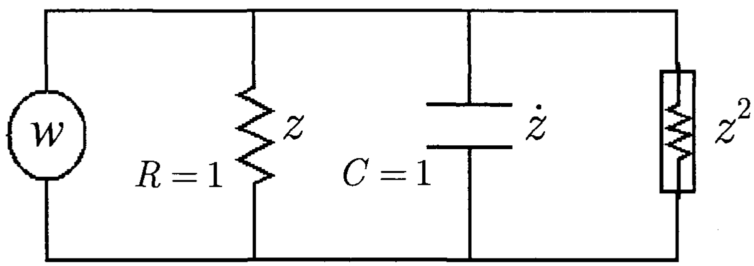

Fig. 6: A nonlinear circuit driven by a noise current source.

technique described in the example above, this nonlinear differential equation can be written algebraically as

$$c + x_0\, c + x_0\, c \sqcup\!\sqcup c = x_1 + y_1.$$

A factorization gives

$$c = (1 + x_0)^{-1} \left(-x_0 c^{\sqcup\!\sqcup\, 2} + (x_1 + y_1) \right).$$

Similarly, c can be obtained by an iterative procedure. First, a c_0 that satisfies $c_0 + x_0 c_0 + x_0 c_0 \sqcup\!\sqcup c_0 = 0$ is $c_0 = 0$. Next,

$$c_1 = (x_1 + y_1)$$

and

$$c_k = -(1 + x_0)^{-1} x_0 \sum_{i=1}^{k-1} c_i \sqcup\!\sqcup c_{k-i}.$$

Calculating the next few c_k's:

$$\begin{aligned} c_2 &= (1 + x_0)^{-1} x_0 \, (c_1 \sqcup\!\sqcup c_1) \\ &= (1 + x_0)^{-1} x_0 \, ((x_1 + y_1) \sqcup\!\sqcup (x_1 + y_1)) \end{aligned}$$

$$
\begin{aligned}
&= 2(1+x_0)^{-1}x_0(x_1+y_1)^2, \\
c_3 &= (1+x_0)^{-1}x_0\left(c_1 \sqcup\!\sqcup\, c_2 + c_2 \sqcup\!\sqcup\, c_1\right) \\
&= 2(1+x_0)^{-1}x_0\left((x_1+y_1) \sqcup\!\sqcup\, (1+x_0)^{-1}x_0(x_1+y_1)^2\right), \\
c_4 &= (1+x_0)^{-1}x_0\left(c_1 \sqcup\!\sqcup\, c_3 + c_2 \sqcup\!\sqcup\, c_2 + c_1 \sqcup\!\sqcup\, c_3\right) \\
&= 2(1+x_0)^{-1}x_0\left(4(x_1+y_1) \sqcup\!\sqcup\, \left((1+x_0)^{-1}x_0\left((x_1+y_1) \sqcup\!\sqcup\, (1+x_0)^{-1}x_0(x_1+y_1)^2\right)\right)\right. \\
&\quad \left. + 4(1+x_0)^{-1}x_0(x_1+y_1)^2 \sqcup\!\sqcup\, (1+x_0)^{-1}x_0(x_1+y_1)^2\right) \\
&\vdots
\end{aligned}
$$

Observe that $c_0, c_1, \ldots$ are all rational series, and thus, they are globally convergent, i.e., each c_i is a series associated with a globally convergent Fliess operator. In the deterministic case, it is well known that rational systems have been used as approximants of more general types of nonlinear systems [32, 33]. Thus, an advantage of this iterative procedure is that it gives an approximating method for Fliess operators associated with locally convergent series through globally convergent (rational) Fliess operators.

□

IV.4 THE OUTPUT PROCESS

From Theorem IV.1.5, $y(t) = F_c[w](t)$ is a well-defined random variable $\forall t \in [0, T]$ when c is globally convergent. The goal of this section is to show that $\{y(t)\}_{t\in[0,T]}$ is a well-defined stochastic process. The next corollary shows the relationship between the set of output processes and the set $\mathcal{UV}^m[0, T]$.

Corollary IV.4.1. *In the context of Theorem* IV.1.5, *the operator*

$$F_c : \mathcal{UV}^m[0,T] \to \widetilde{\mathcal{UV}}^m[0,T],$$

for any $T > 0$.

Proof: By Theorem IV.1.1, the iterated integrals (IV.1.2) and (IV.1.3) can each be written as a sum of a Lebesgue integral and an Itô integral. The sequence of partial sums used in (IV.1.23) can then be written as a sequence of the sum of Lebesgue and Itô integrals. Since it is well known that $L_2^m(\Omega, \mathcal{F}, P)$ with its usual norm is a Hilbert space (modulo the equivalence relations in Definition II.1.4), then Theorem IV.1.5 shows that the limit of this sequence has to be a random vector that can be expressed componentwise as the summation of a Lebesgue integral and an Itô integral. Finally, by Theorems II.3.8 and II.3.13, the limiting random variable as a function of time generates a well-defined Itô process. ∎

Now that the output process y has been identified as an L_2-Itô process, using Theorem II.3.8, there exists an almost surely continuous Itô process y' induced by y. Not surprisingly, this fact implies that there exists some compatibility between the input set and output set in the sense that $\mathcal{UV}^m[0,T] \subset \widetilde{\mathcal{UV}}^m[0,T]$. In other words, any process $y = F_c[w]$ is a well-defined L_2-Itô process. However, the independence of the inputs is not necessarily preserved at the output. In Chapter V, this result will be used to establish the convergence of the interconnection of Fliess operators driven by processes in $\mathcal{UV}^m[0,T]$.

Although it is not easy to write the output process y in its specific Itô form, the next lemma shows how this goal can be done in some situations.

Lemma IV.4.1. *Consider a locally convergent series $c \in \mathbb{R}\langle\langle XY \rangle\rangle$ with growth constants $K, M > 0$. Let $\xi \in XY^k$ be fixed, if $d \in \mathbb{R}\langle\langle XY \rangle\rangle$ is defined as*

$$d = \sum_{\eta \in XY^*} (\xi^{-1}(c), \eta)\, \eta$$

then there exists an $N > 0$ such that $|(d, \eta)| = |(\xi^{-1}(c), \eta)| \leq KM^{|\xi|} \left(MN^{|\xi|}\right)^{|\eta|} |\eta|!$, for all $\eta, \xi \in XY^*$.

Proof: Let $|\xi| = k$. For $N > 0$ large enough, it is not difficult to see that

$$\frac{(n+1)\cdots(n+k)}{(N^k)^n} \leq 1.$$

Thus, $(n+1)\cdots(n+k) \leq (N^k)^n$. As a consequence,

$$\begin{aligned} |(d, \eta)| = |(c, \xi\eta)| &\leq KM^k M^{|\eta|} |\eta|!(n+1)\cdots(n+k) \\ &\leq K' M'^{|\eta|} |\eta|! \end{aligned}$$

with $K' = KM^k$ and $M' = MN^k$. ■

Note that if $c \in \mathbb{R}\langle\langle XY \rangle\rangle$ is globally convergent, then $|(\xi^{-1}(c), \eta)| \leq K'M^{|\eta|}$ where $K' = KM^{|\xi|}$.

Example IV.4.1. Let $c \in \mathbb{R}\langle\langle XY \rangle\rangle$ be a globally convergent series such that if $|\eta|_{x_i, y_i} \neq 0$ for $i \neq 0$ then $(c, \eta) = 0$. It easy to show, since the inputs are trivial, that any $E_\eta[0]$ can be expressed as

$$E_\eta[0](t) = \begin{cases} E_\eta[0](t) & : \ \eta = x_0\eta' \\ \displaystyle\int_0^t E_{\eta'}[0](s)\, dW(s) & : \ \eta = y_0\eta' \text{ and } \eta' = x_0\eta'' \\ \displaystyle\int_0^t E_{\eta'}[0](s)\, dW(s) & : \ \eta = y_0\eta' \text{ and } \eta' = y_0\eta'', \\ \displaystyle +\frac{1}{2}\int_0^t E_{\eta''}[0](s)\, ds & \end{cases}$$

where $\eta', \eta'' \in XY^*$. Now the output $y = F_c[0]$ associated with generating the series c can be written as

$$\begin{aligned}
y(t) &= (c,\emptyset) + \int_0^t F_{x_0^{-1}(c)}[w](s)\,ds + \int_0^t F_{y_0^{-1}(c)}[w](s)\,dW(s) \\
&\quad + \frac{1}{2}\left\langle F_{y_0^{-1}(c)}[w](\cdot), W\right\rangle_{[0,t]} \\
&= (c,\emptyset) + \int_0^t F_{x_0^{-1}(c)}[w](s)\,ds + \int_0^t F_{y_0^{-1}(c)}[w](s)\,dW(s) \\
&\quad + \frac{1}{2}\left\langle (c,y_0) + \int_0^{\cdot} F_{y_0^{-1}(x_0^{-1}(c))}[w](s)\,ds + \oint_0^{\cdot} F_{y_0^{-1}(y_0^{-1}(c))}[w](s)\,dW(s), W\right\rangle_{[0,t]} \\
&= (c,\emptyset) + \int_0^t F_{x_0^{-1}(c)}[w](s)\,ds + \int_0^t F_{y_0^{-1}(c)}[w](s)\,dW(s) \\
&\quad + \frac{1}{2}\left\langle \int_0^{\cdot} F_{y_0^{-1}(y_0^{-1}(c))}[w](s)\,dW(s) + \frac{1}{2}\left\langle F_{y_0^{-1}(y_0^{-1}(c))}[w](\cdot), W\right\rangle_{[0,t]}, W\right\rangle_{[0,t]} \\
&= (c,\emptyset) + \int_0^t F_{x_0^{-1}(c)}[w](s)\,ds + \int_0^t F_{y_0^{-1}(c)}[w](s)\,dW(s) \\
&\quad + \frac{1}{2}\int_0^t F_{y_0^{-1}(y_0^{-1}(c))}[w](s)\,ds + \frac{1}{4}\underbrace{\left\langle\left\langle F_{y_0^{-1}(y_0^{-1}(c))}[w](\cdot), W\right\rangle_{[0,t]}, W\right\rangle_{[0,t]}}_{0} \\
&= (c,\emptyset) + \int_0^t \left(F_{x_0^{-1}(c)}[w](s) + \frac{1}{2}F_{y_0^{-1}(y_0^{-1}(c))}[w](s)\right) ds \\
&\quad + \int_0^t F_{y_0^{-1}(c)}[w](s)\,dW(s), \qquad\qquad \text{(IV.4.1)}
\end{aligned}$$

where $x_0^{-1}(\cdot)$, $y_0^{-1}(\cdot)$ are left-shift operators. From Theorem IV.1.5 and Lemma IV.4.1, the random variables $F_{x_0^{-1}(c)}[0]$, $F_{y_0^{-1}(c)}[0]$ and $F_{y_0^{-1}(y_0^{-1}(c))}[0]$ are well-defined. Hence, the output process y is a well-defined Itô process. $\square$

Observe that, at each instant of time, the Itô process defined by the output process of a Fliess operator is associated with the same generating series. In other words, the generating series can be seen as its series expansion when the words are replaced with their corresponding iterated integral. In this sense, this fact resembles the concept of *analyticity*.

Definition IV.4.1. The Itô process associated with a generating series $c \in \mathbb{R}^{\ell}\langle\langle XY\rangle\rangle$ will be called **analytic**.

Even though it has been established that the output process of a Fliess operator is well-defined when the associated series is either globally convergent or locally convergent and exchangeable, there exists the possibility that two formal power series in $\mathbb{R}\langle\langle XY\rangle\rangle$ might represent the same output process. The next theorem shows that the output process y has a unique generating series.

Theorem IV.4.1. *Let $c, d \in \mathbb{R}^{\ell}\langle\langle XY\rangle\rangle$ be two globally convergent series. Then $F_c = F_d$ on $\mathcal{UV}^m[0,T]$ if and only if $c = d$.*

Proof: If $c = d$, the result follows trivially. The converse result is proved first for words, i.e., one has to establish that if $E_\eta[w](t) = E_\xi[w](t)$ then $\eta = \xi$, $\eta, \xi \in XY^*$. This fact is proven by induction over the length of a word, n. When $n = 0$, the result follows immediately. For $n = 1$, suppose $E_{x_i}[w](t) = E_{x_j}[w](t)$, then by definition

$$\int_0^t u_i(s)\, ds = \int_0^t u_j(s)\, ds$$

or

$$\int_0^t \left(u_i(s) - u_j(s)\right)\, ds = 0.$$

Thus, $u_i = u_j$ a.s., which implies that $x_i = x_j$. Now suppose instead that $E_{x_i}[w](t) = E_{y_j}[w](t)$. Then

$$\oint_0^t v_i(s)\, dW(s) = \int_0^t u_j(s)\, ds.$$

By Corollary II.3.2, this can only occur if $u_i = v_j = 0$. Therefore, $E_{x_i}[w](t) \neq E_{y_j}[w](t)$ for a nonzero $w \in \mathcal{UV}^m[0,T]$. If $E_{y_i}[w](t) = E_{y_j}[w](t)$, then by the linearity

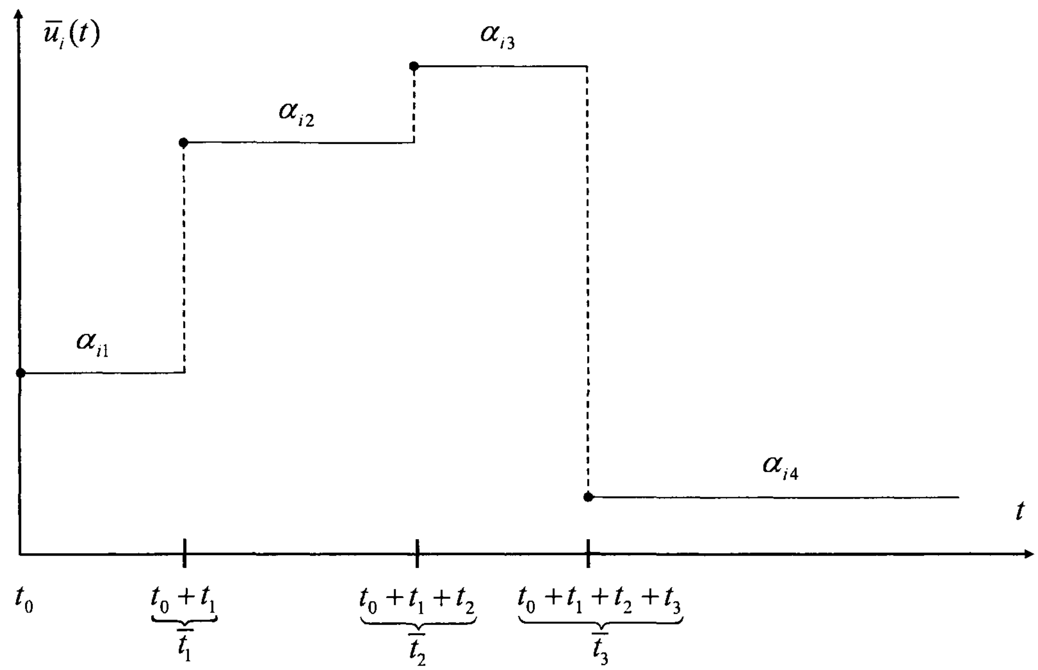

Fig. 7: The i-th component of input $\bar{u}$, $i \neq 0$.

of the Stratonovich integral and Corollary II.3.2 it follows that $v_i = v_j$. Next, assume that the claim holds up to n. If $E_{x_i\eta}[w](t) = E_{x_j\eta}[w](t)$ for $\eta \in XY^n$, then

$$\int_0^t (u_i(s) - u_j(s))\, E_\eta[w](s)\, ds = 0.$$

Given that $E_\eta[w] \neq 0$ for every $t > 0$, then $u_i(s) = u_j(s)$ and $x_i = x_j$. For $E_{x_i\eta}[w](t) = E_{y_j\eta}[w](t)$, it follows that $u_i(s)E_\eta[w](s) = v_j(s)E_\eta[w](s) = 0$. Since $E_\eta[w] \neq 0$, then $u_i(s) = v_j(s) = 0$. Thus, $E_{x_i\eta}[w](s) \neq E_{y_j\eta}[w](s)$ unless $w = 0$. Analogously, using linearity, if $E_{y_i\eta}[w](t) = E_{y_j\eta}[w](t)$, then $v_i(s) = v_j(s)$ and $y_i = y_j$.

Next, consider two series $c, d \in \mathbb{R}^\ell\langle\langle XY\rangle\rangle$. Since $F_c[w] = F_d[w]$ implies $F_{c-d}[w] = 0$, it is sufficient to prove that if $F_c[w] = 0$ for all w then $c = 0$. In [24,63] a uniqueness proof was given for deterministic inputs, where the operator $\frac{\partial}{\partial t_1 \partial t_2 \cdots \partial t_k}$ was applied

to F_c, $c \in \mathbb{R}\langle\langle X\rangle\rangle$ for a piecewise constant input in $B_\infty^m(R)[t_0, t_0+T]$ defined by

$$\bar{u}(t) = \alpha_j \leq R,\ t \in [\bar{t}_{j-1}, \bar{t}_j],$$

with $\alpha_j = [\alpha_{1j} \cdots \alpha_{mj}] \in \mathbb{R}^m$, $\alpha_{0j} = 1$, $\bar{t}_j = \sum_{\ell=0}^{j} t_\ell$ and $t_\ell \geq 0, j, \ell = 1, 2, \ldots, k$; and $\bar{t}_k < T$ (see Fig. 7). The following identity was then employed

$$\left.\frac{\partial^k}{\partial t_1 \partial t_2 \cdots \partial t_k} F_c[\bar{u}](\bar{t}_k)\right|_{\substack{t_j=0^+ \\ j=0,\ldots,k+1}} = \sum_{\xi \in X^k} \alpha_{\xi k}(c, \xi) = 0. \tag{IV.4.2}$$

Unfortunately, stochastic calculus lacks of a notion of a derivative in the usual sense. However, Corollary II.3.2 can be used to overcome this obstacle. Specifically, recall that for a function $F \in \mathcal{C}^2$, the Stratonovich chain rule (I.1.7) established that

$$dF(X(t)) = F(X(0)) + \frac{\partial}{\partial X} F(X(t))u(t)\, dt + \frac{\partial}{\partial X} F(X(t))v(t)\, dW(t),$$

where $X(s) = X(0) + \int_0^t u(s)\, ds + \oint_0^t v(s)\, dW(s)$. If F is the identity function then

$$dX(t) = u(t)\, dt + v(t)\, dW(t).$$

Moreover, from Corollary II.3.2, if $X(t) = 0$ then a.s. $v(s) = u(s) = 0$. Hence $u(t) + v(t) = 0$. One can then define the operator

$$D_t X(t) = u(t) + v(t)$$

such that $X(t) = 0$ implies $D_t X(t) = 0$ for all t. D_t is linear since

$$\begin{aligned} D_t(X_1(t) + X_2(t)) &= u_1(t) + u_2(t) + v_1(t) + v_2(t) \\ &= D_t X_1(t) + D_t X_2(t), \end{aligned}$$

and

$$\begin{aligned} D_t(rX_1(t)) &= D_t\left(\int_0^t r\, u(s)\, ds + \oint_0^t r\, v(s)\, dW(s)\right) \\ &= r(u(t) + v(t)) = rD_t X(t), \end{aligned}$$

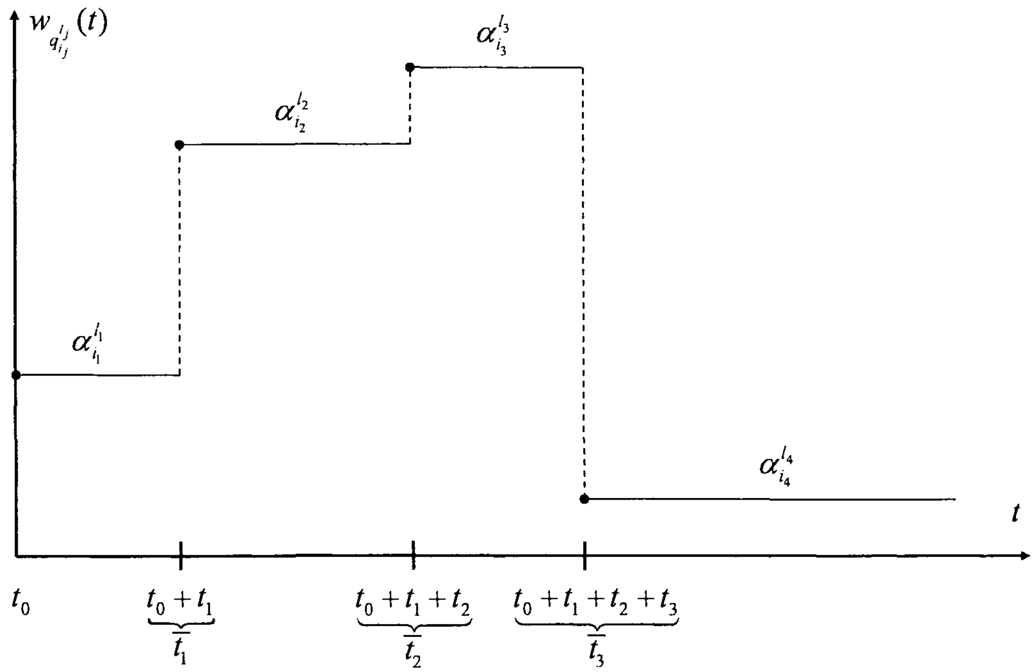

Fig. 8: Piecewise process $w_{q_{i_k}^{l_k}}$.

for every $r \in \mathbb{R}$. Employing this operator and an identity similar to (IV.4.2) are the key to the uniqueness proof.

It is first necessary to show that

$$D_{\bar{t}_1} \cdots D_{\bar{t}_k} F_c[w](\bar{t}_k) \Bigg|_{\substack{t_j=0^+ \\ j=0,\ldots,k+1}} = \sum_{\eta \in XY^k} \alpha_{\eta k}(c, \bar{\xi}) = 0, \tag{IV.4.3}$$

where $\alpha_\eta = \alpha_{(q_{i_k}^{l_k})k} \cdots \alpha_{(q_{i_1}^{l_1})1}$, and $\alpha_{(q_{i_j}^{l_j})j}$ denotes the piecewise part of

$$w_{q_{i_j}^{l_j}}(t) = \begin{cases} u_{i_j}(t), & \text{if } l_j = 1 \\ v_{i_j}(t), & \text{if } l_j = 2 \end{cases}$$

between times $\bar{t}_{j-1}$ and $\bar{t}_j$ (see Fig. 8). An inductive argument over k will prove the identity. In particular, a piecewise constant input is an L_2-Itô process. Consider an

input $w \in \mathcal{UV}^m[0,T]$ with integrands u, v which are piecewise constant. Trivially, $F_c[w]$ can be written as

$$\begin{aligned} F_c[w](t) &= (c,\emptyset) + \int_0^t \sum_{i=0}^m u_i(s) \underbrace{\sum_{\eta\in\mathrm{supp}(x_i^{-1}(c))} (c, x_i\eta) E_\eta[w](s)}_{F_{x_i^{-1}(c)}[w](s)}\, ds \\ &+ \oint_0^t \sum_{i=0}^m v_i(s) \underbrace{\sum_{\eta\in\mathrm{supp}(y_i^{-1}(c))} (c, y_i\eta) E_\eta[w](s)}_{F_{y_i^{-1}(c)}[w](s)}\, dW(s). \qquad \text{(IV.4.4)} \end{aligned}$$

By Corollary II.3.2, equation (IV.4.4) implies

$$\sum_{i=0,l=1}^{m,2} w_{q_i^l}(s) F_{(q_i^l)^{-1}(c)}[w](s) = 0.$$

The same result is obtained by calculating $D_t F_c[w](t)$. Therefore,

$$\begin{aligned} D_{\bar t_1} F_c[w](\bar t_1)\Big|_{t_1=0^+} &= \sum_{i=0,l=1}^{m,2} w_{q_i^l}(\bar t_1) F_{(q_i^l)^{-1}(c)}[w](\bar t_1)\Big|_{t_1=0^+} \\ &= \sum_{\substack{q_i^l\in XY \\ \eta\in XY^*}} w_{q_i^l}(t_0^+)(c, q_i^l\eta) E_\eta[w](t_0) \\ &= \sum_{q_i^l\in XY} \alpha_{q_i^l 1}(c, q_i^l), \end{aligned}$$

where $E_\eta[w](t_0) = 1$ if $\eta = \emptyset$ or 0 if $\eta \neq \emptyset$. Since by assumption $F_c[w] = 0$ then $D_{\bar t_1} F_c[w](\bar t_1) = 0$. Hence, $\sum_{\eta\in XY} \alpha_{q_i^l 1}(c, q_i^l) = 0$. For $F_c[w](\bar t_2)$, one can follow a similar procedure

$$\begin{aligned} D_{\bar t_2} F_c[w](\bar t_2)\Big|_{t_2=0^+} &= \int_0^{\bar t_2} \sum_{i=0,l=1}^{m,2} \sum_{j=0}^m w_{q_i^l}(\bar t_2) u_{x_j}(s) F_{(q_i^l)^{-1}(x_j^{-1}(c))}[w](\bar t_2)\, ds \Big|_{t_2=0^+} \\ &+ \oint_0^{\bar t_2} \sum_{i=0,l=1}^{m,2} \sum_{j=0}^m w_{q_i^l}(\bar t_2) v_{y_j}(s) F_{(q_i^l)^{-1}(y_j^{-1}(c))}[w](\bar t_2)\, dW(s) \Big|_{t_2=0^+} \\ &= \int_0^{t_0+t_1} \sum_{i=0,l=1}^{m,2} \sum_{j=0}^m w_{q_i^l}(t_0+t_1^+) u_j(s) F_{(q_i^l)^{-1}(x_j^{-1}(c))}[w](s)\, ds \\ &+ \oint_0^{t_0+t_1} \sum_{i=0,l=1}^{m,2} \sum_{j=0}^m w_{q_i^l}(t_0+t_1^+) v_j(s) F_{(q_i^l)^{-1}(y_j^{-1}(c))}[w](s)\, dW(s). \end{aligned}$$

Observe that $\bar{u}_i(t_0+t_1^+) = \alpha_{\bar{x}_i 2}$. One can then apply $D_{\bar{t}_1}$ to $D_{\bar{t}_2}F_c[w](\bar{t}_2)\Big|_{t_2=0^+}$, which amounts to

$$
\begin{aligned}
D_{\bar{t}_1}&\left(D_{\bar{t}_2}F_c[w](\bar{t}_2)\Bigg|_{t_2=0^+}\right)\Bigg|_{t_1=0^+} \\
&= \sum_{i=0,l_1=1}^{m,2}\sum_{j=0,l_2=1}^{m,2} \alpha_{q_i^{l_1}2} w_{q_j^{l_2}}(t_0+t_1) F_{(q_i^{l_1})^{-1}((q_j^{l_2})^{-1}(c))}[w](t_0+t_1)\Bigg|_{t_1=0^+} \\
&= \sum_{i=0,l_1=1}^{m,2}\sum_{j=0,l_2=1}^{m,2}\sum_{\eta\in XY^*} \alpha_{q_i^{l_1}2}\alpha_{q_j^{l_2}1}(c, q_i^{l_1}q_j^{l_2}\eta)E_\eta[w](t_0) \\
&= \sum_{\substack{q_i^{l_1},q_j^{l_2}\in\bar{X} \\ \eta\in XY^*}} \alpha_{q_i^{l_1}2}\alpha_{q_j^{l_2}1}(c, q_i^{l_1}q_j^{l_2}\eta)E_\eta[w](t_0) \\
&= \sum_{\eta\in XY^2} \alpha_{\eta 2}(c,\eta),
\end{aligned}
$$

where $\sum_{\eta\in XY^2}\alpha_{\eta 2}(c,\eta) = 0$ since $D_{\bar{t}_2}F_c[w](\bar{t}_2) = 0$. Now assume the identity in question is valid up to $k \geq 0$. Then

$$
\begin{aligned}
D_{\bar{t}_1}&\cdots D_{\bar{t}_k}\left(D_{\bar{t}_{k+1}}F_c[w](\bar{t}_{k+1})\Bigg|_{t_{k+1}=0^+}\right)\Bigg|_{\substack{t_j=0^+ \\ j=0,\ldots,k}} \\
&= D_{\bar{t}_1}\cdots D_{\bar{t}_k}\left(\sum_{i_{k+1}=0,l_{k+1}=1}^{m,2} w_{q_{i_{k+1}}^{l_{k+1}}}(\bar{t}_{k+1}) F_{\left(q_{i_{k+1}}^{l_{k+1}}\right)^{-1}(c)}[w](\bar{t}_{k+1})\Bigg|_{t_{k+1}=0^+}\right)\Bigg|_{\substack{t_j=0^+ \\ j=0,\ldots,k}} \\
&= \sum_{q_{i_{k+1}}^{l_{k+1}}\in XY} \alpha_{\left(q_{i_{k+1}}^{l_{k+1}}\right)k+1} D_{\bar{t}_1}\cdots D_{\bar{t}_k} F_{\left(q_{i_{k+1}}^{l_{k+1}}\right)^{-1}(c)}[w](\bar{t}_k)\Bigg|_{\substack{t_j=0^+ \\ j=0,\ldots,k}},
\end{aligned}
$$

where the fact that $w_{q_{i_{k+1}}^{l_{k+1}}}(\bar{t}_{k+1})\Big|_{t_{k+1}=0^+} = \alpha_{\left(q_{i_{k+1}}^{l_{k+1}}\right)k+1}$ has been used. By the induction hypothesis,

$$
\begin{aligned}
D_{\bar{t}_1}\cdots D_{\bar{t}_{k+1}}F_c[w](\bar{t}_{k+1})\Bigg|_{\substack{t_j=0^+ \\ j=0,\ldots,k+1}} &= \sum_{q_{i_{k+1}}^{l_{k+1}}\in XY} \alpha_{\left(q_{i_{k+1}}^{l_{k+1}}\right)k+1}\left(\sum_{\eta\in XY^k}\alpha_{\eta k}\left(\left(q_{i_{k+1}}^{l_{k+1}}\right)^{-1}(c),\eta\right)\right) \\
&= \sum_{\eta\in XY^{k+1}} \alpha_{\eta k+1}(c,\eta).
\end{aligned}
\tag{IV.4.5}
$$

Again, since $F_c[w] = 0$, then $\sum_{\eta\in XY^{k+1}}\alpha_{\eta k+1}(c,\eta) = 0$.

Observe that equation (IV.4.5) is a polynomial in α_{ij}. Define the following operator

$$\frac{\partial^k}{\partial\alpha_{\ell_k k}^{l_k}\cdots\alpha_{\ell_1 1}^{l_1}}\alpha_{i_k k}^{j_k}\cdots\alpha_{i_1 1}^{j_1}$$
$$=\begin{cases} 1 & : \ \ell_s = i_s,\ l_s = j_s,\ s = 1,\ldots,k \\ 0 & : \ \text{otherwise.}\end{cases}$$

Applying this operator to (IV.4.3) gives

$$\frac{\partial^k}{\partial\alpha_{\ell_k k}^{l_k}\cdots\alpha_{\ell_1 1}^{l_1}}\left(\sum_{q_{i_k}^{j_k}\cdots q_{i_1}^{j_1}\in XY^k}\alpha_{i_k k}^{j_k}\cdots\alpha_{i_1 1}^{j_1}(c, q_{i_k}^{j_k}\cdots q_{i_1}^{j_1})\right) = (c, q_{\ell_k}^{l_k}\cdots q_{\ell_1}^{l_1}) = 0.$$

Repeating this procedure for any $\eta = q_{i_k}^{j_k}\cdots q_{i_1}^{j_1} \in XY^k$, $k \geq 0$, it follows that $(c,\eta) = 0$. Thus $c = 0$ as desired. ■

IV.5 CHEN SERIES AND FLIESS OPERATORS

An important class of formal power series are the *Chen series*. Basically, a Chen series is used to represent all possible paths, say $C_{[t_0,t_1]}$, from the interval $[t_0, t_1]$ to $\mathbb{R}^{m+1}$ made by a set of $m+1$ input signals [7,22]. Chen series are also closely related to Fliess operators driven by deterministic inputs. That is, a Fliess operator can be written as the inner product of its generating series and the Chen series, expressed as an exponential, generated by its input. In the stochastic setting, Chen series have been used, for example, to study discretization schemes for stochastic differential equations of the form

$$dz(t) = f_0(z(t)) + \sum_{i=1}^{m} f_i(z(t))\, dW_i(t),$$

where each $f_i \in C_b^\infty$, as well as the local structure of the general solution of stochastic differential equations [3, 6, 59]. In other words, the interpretation of Chen series as input paths can actually be extended to Brownian paths [1,3,6,22] and even to rough paths [43]. This last fact makes viable the construction of a Chen series using inputs in $\mathcal{UV}^m[0,T]$. The purpose of this section is to show that Fliess operators with such stochastic inputs are also related to a class of Chen series. The treatment here is based on the deterministic case as presented in [24]. Some preliminaries are needed first.

Definition IV.5.1. Let $\mathcal{L}$ be an $\mathbb{R}$-vector space with and $\mathbb{R}$-bilinear mapping

$$\mathcal{L} \times \mathcal{L} \to \mathcal{L} : (x, y) \mapsto [x, y],$$

satisfying the identities:

i. $[x, x] = 0$,

ii. $[[x, y], z] + [[y, z], x] + [[z, x], y] = 0.$

for $x, y, z \in \mathcal{L}$. The operation $[\cdot, \cdot]$ is called the **Lie Bracket** on $\mathcal{L}$, and the space $\mathcal{L}$ and this product is called a **Lie algebra**.

Given any associative $\mathbb{R}$-algebra, $\mathcal{A}$, with multiplicative identity element 1, there is an associated Lie algebra whose bracket operation is

$$[x, y] = xy - yx.$$

For the alphabet XY, the smallest subset of $\mathbb{R}\langle XY\rangle$ containing XY which is closed under the Lie bracket operation forms a Lie algebra denoted $\mathcal{L}(XY)$.

Definition IV.5.2. Any polynomial in $\mathcal{L}(XY)$ is called a **Lie polynomial**.

Definition IV.5.3. A series $d \in \mathbb{R}\langle\langle XY \rangle\rangle$ is called a **Lie series** if it can be decomposed as $d = \sum_{n\geq 1} p_n$, where each p_n is a Lie polynomial whose support resides in XY^n.

For any series $d \in \mathbb{R}\langle\langle XY \rangle\rangle$ such that $(d, \emptyset) = 0$, the exponential $\exp(d)$ is defined by the series

$$\exp(d) = \sum_{n=0}^{\infty} \frac{d^n}{n!},$$

and $\log(d)$ is defined by

$$\log(1+d) = \sum_{n=0}^{\infty} (-1)^{n-1} \frac{d^n}{n}.$$

Here $\log(\exp(d)) = d$ and $\exp(\log(1+d)) = 1 + d$.

Definition IV.5.4. A series is called an **exponential Lie series** when $c = \exp(d)$, where d is a Lie series.

A characterization of exponential Lie series is given in the next theorem.

Theorem IV.5.1. [49] *(Ree's Criterion) c is an exponential Lie series if and only if $(c, \xi ⧢ \nu) = (c, \xi)(c, \nu)$ for all $\xi, \nu \in XY^*$.*

Example IV.5.1. Suppose $XY = \{y_0\}$, $c \in \mathbb{R}\langle\langle XY \rangle\rangle$ with $(c, \emptyset) = 1$ and $(c, \xi ⧢ \nu) = (c, \xi)(c, \nu)$ for all $\xi, \nu \in XY^*$. Using the identity

$$y_0^i ⧢ y_0^j = \binom{i+j}{i} y_0^{i+j},$$

it follows that

$$\begin{aligned} (c, y_0^i ⧢ y_0^j) &= \binom{i+j}{i} (c, y_0^{i+j}) \\ &= (c, y_0^i)(c, y_0^j). \end{aligned}$$

Setting $i = n \geq 0$ and $j = 1$ yields the recursion

$$(c, y_0^{n+1}) = \left[\frac{(c, y_0)}{n+1}\right](c, y_0),$$

so that

$$(c, y_0^n) = \frac{(c, y_0)^n}{n!}, \; n \geq 0.$$

Thus, as predicted by Theorem IV.5.1, it follows that c is the exponential Lie series

$$\begin{aligned} c &= \sum_{n \geq 0} \frac{(c, y_0)^n}{n!} = \exp((c, y_0)y_0) \\ &= \exp(p), \end{aligned}$$

where $p = (c, y_0)y_0 \in \mathcal{L}(XY)$. □

Definition IV.5.5. For any $T > 0$, $w \in \mathcal{UV}^m[0, T]$ and $t \in [0, T]$, the Chen series associated with a formal power series in $\mathbb{R}^\ell\langle\langle XY \rangle\rangle$ is defined as

$$P[w](t) = \sum_{\eta \in XY^*} \eta \, E_\eta[w](t).$$

One can see from Theorem IV.5.1 that for any $t \in [0, T]$ and $\xi, \nu \in XY^*$

$$\begin{aligned} (P[u](t), \xi \sqcup\!\sqcup \nu) &= \sum_{\eta \in XY^*} E_\eta[u](t)(\eta, \xi \sqcup\!\sqcup \nu) \\ &= E_{\xi \sqcup\!\sqcup \nu}[u](t) \\ &= E_\xi[u](t) E_\nu[u](t) \\ &= \left(\sum_{\bar{\xi} \in XY^*} \bar{\xi} \, E_{\bar{\xi}}[u](t), \xi\right) \left(\sum_{\bar{\nu} \in XY^*} \bar{\nu} \, E_{\bar{\nu}}[u](t), \nu\right) \\ &= (P[u](t), \xi)\,(P[u](t), \nu)\,. \end{aligned}$$

Therefore, $P[u](t)$ is an exponential Lie series. This type of series is directly related to Fliess operators with inputs on $\mathcal{UV}^m[0, T]$ in the sense that for any Fliess operator

F_c, $c \in \mathbb{R}^\ell\langle\langle XY\rangle\rangle$, one can write

$$\begin{aligned} y(t) &= F_c[w](t) \\ &= \sum_{\eta\in XY^*} (c,\eta)E_\eta[w](t) \\ &= \left(c, \sum_{\eta\in XY^*} \eta E_\eta[w](t)\right) \\ &= (c, P[w](t)) \\ &= (c, \exp(U)), \end{aligned}$$

where $U = \log(P[u])$.

Example IV.5.2. The Fliess operator driven by $w_1 = \int_0^t u_1(s)\,ds + \oint_0^t v_1(s)\,dW(s) \in \mathcal{U}\mathcal{V}^m[0,T]$ associated with the series $c = (x_1+y_1)^* = (1-x_1-y_1)^{-1} = \sum_{k\geq 0}(x_1+y_1)^k$ is

$$y(t) = F_{(x_1+y_1)^*}[w](t) = \sum_{k\geq 0} E_{(x_1+y_1)^k}[w](t).$$

By identity (IV.1.31),

$$\begin{aligned} y(t) &= \sum_{k\geq 0} \frac{1}{k!} E_{(x_1+y_1)^{\sqcup\!\sqcup k}}[w](t) \\ &= \sum_{k\geq 0} \frac{1}{k!} E^k_{(x_1+y_1)}[w](t) \\ &= \exp\left(\int_0^t u_1(s)\,ds + \oint_0^t v_1(s)\,dW(s)\right) \\ &= \exp\left(w_1(t)\right). \end{aligned}$$

Hence, tihs Fliess operator can be written as an exponential. □

Another characteristic of Chen series for deterministic inputs is that they satisfy a differential equation [58]. To conclude this section, the next theorem shows that

the Chen series generated by stochastic inputs in $\mathcal{UV}^m[0,T]$ also satisfy a stochastic differential equation.

Theorem IV.5.2. *For any $T > 0$, let $w \in \mathcal{UV}^m[0,T]$. Then the corresponding Chen series $P[w](t)$ satisfies the stochastic differential equation*

$$dP[u](t) = \left(\sum_{i=0}^{m} x_i u_i(t)\,dt + y_i v_i(t)\,dW(t)\right) P[u](t). \tag{IV.5.1}$$

Proof: Observe that,

$$\begin{aligned}
dP[u](t) &= \sum_{\eta \in XY^*} \eta\, dE_\eta[u](t) \\
&= \sum_{\eta \in XY^*} \eta \left(\sum_{i=0}^{m} u_i(t) E_{x_i^{-1}(\eta)}[u](t)\,dt + v_i(t) E_{y_i^{-1}(\eta)}[u](t)\,dW(t)\right) \\
&= \left(\sum_{i=0}^{m} x_i u_i(t)\,dt + y_i v_i(t)\,dW(t)\right) \sum_{\eta \in XY^*} \eta E_\eta[u](t) \\
&= \left(\sum_{i=0}^{m} x_i u_i(t)\,dt + y_i v_i(t)\,dW(t)\right) P[u](t).
\end{aligned}$$

■

It is important to point out that equation (IV.5.1) is strictly valid only in its integral form, i.e.,

$$P[u](t) = \left(\sum_{i=0}^{m} x_i \int_0^t u_i(s) P[w](s)\,ds + \sum_{i=0}^{m} y_i \oint_0^t v_i(s) P[w](s)\,dW(s)\right).$$

CHAPTER V

INTERCONNECTIONS OF FLIESS OPERATORS

In this chapter, the theory describing the interconnection of Fliess operators is presented, and, in the process, dissertation problems (*iv*) and (*v*) are addressed. In the first section, the existing formalism for the parallel, product and cascade interconnection of systems driven by deterministic inputs is summarized. Since an extension of the shuffle algebra has been developed in Chapter IV for the language XY^*, an extension of the composition product can be also developed for XY^*. Using this concept, an analogue of the formal Fliess operator described in [24, 30] is developed for formal stochastic input processes. Next, the parallel, product and cascade connections are defined algebraically for formal Fliess operators and for Fliess operators with inputs in $\mathcal{UV}^m[0,T]$. The chapter is concluded giving sufficient conditions for the global and local convergence of the parallel, product and cascade connection of Fliess operators.

V.1 PRELIMINARIES

The four elementary interconnections used in system theory are the parallel, product, cascade and feedback connections. The focus of this dissertation is on the *non-recursive* connections, which are the parallel, product and cascade connections. In the deterministic case, the parallel connection is trivial. The product connection was analyzed by Fliess [17] and Wang [63]. The cascade connection was analyzed

by Ferfera [13], and Gray and Li [28]. Gray and Wang showed in [29] that Fliess operators always have the form

$$F_c : B_p^m(R)[t_0,\, t_0+T] \to B_q^\ell(S)[t_0,\, t_0+T],$$

provided that $c \in \mathbb{R}^\ell\langle\langle X\rangle\rangle$ is locally convergent and $R, S, T > 0$ are sufficiently small. This fact makes the interconnections of analytic nonlinear input-output systems well-defined in the deterministic setting. The next theorem describes the generating series of the three interconnections under study.

Theorem V.1.1. [13, 24, 28] *Let $c, d \in \mathbb{R}^m\langle\langle X\rangle\rangle$. The generating series for the parallel, product and cascade interconnections are given by*

$$\begin{aligned} F_c + F_d &= F_{c+d} \\ F_c \cdot F_d &= F_{c \,\sqcup\!\sqcup\, d} \\ F_c \circ F_d &= F_{c \circ d}. \end{aligned}$$

The next theorem states that the series operations $+$, $\sqcup\!\sqcup$ and $\circ$ preserve local convergence.

Theorem V.1.2. [28] *Suppose $c, d \in \mathbb{R}^m\langle\langle X\rangle\rangle$ are locally convergent. Then $c + d$, $c \,\sqcup\!\sqcup\, d$ and $c \circ d$ are also locally convergent.*

The stochastic counterpart of these results is presented next.

V.2 THE STOCHASTIC CASE

In Chapter IV, it was shown that Fliess operators driven by inputs from $\mathcal{UV}^m[0,T]$ have the form

$$F_c : \mathcal{UV}^m[0,T] \to \widetilde{\mathcal{UV}}^m[0,T], \tag{V.2.1}$$

when c is a globally convergent series, and

$$F_c : \mathcal{UV}^m[0,T] \to \widetilde{\mathcal{UV}}^m[0,\tau_R], \tag{V.2.2}$$

when c is locally convergent. Unlike the deterministic situation, this is *not* enough to establish the well-posedness of the corresponding interconnections. For example, the output process of a Fliess operator with stochastic inputs does *not* have the independence properties that inputs in $\mathcal{UV}^m[0,T]$ have. Therefore, the outputs *cannot* drive a second Fliess operator. In this situation, it is first convenient to consider the interconnection of Fliess operators when there are *no convergence* assumptions.

V.2.1 Formal Fliess operators

The main objective of this section is to define a class Fliess operators where the associated generating series is independent of any assumptions concerning global or local convergence. The main idea here is to obtain clues about the generating series of the interconnection of Fliess operators free of any convergence requirement, and then find a correspondence to the case where Fliess operators are convergent. This latter point of view motivates the definition of a formal stochastic process.

Definition V.2.1. Let $c_w \in \mathbb{R}^m\langle\langle X_0Y_0\rangle\rangle$. A **formal stochastic process** w is defined by

$$w(t) = \sum_{\eta \in X_0Y_0^*} (c_w, \eta) E_\eta[0](t). \tag{V.2.3}$$

The set formed by all formal stochastic processes is denoted by $\mathscr{W}$.

There is no input in $\mathcal{UV}^m[0,T]$ playing a role in (V.2.3) since the iterated integrals only utilize letters from X_0Y_0. It is understood that for any $w \in \mathscr{W}$ there exist a

corresponding generating series $c_w \in \mathbb{R}\langle\langle X_0 Y_0\rangle\rangle$, and since c_w is arbitrary, w is simply a formal summation of iterated integrals. Note also that c_w can be written as

$$c_w = (c_w, \emptyset) + x_0 \underbrace{\sum_{\xi \in X_0 Y_0{}^*} (x_0^{-1}(c_w), \xi)\xi}_{c_w^1} + y_0 \underbrace{\sum_{\xi \in X_0 Y_0{}^*} (y_0^{-1}(c_w), \xi)\xi}_{c_w^2},$$

where $(c_w^1, \xi) = (c_w, x_0\xi)$ and $(c_w^2, \xi) = (c_w, y_0\xi)$ for all $\xi \in X_0 Y_0{}^*$. Trivially, $c_w^1, c_w^2 \in \mathbb{R}^m\langle\langle X_0 Y_0\rangle\rangle$, and they are called the *drift component* and the *diffusion component* of c_w, respectively. This also implies that c_w^1 and c_w^2 define the formal processes w^1 and w^2, which are the formal drift and diffusion integrands of w, respectively. A formal input here can be understood as an exogenous input process being generated from a exosystem where $w = 0$, i.e., $u = v = 0$.

In Example IV.4.1, it was observed that the output of a Fliess operator associated with $c \in \mathbb{R}\langle\langle X_0 Y_0\rangle\rangle$ can be formally written in Itô form as

$$\begin{aligned} y(t) &= F_c[w](t) \\ &= (c, \emptyset) + \int_0^t \left(F_{x_0^{-1}(c)}[w](s) + \frac{1}{2} F_{y_0^{-1}(y_0^{-1}(c))}[w](s) \right) ds \\ &\quad + \int_0^t F_{y_0^{-1}(c)}[w](s)\, dW(s), \end{aligned} \tag{V.2.4}$$

and when c is globally convergent y is an L_2-Itô process. According to this representation, there exists a relationship between formal processes and the set $\widetilde{\mathcal{UV}}^m[0,T]$ as described below.

Theorem V.2.1. *Let $w \in \mathscr{W}$. If c_w is globally convergent then $w \in \widetilde{\mathcal{UV}}^m[0,T]$. On the other hand, if w is ordered in the sense that*

$$w(t) = \sum_{j=0}^{\infty} \sum_{k=0}^{j} \sum_{\eta \in X_0^k Y_0^{j-k}} (c_w, \eta) E_\eta[0](t), \tag{V.2.5}$$

*and c_w is locally convergent then $w \in \widetilde{\mathcal{UV}}^m[0,\tau_R]$, where $\tau_R = \inf\{t \in [0,T] : |W(t)| = R\}$. In particular, if c_w is exchangeable and locally convergent, then $w \in \widetilde{\mathcal{UV}}^m[0,\tau_R]$ regardless of the ordering implied in (*V.2.5*).*

Proof: Since w is a formal process, it is associated to a series $c_w \in \mathbb{R}^m\langle\langle X_0Y_0\rangle\rangle$. Next, similar to (V.2.4), w can be formally written in Itô form. The same can be done for the formal drift and diffusion integrands of w associated with c_w^1 and c_w^2, respectively. Now, if the formal inputs w^1 and w^2 correspond to c_w^1 and c_w^2, respectively, and $\max\{|(c_w^1,\eta)|,|(c_w^2,\eta)|\} \leq KM^{|\eta|}$ for all $\eta \in X_0Y_0{}^*$, then, by Theorem IV.1.5 and Lemma IV.4.1, the formal process $w \in \widetilde{\mathcal{UV}}^m[0,T]$. On the other hand, if $\max\{|(c_w^1,\eta)|,|(c_w^2,\eta)|\} \leq KM^{|\eta|}\,|\eta|!$ for all $\eta \in X_0Y_0{}^*$ then by Theorem IV.3.1 and Lemma IV.4.1 there exist $R \in \mathbb{R}$ and a stopping time $\tau_R = \inf\{t \in [0,T] : |W(t)| = R\}$ such that $w \in \widetilde{\mathcal{UV}}^m[0,\tau_R]$ when w is taken as in (V.2.5). If c_w is an exchangeable series, the order in (V.2.5) is irrelevant. Thus $w \in \widetilde{\mathcal{UV}}^m[0,\tau_R]$. ■

In order to define a formal Fliess operator, an extension of the composition product definition has to be formulated over XY^*. Without loss of generality, any $\eta \in XY^*$ can be written as

$$\eta = \eta_k q_{i_k}^{l_k}\eta_{k-1}q_{i_{k-1}}^{l_{k-1}}\cdots\eta_1 q_{i_1}^{l_1}\eta_0, \tag{V.2.6}$$

where $\eta_i \in X_0Y_0{}^*$, $q_{i_j}^{l_j} = x_{i_j}$ when $l_j = 1$, $q_{i_j}^{l_j} = y_{i_j}$ when $l_j = 2$, and in both cases $i_j \in \{1,\ldots,m\}$. In addition, the set of words $\{\bar\eta_k,\ldots,,\bar\eta_0\}$ such that $\bar\eta_{j+1} = \eta_{j+1}q_{i_{j+1}}^{l_{j+1}}\bar\eta_j$ and $i_{j+1} \neq 0$ are called the *right factors* of η. The composition product of a series with a word is defined below.

Definition V.2.2. The **composition product** is

$$\eta \circ d = \begin{cases} \eta & : \quad |\eta|_{x_i, y_i} = 0, \ \forall\, i \neq 0 \\ \eta' q_0^j [d_i^j \sqcup\!\sqcup (\bar{\eta} \circ d)] & : \quad \eta = \eta' q_i^j \bar{\eta}, \ i \neq 0, \ j \in \{1, 2\}, \ \eta' \in X_0 {Y_0}^*, \end{cases}$$

where $\bar{\eta} \in XY^*$ is a right factor, $d_i^j : \xi \mapsto (d, \xi)_i^j$, $(d, \xi)_i^j$ is the ith component of $(d, \xi)^j$, $j = 1$ represents drift coefficients and $j = 2$ represents diffusion coefficients. Furthermore, the composition of a series $c \in \mathbb{R}^\ell\langle\langle XY \rangle\rangle$ with d is

$$c \circ d = \sum_{\eta \in XY^*} (c, \eta) \eta \circ d.$$

It is important to observe that if $Y = \emptyset$, then this composition product reduces to the usual composition product on $\mathbb{R}^m\langle\langle X \rangle\rangle$. The theorem below says that the composition product on XY^* is summable.

Theorem V.2.2. *Let $d \in \mathbb{R}^m\langle\langle XY \rangle\rangle$; then the family of series $\{\eta \circ d \ : \ \eta \in XY^*\}$ is locally finite and therefore summable.*

Proof: Given that any word in XY^* can be written as in (V.2.6), it follows that

$$\mathrm{ord}(\eta \circ d) = n_0 + k + \sum_{j=1}^{k} n_j + \mathrm{ord}(d_{i_j}^{l_j}) = |\eta| + \sum_{j=1}^{|\eta| - |\eta|_{x_0, y_0}} \mathrm{ord}(d_{i_j}^{l_j}).$$

Thus, for any $\xi \in XY^*$,

$$\begin{aligned} I_d(\xi) &\triangleq \{\eta \in XY^* \ : \ (\eta \circ d) \neq 0\} \\ &\subset \{\eta \in XY^* \ : \ \mathrm{ord}(\eta \circ d) \leq |\xi|\} \\ &= \left\{ \eta \in XY^* \ : \ |\eta| + \sum_{j=1}^{|\eta| - |\eta|_{x_0, y_0}} \mathrm{ord}(d_{i_j}^{l_j}) \leq |\xi| \right\}. \end{aligned}$$

One can see that the latter set is finite, ensuring $I_d(\xi)$ is finite, which in turn implies summability. ■

Now that all the necessary concepts are available, the definition of a formal Fliess operator can be stated.

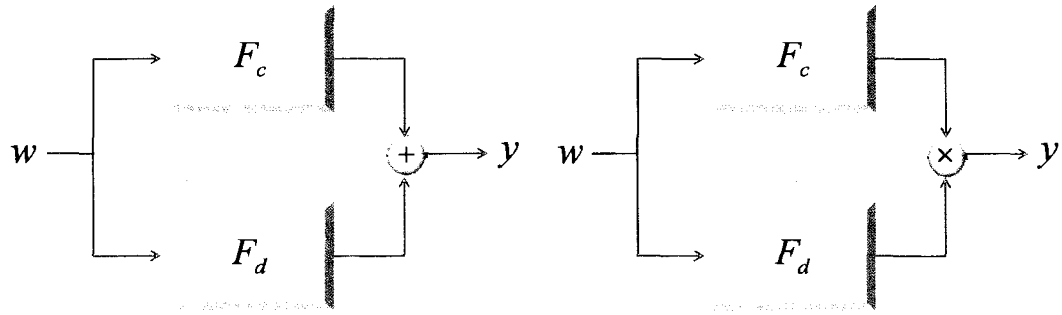

Fig. 9: Parallel and product connections of Fliess operators.

Definition V.2.3. The class of **formal Fliess operators** on $\mathbb{R}^m\langle\langle X_0Y_0\rangle\rangle$ is the collection of mappings

$$\mathscr{F} \triangleq \{c\circ : \mathbb{R}^m\langle\langle X_0Y_0\rangle\rangle \to \mathbb{R}^\ell\langle\langle X_0Y_0\rangle\rangle : c_w \mapsto c_y = c\circ c_w, c \in \mathbb{R}^\ell\langle\langle XY\rangle\rangle\}.$$

The operator $c\circ$ is a *formal operator* in that it acts on a formal input, i.e., one that has a series representation.

V.2.2 The parallel, product and cascade interconnections of formal Fliess operators

The parallel and product connections are studied first. It is helpful to recall the meaning of these interconnections. Consider two input-output maps F_c and F_d (see Figure 9) both driven by the input w and having outputs y_c, y_d, respectively. The parallel connection is simply the addition of the two outputs, i.e.,

$$y = y_c + y_d = F_c[w] + F_d[w].$$

The product connection is the multiplication of the two outputs, i.e.,

$$y = y_c \cdot y_d = F_c[w] \cdot F_d[w].$$

Using Definitions IV.2.1 and IV.2.3, one can characterize the parallel and product connections assuming inputs from $\mathscr{W}$.

Theorem V.2.3. *Let* $c, d \in \mathbb{R}^\ell\langle\langle XY\rangle\rangle$*. Then*

$$c \circ c_w + d \circ c_w = (c + d) \circ c_w, \quad \forall c_w \in \mathbb{R}\langle\langle X_0 Y_0\rangle\rangle.$$

Proof: The parallel connection is the addition of two formal Fliess operators. Therefore, from the left linearity of the composition product,

$$c \circ c_w + d \circ c_w \quad = \quad (c + d) \circ c_w.$$

■

Given that the formal Fliess operators $c\circ$ and $d\circ$ both act on the series c_w, the product connection of formal Fliess operators is represented by the shuffle product of the output series, i.e., $(c \circ c_w) \sqcup\!\sqcup (d \circ c_w)$.

Theorem V.2.4. *Let* $c, d \in \mathbb{R}^\ell\langle\langle XY\rangle\rangle$*. Then*

$$(c \circ c_w) \sqcup\!\sqcup (d \circ c_w) = (c \sqcup\!\sqcup d) \circ c_w, \quad \forall c_w \in \mathbb{R}\langle\langle X_0 Y_0\rangle\rangle. \tag{V.2.7}$$

Proof: It is known that the product of two iterated integrals is represented by the shuffle product of words in XY^*. Note that

$$(c \sqcup\!\sqcup d) \circ e = \sum_{\eta,\xi \in XY^*} (c, \eta)(d, \xi)\,(\eta \sqcup\!\sqcup \xi) \circ e, \tag{V.2.8}$$

and

$$(c \circ e) \sqcup\!\sqcup (d \circ e) = \sum_{\eta,\xi \in XY^*} (c,\eta)(d,\xi) \left((\eta \circ e) \sqcup\!\sqcup (\xi \circ e)\right). \tag{V.2.9}$$

Therefore, it is sufficient to show that

$$(\eta \sqcup\!\sqcup \xi) \circ e = (\eta \circ e) \sqcup\!\sqcup (\xi \circ e) \tag{V.2.10}$$

for all $\eta, \xi \in XY^*$. This is proven by induction over $|\eta| + |\xi| = n$. If $|\eta| + |\xi| = 0$ then the identity is satisfied trivially. Consider $\eta = q_{j_1}^{l_1}\eta'$, $\xi = q_{j_2}^{l_2}\xi'$, and $\eta', \xi' \in XY^*$ such that $|\eta| + |\xi'| = n - 1$ and $|\eta'| + |\xi| = n - 1$. Now assume that (V.2.10) holds up to $n - 1$, and calculate the identity for $|\eta| + |\xi| = n$. Then

$$(\eta \sqcup\!\sqcup \xi) \circ e = \left(q_{j_1}^{l_1}(\eta' \sqcup\!\sqcup \xi)\right) \circ e + \left(q_{j_2}^{l_2}(\eta \sqcup\!\sqcup \xi')\right) \circ e.$$

By the inductive step and Definitions IV.2.3, V.2.2, if $j_1 = j_2 = 0$ then

$$\begin{aligned}
(\eta \sqcup\!\sqcup \xi) \circ e &= q_0^{l_1}\left((\eta' \sqcup\!\sqcup \xi) \circ e\right) + q_0^{l_2}\left((\eta \sqcup\!\sqcup \xi') \circ e\right) \\
&= q_0^{l_1}\left((\eta' \circ e) \sqcup\!\sqcup (\xi \circ e)\right) + q_0^{l_2}\left((\eta \circ e) \sqcup\!\sqcup (\xi' \circ e)\right) \\
&= q_0^{l_1}\left((\eta' \circ e) \sqcup\!\sqcup q_0^{l_2}(\xi' \circ e)\right) + q_0^{l_2}\left(q_0^{l_1}(\eta' \circ e) \sqcup\!\sqcup (\xi' \circ e)\right) \\
&= \left(q_0^{l_1}(\eta' \circ e)\right) \sqcup\!\sqcup \left(q_0^{l_2}(\xi' \circ e)\right) \\
&= (\eta \circ e) \sqcup\!\sqcup (\xi \circ e).
\end{aligned}$$

If $j_1 = 0$ and $j_2 \neq 0$ then, by the commutativity of the shuffle product,

$$\begin{aligned}
(\eta \sqcup\!\sqcup \xi) \circ e &= q_0^{l_1}\left((\eta' \sqcup\!\sqcup \xi) \circ e\right) + q_0^{l_2}\left(d_{j_2}^{l_2} \sqcup\!\sqcup \left((\eta \sqcup\!\sqcup \xi') \circ e\right)\right) \\
&= q_0^{l_1}\left((\eta' \circ e) \sqcup\!\sqcup (\xi \circ e)\right) + q_0^{l_2}\left(d_{j_2}^{l_2} \sqcup\!\sqcup \left((\eta \circ e) \sqcup\!\sqcup (\xi' \circ e)\right)\right) \\
&= q_0^{l_1}\left((\eta' \circ e) \sqcup\!\sqcup q_0^{l_2}\left(d_{j_2}^{l_2} \sqcup\!\sqcup (\xi' \circ e)\right)\right) \\
&\quad + q_0^{l_2}\left(q_0^{l_1}(\eta' \circ e) \sqcup\!\sqcup d_{j_2}^{l_2} \sqcup\!\sqcup (\xi' \circ e)\right)
\end{aligned}$$

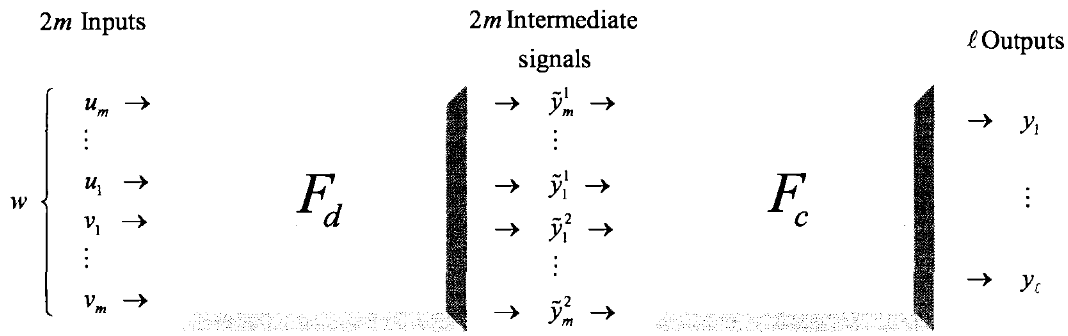

Fig. 10: Cascade connection of Fliess operators.

$$
\begin{aligned}
&= q_0^{l_1}(\eta' \circ e) \sqcup\!\sqcup q_0^{l_2}\left(d_{j_2}^{l_2} \sqcup\!\sqcup (\xi' \circ e)\right) \\
&= (\eta \circ e) \sqcup\!\sqcup (\xi \circ e).
\end{aligned}
$$

If $j_1 = j_2 \neq 0$ then, by again the commutativity of the shuffle product,

$$
\begin{aligned}
(\eta \sqcup\!\sqcup \xi) \circ e &= q_0^{l_1}\left(d_{j_1}^{l_1} \sqcup\!\sqcup ((\eta' \sqcup\!\sqcup \xi) \circ e)\right) + q_0^{l_2}\left(d_{j_2}^{l_2} \sqcup\!\sqcup ((\eta \sqcup\!\sqcup \xi') \circ e)\right) \\
&= q_0^{l_1}\left(d_{j_1}^{l_1} \sqcup\!\sqcup ((\eta' \circ e) \sqcup\!\sqcup (\xi \circ e))\right) + q_0^{l_2}\left(d_{j_2}^{l_2} \sqcup\!\sqcup ((\eta \circ e) \sqcup\!\sqcup (\xi' \circ e))\right) \\
&= q_0^{l_1}\left(d_{j_1}^{l_1} \sqcup\!\sqcup (\eta' \circ e) \sqcup\!\sqcup q_0^{l_2}\left(d_{j_2}^{l_2} \sqcup\!\sqcup (\xi' \circ e)\right)\right) \\
&\quad + q_0^{l_2}\left(q_0^{l_1}\left(d_{j_1}^{l_1} \sqcup\!\sqcup (\eta' \circ e)\right) \sqcup\!\sqcup d_{j_2}^{l_2} \sqcup\!\sqcup (\xi' \circ e)\right) \\
&= \left(q_0^{l_1}\left(d_{j_1}^{l_1} \sqcup\!\sqcup (\eta' \circ e)\right)\right) \sqcup\!\sqcup \left(q_0^{l_2}\left(d_{j_2}^{l_2} \sqcup\!\sqcup (\xi' \circ e)\right)\right) \\
&= (\eta \circ e) \sqcup\!\sqcup (\xi \circ e).
\end{aligned}
$$

This result, along with (V.2.8) and (V.2.9), completes the proof. ∎

The cascade connection needs a special treatment. It is represented by the composite system, $F_c \circ F_d$, with input u_d and output y_c (see Figure 10). In other words,

the output of the system F_d is connected directly to the input of the system F_c, i.e.,

$$y_c = (F_c \circ F_d)\,[u_d] = F_c[F_d[u_d]].$$

Since these input-output maps have been described as functional expansions, the cascade connection will follow the rules of the usual functional composition.

Observe in Figure 10 that when two Fliess operators are connected in cascade fashion, an intermediate signal, $\tilde{y}$, is generated. In order to properly establish this interconnection, it will be desirable that the intermediate signal behaves as an input of the next system. Moreover, $\tilde{y}$ needs to have $2m$ components since it is assumed that the second system has $2m$ inputs. That is, one is expecting two sets of outputs: one set to serve as drift inputs and the other to serve as diffusion inputs for F_c. In other words, the associated series d must be in $d \in \mathbb{R}^{2m}\langle\langle XY\rangle\rangle$. For example, if $\tilde{y}$ has ℓ components then $\tilde{y}$ is divided, depending on how noisy are its components, into two groups: m_1 drift inputs and m_2 diffusion inputs labeled with superscripts 1 and 2, respectively. Since one desires $\ell = 2m$, then $m = \max\{m_1, m_2\}$ and the group with fewer components is completed with zeros. For the cascade of formal Fliess operators, the compatibility between inputs and outputs is not a problem given that by Definition V.2.3, $c_w \in \mathbb{R}^m\langle\langle X_0Y_0\rangle\rangle$ and $c_{\tilde{y}} \in \mathbb{R}^{2m}\langle\langle X_0Y_0\rangle\rangle$ when $\ell = 2m$ is chosen. Therefore, $c_{\tilde{y}}$ can be the input of a second formal Fliess operator. Using Definition V.2.2, one can characterize the cascade connection assuming inputs from $\mathscr{W}$.

Theorem V.2.5. *Let $c \in \mathbb{R}^\ell\langle\langle X_0Y_0\rangle\rangle$ and $d \in \mathbb{R}^{2m}\langle\langle XY\rangle\rangle$. Then*

$$c \circ (d \circ c_w) = (c \circ d) \circ c_w, \quad \forall c_w \in \mathbb{R}^{2m}\langle\langle X_0Y_0\rangle\rangle.$$

Proof: First, observe that

$$c \circ (d \circ c_w) = \sum_{\eta \in XY^*} (c, \eta) \eta \circ (d \circ c_w).$$

Therefore, it is sufficient to show

$$(\eta \circ d) \circ c_w = \eta \circ (d \circ c_w) \tag{V.2.11}$$

for all $\eta \in XY^*$. Now consider that any word in XY^* can be written as in (V.2.6). The proof is done by induction over the numbers of right factors in η, said $\{\bar{\eta}_k, \bar{\eta}_{k-1} \ldots, \bar{\eta}_0\}$. If $k = 0$ then $\eta = \bar{\eta}_0 \in X_0 {Y_0}^*$. By Definition V.2.2,

$$(\bar{\eta}_0 \circ d) \circ c_w = \bar{\eta}_0 \circ c_w = \bar{\eta}_0 = \bar{\eta}_0 \circ d \circ c_w.$$

Now suppose identity (V.2.11) holds up to the k-th right factor of η. Using (V.2.7), for $\bar{\eta}_{k+1} = \eta_{k+1} q_{i_{k+1}}^{l_{k+1}} \bar{\eta}_k$,

$$\begin{aligned}
(\bar{\eta}_{k+1} \circ d) \circ c_w &= \left(\eta_{k+1} q_{i_{k+1}}^{l_{k+1}} \bar{\eta}_k \circ d \right) \circ c_w \\
&= \left(\eta_{k+1} q_0^{l_{k+1}} \left[d_{i_{k+1}}^{l_{k+1}} \sqcup\!\sqcup (\bar{\eta}_k \circ d) \right] \right) \circ c_w \\
&= \eta_{k+1} q_0^{l_{k+1}} \left(\left[d_{i_{k+1}}^{l_{k+1}} \sqcup\!\sqcup (\bar{\eta}_k \circ d) \right] \circ c_w \right) \\
&= \eta_{k+1} q_0^{l_{k+1}} \left(\left(d_{i_{k+1}}^{l_{k+1}} \circ c_w \right) \sqcup\!\sqcup ((\bar{\eta}_k \circ d) \circ c_w) \right) \\
&= \eta_{k+1} q_0^{l_{k+1}} \left(\left(d^{l_{k+1}} \circ c_w \right)_{i_{k+1}} \sqcup\!\sqcup (\bar{\eta}_k \circ (d \circ c_w)) \right) \\
&= \eta_{k+1} q_{i_{k+1}}^{l_{k+1}} \bar{\eta}_k \circ (d \circ c_w) \\
&= \bar{\eta}_{k+1} \circ (d \circ c_w) .
\end{aligned}$$

Hence,

$$\begin{aligned}
c \circ (d \circ c_w) &= \sum_{\eta \in XY^*} (c, \eta) \eta \circ (d \circ c_w) \\
&= \sum_{\eta \in XY^*} (c, \eta) (\eta \circ d) \circ c_w \\
&= (c \circ d) \circ c_w,
\end{aligned}$$

which completes the proof. ■

Generally speaking, Theorems V.2.4 and V.2.5 show that the composition product over $\mathbb{R}\langle\langle XY\rangle\rangle$ is distributive from the right with respect to the shuffle product and associative. These two properties were originally proven by Ferfera in [13] for the composition product over $\mathbb{R}^{\ell}\langle\langle X\rangle\rangle$. In conclusion, the parallel, product and cascade connections of formal Fliess operators are characterized by the operations $+$, $\sqcup\!\sqcup$, and $\circ$ on $\mathbb{R}\langle\langle XY\rangle\rangle$. In the next section, these three operations are going to be related to the interconnections of Fliess operators with inputs from $\mathcal{UV}^m[0,T]$.

V.2.3 Convergence of the interconnections of Fliess Operators

In light of (V.2.1) and (V.2.2), one needs more than an algebraic notion of well-posed of an interconnection when inputs from $\mathcal{UV}^m[0,T]$ are considered. For example, the cascade connection requires the convergence of $w, \tilde{y}$ and y in addition to $\tilde{y}$ belonging to $\mathcal{UV}^m[0,T]$. Unfortunately, from Corollary IV.4.1, it has been shown that $\tilde{y}$ resides in $\widetilde{\mathcal{UV}}^m[0,T]$, which is a set including $\mathcal{UV}^m[0,T]$. However, one can utilize the concepts of rationality and local convergence of formal power series to develop situations where the interconnections of Fliess operators produce a well-defined output process. This section begins with the description of the interconnections of Fliess operators with inputs from $\mathcal{UV}^m[0,T]$.

Theorem V.2.6. *Let $c, d \in \mathbb{R}^{\ell}\langle\langle XY\rangle\rangle$ be locally convergent series and $w \in \mathcal{UV}^m[0,T]$. Then*

$$\begin{aligned} F_c[w] + F_d[w] &= F_{c+d}[w], \\ F_c[w] \cdot F_d[w] &= F_{c \sqcup\!\sqcup d}[w]. \end{aligned}$$

Proof: The parallel connection is the addition of two Fliess operators. Therefore,

$$\begin{aligned}
F_c[w](t) + F_d[w](t) &= \sum_{\eta \in XY^*} (c,\eta) E_\eta[w](t) + \sum_{\eta \in XY^*} (d,\eta) E_\eta[w](t) \\
&= \sum_{\eta \in XY^*} (c+d,\eta) E_\eta[w](t) \\
&= F_{c+d}[w](t).
\end{aligned}$$

The product connection is the product of two Fliess operators. Therefore

$$\begin{aligned}
F_c[w](t) \cdot F_d[w](t) &= \sum_{\eta \in XY^*} (c,\eta) E_\eta[w](t) \cdot \sum_{\xi \in XY^*} (d,\xi) E_\xi[w](t) \\
&= \sum_{\eta,\xi \in XY^*} (c,\eta)(d,\xi) E_{\eta \sqcup\!\sqcup \xi}[w](t) \\
&= F_{c \sqcup\!\sqcup d}[w](t).
\end{aligned}$$

■

Here, the minimum requirements from Theorem IV.3.1 for the convergence of $F_c[w]$ and $F_d[w]$ have been assumed. Hence, $F_{c+d}[w]$ and $F_{c \sqcup\!\sqcup d}[w]$ are infinite summations of well-defined iterated integrals, since $E_{\eta \sqcup\!\sqcup \xi}[w]$ is well-defined for any $w \in \mathcal{UV}^m[0,T]$. Note that the generating series of the parallel and product connections of Fliess operators on $\mathcal{UV}^m[0,T]$ and formal Fliess operators coincide. To study the convergence of $F_{c+d}[w]$ and $F_{c \sqcup\!\sqcup d}[w]$, it is necessary to study the properties of $c+d$ and $c \sqcup\!\sqcup d$ on $\mathbb{R}^\ell\langle\langle XY \rangle\rangle$. The next theorem is a direct consequence of its counterpart for series in $\mathbb{R}\langle\langle X \rangle\rangle$ [24, 63].

Theorem V.2.7. *If $c, d \in \mathbb{R}^\ell\langle\langle XY \rangle\rangle$ are globally convergent then $c+d$ and $c \sqcup\!\sqcup d$ are globally convergent. Moreover, if $c, d \in \mathbb{R}^\ell\langle\langle XY \rangle\rangle$ are locally convergent then $c+d$ and $c \sqcup\!\sqcup d$ are locally convergent.*

It then follows in the next corollary that the parallel and product connections are

well-defined and produce a well-defined L_2-Itô process over $[0,T]$, $T>0$, when the generating series involved are globally convergent.

Corollary V.2.1. *Let* $c \in \mathbb{R}^{\ell}\langle\langle XY\rangle\rangle$ *and* $d \in \mathbb{R}^{\ell}\langle\langle XY\rangle\rangle$ *be globally convergent series. For any* $w \in \mathcal{UV}^m[0,T]$*, the operators* F_{c+d} *and* $F_{c \sqcup\!\sqcup\, d}$ *produce well-defined* L_2*-Itô processes over* $[0,T]$.

Proof: Since c and d are globally convergent, by Theorem V.2.7, $c+d$ and $c \sqcup\!\sqcup\, d$ are globally convergent. Thus, by Theorem IV.1.5 and Corollary IV.4.1, the operators F_{c+d} and $F_{c \sqcup\!\sqcup\, d}$ converge and produce well-defined L_2-Itô output processes over $[0,T]$, $T>0$. ∎

For c and d locally convergent series, $F_{c+d}[w]$ and $F_{c \sqcup\!\sqcup\, d}[w]$ also converge conditionally over a nonzero stochastic time interval.

Corollary V.2.2. *Let* $c, d \in \mathbb{R}^{\ell}\langle\langle XY\rangle\rangle$ *be locally convergent series. For any* $w \in \mathcal{UV}^m[0,T]$*, there exist an* $R>0$ *and a stopping time* τ_R *such that* F_{c+d} *and* $F_{c \sqcup\!\sqcup\, d}$*, respectively, produce* L_2*-Itô processes over* $[0,\tau_R]$ *assuming the order of summation defined in* (IV.3.5).

Proof: Since c and d are locally convergent, by Theorem V.2.7, $c+d$ and $c \sqcup\!\sqcup\, d$ are locally convergent. Then, by Theorem IV.3.1, the operators F_{c+d} and $F_{c \sqcup\!\sqcup\, d}$ define L_2-Itô processes over the stochastic interval of time $[0,\tau_R]$, where τ_R is defined in (IV.3.4) and when the series defining F_{c+d} and $F_{c \sqcup\!\sqcup\, d}$ are added in the order described in (IV.3.5). ∎

Observe that if $c+d$ and $c \sqcup\!\sqcup\, d$ were exchangeable series then Corollary V.2.2 is valid unconditionally. In particular, the local convergence of the parallel connection

of Fliess operators is well-posed for c and d exchangeable series.

Corollary V.2.3. *Let $c, d \in \mathbb{R}^\ell\langle\langle XY\rangle\rangle$ be locally convergent and exchangeable. For any $w \in \mathcal{UV}^m[0,T]$, the Fliess operator $F_{c+d}[w]$ converges over a stochastic interval of time regardless of the order introduced in* (IV.3.5).

Proof: From Corollaries IV.3.1 and V.2.2, it is only sufficient to show that $c+d$ is exchangeable. Fix $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^{m+1}$; then, from Definition IV.3.2, $c_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ represent the image under c for words in $\boldsymbol{L}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ and $d_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ represent the image under d for words in $\boldsymbol{L}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$. It is easy to see that $(c+d)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = c_{\boldsymbol{\alpha},\boldsymbol{\beta}} + d_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ for all $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^{m+1}$. Thus, $c+d$ is exchangeable. ■

The cascade connection is again treated again separately. Below, the relationship between the cascade connection of Fliess operators on $\mathcal{UV}^m[0,T]$ and the composition product on $\mathbb{R}\langle\langle XY\rangle\rangle$ is obtained.

Theorem V.2.8. *Let $c \in \mathbb{R}^\ell\langle\langle XY\rangle\rangle$ and $d \in \mathbb{R}^{2m}\langle\langle XY\rangle\rangle$ be locally convergent, and $w \in \mathcal{UV}^m[0,T]$. Then*

$$(F_c \circ F_d)\,[w] = F_{c\circ d}[w].$$

Proof: Since $F_d[w]$ is a convergent operator, the iterated integrals in F_c defined using the components of $F_d[w]$ as integrands are well-defined. It is shown first that $F_\eta \circ F_d = F_{\eta\circ d}$ for any $\eta \in XY^*$ using induction over $k = |\eta|_{x_i} + |\eta|_{y_i}$, $i \in \{1,\ldots,m\}$. Let $\eta \in XY^*$ be written as in (V.2.6) and consider the set of right factors $\{\bar{\eta}_k, \bar{\eta}_{k-1}, \ldots, \bar{\eta}_0\}$ of η. Clearly,

$$(F_{\bar{\eta}_0} \circ F_d[w])\,(t) = E_{\bar{\eta}_0}[w](t) = F_{\bar{\eta}_0}[w](t) = F_{\bar{\eta}_0\circ d}[w](t).$$

Now assume that $(F_{\bar{\eta}_k} \circ F_d[u])(t) = F_{\bar{\eta}_k \circ d}[u](t)$ holds up to some fixed $k > 0$. Then by Definition V.2.2

$$
\begin{aligned}
&(F_{\bar{\eta}_{j+1}} \circ F_d[w])(t) \\
&= E_{\eta_{j+1} q_{i_{j+1}}^{l_{j+1}} \bar{\eta}_j}[F_d[w]](t) \\
&= \underbrace{\int_0^t \cdots \int_0^{t_2}}_{n_{j+1}+1 \text{ times}} F_{d_{i_{j+1}}^{l_{j+1}}}[u](t_1) E_{\bar{\eta}_j}[F_d[w]](t_1)\, dq_{i_{j+1}}^{l_{j+1}}(t_1) \cdots dq_{i_{n_{j+1}+1}}^{l_{n_{j+1}+1}}(t_{n_{j+1}+1}) \\
&= \underbrace{\int_0^t \cdots \int_0^{t_2}}_{n_{j+1}+1 \text{ times}} F_{d_{i_{j+1}}^{l_{j+1}} \sqcup\!\sqcup (\bar{\eta}_j \circ d)}[w](t_1)\, dq_{i_{j+1}}^{l_{j+1}}(t_1) \cdots dq_{i_{n_{j+1}+1}}^{l_{n_{j+1}+1}}(t_{n_{j+1}+1}) \\
&= F_{\eta_{j+1}\left[d_{i_{j+1}}^{l_{j+1}} \sqcup\!\sqcup (\bar{\eta}_j \circ d)\right]}[w](t) \\
&= F_{\eta_{j+1} \circ d}[w](t).
\end{aligned}
$$

Finally,

$$
\begin{aligned}
(F_c \circ F_d[w])(t) &= \sum_{\eta \in XY^*} (c, \eta) E_\eta[F_d[w]](t) = \sum_{\eta \in XY^*} (c, \eta) F_{\eta \circ d}[w](t) \\
&= \sum_{\eta \in XY^*} (c, \eta) \left[\sum_{\nu \in XY^*} (\eta \circ d, \nu) E_\nu[w](t) \right] \\
&= \sum_{\eta \in XY^*} \left[\sum_{\nu \in XY^*} (c, \eta)(\eta \circ d, \nu) \right] E_\nu[w](t) \\
&= \sum_{\eta \in XY^*} (c \circ d, \nu) E_\nu[w](t) \\
&= F_{c \circ d}[w](t).
\end{aligned}
$$

Note that the generating series of the cascade connection of Fliess operators in $\mathcal{UV}^m[0,T]$ and formal Fliess operators coincide. ∎

Even though $F_c[w]$ and $F_d[w]$ are convergent, the composite operator $F_{c \circ d}[w]$ is just an infinite summation of iterated integrals. In addition to globally and locally convergent formal power series, the concept of rationality plays an important role

in the convergence analysis of the cascade of Fliess operators on $\mathcal{UV}^m[0,T]$. For global convergence, an analogue to Ferfera's condition presented in Theorem III.2.5 is needed for series in $\mathbb{R}\langle\langle XY\rangle\rangle$. Fortunately, the rationality of the composition product on $\mathbb{R}\langle\langle XY\rangle\rangle$ can be treated in a manner completely analogous to the case for $\mathbb{R}\langle\langle X\rangle\rangle$. From Definitions III.2.5 and V.2.2, the definition of input limited can be directly extended to series in $\mathbb{R}\langle\langle XY\rangle\rangle$.

Definition V.2.4. A series $c \in \mathbb{R}\langle\langle XY\rangle\rangle$ is **limited relative to x_i and y_i** if there exists an integer $\mathcal{N}_i \geq 0$ such that

$$\sup_{\eta\in\mathrm{supp}(c)} |\eta|_{x_i,y_i} = \mathcal{N}_i < \infty.$$

If c is limited relative to x_i and y_i for every $i = 1,\ldots,m$ then c is said to be **input-limited**. In such cases, let $\mathcal{N}_c := \max_i \mathcal{N}_i$. A series $c \in \mathbb{R}^\ell\langle\langle XY\rangle\rangle$ is input-limited if each component series, c_j, is input-limited for $j = 1,\ldots,\ell$. In this case, $\mathcal{N}_c := \max_j \mathcal{N}_{c_j}$.

Theorem V.2.9. *Let $c \in \mathbb{R}^\ell\langle\langle XY\rangle\rangle$ and $d \in \mathbb{R}^{2m}\langle\langle XY\rangle\rangle$. If c and d are rational series then $c\circ d$ is rational if c is input-limited.*

A slight extension of Lemma III.2.1 is presented next.

Lemma V.2.1. *Let $c \in \mathbb{R}^\ell\langle\langle XY\rangle\rangle$ be a rational series with a linear representation (μ,γ,λ). Let $N_i^1 \triangleq \mu(x_i) \in \mathbb{R}^{n\times n}$ and $N_i^2 \triangleq \mu(y_i) \in \mathbb{R}^{n\times n}$, $i = 0,1,\ldots,m$. Then for any $d \in \mathbb{R}^{2m}\langle\langle XY\rangle\rangle$ it follows that*

$$c\circ d = \sum_{\eta\in\widehat{XY}^*} \lambda D_\eta((N_0^1 x_0 + N_0^2 y_0)^*)\gamma,$$

where $\widehat{XY} \triangleq \{x_1, \ldots, x_m, y_1, \ldots, y_m\}$, *and the set of operators* $\{D_\eta : \eta \in \widehat{XY}^*\}$ *is the monoid under composition uniquely specified by*

$$D_{q_i^j} : \mathbb{R}^{n\times n}\langle\langle XY\rangle\rangle \to \mathbb{R}^{n\times n}\langle\langle XY\rangle\rangle : E \mapsto (N_0^1 x_0 + N_0^2 y_0)^* q_0^j N_i^j (d_i^j \sqcup\!\sqcup E)$$

with $D_\emptyset$ *equivalent to the identity map. (The shuffle product above is defined componentwise.)*

Proof: Without loss of generality, assume $\ell = 1$. Define $N_\eta = N_0^{l_k} \cdots N_0^{l_1}$ for $\eta = q_0^{l_k} \cdots q_0^{l_1} \in XY^*$. Directly from the definition of the composition product observe that

$$\begin{aligned} c \circ d &= \sum_{k\geq 0} \sum_{\substack{i_1,\ldots,i_k=1 \\ l_1,\ldots,l_k=1}}^{m,2} \sum_{\eta_0,\ldots,\eta_k \in X_0 Y_0^*} \lambda N_{\eta_k} N_{i_k}^{l_k} N_{\eta_{k-1}} N_{i_{k-1}}^{l_{k-1}} \cdots N_{\eta_1} N_{i_1}^{l_1} N_{\eta_0}^{n_0} \gamma \cdot \\ &\qquad \eta_k q_{i_k}^{l_k} \eta_{k-1} q_{i_{k-1}}^{l_{k-1}} \cdots \eta_1 q_{i_1}^{l_1} \eta_0 \circ d \\ &= \sum_{k\geq 0} \sum_{\substack{i_1,\ldots,i_k=1 \\ l_1,\ldots,l_k=1}}^{m,2} \sum_{\eta_0,\ldots,\eta_k \in X_0 Y_0^*} \lambda N_{\eta_k} N_{i_k}^{l_k} N_{\eta_{k-1}} N_{i_{k-1}}^{l_{k-1}} \cdots N_{\eta_1} N_{i_1}^{l_1} N_{\eta_0}^{n_0} \gamma \cdot \\ &\qquad \eta_k q_0^{l_k} \left[d_{i_k}^{l_k} \sqcup\!\sqcup \left[\eta_{k-1} q_0^{l_{k-1}} \left[d_{i_{k-1}}^{l_{k-1}} \sqcup\!\sqcup \cdots \sqcup\!\sqcup \left[\eta_1 q_0^{l_1} \left[d_{i_1}^{l_1} \sqcup\!\sqcup \eta_0\right]\right] \cdots \right]\right]\right]. \end{aligned}$$

From the bilinearity and continuity of the shuffle product (in the ultrametric sense), it follows that

$$\begin{aligned} c \circ d &= \sum_{k\geq 0} \sum_{\substack{i_1,\ldots,i_k=1 \\ l_1,\ldots,l_k=1}}^{m,2} \lambda \left(\sum_{\eta_k \in X_0 Y_0^*} N_{\eta_k} \eta_k\right) q_0^{l_k} N_{i_k}^{l_k} \left[d_{i_k}^{l_k} \sqcup\!\sqcup \cdot \right. \\ &\qquad \left[\left(\sum_{\eta_{k-1} \in X_0 Y_0^*} N_{\eta_{k-1}} \eta_{k-1}\right) q_0^{l_{k-1}} N_{i_{k-1}}^{l_{k-1}} \left[d_{i_{k-1}}^{l_{k-1}} \sqcup\!\sqcup \cdots \sqcup\!\sqcup \right.\right. \\ &\qquad \left.\left.\left.\left[\left(\sum_{\eta_1 \in X_0 Y_0^*} N_{\eta_1} \eta_1\right) q_0^{l_1} N_{i_1}^{l_1} \left[d_{i_1}^{l_1} \sqcup\!\sqcup \left(\sum_{\eta_0 \in X_0 Y_0^*} N_{\eta_0} \eta_0\right)\right]\right] \cdots \right]\right]\right] \gamma \\ &= \sum_{k\geq 0} \sum_{\substack{i_1,\ldots,i_k=1 \\ l_1,\ldots,l_k=1}}^{m,2} \lambda (N_0^1 x_0 + N_0^2 y_0)^* q_0^{l_k} N_{i_k}^{l_k} \left[d_{i_k}^{l_k} \sqcup\!\sqcup \left[(N_0^1 x_0 + N_0^2 y_0)^* q_0^{l_{k-1}} N_{i_{k-1}}^{l_{k-1}} \cdot \right.\right. \end{aligned}$$

$$\left[d_{i_{k-1}}^{l_{k-1}} \sqcup\!\sqcup \cdots \sqcup\!\sqcup \left[(N_0^1 x_0 + N_0^2 y_0)^* q_0^{l_1} N_{i_1}^{l_1} \left[d_{i_1}^{l_1} \sqcup\!\sqcup (N_0^1 x_0 + N_0^2 y_0)^*\right]\right] \cdots \right]\right]\right] \gamma.$$

Finally, applying the definition of D_η,

$$\begin{aligned} c \circ d &= \sum_{k \geq 0} \sum_{q_{i_k}^{l_k} \cdots q_{i_1}^{l_1} \in \widehat{XY}^k} \lambda D_{q_{i_k}^{l_k}} D_{q_{i_{k-1}}^{l_{k-1}}} \cdots D_{q_{i_1}^{l_1}} ((N_0^1 x_0 + N_0^2 y_0)^*) \gamma \\ &= \sum_{\eta \in \widehat{XY}^*} \lambda D_\eta ((N_0^1 x_0 + N_0^2 y_0)^*) \gamma, \end{aligned}$$

and the lemma is proved. ■

Proof of Theorem V.2.9: Since c is input-limited, it follows from Lemma III.2.1 that

$$c \circ d = \sum_{k=0}^{N_c} \sum_{\eta \in \widehat{XY}^k} \lambda D_\eta ((N_0^1 x_0 + N_0^2 y_0)^*) \gamma.$$

Clearly, each operator D_η is mapping a rational series to another rational series as it involves only a finite number of rational operations. Therefore, for any integer $k \geq 0$ the formal power series

$$\sum_{\eta \in \widehat{XY}^k} \lambda D_\eta ((N_0^1 x_0 + N_0^2 y_0)^*) \gamma$$

is again rational since the summation is finite. Thus, $c \circ d$ must be rational. ■

By a direct application of Theorem V.2.9, a sufficient condition for the convergence of the cascade of two rational Fliess operators is described next.

Corollary V.2.4. *Let $c \in \mathbb{R}^\ell\langle\langle XY \rangle\rangle$ and $d \in \mathbb{R}^{2m}\langle\langle XY \rangle\rangle$ be rational series. If c is input-limited, then for any $w \in \mathcal{UV}^m[0,T]$, $T > 0$, the output of a cascade of two Fliess operators, $F_{c \circ d}[w]$, converges in the mean square sense to a well-defined L_2-Itô output process.*

Proof: Since c is input-limited, the series $c \circ d$ is rational by Theorem V.2.9. Then, by Theorem IV.1.5 and Corollary IV.4.1, the operator $F_c[F_d[w]] = F_{c \circ d}[w]$ defines a well-defined L_2-Itô output process. ■

One can use Theorem V.2.9 to establish a condition for the analyticity of $F_c[w]$ in the rational case.

Theorem V.2.10. *Consider an operator F_c with $c \in \mathbb{R}^\ell\langle\langle XY \rangle\rangle$ rational and input-limited. Select an input $w \in \mathcal{UV}^m[0,T]$ which is an analytic Itô process described by the rational series $c_w \in \mathbb{R}^m\langle\langle X_0 Y_0 \rangle\rangle$. Then the output process $y = F_c[w]$ is also analytic and has the generating series $c_y = c \circ c_w \in \mathbb{R}\langle\langle X_0 Y_0 \rangle\rangle$.*

Proof: Because c is input-limited, the series $c \circ c_w$ is rational. Then, by Corollaries IV.4.1, V.2.4 and Theorem IV.1.5, the process $y(t) = F_{c \circ c_w}[w](t)$ is a well-defined Itô output process for all $t \in [0,T]$ and has the series representation $c \circ c_w$. ■

Example V.2.1. Consider the input-limited series $c = (1 - x_0)^{-1}(x_1 + y_1)^2$ and $d = (1 - y_1)^{-1}$. From the last theorem, the composition product $c \circ d$ must be rational since c is input-limited, i.e., $\max\left\{ \max_{\eta \in XY^*}\{|\eta|_{x_1}\}, \max_{\eta \in XY^*}\{|\eta|_{y_1}\} \right\} \leq 2$. To confirm that $c \circ d$ is rational, using standard formal power series tools, the composition of c and d can be computed as

$$\begin{aligned}
c \circ d &= (1 - x_0)^{-1}(x_1 + y_1)^2 \circ (1 - y_1)^{-1} \\
&= (1 - x_0)^{-1}\left(x_1^2 + x_1 y_1 + y_1 x_1 + y_1^2\right) \circ (1 - y_1)^{-1} \\
&= (1 - x_0)^{-1} x_0 \left((1 - y_1)^{-1} \sqcup\!\sqcup (x_1 \circ (1 - y_1)^{-1})\right) \\
&\quad + (1 - x_0)^{-1} x_0 \left((1 - y_1)^{-1} \sqcup\!\sqcup (y_1 \circ (1 - y_1)^{-1})\right) \\
&\quad + (1 - x_0)^{-1} y_0 \left((1 - y_1)^{-1} \sqcup\!\sqcup (x_1 \circ (1 - y_1)^{-1})\right)
\end{aligned}$$

$$+(1-x_0)^{-1}y_0\left((1-y_1)^{-1} ⧢ (y_1 \circ (1-y_1)^{-1})\right)$$

$$\begin{aligned}
= \quad & (1-x_0)^{-1}x_0\left((1-y_1)^{-1} ⧢ (x_0(1-y_1)^{-1} ⧢ \left(\emptyset \circ (1-y_1)^{-1}\right)\right) \\
& +(1-x_0)^{-1}x_0\left((1-y_1)^{-1} ⧢ (y_0(1-y_1)^{-1} ⧢ \left(\emptyset \circ (1-y_1)^{-1}\right)\right) \\
& +(1-x_0)^{-1}y_0\left((1-y_1)^{-1} ⧢ (x_0(1-y_1)^{-1} ⧢ \left(\emptyset \circ (1-y_1)^{-1}\right)\right) \\
& +(1-x_0)^{-1}y_0\left((1-y_1)^{-1} ⧢ (y_0(1-y_1)^{-1} ⧢ \left(\emptyset \circ (1-y_1)^{-1}\right)\right) \\
= \quad & (1-x_0)^{-1}x_0\left((1-y_1)^{-1} ⧢ x_0(1-y_1)^{-1}\right) \\
& +(1-x_0)^{-1}x_0\left((1-y_1)^{-1} ⧢ y_0(1-y_1)^{-1}\right) \\
& +(1-x_0)^{-1}y_0\left((1-y_1)^{-1} ⧢ x_0(1-y_1)^{-1}\right) \\
& +(1-x_0)^{-1}y_0\left((1-y_1)^{-1} ⧢ y_0(1-y_1)^{-1}\right).
\end{aligned}$$

The latter summands can be determined directly as

$$\begin{aligned}
& q_0^j(1-y_1)^{-1} ⧢ (1-y_1)^{-1} \\
& \quad = \quad q_0^j(1-y_1)^{-1} ⧢ (1+y_1(1-y_1)^{-1}) \\
& \quad = \quad q_0^j(1-y_1)^{-1} + q_0^j(1-y_1)^{-1} ⧢ y_1(1-y_1)^{-1} \\
& \quad = \quad q_0^j(1-y_1)^{-1} + q_0^j\left((1-y_1)^{-1} ⧢ y_1(1-y_1)^{-1}\right) + y_1\left(q_0^j(1-y_1)^{-1} ⧢ (1-y_1)^{-1}\right) \\
& \quad = \quad q_0^j\left((1-y_1)^{-1} ⧢ \left(y_1(1-y_1)^{-1}+1\right)\right) + y_1\left(q_0^j(1-y_1)^{-1} ⧢ (1-y_1)^{-1}\right) \\
& \quad = \quad q_0^j\left((1-y_1)^{-1} ⧢ (1-y_1)^{-1}\right) + y_1\left(q_0^j(1-y_1)^{-1} ⧢ (1-y_1)^{-1}\right) \\
& \quad = \quad q_0^j(1-2y_1)^{-1} + y_1\left(q_0^j(1-y_1)^{-1} ⧢ (1-y_1)^{-1}\right).
\end{aligned}$$

Factoring $q_0^j(1-y_1)^{-1} ⧢ (1-y_1)^{-1}$ from the left-hand and right-hand sides gives

$$(1-q_0^j)\left(q_0^j(1-y_1)^{-1} ⧢ (1-y_1)^{-1}\right) \quad = \quad q_0^j(1-2y_1)^{-1}.$$

In which case,

$$q_0^j(1-y_1)^{-1} ⧢ (1-y_1)^{-1} = (1-y_1)^{-1}q_0^j(1-2y_1)^{-1}. \tag{V.2.12}$$

Using identity (V.2.12), the composition $c \circ d$ is computed as

$$\begin{aligned} c \circ d \;=\; & (1-x_0)^{-1}x_0\left((1-y_1)^{-1}x_0(1-2y_1)^{-1}\right) \\ & +(1-x_0)^{-1}x_0\left((1-y_1)^{-1}y_0(1-2y_1)^{-1}\right) \\ & +(1-x_0)^{-1}y_0\left((1-y_1)^{-1}x_0(1-2y_1)^{-1}\right) \\ & +(1-x_0)^{-1}y_0\left((1-y_1)^{-1}y_0(1-2y_1)^{-1}\right). \end{aligned}$$

Observe that $c \circ d$ is a rational series. Thus, the corresponding Fliess operator $F_{c \circ d}[w]$ is the input-output map of a bilinear system. □

Now it is only left to establish under what condition the cascade of Fliess operators on $\mathcal{UV}^m[0,T]$ having locally convergent generating series produce an L_2-Itô process. In this setting, consider the next theorem, which is a direct consequence of its counterpart for series in $\mathbb{R}\langle\langle X\rangle\rangle$ [24, 28].

Theorem V.2.11. *Let $c \in \mathbb{R}^{\ell}\langle\langle XY\rangle\rangle$ and $d \in \mathbb{R}^{2m}\langle\langle XY\rangle\rangle$. If c and d are locally convergent then $c \circ d$ is also locally convergent.*

Corollary V.2.5. *Let $c \in \mathbb{R}^{\ell}\langle\langle XY\rangle\rangle$ and $d \in \mathbb{R}^{2m}\langle\langle XY\rangle\rangle$ be locally convergent. For any $w \in \mathcal{UV}^m[0,T]$, there exist an $R > 0$ and a stopping time τ_R such that $F_{c \circ d}[w]$ converges to an L_2-Itô output processes over $[0,\tau_R]$ assuming the order of summation defined in* (IV.3.5).

Proof: First observe that F_c and F_d, independently, are each only well-defined on $\mathcal{UV}^m[0,T]$ when the summations are ordered as in (IV.3.5). By Theorem V.2.11, regardless of any order for c and d, the series $c \circ d$ have locally convergent coefficients. Hence, the operator $F_{c \circ d}[w]$ is again defined in the same sense. Therefore,

the ordered summation of $F_{cod}[w]$ converges to an L_2-Itô output process over the stochastic interval of time $[0, \tau_R]$, where τ_R is defined in (IV.3.4). ∎

From the last theorem, once again a conditional convergence result is obtained when locally convergent series are involved. Earlier, the property of exchangeability was added to remove this conditionality (see Theorem IV.3.1). However, in this particular theorem, exchangeability cannot be used since it is not known at present whether the composition product of exchangeable series is exchangeable.

CHAPTER VI

CONCLUSIONS AND FUTURE RESEARCH

In this final chapter, the main contributions and conclusions of this dissertation are summarized. Then some future research topics are described.

VI.1 MAIN CONCLUSIONS

This dissertation was focused on the solution of five problems.

The first problem involved defining a class of L_2-Itô processes that formed a set of admissible inputs to a Fliess operator. The proposed classes of inputs were given in Definitions IV.1.1 and IV.3.1. In these definitions three sets were defined: $\mathcal{UV}^m[0,T]$, $\widetilde{\mathcal{UV}}^m[0,T]$ and $\mathcal{UV}^m[0,\tau_R]$. The first is a subset of the second, and it formed the set of inputs to be used for driving a Fliess operator. The latter set is the restriction of $\mathcal{UV}^m[0,T]$ to the nonzero stochastic time interval $[0,\tau_R]$. In particular, any element of these sets is an L_2-Itô process and can be written as the sum of a Lebesgue and a Stratonovich integral whose integrands are themselves L_2-Itô processes.

The second problem was to define a Fliess operator over $\mathcal{UV}^m[0,T]$ and $\mathcal{UV}^m[0,\tau_R]$, and provide the necessary conditions under which the operator converges to produce a well-defined stochastic output process over a nonzero interval of time. The proposed definition of a Fliess operator driven by inputs from $\mathcal{UV}^m[0,T]$ was given in Definition IV.1.2. For this definition, an iterated integral comprised by Lebesgue and Stratonovich integrals was introduced in (IV.1.2) and (IV.1.3). To establish the convergence of this Fliess operator, it was natural to take limits in

the mean square sense. The fact that Stratonovich integrals now played a role in the Fliess operator definition added a level of complication in the calculation of an L_2 upper bound for these Lebesgue-Stratonovich iterated integrals in that they do not satisfy the isometry property. Then, identity (IV.1.4) was used to transform Stratonovich integrals into Itô integrals so that Theorems IV.1.2, IV.1.19 and IV.1.4 gave the L_2 upper bounds needed. The main idea for the global convergence of Fliess operators, Theorem IV.1.5, was to use the concept of globally convergent series in combination with the notion of a Cauchy series to prove the mean square convergence of a Fliess operator. This also meant that the limit is a well-defined L_2 bounded random variable. In addition, using properties of the shuffle product, a conditional local convergence result for these operators was given in Theorem IV.3.1. However, if the generating series of the Fliess operator was an exchangeable series then Theorem IV.3.1 produced a well-defined stochastic process defined over a nonzero stochastic interval of time independent of the order defined in Theorem IV.3.1.

The third problem was to characterize the set of output processes giving their main properties and describing in what sense there is compatibility between the input class and the output class. A characterization of the output process was given in Corollary IV.4.1. Specifically, this corollary showed that an input in $\mathcal{UV}^m[0,T]$ maps to an output in $\widetilde{\mathcal{UV}}^m[0,T]$. This meant that the output set $\widetilde{\mathcal{UV}}^m[0,T]$ satisfies all the properties of $\mathcal{UV}^m[0,T]$ with the exception of the independence properties presented in Definition IV.1.2. One critical problem related to this fact is that the output process generated by a Fliess operator might not, in general, be suitable for driving another Fliess operator. It was also shown in Theorem IV.4.1 that two different generating series cannot represent the same output process. Finally, an

Itô process generated by a Fliess operator was called analytic since it resembles the notion of analyticity in the usual sense.

The fourth problem was to characterize the generating series for the parallel, product and cascade interconnections of Fliess operators for formal input processes and for inputs from $\mathcal{UV}^m[0,T]$. The proposed solution to this problem was to first define the set of formal inputs in Definition V.2.1 and the class of formal Fliess operators in Definition V.2.3. Then the characterization of the nonrecursive interconnections (parallel, product and cascade) were presented in Theorems V.2.3, V.2.4 and V.2.5. In these theorems, it was shown that the operations $+$, $\sqcup\!\sqcup$ and $\circ$ on $\mathbb{R}\langle\langle XY\rangle\rangle$ characterize the parallel, product and cascade connections, respectively. Theorems V.2.6 and V.2.8 showed that the same operations also characterize their respective connections for inputs in $\mathcal{UV}^m[0,T]$.

Finally, the fifth problem was to provide conditions under which these interconnections are well-defined in the sense that they produce a well-defined output process. It was discovered that the algebraic operations $+$, $\sqcup\!\sqcup$ and $\circ$ on $\mathbb{R}\langle\langle XY\rangle\rangle$ behaves similarly as on $\mathbb{R}\langle\langle X\rangle\rangle$. Then, Theorem V.2.7 showed that the $+$ and $\sqcup\!\sqcup$ operations preserve global convergence, and Theorems V.2.7 and V.2.11 showed that $+$, $\sqcup\!\sqcup$ and $\circ$ operations preserve local convergence. In Corollaries V.2.1, V.2.2 and V.2.3, the convergence of the parallel and product connections was presented. For the composition product, Ferfera's sufficient condition for the rationality was extended to $\mathbb{R}\langle\langle XY\rangle\rangle$ in Theorem V.2.9. Using the previous theorem, it was shown in Corollary V.2.4 that under Ferfera's condition the cascade of two Fliess operators on $\mathcal{UV}^m[0,T]$ produces a well-defined L_2-Itô output process. Finally, Corollary V.2.5 showed that the composition of two locally convergent Fliess operators generates a convergent

output when ordered in the form presented in Theorem IV.3.1. One problem with this theorem is that if the summation of the composite system is taken in a different order, the system might not yield a well-defined output process.

VI.2 FUTURE RESEARCH

Some interesting future problems related to this dissertation are listed below.

1. It was observed in Corollary IV.4.1 that the output process of a Fliess operator belongs to $\widetilde{\mathcal{UV}}^m[0,T]$. If the input set were $\widetilde{\mathcal{UV}}^m[0,T]$ then there would exist full compatibility between inputs and outputs. Thus, the signal $\tilde{y}$ (see Figure 10) can be used to drive a second Fliess operator. One could then ask: Is $\widetilde{\mathcal{UV}}^m[0,T]$ a more natural input space for the interconnection of Fliess operators driven by stochastic inputs?

2. Another open problem concerns the absolute convergence of Fliess operators when the generating series is locally convergent. The structure of an affine state space system is independent of the type of inputs (deterministic or stochastic) used to drive it. Therefore, there are plenty of systems described by locally convergent series that can be driven by stochastic processes. Then, the following questions can be formulated: Under what conditions does a Fliess operator with stochastic inputs converge absolutely to a well-defined output process? Are the conditions associated with $\mathcal{UV}^m[0,T]$ enough to achieve convergence?

3. It was noted in Chapter V that if the composition product of two exchangeable series is exchangeable then the cascade of two Fliess operators with inputs from

$\mathcal{UV}^m[0,T]$ will produce a well-defined output process. But, under what conditions is exchangeability preserved under the composition product? What other properties of exchangeable series can be exploited in the stochastic setting?

4. It was shown in Example IV.1.4 that inputs from $\mathcal{UV}^m[0,T]$ are not adequate to model switched systems. It was conjectured that stochastic processes such as Poisson processes (Lévy processes) would be more suitable for this purpose. Thus, how can the theory developed in this dissertation be adapted to consider inputs such as Poisson processes?

BIBLIOGRAPHY

[1] G. B. Arous, "Flots et séries de Taylor stochastiques," *Probability Theory Related Fields*, vol. 81, no. 1, pp. 29–77, 1989.

[2] R. Ash, *Probability and Measure Theory*, 2nd ed. San Diego, CA: Academic Press, 2000.

[3] F. Baudoin, *An Introduction to the Geometry of Stochastic Flows.* London: Imperial College Press, 2004.

[4] J. Berstel and C. Reutenauer, *Rational series and their languages.* Berlin: Springer-Verlag, 1988.

[5] R. Brown, "A brief account of microscopical observations made in the months of June, July and August, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies," *Philosophical Magazine*, vol. 4, pp. 161–173, 1828.

[6] F. Castell, "Asymptotic expansion of stochastic flows," *Probability Theory and Related Fields*, vol. 96, no. 2, pp. 225–239, 1993.

[7] K. T. Chen, "Iterated path integrals," *Bulletin of the American Mathematical Society*, vol. 83, pp. 831–879, 1977.

[8] K. L. Chung and R. J. Williams, *Introduction to Stochastic Integration.* Boston: Birkhäuser, 1990.

[9] L. A. Duffaut Espinosa, W. S. Gray, and O. R. González, "On the bilinearity of cascaded bilinear systems," in *Proceedings 46th IEEE Conference on Decision and Control*. New Orleans, Louisiana, 2007, pp. 5581–5587.

[10] L. A. Duffaut Espinosa, W. S. Gray, and O. R. González, "On Fliess operators driven by L_2-Itô random processes," in *Proceedings 48th IEEE Conference on Decision and Control*. Shanghai, China, 2009, to appear.

[11] A. Einstein, "Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen," *Annalen der Physik*, vol. 322, no. 8, pp. 549–560, 1905.

[12] M. Émery, "On two transfer principles in stochastic differential geometry," *Séminaire de probabilités de Strasbourg*, vol. 24, pp. 407–441, 1990.

[13] A. Ferfera, *Combinatoire du monoïde libre appliquée à la composition et aux variations de certaines fonctionnelles issues de la théorie des systèmes*, doctoral dissertation, University of Bordeaux I, 1979.

[14] ——, "Combinatoire du monoïde libre et composition de certains systèmes non linéaires," *Astérisque*, vol. 75-76, pp. 87–93, 1980.

[15] M. Fliess, "Matrices de Hankel," *Journal de Mathématiques Pures et Appliquées*, vol. 53, pp. 197–224, 1974.

[16] ——, "Intégrales itérées de K.T. Chen, bruit blanc Gaussien et filtrage non linéaire," *C. R. Acad. Sc. Paris*, vol. A-284, pp. 459–462, 1977.

[17] ——, "Fonctionnelles causales non linéaires et indéterminées non commutatives," *Bulletin de la Société Mathématique de France*, vol. 109, pp. 3–40, 1981.

[18] ——, "Stability of an elementary type of stochastic differential equation driven by multidimensional noises," *Stochastics*, vol. 14, pp. 205–213, 1981.

[19] ——, "Réalisation locale des systèmes non linéaires, algèbres de Lie filtrées transitives et séries génératrices non commutatives," *Inventiones Mathematicae*, vol. 71, pp. 521–537, 1983.

[20] M. Fliess, M. Lamnabhi, and F. Lamnabhi-Lagarrigue, "An algebraic approach to nonlinear functional expansions," *IEEE Trans. Circuits Systems*, vol. CAS-30, pp. 554–570, 1983.

[21] M. Fliess and F. Lamnabhi-Lagarrigue, "Applications of a new functional expansion to the cubic anharmonic oscillator," *Journal of Mathematical Physics*, vol. 23, pp. 495–502, 1981.

[22] M. Fliess and D. Normand-Cyrot, "Algèbres de Lie nilpotentes, formule de Baker-Campbell-Hausdorff et intégrales itérées de K.T. Chen," *Séminaire de Probabilités de Strasbourg*, vol. 16, pp. 257–267, 1982.

[23] J. Gaines, "The algebra of iterated stochastic integrals," *Stochastic and Stochastics Reports*, vol. 49, pp. 169–179, 1994.

[24] W. S. Gray, *Formal Power Series Methods for Nonlinear Systems Analysis*, in preparation.

[25] ——, "A unified approach to generating series for nonlinear cascade systems," in *Proc. 48th IEEE Conference on Decision and Control*, 2009, Shanghai, China, to appear.

[26] W. S. Gray, H. Herencia Zapana, L. A. Duffaut Espinosa, and O. R. González, "On cascades of bilinear systems and generating series of weighted Petri nets," in *Proc. 47th IEEE Conference on Decision and Control*, Cancun, Mexico, 2008, pp. 4115–4120.

[27] ——, "Bilinear systems interconnections and generating series of weighted Petri nets," *Systems and Control Letters*, to appear.

[28] W. S. Gray and Y. Li, "Generating series for interconnected analytic nonlinear systems," *SIAM Journal on Control and Optimization*, vol. 44, pp. 646–672, 2005.

[29] W. S. Gray and Y. Wang, "Fliess operators on L_p-spaces: Convergence and continuity," *Systems and Control Letters*, vol. 46, no. 2, pp. 67–74, 2002.

[30] ——, "Formal Fliess operators with applications to feedback interconnections," in *Procedings 18th International Symposium on the Mathematical Theory of Networks and Systems*, Blacksburg, Virginia, 2008.

[31] ——, "Noncausal Fliess operators and their shuffle algebra," *International Journal of Control*, vol. 81, no. 3, pp. 344–357, 2008.

[32] C. Hespel, "Truncated bilinear approximants: Carleman, finite Volterra, Padé-type, geometric and structural automata," in *Algebraic Computing in Control*,

G. Jacob and F. Lamnabhi-Lagarrigue, Eds., Springer-Verlag, Berlin, pp. 264–278, 1991.

[33] C. Hespel and G. Jacob, "Approximation of nonlinear dynamic systems by rational series," *Theoretical Computer Science*, vol. 79, no. 1, pp. 151–162, 1991.

[34] Y. Z. Hu and P. A. Meyer, "Sur les intégrales multiples de Stratonovitch," *Séminaire de probabilités de Strasbourg*, vol. 22, pp. 72–81, 1988.

[35] A. Isidori, *Nonlinear Control Systems*, 3rd ed. London: Springer-Verlag, 1995.

[36] T. Kailath, *Linear Systems.* Englewood Cliffs, NJ: Prentice-Hall Inc., 1971.

[37] R. E. Kalman, P. L. Falb, and M. A. Arbib, *Topics in Mathematical System Theory.* New York, NY: McGraw-Hill Inc., 1969.

[38] F. C. Klebaner, *Introduction to Stochastic Calculus with Applications.* London: Imperial College Press, 1998.

[39] P. E. Kloeden and E. Platen, *Numerical Solutions of Stochastic Differential Equations.* Heidelberg: Springer-Verlag, 1992.

[40] E. Kreyszig, *Introductory Functional Analysis with Application.* New York, NY: Wiley, 1978.

[41] D. F. Kuznetsov, "Expansion of the Stratonovich multiple stochastic integrals based on the Fourier multiple series," *Journal of Mathematical Sciences*, vol. 109, no. 6, pp. 2148–2165, 2002.

[42] F. Lamnabhi-Lagarrigue and M. Lamnabhi, "Algebraic computation of the statistics of the solutions of some nonlinear stochastic differential equations," in *Computer Algebra*, ser. Lectures Notes in Computer Science, G. Goos and J. Hartmanis, Eds., Springer-Verlag, Berlin, vol. 162, pp. 55–67, 1983.

[43] T. Lyons, "Differential equations driven by rough signals," *Revista Matemática. Iberoamericana*, vol. 14, no. 2, pp. 215–310, 1998.

[44] P. Malliavin, "Stochastic calculus of variations and hypoelliptic operators," in *Proceedings of the International Symposium on Stochastic Differential Equations*, K. Itô, Ed., Wiley, New York, pp. 195–263, 1978.

[45] P. A. Meyer, "Sur deux estimations d'intégrales multiples," *Séminaire de probabilités de Strasbourg*, vol. 25, pp. 425–426, 1991.

[46] L. Nielsen, *Pricing and Hedging of Derivative Securities.* Oxford, NY: Oxford University Press, 1999.

[47] B. Øksendal, *Stochastic Differential Equations*, 5th ed. Berlin: Springer, 1998.

[48] P. E. Protter, *Stochastic Integration and Differential Equations*, 2nd ed. Berlin: Springer, 2005.

[49] R. Ree, "Lie elements and an algebra associated with shuffles," *The Annals of Mathematics*, vol. 68, no. 2, pp. 210–220, 1958.

[50] E. Riccomagno, *Calcolo Simbolico per Processi Stocastici*, master thesis, Università di Genova, 1993.

[51] ——, "Rational non-commutative formal power series and iterated integral representation of a class of Itô processes," *Bollettino dell'Unione Matematica Italiana*, vol. 10B, no. 1, pp. 25–50, 1996.

[52] W. J. Rugh, *Nonlinear System Theory: The Volterra/Wiener Approach.* Baltimore, MD: The Johns Hopkins University Press, 1981.

[53] M. P. Schützenberger, "On the definition of a family of automata," *Information and Control*, vol. 4, pp. 245–270, 1961.

[54] S. E. Shreve, *Stochastic Calculus for Finance II.* New York, NY: Springer, 2004.

[55] M. J. Steele, *Stochastic Calculus and Financial Applications.* New York, NY: Springer, 2003.

[56] H. J. Sussmann, "On the gap between deterministic and stochastic ordinary differential equations," *The Annals of Probability*, vol. 6, no. 1, pp. 19–41, 1978.

[57] ——, "Lie brackets and local controllability: A sufficient condition for scalar-input systems," *SIAM Journal on Control and Optimization*, vol. 21, pp. 686–713, 1983.

[58] ——, "A product expansion for the Chen series," *Theory and Applications of Nonlinear Control Systems*, pp. 323–335, 1986.

[59] ——, "Product expansions of exponential Lie series and the discretization of stochastic differential equations," in *Stochastic Differential Systems, Stochastic Control and Applications*, W. H. Fleming and P. L. Lions, Eds., Springer-Verlag, New York, vol. 10, pp. 563–582, 1988.

[60] V. Volterra, *Leçons sur les Fonctions de Lignes.* Paris: Gauthier-Villars, 1913.

[61] Y. Wang and E. D. Sontag, "On two definitions of observation spaces," *Systems and Control Letters*, vol. 13, pp. 279–289, 1989.

[62] ——, "Generating series and nonlinear systems: analytic aspects, local realizability and I/O representations," *Forum Mathematicum*, vol. 4, pp. 299–322, 1992.

[63] Y. Wang, *Algebraic Differential Equations and Nonlinear Control Systems*, doctoral dissertation, Rutgers University, 1990.

[64] N. Wiener, *Nonlinear Problems in Random Theory.* Cambridge, MA: The MIT Press, 1966.

VITA

Luis A. Duffaut Espinosa
Department of Electrical and Computer Engineering
Old Dominion University
Norfolk, Virginia 23509 USA

EDUCATION

- M.Sc. Mathematics 2007, Pontificia Universidad Católica del Perú, Lima, Perú
- B.Sc. Physics 2003, Universidad Nacional de Ingeniería, Lima, Perú

PUBLICATIONS

JOURNAL PUBLICATIONS:

1. W. S. Gray, H. Herencia-Zapana, **L. A. Duffaut Espinosa**, and O. R. González, "Bilinear systems interconnections and generating series of weighted Petri nets," *Systems and Control Letters*, 2009, to appear.

SELECTED CONFERENCE PUBLICATIONS:

1. **L. A. Duffaut Espinosa**, W. S. Gray, and O. R. González, "On Fliess operators driven by L_2-Itô random processes," *Proc. 48th IEEE Conference on Decision and Control*, Shanghai, China, 2009, to appear.
2. **L. A. Duffaut Espinosa**, W. S. Gray, and O. R. González, "Growth bounds for iterated integrals of L_2-Itô random processes," *Proc. 41st IEEE Southeastern Symposium on System Theory*, Tullahoma, Tennessee, 2009, pp. 223-229.
3. W. S. Gray, H. Herencia-Zapana, **L. A. Duffaut Espinosa**, and O. R. González, "On cascades of bilinear systems and generating series of weighted Petri nets," *Proc. 47th IEEE Conference on Decision and Control*, Cancun, Mexico, 2008, pp. 4115-4120.
4. **L. A. Duffaut Espinosa**, W. S. Gray, and O. R. González, "On the bilinearity of cascaded bilinear systems," *Proc. 46th IEEE Conference on Decision and Control*, New Orleans, Louisiana, 2007, pp. 5581-5587.

Typeset using LaTeX.