



Mean Square Analytic Solutions of Random Linear Models

PhD Thesis

Presented by: **Gema Calbo Sanjuán**

Advisers: **Dr. Lucas Antonio Jódar Sánchez**
Dr. Juan Carlos Cortés López

Universidad Politécnica de Valencia, summer 2010

Departamento de Matemática Aplicada

Dr. Lucas Antonio Jódar Sánchez and Juan Carlos Cortés López, full professors at the Technical University of Valencia,

CERTIFY that the present thesis entitled *Mean Square Analytic Solutions of Random Linear Models* has been directed under our supervision in the Applied Mathematics Department of the Technical University of Valencia by Gema Calbo Sanjuán. It constitutes her thesis dissertation to obtain the PhD degree in Mathematics.

In compliance with the current legislation, we authorize the presentation of this memory signing the present certificate.

In Valencia, September 2010.

Dr. Lucas Antonio Jódar Sánchez

Dr. Juan Carlos Cortés López

*“Individual behaviour may be erratic,
but aggregate behaviour is often quite predictable”*

C.R. MacCluer

To my husband
Juan Carlos

and

To my parents
Lauí and Tere

Thanks

Firstly, thanks to God for allowing me to write this dissertation.

I am grateful to my advisers, Lucas Jódar and Juan Carlos Cortés, for their continuous help, encouragement and support provided along these years.

Thanks to the colleagues of the “Multidisciplinary Institute of Mathematics” at the Technical University of Valencia for their hospitality during the annual conferences on multidisciplinary mathematics in which they allowed me to participate during the last four years. This has been a valuable source of inspiration both human and academic.

Finally, thanks to my family for their constant and unconditional support throughout this project.

Contents

Summary	7
Spanish summary (Resumen)	9
Valencian summary (Resum)	11
1 Introduction	17
2 Preliminaries	25
2.1 Probability, random variables and stochastic processes	26
2.1.1 Probability space	26
2.1.2 Random variables	27
2.1.3 Stochastic processes	36
2.2 Mean square calculus	39
2.2.1 Mean square continuity	44
2.2.2 Mean square differentiation	46
2.2.3 Mean square and mean fourth calculus	49
2.2.4 Random variables of class p	52
3 Mean square analytic solutions of coupled linear difference models with random initial condition and source term	55
3.1 Introduction	56
3.2 Preliminaries	58

Contents

3.3	Mean square solution of random matrix linear difference equations	63
3.4	Mean square boundedness of solutions	68
3.5	Statistical functions of the mean square solution	71
3.6	Examples	76
3.7	Conclusions	82
4	Mean square analytic solution of coupled linear differential models with random initial condition and source term	83
4.1	Introduction	84
4.2	Preliminaries	85
4.3	Constructing a random power series solution	92
4.4	Statistical functions of the mean square approximating solution	96
4.5	Examples	100
4.6	Conclusions	105
5	Mean square power series solutions of first-order linear differential equations with multiplicative and additive randomness	107
5.1	Introduction	108
5.2	The case where $B(t)$ is null	111
5.3	Solving the general case	114
5.4	Computing the main statistical functions of the solution	119
5.5	Examples	123
5.6	Conclusions	134
6	Trigonometric stochastic processes solutions of linear second-order random differential equations	137
6.1	Introduction	138
6.2	Preliminaries	139
6.3	Trigonometric processes solutions of random differential equations	142
6.4	Statistical properties of the solution process. An example . . .	145
6.5	Conclusions	149
7	Mean square power series solutions of random Hermite differential equations	151
7.1	Introduction	152

7.2	Preliminaries	153
7.3	A crucial inequality	154
7.4	Solving random Hermite differential equations	157
7.5	Introducing the random Hermite polynomials	160
7.6	Statistical functions of the mean square random power series solution	163
7.7	Examples	166
7.8	Conclusions	170
	References	173
	List of Tables	180
	List of Figures	181

Summary

The aim of this PhD project is to develop analytical-numerical techniques for solving random initial value problems based on linear difference and differential equations and systems in the mean square sense.

Regarding the study made on difference equations, see Chapter 3, we extend to random context some of the main results that in the deterministic case are known for solving such equations and to study the asymptotic behavior of their solution stochastic processes.

With respect to differential equations should be emphasized that the unifying element of the study in this piece is the extension to random scenario of the so-called Fröbenius method for computing the solution of differential equations as power series expansions. Chapters 4-7 deal with problems of both scalar and matrix type of first and second order, where the randomness enters into the model through the initial conditions and coefficients, and moreover uncertainty in the latter case, is considered in additive and multiplicative way. The problems based on random differential equations allow us to introduce relevant stochastic processes such as the exponential process (see Chapter 5) or the trigonometric sine and cosine processes as well as some of their basic algebraic properties (see Chapter 6). The final chapter is devoted to study the Hermite differential equation with random coefficients and, under certain conditions, its solutions are obtained as finite random series that allow us to define the random Hermite polynomials. Besides obtaining the solutions by

Summary

means of mean square convergent random power series, for each of the considered problems, we provide approximations of the main statistical functions of the solution process, such as the mean and variance (or in the vector case, the so-called covariance matrix). These approximations are compared through illustrative examples with those obtained by other available methods.

The AMS (2010) thematic classification of this dissertation is: 65C20, 60H35, 65N12.

Resumen

El objetivo de este proyecto de tesis doctoral es el desarrollo de técnicas analítico-numéricas para resolver en media cuadrática problemas de valores iniciales de ecuaciones y sistemas de ecuaciones en diferencias y diferenciales aleatorias de tipo lineal.

Respecto del estudio aportado sobre ecuaciones en diferencias, véase Capítulo 3, se extienden al contexto aleatorio algunos de los principales resultados que en el caso determinista se conocen para resolver este tipo de ecuaciones así como para estudiar el comportamiento asintótico de su solución.

En lo que se refiere a las ecuaciones diferenciales hay que señalar que el elemento unificador del estudio realizado en esta memoria es la extensión al escenario aleatorio del método de Fröbenius para la búsqueda de soluciones de ecuaciones diferenciales en forma de desarrollos en serie de potencias. A largo de los Capítulos 4-7 se abordan problemas tanto de tipo escalar como de tipo matricial tanto de primer como de segundo orden, donde la aleatoriedad se introduce en los modelos a través de las condiciones iniciales y los coeficientes, siendo además la incertidumbre en este último caso, considerada tanto de forma aditiva como multiplicativa. Los problemas basados en ecuaciones diferenciales aleatorias tratados permiten introducir procesos estocásticos importantes como son el proceso exponencial (véase Capítulo 5), los procesos trigonométricos seno y coseno y algunas de sus propiedades algebraicas básicas (véase Capítulo 6). En el último capítulo se estudia la ecuación

Spanish summary (Resumen)

diferencial de Hermite con coeficientes aleatorios y, bajo ciertas condiciones, se obtienen soluciones en forma de serie aleatoria finita que definen los polinomios de Hermite aleatorios.

Además de obtener las soluciones en forma de serie aleatoria convergente en el sentido estocástico de la media cuadrática, para cada uno de los problemas tratados se calculan aproximaciones de las principales propiedades estadísticas del proceso solución, tales como la media y la varianza (o en el caso vectorial, la matriz de varianzas-covarianzas). Estas aproximaciones se comparan a través de ejemplos ilustrativos con las obtenidas por otros métodos disponibles en la literatura.

La clasificación temática de esta memoria atendiendo a la clasificación por áreas de la AMS (2010) es: 65C20, 60H35, 65N12.

Resum

L'objectiu d'aquest projecte de tesi doctoral és el desenvolupament de tècniques analítico-numèriques per a resoldre en mitjana quadràtica problemes de valors inicials d'equacions i sistemes d'equacions en diferències i diferencials aleatòries de tipus lineal.

Atenent l'estudi aportat sobre equacions en diferències, consulteu Capítol 3, s'estenen al context aleatori alguns dels principals resultats que es coneixen en el cas determinista per a resoldre aquest tipus d'equacions i també per a estudiar el comportament asimptòtic de la seua solució.

Pel que fa a les equacions diferencials cal assenyalar que l'element unificador de l'estudi realitzat en aquesta memòria és l'extensió a l'escenari aleatori del mètode de Fröbenius per a la cerca de solucions d'equacions diferencials en forma de desenvolupaments en sèrie de potències. A llarg dels Capítols 4-7 s'estudien problemes tant de tipus escalar com de tipus matricial de primer i de segon ordre, on l'aleatorietat s'introdueix en el model a través de les condicions inicials i els coeficients, sent a més a més l'aleatorietat en aquest últim cas, considerada tant de forma additiva com multiplicativa. Els problemes basats en equacions diferencials aleatòries tractats permeten introduir processos estocàstics importants com són el procés exponencial (consulteu el Capítol 5), els processos trigonomètrics sinus i cosinus i algunes de les seues propietats algebraïques bàsiques (consulteu el Capítol 6). En l'últim capítol s'estudia l'equació diferencial d'Hermite amb coeficients aleatoris i, sota determinades

Valencian summary (Resum)

condicions, s'obtenen solucions en forma de sèrie aleatòria finita que defineixen els polinomis d'Hermite aleatoris.

A més a més d'obtenir solucions en forma de sèrie aleatòria convergent en el sentit estocàstic de la mitjana quadràtica, per a cadascun dels problemes tractats es calculen aproximacions de les principals propietats estadístiques del procés solució, com la mitjana i la variància (o en el cas vectorial, la matriu de variàncies-covariàncies). Aquestes aproximacions es comparen mitjançant d'exemples il·lustratius amb les obtingudes per altres mètodes que poden trobar-se en la literatura.

La classificació temàtica d'aquesta memòria atenent a la classificació per àrees de l'AMS (2010) és: 65C20, 60H35, 65N12.

GLOSSARY

A	special notation for writing a matrix A
a.s	almost surely
d.f.	distribution function
c.f.	characteristic function
i.i.d.	identical independently distributed
i.v.p.	initial value problem
l.i.m.	mean square limit or limit in mean
m.s.	mean square
p.d.f.	probability distribution function
p.f.	probability function
r.v.	random variable
2-r.v.	second order random variable
s.p.	stochastic process
2-s.p.	second order stochastic process
2-v.s.p.	second order vector stochastic process
w.p.	with probability
C	set of complex numbers
$E[X] = \mu_X$	expectation or mean or average of random variable X

$E^d[X]$	expectation of random variable X computed from dishonest method
$\mu_X^m(t)$	expectation of random variable $X(t)$ computed from Monte Carlo method by using m simulations
L_2	set of second order scalar random variables
L_2^r	set of second order vector random variables of dimension r
L_4	set of fourth order scalar random variables
\mathbb{N}	set of nonnegative integers
\mathbb{R}	set of real numbers
$\mathbb{R}^{r \times r}$	set of matrices of size $r \times r$ with real coefficients
T	superscript it denotes the transpose for vectors/matrices
\mathcal{T}	time-space of an arbitrary continuous stochastic process
$\text{Var}[X] = \sigma_X^2$	variance of random variable X
σ_X	standard deviation of r.v. X
$\sigma_X^m(t)$	standard deviation of random variable $X(t)$ computed from Monte Carlo method by using m simulations
$\sigma(\mathbf{A})$	spectrum or set of eigenvalues of matrix \mathbf{A}
$\rho(\mathbf{A})$	spectral radius of matrix \mathbf{A}
$A \sim \text{Be}(\alpha; \beta)$	A has a Beta distribution of parameters $\alpha, \beta > 0$
$A \sim \text{Bi}(n; p)$	A has a Binomial distribution of parameters $n \in \mathbb{N}, p > 0$
$A \sim \text{Ga}(r; \alpha)$	A has a Gamma distribution of parameters $r, \alpha > 0$
$A \sim \text{N}(\mu; \sigma)$	A has a Gaussian distribution of parameters $\mu \in \mathbb{R}, \sigma > 0$

$A \sim \text{Un}([a, b])$	A has a Uniform distribution on the interval $[a, b]$
$A \sim \text{Po}(\lambda)$	A has a Poisson distribution of parameter $\lambda > 0$
$A \sim \text{Exp}(\lambda)$	A has an Exponential distribution of parameter $\lambda > 0$
$\ \mathbf{A}\ _\infty$	infinite norm of matrix \mathbf{A}
$\ \vec{v}\ _\infty$	infinite norm of vector \vec{v}
$\ \mathbf{A}\ _p$	p norm of matrix \mathbf{A}
$\ \vec{v}\ _p$	p norm of vector \vec{v}
$\ X\ _2$	2-norm of random variable X
$\ \vec{X}\ _r$	r -norm of vector random variable $\vec{X} = (X_1, \dots, X_r)^T$
$\frac{d^n}{dt^n}$	ordinary derivative of order n with respect to t
$X^{(n)}(t)$	mean square derivative of order n of $X(t)$
\vec{X}	special notation for writing a vector X
$\dot{\vec{X}}(t)$	first mean square derivative of $\vec{X}(t)$
$\ddot{\vec{X}}(t)$	second mean square derivative of $\vec{X}(t)$,
$X_n \xrightarrow[n \rightarrow \infty]{d} X$	X_n converges in distribution to X
$X_n \xrightarrow[n \rightarrow \infty]{\text{m.s.}} X$	X_n converges in the mean square sense to X
$\text{Cov}[X, Y]$	covariance of random variables X, Y
$\Gamma_X(t, s)$	correlation function of stochastic process $X(t)$
$\mathbf{\Lambda}_{\vec{X}}$	covariance matrix of random vector \vec{X}

- $\Gamma_{\vec{X}}$ correlation matrix of random vector \vec{X}
- $\Lambda_{\vec{X},\vec{Y}}$ cross-covariance matrix of random vectors \vec{X} and \vec{Y}
- $\Gamma_{\vec{X},\vec{Y}}$ cross-correlation matrix of random vectors \vec{X} and \vec{Y}
- ☒ end of the proof of a mathematical result

Chapter 1

Introduction

Numerous physical, chemical and biological phenomena, which are very important for scientific and technological progress, have been traditionally formulated through mathematical models based on ordinary or partial differential equations, where the data (initial conditions, source term and/or coefficients) are expressed by means of numerical values or deterministic functions. Nevertheless, scientists really fix these data from measurements, which always are subject to error. Depend on the quality of these measurements (which frequently can take a lot of time and incur high costs), the results obtained from the model to be satisfactory. In addition to measurement errors, we must consider the random character of complex external factors that can affect the system, such as pressure, temperature and humidity in Meteorology; the composition of the land in Seismology; investor tendency and economical policy of countries and companies in Finance; the environmental and genetical factors in Epidemiology; etc.

These circumstances make it is more advisable to consider the data as random magnitudes; if what is to be measured is a magnitude functionally independent of others values, it would be better to consider it as a random variable; when a dependency exists with respect to other magnitudes, such as time, space, etc., it would be more advisable to interpret it, not as a function

but as a stochastic process. The consideration of these facts leads to the reformulation of the traditional deterministic models, which, in order to improve them, must be replaced by random models. Thus, this is the main purpose of this piece: the development of analytical and numerical-analytical techniques to solve random linear ordinary differential and difference equations, whose formulation and resolution is of interest in different scientist and technical fields such as Epidemiology, Finance, Chemistry, Seismology, Physics, Biology, Engineering, etc.

Taking into consideration the random effects in differential equations leads mainly to two distinct classes of equations, for which the solution processes have differentiable and non-differentiable sample paths. Both require fundamentally different methods of analysis.

The first class arises when an ordinary differential equation has random coefficients, a random initial value or is forced by a fairly regular stochastic process, or when some combination of these holds. One of the methods to deal with these types of differential equations is based on the so-called **sample calculus**, where these equations are solved sample path by sample path as ordinary differential equations. The sample paths of the solution processes are then differentiable functions. As an example, consider the linear random differential equation

$$\dot{x}(t) = a(\omega)x(t) + b(t, \omega),$$

where the forcing process b is continuous in t for each ω . For an initial value $x_0(\omega)$ at $t = 0$, the solution is given by

$$x(t, \omega) = \exp(a(\omega)t) \left(x_0(\omega) + \int_0^t \exp(-a(\omega)s)b(s, \omega)ds \right).$$

Its sample paths are clearly differentiable functions of t . Many models of interest, such as those mathematical models to describe the trajectories of the spacecrafts in Aerospace Engineering, demand smoothness (differentiability) in order to guarantee accurate navigation (see [55]). In this scenario, the techniques developed in the sample calculus are then adequate for dealing with these types of problems.

The main disadvantage of previous approach is that properties such as continuity or differentiability of the trajectories, are not satisfied in many other applications of interest. This leads us to the second class; when the forcing term is an irregular stochastic process such as Gaussian white noise, one presents

the so-called **stochastic differential equations**. These equations are then written symbolically as stochastic differentials, but are interpreted as integral equations with Itô or Stratonovich stochastic integrals. In general their solutions inherit the non-differentiability of sample paths from the Wiener process in the stochastic integrals. In many applications, such equations result from the incorporation of either internally or externally originating random fluctuations in the dynamical description of a system. As an example of the former is the molecular bombardment of a speck of dust on a water surface which results in Brownian motion. The intensity of this bombardment does not depend on state variables, for instance the position and the velocity of the speck. Taking X_t as one of the components of the velocity of the particle, Langevin wrote the equation

$$\frac{dX_t}{dt} = -aX_t + b\xi_t,$$

for the acceleration of the particle. This is the sum of the holding-back or retarding frictional force depending on velocity and molecular forces represented by a white noise process ξ_t , whose intensity b is independent of the velocity. Here a and b are positive constants. Then the Langevin equation is represented as a stochastic differential

$$dX_t = -aX_t dt + b dW_t,$$

that is a stochastic integral equation

$$X_t = X_{t_0} - \int_{t_0}^t aX_s ds + \int_{t_0}^t b dW_s,$$

where the second integral is an Itô integral. Similar equations arise from electrical systems where X_t is the current and ξ_t represents thermal noise, [47], [3], [49], [29].

From the seventies, stochastic differential equations have polarized a great part of the attention of the research, relegating **random differential equations** to a second plane, where randomness is considered through processes different to those of white noise (see [56]), which allows us to deal with a great variety of problems in the applications (see [56], [57]). The solution of these differential equations can be faced by means of the **mean square calculus** which differing remarkably with respect to sample calculus, does not require the sample solution of each deterministic differential equation obtained by fixing ω to have differentiable properties. This makes the application of random differential equations to a wide variety of situations possible. What is more,

mean square calculus is a well-developed set of results and techniques that is simple and attractive because it follows ordinary deterministic procedures generally. This fact can be considered as an important advantage with respect to Itô calculus, where even the solution of elementary stochastic differential equations can become either very difficult or simply impossible. As in the deterministic case, in practice most of the random differential equations can not be solved, that is, often it will be impossible to obtain an explicit representation of their (exact) solution, hence random numerical methods are required. Once this goal has been achieved, and as an important difference with respect to the deterministic scenario, we will focus on computing the main statistical functions associated to the solution stochastic process such as average, standard deviation, higher moments or probability density function. In this way, mean square calculus is attractive because useful tools in this sense are available.

In dealing with random/stochastic differential equations, apart from above exposition, other relevant approaches deserve to be mentioned. The following list does not constitute an exhaustive list:

- Monte Carlo method. Monte Carlo method described in [28] is the classical and most used technique for approximating expected values of quantities of interest depending on the solution of differential equations with random inputs. These expected values include not only the mean or average but also higher moments and standard deviation. The algorithm approximates the desired expectation by a sample average of independent identically distributed (i.i.d.) realizations. When solving differential equations with random inputs, this method implies the solution of one deterministic differential equation for each realization of the input parameters. This makes the method simple to implement, allows for maximal code reusability and it is straightforward to parallelize. Its numerical error is approximately $\mathcal{O}(1/\sqrt{M})$, where M is the number of realizations. The advantage of using this approach is that the rate does not deteriorate with respect to the number of random variables in the problem, making the method very attractive for problems with large dimensional random inputs. As a main inconvenient of this approach we underline that it is expensive and it waste the potential possibilities that the structure of the ordinary differential equation could have: linearity, autonomy, etc. Based on the above information, Monte Carlo method is used in this piece in several examples with the aim of comparing the numerical results obtained by means of the new methods here proposed.

- Dishonest method. For a given random differential equation whose coefficients are random variables and/or stochastic processes, this approach, in a first step, constructs deterministic differential equations taking the expectation operator directly and using its algebraic properties such as linearity, and, in a second step, deterministic techniques to solve them. The application of this approach usually requires us to assume weak correlation or stationary hypotheses on the data and the unknown that are not always satisfied (this justifies, the denomination of the method). In spite of these hypotheses being possibly very difficult to justify in many cases, the method provides an answer that can be acceptable in many situations (see, [38, p.148], [46]).
- Method of Moments. This is based on obtaining the moments of the solution stochastic process. The procedure for doing this begins with the derivation of a set of differential equations satisfied by these moments. This system can be established in different ways, but usually most of the methods lead to an infinite hierarchy of differential equations for the moments. Then, in order to obtain moment solutions, some closure approximations must be used to truncate this hierarchy. The paper [65] represents one of the most outstanding and pioneering references of this method.
- Perturbation method. This method is applied to solve nonlinear random differential equations where uncertainty enters in the equation through a small parameter. The approach is based on the assumption that the solution stochastic process can be expanded in the powers of the parameter. Then, by making this series satisfy the equation and equating terms of the same power of the parameter one obtains a system of random linear differential equations that are easier to solve. From the computational viewpoint, however, the statistical properties of higher order to mean are, in general, difficult to compute because random parts rapidly become exceedingly complex as a function of random data [56, p.207].
- Wiener-Hermite expansions, polynomial chaos and related methods. In this approach each random variable involved in the random differential equation to be solved is projected on the space of the so-called polynomial chaoses by means an infinite series whose coefficients are referred as to polynomial chaoses. The polynomial chaoses can be arranged in a sequence of a properly chosen basis functions of random variable vector in such a way that the unknown of the random differential equation is also represented by means of an infinite series. In the case that the random

variable vector is defined by standard Gaussian random variables (whose number of component defines the dimension of the chaos) and the basis functions of this series are taken as the Hermite polynomials, the method above leads to the so-called Wiener-Hermite chaos expansion. Then a Galerkin projection, using the orthogonality of the basis functions, together with truncation of the infinite series leads to a system of ordinary differential equations governing the time evolution of the chaos coefficients of the solution. This system is then integrated in time using an appropriate numerical method, and the chaos coefficients thus obtained can be subsequently used to compute quantities of interest such as mean or variance. For details see, for instance, [54], [41], [13]. In the last few years, taking advantage of the Wiener-Hermite expansion, some authors have proposed interesting modified versions of this approach by using some mixture of it with other techniques, such as perturbation method and/or the so-called homotopy method. Interesting contributions in this sense can be found in [27], [23], [25].

- Transformation method. This is a nice direct approach that allows us to compute the probability density functions associated to the solution stochastic process of a random differential equation by means basic techniques. Then, the main statistical properties such as mean or variance can be also computed from it. In [24] one can find the introduction of this technique as well as wide variety of random problems where it is applied. In [44] one generalizes the transformation method.
- Further methods. As we have previously emphasized, the list above does not attempt to be exhaustive. In addition to previous methods, other ones can be found in technical literature. Some of them are stochastic linearization (see [56, p.209], [58]), decomposition method [1], stochastic finite element [26], [30] or the use of Fokker-Plank equation [56, p.180].

We conclude this introduction by summarizing the contents of this dissertation. First at all, we wish to emphasize that the main goal of this dissertation is the development of random analytic and analytic-numerical methods to solve random linear difference and differential equations in the mean square sense. The unifying method in this dissertation is the extension of the powerful Fröbenius analytic method to the random framework. By means of this method we will face random linear differential models in chapters 4-7. We also remark that once the considered random problem has been solved in the mean square sense by using a Fröbenius method, we will be interested in com-

puting its main statistical functions such as mean or average, standard deviation/variance, correlation and/or covariance functions since we will consider, not only scalar problems, but also matrix versions. Chapter 2 is concerned with the fundamental principles and results in probability, random variables and stochastic processes that will be required throughout these pages. Chapter 3 deals with the construction of random discrete solutions of coupled linear difference equations incorporating uncertainty into both the initial condition and the source term. Chapter 4 deals with the construction of a random power series solution of coupled linear differential equations containing uncertainty in both, initial condition and source term. Whereas in the two previous chapters, apart from initial conditions, only additive uncertainty is considered, the following chapters are devoted to solve random differential equations where uncertainty enters in the models in a multiplicative way. Chapter 5 considers the first-order linear case, that as an especial case allows us to introduce the exponential stochastic process. Chapter 6 is addressed to studying an important second-order random linear differential equation that permits the introduction of some relevant trigonometric stochastic processes as well as extending algebraic properties from the scalar framework to the stochastic one. This dissertation ends by studying in chapter 7 the so-called random Hermite differential equation and obtaining conditions under which the solutions generated by a Fröbenius method become the random Hermite polynomials.

Chapter 2

Preliminaries

This chapter is concerned with the fundamental principles and results in probability, random variables and stochastic processes. Since the reader is assumed to be familiar with the basic concepts in these subjects, the chapter is written in a direct way in order to provide a brief introduction and a summary of results useful in later development. In particular, we will stress our presentation in developing the main results belonging to the so-called mean square calculus, a special calculus that can be considered the counterpart of the ordinary calculus for handling real/complex functions but that allows us to manage satisfactorily stochastic processes. A brief introduction to the so-called fourth calculus together with its links with the mean square calculus will be a clue to establish certain basic properties that will be required throughout this dissertation. The set of the random variables of class p will be also introduced in this chapter. They will play a relevant role in chapters 5 and 6, where random differential equations with multiplicative uncertainty will be considered. The reader is encouraged to check classical references like [56], [51], [48], [66], [21], for further details.

2.1 Probability, random variables and stochastic processes

2.1.1 Probability space

The principal concept of probability theory is the probability space, which can be represented mathematically as an ordered triple (Ω, \mathcal{A}, P) , consisting of the sample space, Ω , a σ -algebra \mathcal{A} of subsets of Ω called events, and a real-valued set function P defined on \mathcal{A} called a probability. Ω can be an arbitrary set, whereas \mathcal{A} and P must satisfy the following properties. Firstly, for \mathcal{A}

- i) $\emptyset \in \mathcal{A}$.
- ii) If $A \in \mathcal{A}$ then $A^C \in \mathcal{A}$ (where A^C denotes the complement of A in Ω).
- iii) If $A_1, A_2, \dots \in \mathcal{A}$ then $A = \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

With respect to P

- iv) If $A \in \mathcal{A}$ then $P(A) \geq 0$.
- v) If $A_1, A_2, \dots \in \mathcal{A}$ is a family of disjoint sets, i.e., $A_i \cap A_j = \emptyset$ for $i \neq j$ then

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

- vi) $P(\Omega) = 1$.

Properties *i*)-*iii*) define \mathcal{A} as a σ -algebra of subsets of Ω , while properties *iv*)-*v*) determine that P is a measure and, *vi*) defines P as a probability (measure).

If $A \in \mathcal{A}$, $P(A)$, the probability of A can be interpreted as a measure of the likelihood of occurrence of the event A in a experiment whose possible outcomes are the elements of A . If, $P(A) = 0$, A is virtually impossible, and A is called a P -null set. If $P(A) = 1$, A is a certainty event and it is said that A occurs with probability 1 (w.p. 1) or almost surely (a.s.).

2.1.2 Random variables

Now let $(\Omega_1, \mathcal{A}_1)$ and $(\Omega_2, \mathcal{A}_2)$ be measurable spaces, that is, \mathcal{A}_i is a σ -algebra of subsets of Ω_i , $i = 1, 2$. If X is a mapping of Ω_1 into Ω_2 , denoted by $X : \Omega_1 \rightarrow \Omega_2$, then X is measurable if for each $A \in \mathcal{A}_2$,

$$\{\omega : X(\omega) \in A\} = X^{-1}(A) \in \mathcal{A}_1.$$

The collection $\mathcal{A}(X) = \{X^{-1}(A) : A \in \mathcal{A}_2\}$ is a σ -algebra in Ω_1 , called the σ -algebra generated by X and it is the smallest such σ -algebra such that X is measurable. The particular interest will be the cases of $\Omega_2 = \mathbb{R}$ (\mathbb{R} denotes the set of all real numbers); $\Omega_2 = \mathbb{R}^n$ (\mathbb{R}^n denotes the set of n -dimensional vectors with real components); $\Omega_2 = \mathbb{R}^{n \times m}$ ($\mathbb{R}^{n \times m}$ denotes the set of $n \times m$ matrices with real entries). In these cases \mathcal{A}_2 is usually taking to be the σ -algebra of Borel sets of the appropriate dimensions: the σ -algebra generated by the set of intervals (rectangles, etc) and denoted by \mathcal{B} , \mathcal{B}^n or $\mathcal{B}^{n \times m}$. Measurable functions $X : \Omega_1 \rightarrow \Omega_2$ are called random variables, random vectors, and random matrices, respectively, in each of these cases; X is a random vector (resp. matrix) if and only if each component X_i (resp. X_{ij}) is a random variable. Throughout this dissertation a vector and a matrix random variable will be denoted by \vec{X} and \mathbf{X} , respectively. A function $X : \Omega_1 \rightarrow \mathbb{R}$ is a real random variable (r.v.) if, and only if,

$$X^{-1}((-\infty, a]) = \{\omega : X(\omega) \leq a\} \in \mathcal{A}_1, \quad \forall a \in \mathbb{R}.$$

Example 2.1. *The indicator function I_A of a set $A \subseteq \mathcal{A}_1$ defined by*

$$I_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{if } \omega \notin A, \end{cases}$$

is a r.v. if and only if $A \in \mathcal{A}_1$, that is, if, and only if, A is an event or A is measurable.

If the measurable space (Ω, \mathcal{A}) represents an experiment, a r.v. X defined on (Ω, \mathcal{A}) can be considered as an abstraction of information from the possible outcomes. This abstraction can be very coarse as in the case of an indicator function or a constant function. However, the measurability requirement means that only those subsets of Ω that are events can be distinguished by

different values of X . This restriction indicates how discerning X can be, in that the information made available by X about the experiment cannot exceed that provided by the full σ -algebra \mathcal{A} which describes all possible outcomes of the experiment.

The usual algebraic operations of the analysis are, in general, preserved for the case that we manage r.v.'s. For example, sums, differences, products, and, when they exist, quotients, extrema, and limits of r.v.'s are also r.v.'s. This situation can be exploited to determine probabilities of compound events from the probabilities of simpler constituent elements.

Given a random experiment with its associated r.v. X and given a real number x , let us consider the probability of the event $\{x : X(\omega) \leq x\}$ or simply $P(X \leq x)$. This probability is a number, clearly dependent upon x . The function

$$F_X(x) = P(X \leq x),$$

is defined as the distribution function (d.f.) of X . The d.f. always exists. By definition it is a nonnegative, continuous to the right and nondecreasing function of the real variable x . Moreover, it satisfies

$$F_X(-\infty) = 0, \quad F_X(+\infty) = 1.$$

A r.v. X is called a continuous r.v. if its associated d.f. is continuous and differentiable almost everywhere. It is a discrete r.v. when the d.f. assumes the form of a staircase with a finite or countable finite jumps. For a continuous r.v. X , the derivative

$$f_X(x) = \frac{dF_X(x)}{dx},$$

exists and it is called the probability density function (p.d.f.) of the r.v. X . This function satisfies the following properties:

$$f_X(x) \geq 0, \quad \int_a^b f_X(x)dx = F_X(b) - F_X(a), \quad \int_{-\infty}^{\infty} f_X(x)dx = 1. \quad (2.1)$$

On the other hand, when X is a r.v. that only takes a finite number of values, say, x_1, x_2, \dots, x_N , its p.d.f. is usually called the probability function (p.f.) or the mass probability function and it is defined by

$$f_X(x) = \sum_{j=1}^N p_j \delta(x - x_j), \quad (2.2)$$

where $\delta(x)$ is the Dirac delta function and

$$p_j = P(X = x_j), \quad j = 1, 2, \dots, N.$$

This definition can be extended in a straightforward way for the case that X takes an infinite countable number of values or it is a mixed r.v., that is, a r.v. that takes values in an interval as well as in a discrete (finite or infinite countable) set of values. This definition of $f_X(x)$ is consistent in the sense that it has all the properties indicated by expressions (2.1), and the distribution function $F_X(x)$ is recoverable from equation (2.2) by integration.

In order to introduce later the concept of independent r.v.'s that will play an important role in this dissertation, we need introduce previously the concept of d.f. and p.d.f. for two r.v.'s (for the extension to more than two r.v.'s, see for instance, the reference [56, p.13]). Given two r.v.'s X and Y , then they induce a probability distribution in the Euclidean plane \mathbb{R}^2 which is described by the joint distribution function

$$F_{XY}(x, y) = P(\{\omega \in \Omega : X(\omega) \leq x, Y(\omega) \leq y\}).$$

The function $F_{XY}(x, y)$ is nonnegative, nondecreasing, and continuous at the right with respect to each of the real variables x and y . Moreover it satisfies the following properties:

$$F_{XY}(-\infty, -\infty) = F_{XY}(-\infty, y) = F_{XY}(x, -\infty) = 0,$$

$$F_{XY}(+\infty, +\infty) = 1, \quad F_{XY}(x, +\infty) = F_X(x), \quad F_{XY}(+\infty, y) = F_Y(y),$$

where $F_X(x)$ and $F_Y(y)$ denote the individual d.f. of r.v.'s X and Y , respectively, but in the context of several r.v.'s, each (individual) d.f. is commonly referred to as the marginal distribution function. If the second-order partial derivative

$$\frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y},$$

exists, we denote it by

$$f_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y},$$

and it is called the joint probability density function of X and Y . This function satisfies similar properties that we have enunciated previously for an individual

p.d.f. in (2.1), that is, it is nonnegative and

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y) dx dy = 1, \quad \int_{-\infty}^x \int_{-\infty}^y f_{XY}(u, v) du dv = F_{XY}(x, y),$$

$$\int_{-\infty}^{\infty} f_{XY}(x, y) dy = f_X(x), \quad \int_{-\infty}^{\infty} f_{XY}(x, y) dx = f_Y(y).$$

In this framework, p.d.f. $f_X(x)$ and $f_Y(y)$ are called the marginal density functions of r.v. X and Y , respectively.

Two r.v.'s X and Y are said to be independent if their joint p.d.f. $f_{XY}(x, y)$ can be factorized as the product of each individual p.d.f. $f_X(x)$ and $f_Y(y)$, that is,

$$f_{XY}(x, y) = f_X(x)f_Y(y). \quad (2.3)$$

The extension of the previous definition to the case of more than two r.v.'s is straightforward. A family of n r.v.'s X_1, X_2, \dots, X_n are mutually independent if, and only if,

$$f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = f_{X_1}(x_1)f_{X_2}(x_2) \cdots f_{X_n}(x_n).$$

An important result concerning independence is the following:

Proposition 2.1. ([34]) *If X and Y are independent r.v.'s, then for any (measurable) functions f and g , the r.v.'s $f(X)$ and $g(Y)$ are also independent.*

In practice a r.v. tries to model the uncertainty of some phenomenon. Some of the most useful information concerning a r.v. is revealed by deterministic quantities called moments that capture its behavior. Given a r.v. X , the n -th moment with respect to the origin is defined by

$$\alpha_n = E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x) dx, \quad n = 0, 1, 2, \dots, \quad (2.4)$$

whenever the following integral exists and satisfies

$$\int_{-\infty}^{\infty} |x|^n f_X(x) dx < +\infty.$$

The symbol $E[\]$ in (2.4) is usually called the expectation operator. Obviously, $\alpha_0 = 1$. The first moment $\alpha_1 = E[X]$ is particularly important because it gives

the statistical average of the r.v. X and usually it is referred to as the mean, the mathematical expectation or the statistical average. In the following it will denoted by α_1 , μ_X or $E[X]$, indistinctively.

An important feature about the existence of the moments with respect to the origin is given in the following result:

$$\text{If } E[|X|^p] < +\infty \text{ then } E[X^q] < +\infty, \forall q \leq p, \quad (2.5)$$

where the term $E[|X|^p]$ is called the p -th absolute moment of r.v. X with respect to the origin. In the sequel, we will assume the existence of the mean of a r.v. every time we consider its mean.

Given a collection of r.v.'s X_1, X_2, \dots, X_n and constants a_1, a_2, \dots, a_n , the expectation operator is linear, that is, it satisfies

$$E \left[\sum_{i=1}^n a_i X_i \right] = \sum_{i=1}^n a_i E[X_i]. \quad (2.6)$$

The central moments of a r.v. X are the moments of X with respect to its mean. Hence, the n -th central moment of X is defined as

$$\mu_n = E[(X - \mu)^n] = \int_{-\infty}^{\infty} (x - \mu)^n f_X(x) dx, \quad n = 0, 1, 2, \dots$$

Obviously, $\mu_0 = 1$ and $\mu_1 = 0$. The second central moment measures the spread or dispersion of the r.v. X about its mean. It is usually called the variance and in the following will be denoted by μ_2 , σ_X^2 or $\text{Var}[X]$. This statistical parameter can be expressed in terms of the first and second moment with respect to the origin as follows

$$\text{Var}[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2 = \alpha_2 - (\alpha_1)^2. \quad (2.7)$$

In practice, the dispersion of a r.v. X is given in the same units that r.v. X . It is done by means the square root of the variance which is called standard deviation and it will be denoted by σ_X .

Tables 2.1 and 2.2 collect, respectively, the mean, the variance and, the moments with respect to the origin, α_n , of some relevant r.v.'s that will appear throughout this dissertation.

Distribution	Notation	Mean	Variance
Uniform	$X \sim \text{Un}([a, b])$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Beta	$X \sim \text{Be}(\alpha; \beta), \alpha, \beta > 0$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
Gaussian	$X \sim \text{N}(\mu; \sigma), \sigma > 0$	μ	σ^2
Exponential	$X \sim \text{Exp}(\lambda), \lambda > 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Gamma	$X \sim \text{Ga}(r; a), r, a > 0$	$\frac{r}{a}$	$\frac{r}{a^2}$

Table 2.1: Some relevant r.v.'s and their mean and variance

Distribution	Notation	$\alpha_n = \text{E}[X^n]$
Uniform	$X \sim \text{Un}([0, 1])$	$\frac{1}{n+1}$
Beta	$X \sim \text{Be}(\alpha; \beta), \alpha, \beta > 0$	$\frac{(\alpha+n-1)!(\alpha+\beta-1)!}{(\alpha+\beta+n-1)!(\alpha-1)!}$
Gaussian	$X \sim \text{N}(0; \sigma), \sigma > 0$	0 if $n = 2k+1$, $\frac{(2k)!}{2^k k!} \sigma^{2k}$ if $n = 2k$
Exponential	$X \sim \text{Exp}(\lambda), \lambda > 0$	$\frac{n!}{\lambda^n}$
Gamma	$X \sim \text{Ga}(r; a), r, a > 0$	$\frac{(r+n-1)!}{a^n (r-1)!}$

Table 2.2: Some relevant r.v.'s and their moments with respect to the origin

It is worthwhile pointing out that an Exponential r.v. $X \sim \text{Exp}(\lambda)$ is a special case of a Gamma r.v. of parameters $r = 1$ and $a = \lambda$. If $\alpha = \beta = 1$, a Beta r.v. becomes a Uniform r.v. on the interval $[0, 1]$.

The transformation of a r.v. into another one will play a central role in the examples throughout this dissertation. For the resulting r.v., it is convenient to have a direct way to compute its statistical properties like expectation or variance in terms of the information available for the original r.v. The following result shows how can be calculated the expectation of a function of a continuous r.v. X , say $g(X)$, with p.d.f. $f_X(x)$ (an analogous expression can be written for the discrete case replacing the integral by a sum), without first determining its distribution

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x)f_X(x)dx. \tag{2.8}$$

In particular, if $g(x) = x$, (2.8) becomes the mean $E[X]$ and, if $g(x) = x^2$ taking into account (2.7) we can obtain the variance σ_X^2 .

The moments of two or more r.v.'s are defined in a similar way. For the case of two r.v.'s X, Y , the joint moments of order (n, m) with respect to the origin are defined by

$$\alpha_{nm} = E[X^n Y^m] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^n y^m f_{XY}(x, y) dx dy, \quad n, m = 0, 1, 2, \dots,$$

whenever they exist. It is seen that $\alpha_{10} = E[X]$ and $\alpha_{01} = E[Y]$. In the particular case that X and Y are independent r.v.'s, by (2.3) one gets

$$\begin{aligned} \alpha_{nm} &= E[X^n Y^m] \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^n y^m f_X(x) f_Y(y) dx dy \\ &= \left(\int_{-\infty}^{\infty} x^n f_X(x) dx \right) \left(\int_{-\infty}^{\infty} y^m f_Y(y) dy \right) \\ &= E[X^n] E[Y^m] = \alpha_n \alpha_m. \end{aligned} \tag{2.9}$$

The joint central moments of the r.v.'s X and Y of order (n, m) , μ_{nm} are given by

$$\begin{aligned}\mu_{nm} &= \text{E}[(X - \text{E}[X])^n (Y - \text{E}[Y])^m] \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)^n (y - \mu_Y)^m f_{XY}(x, y) dx dy, \quad n, m = 0, 1, 2, \dots\end{aligned}$$

In this case $\mu_{20} = \text{Var}[X]$ and $\mu_{02} = \text{Var}[Y]$. The particular case

$$\mu_{11} = \text{E}[(X - \text{E}[X])(Y - \text{E}[Y])] = \text{E}[XY] - \text{E}[X]\text{E}[Y], \quad (2.10)$$

is called the covariance between X and Y and in advance it will be denoted by $\text{Cov}[X, Y]$. The term $\text{E}[XY]$ is usually referred to as the correlation between X and Y . Note that given a r.v. X with mean $\text{E}[X] = \mu_X$ then the r.v. $Z = X - \mu_X$ satisfies that $\text{E}[Z] = 0$. Therefore without loss of generality one can assume that any r.v. X (with finite expectation) is centered in its mean and then $\text{E}[X] = 0$. In this case, its covariance coincides with its correlation. From (2.10), if $X = Y$ then it is seen that covariance becomes the variance of X . The covariance $\text{Cov}[X, Y]$ has great importance in the statistical analysis of two r.v.'s X, Y because it provides a measure of their statistical linear interdependency in the sense that its value is a measure of accuracy with which one r.v. can be approximated by a linear function of the other [56, p.18].

Two r.v.'s X and Y are said to be uncorrelated if $\mu_{11} = 0$ or equivalently from (2.10) if $\text{E}[XY] = \text{E}[X]\text{E}[Y]$. Note that from (2.9) with $n = 1$, if X and Y are independent r.v.'s with finite second moments with respect to the origin then they are uncorrelated, however we point out that the converse of this statement is not true in general, but an important exception is the case where X and Y are Gaussian r.v.'s.

In the following chapters we will require to compute the covariance between two r.v.'s which are defined by means linear combinations of other r.v.'s. Let X_1, X_2, \dots, X_n and Y_1, Y_2, \dots, Y_m be two families of r.v.'s and a_1, a_2, \dots, a_n and b_1, b_2, \dots, b_m arbitrary constants, then

$$\text{Cov} \left[\sum_{i=1}^n a_i X_i, \sum_{j=1}^m b_j Y_j \right] = \sum_{i=1}^n \sum_{j=1}^m a_i b_j \text{Cov}[X_i, Y_j], \quad (2.11)$$

where

$$\text{Cov}[X_i, Y_j] = \text{E}[X_i Y_j] - \text{E}[X_i]\text{E}[Y_j].$$

In the particular case that (2.11) is applied to $Y_j = X_i$ and $b_j = a_i$ with $m = n$, one gets a general formula for the variance of a linear combination of r.v.'s

$$\text{Var} \left[\sum_{i=1}^n a_i X_i \right] = \text{Cov} \left[\sum_{i=1}^n a_i X_i, \sum_{j=1}^n a_j X_j \right] = \sum_{i=1}^n \sum_{j=1}^n a_i a_j \text{Cov} [X_i, X_j].$$

Even more, if X_1, X_2, \dots, X_n are set of mutually independent r.v.'s then $\text{Cov} [X_i, X_j] = 0$ if $i \neq j$ and previous expression yields

$$\text{Var} \left[\sum_{i=1}^n a_i X_i \right] = \sum_{i=1}^n a_i a_i \text{Cov} [X_i, X_i] = \sum_{i=1}^n a_i^2 \text{Var} [X_i].$$

The characteristic function of a r.v. X is defined by

$$\phi_X(u) = \text{E} [e^{iuX}] = \int_{-\infty}^{\infty} e^{iux} f_X(x) dx,$$

where $f_X(x)$ denotes the p.d.f. of X , $i = \sqrt{-1}$ is the imaginary unit and u is an arbitrary real number. For two (or more) r.v.'s, say, X, Y , it can be extended as follows

$$\phi_{X,Y}(u, v) = \text{E} [e^{iuX+ivY}] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(ux+vy)} f_{X,Y}(x, y) dx dy,$$

where $f_{X,Y}(x, y)$ denotes the joint p.d.f. of X and Y , u and v are arbitrary real numbers. This function always exists and it allows to establish important properties. For instance, X, Y are independent r.v.'s if, and only if, $\phi_X(u)\phi_Y(v) = \phi_{X,Y}(u, v)$.

In closing this section we collect a set of inequalities involving the moments with respect to the origin of r.v.'s that will be used throughout this dissertation. If X and Y are r.v.'s with finite second moments then Schwarz inequality assures:

$$\text{E} [|XY|] \leq (\text{E} [X^2])^{1/2} (\text{E} [Y^2])^{1/2}. \quad (2.12)$$

A generalization of Schwarz inequality is known as the Hölder inequality:

$$\text{E} [|XY|] \leq \text{E} [|X|^n]^{1/n} \text{E} [|Y|^m]^{1/m}, \quad n, m > 1: \frac{1}{n} + \frac{1}{m} = 1.$$

Given a r.v. X such that $\text{E} [|X|^n] < +\infty$ then

$$(\text{E} [|X|^m])^{1/m} \leq (\text{E} [|X|^n])^{1/n}, \quad 0 \leq m \leq n. \quad (2.13)$$

Note that in this case, $E[|X|^m] < \infty$ holds. This result is called the Liapounov's inequality (see, [48, p.157]).

Let f be a convex function on \mathbb{R} . If $E[|X|]$ and $E[|f(X)|]$ are finite, then Jensen inequality assures

$$f(E[X]) \leq E[f(X)]. \quad (2.14)$$

In the particular case that $f(x) = x^2$, Jensen inequality (2.14) guarantees that the variance given by (2.7) is nonnegative.

In dealing with the moment of a binomial of r.v.'s the following inequalities are convenient. First, we enunciate the so-called c_s -inequality which is useful for bounding the absolute moments of a binomial expression in terms of the absolute moments of both summands. Note that it also establishes that if s -th absolute moments of X and Y are finite then the s -th absolute moment of $X + Y$ does.

$$E[|X + Y|^s] \leq c_s (E[|X|^s] + E[|Y|^s]), \quad c_s = \begin{cases} 1 & \text{if } s \leq 1, \\ 2^{s-1} & \text{if } s \geq 1. \end{cases} \quad (2.15)$$

The second result is called the Minkowski inequality

$$(E[|X + Y|^p])^{1/p} \leq (E[|X|^p])^{1/p} + (E[|Y|^p])^{1/p}, \quad p > 1.$$

Another important inequality that involves a family of r.v.'s is the following

$$E \left[\left| \frac{1}{n} \sum_{j=1}^n X_j \right|^p \right] \leq \frac{1}{n} \sum_{j=1}^n E[|X_j|^p].$$

In all the above results we assume that the indicated moments exist.

2.1.3 Stochastic processes

Random variables are adequate for describing results of random experiments which assume scalar or vector values in a given trial. In many physical applications, however, the outcomes of a random experiment are represented by functions depending upon a parameter. These outcomes are then described by a random function $X(t)$, where t is a parameter assuming values in a reference set \mathcal{T} , usually called the index set. A typical example of random

experiments giving rise to random functions is found in the theory of Brownian motion, where each coordinate of a particle executing Brownian motion exhibits random behavior as a function of time. Other examples of random functions appearing in physical contexts are thermal noise, earthquake motion, epidemics, etc.

A stochastic process (s.p.) $\{X(t) : t \in \mathcal{T}\}$ is a collection of r.v.'s, that is, for each $t \in \mathcal{T}$, $X(t)$ is a r.v. Depending on the index set \mathcal{T} one classifies the s.p. as follows: If \mathcal{T} is a countable set, the s.p. is said to be a discrete-time s.p. For instance, a sequence of r.v.'s $\{X_n : n = 0, 1, 2, \dots\}$ is a discrete-time s.p. indexed by the nonnegative integers. If \mathcal{T} is an interval of the real line, the s.p. is said to be a continuous-time s.p. For instance, $\{X(t) : t \geq 0\}$ is a continuous-time s.p. indexed by the nonnegative real numbers. Throughout this dissertation, \mathcal{T} will denote an interval.

Note that, every s.p. is defined not only on an index set \mathcal{T} , but also on a probability space, say, (Ω, \mathcal{A}, P) . Then, it is more correct to denote it as $\{X(t; \omega) : t \in \mathcal{T}, \omega \in \Omega\}$, that is, as a function of two variables although the usual notation suppresses the probability space variable ω (sometimes called as hidden parameter) for convenience. For each fixed $t \in \mathcal{T}$, $X(t; \cdot)$ denotes a r.v. on the probability space (Ω, \mathcal{A}, P) ; while for each fixed $\omega \in \Omega$, $X(\cdot; \omega)$ corresponds to a real-valued function defined on \mathcal{T} . The latter is called a sample path, trajectory, or realization of the process. The theory of stochastic processes relates these inherent random and sample path structures.

Since we have emphasized previously, at a fixed $t \in \mathcal{T}$, a s.p. $\{X(t) : t \in \mathcal{T}\}$ is a r.v. Hence, another characterization of a s.p. is to regard it as a family of r.v.'s, say, $X(t_1), X(t_2), \dots$, depending upon a parameter $t \in \mathcal{T}$. The totality of all the r.v.'s defines the s.p. $\{X(t) : t \in \mathcal{T}\}$. For discrete-time s.p.'s, this set of r.v.'s is finite or countably finite. For continuous-time s.p.'s, the number is of non-countable infinite. Thus, we see that the mathematical description of a s.p. is considerably more complicated than that of a r.v. It, in fact, is equivalent to the mathematical description of an infinite and generally non-countable number of r.v.'s. Then a s.p. $\{X(t) : t \in \mathcal{T}\}$ is defined by a family of joint distribution or density functions, taking into account that we are now dealing with infinitely many r.v.'s. More exactly, if to every finite set $\{t_1, t_2, \dots, t_n\}$ of $t \in \mathcal{T}$, there corresponds a set of r.v.'s $X_1 = X(t_1), X_2 = X(t_2), \dots, X_n =$

$X(t_n)$, have a well-defined joint probability distribution function

$$\begin{aligned} & F_{X_1, \dots, X_n}(x_1, t_1; x_2, t_2; \dots; x_n, t_n) \\ &= P(\{\omega \in \Omega: X_1(\omega) \leq x_1 \cap X_2(\omega) \leq x_2 \cap \dots \cap X_n(\omega) \leq x_n\}), \quad n \geq 1, \end{aligned} \tag{2.16}$$

then this family of joint distribution functions defines a s.p. $\{X(t) : t \in \mathcal{T}\}$. Usually, we will use the following shorten notation

$$F_n(x_1, t_1; x_2, t_2; \dots; x_n, t_n) = F_{X_1, \dots, X_n}(x_1, t_1; x_2, t_2; \dots; x_n, t_n).$$

It is called the n -th distribution function of the s.p. $\{X(t) : t \in \mathcal{T}\}$. We stress that the collection of distributions functions is not arbitrary; they must satisfy the Kolmogorov compatibility conditions:

i) For $m > n$:

$$F_m(x_1, t_1; \dots; x_n, t_n; +\infty, t_{n+1}; \dots; +\infty, t_m) = F_n(x_1, t_1; \dots; x_n, t_n)$$

ii) The joint distribution function given by (2.16) is invariant under an arbitrary permutation of the index $1, 2, \dots, n$, i.e.,

$$F_n(x_1, t_1; x_2, t_2; \dots; x_n, t_n) = F_n(x_{i_1}, t_{i_1}; x_{i_2}, t_{i_2}; \dots; x_{i_n}, t_{i_n}),$$

where $\{i_1, i_2, \dots, i_n\}$ is an arbitrary permutation of $\{1, 2, \dots, n\}$.

Given the n -th distribution function, say, $F_n(x_1, t_1; x_2, t_2; \dots; x_n, t_n)$ of an arbitrary s.p. $\{X(t) : t \in \mathcal{T}\}$, its associated n -th joint density function, assuming that it exists, is defined as

$$f_n(x_1, t_1; x_2, t_2; \dots; x_n, t_n) = \frac{\partial^n F_{X_1, \dots, X_n}(x_1, t_1; \dots; x_n, t_n)}{\partial x_1 \cdots \partial x_n}.$$

As in the case of r.v.'s, some of the most important properties of a s.p. are characterized by its moments, mainly, those of the first and second order. In the sequel, the existence of density functions shall be assumed. In terms of its first density function $f_1(x, t)$, the n -th moment of a s.p. $\{X(t) : t \in \mathcal{T}\}$ at a fixed $t \in \mathcal{T}$ is defined by

$$\alpha_n(t) = \mathbb{E}[(X(t))^n] = \int_{-\infty}^{+\infty} (x(t))^n f_1(x, t) dx.$$

The first moment $\alpha_1(t)$, represents the mean of the s.p. at t . It will be sometimes denoted by $\mu_X(t)$ or $E[X(t)]$. $\alpha_2(t)$ represents the mean square value of the s.p. at t . The n -th central moment of $\{X(t) : t \in \mathcal{T}\}$ at a given $t \in \mathcal{T}$ is

$$\mu_n(t) = E[(X^n(t) - \mu_X(t))^n] = \int_{-\infty}^{+\infty} (x(t) - \mu_X(t))^n f_1(x, t) dx.$$

Of particular interest is $\mu_2(t)$, that represents the variance of the s.p. at t , and it will be denoted by $\sigma_{X(t)}^2$ or $\text{Var}[X(t)]$.

2.2 Mean square calculus

In these pages we are mainly interested in the study of random differential equations, that is, differential equations containing uncertainty in their formulation. In dealing with this goal there mainly exist two different approaches. The first one, namely, sample-function approach is based upon the solution s.p. $X(t; \omega)$ of a random differential equation can be considered as a collection of individual deterministic functions $X(t; \cdot)$. Each function or realization $X(t; \cdot)$ of the solution s.p. comes from the unique solution trajectory of the deterministic sample differential equation that one obtains for a fixed $\omega \in \Omega$ if the problem is well-posed. Then the sample treatment of random differential equation depends heavily on the theory of deterministic differential equations in real analysis. Hence, all s.p.'s involved in the random differential equations are assumed to possess continuous trajectories. This means, quite naturally, that physical random systems described in terms of random differential equations should have continuous realizations [56, Appendix A]. On the other hand, the collection of trajectories can be investigated with the tools of probability theory if it possess certain measurability properties in order that they define rigorously a s.p.

In this dissertation we are interested in a second approach usually referred to as mean square approach because it requires to introduce previously a special calculus (called the mean square calculus, in short, m.s. calculus) to manage certain class of s.p.'s (called second order s.p.'s). In particular, since we are interested in solving random differential equations, it is clear that we will need to handle derivatives and integrals of s.p.'s, hence we speak of limiting operations.

For that it is convenient to have a Borel field structure for guaranteeing closure under countable set operations, because such limits would lead to quantities to which probabilities might not be readily assigned. The mean square calculus is important for several practical reasons. First of all, its importance lies in the fact that simple yet powerful and well-developed methods are available. Secondly, the development of m.s. calculus and its applications to physical problems follows in a broad outline the same steps in considering calculus of deterministic functions. Even more, many important results belonging to m.s. calculus can be stated in terms of the (deterministic) correlation function (or equivalently, the covariance function) associated to a s.p., then we can take advantage of the numerous results available from the well-developed differential and integral ordinary calculus. As an important difference with respect to sample-function approach, it is worth to point out that m.s. approach do not require to impose regular behavior conditions to the process trajectories like continuity. Another important reason for studying m.s. calculus is that, for the important case of Gaussian processes, m.s. properties lead to properties on the sample function level [48, p.485].

For the sake of clarity in the presentation, in this section we recall some concepts, notations and results related to m.s. calculus that may be found in [56, chap.4]. Let (Ω, \mathcal{A}, P) be a probability space, here we are interested in real r.v.'s $X : \Omega \rightarrow \mathbb{R}$ whose second moment with respect to the origin is finite, that is, $E[X^2] < +\infty$. These r.v.'s are called of second order and in the following they will be denoted by (2-r.v.'s). From Schwarz inequality (2.12) and property (2.6) it follows directly

$$E[(X + Y)^2] < +\infty, \quad E[(cX)^2] = c^2 E[X^2] < +\infty, \quad \forall c \in \mathbb{R}.$$

Hence the class of all 2-r.v.'s on a probability space constitute a real linear space if all equivalent r.v.'s are identified by the the following equivalence binary relationship: two r.v.'s X and Y are said to be equivalent if, and only if, $P(X \neq Y) = 0$.

If we define

$$\langle X, Y \rangle = E[XY], \tag{2.17}$$

by Schwarz inequality (2.12) and Jensen inequality (2.14), one gets $|\langle X, Y \rangle| < \infty$. It is easy to show that $\langle \rangle$ given by (2.17) defines an inner product when X and Y are 2-r.v.'s. This inner product automatically defines a norm

$$\|X\|_2 = +\sqrt{\langle X, X \rangle} = (E[X^2])^{1/2}, \tag{2.18}$$

and a metric

$$d(X, Y) = \|X - Y\|_2 = +\sqrt{\mathbb{E}[(X - Y)^2]}.$$

The linear vector space of 2-r.v.'s with this inner product, the norm and the distance defined above is called a L_2 -space. The convergence that above norm defines in L_2 is referred to as mean square (m.s.) convergence. We say that a sequence of 2-r.v.'s $\{X_n : n \geq 0\}$ is m.s. convergent to X if

$$\lim_{n \rightarrow \infty} \|X_n - X\|_2 = \lim_{n \rightarrow \infty} \left(\mathbb{E} [(X_n - X)^2] \right)^{1/2} = 0.$$

This type of convergence is often expressed by

$$\text{l.i.m.}_{n \rightarrow \infty} X_n = X \quad \text{or} \quad X_n \xrightarrow[n \rightarrow \infty]{\text{m.s.}} X.$$

A sequence $\{X_n : n \geq 0\}$ in L_2 is said to be a Cauchy or fundamental sequence in the m.s. sense if

$$\|X_n - X_m\|_2 \xrightarrow[n, m \rightarrow \infty]{} 0.$$

It can be proven that a sequence $\{X_n : n \geq 0\}$ of r.v.'s is m.s. convergent to a r.v. X as $n \rightarrow \infty$ if, and only if, the sequence is m.s. fundamental. This result means that the L_2 is an inner product space which, as a metric space, is complete. Then by definition, L_2 is a Hilbert space.

The following desirable property motivates in certain sense the use of m.s. convergence:

Lemma 2.1. (*[56, p.76]*) *Let $\{X_n : n \geq 0\}$ be a sequence of 2-r.v.'s m.s. convergent to the 2-r.v. X , i.e.,*

$$X_n \xrightarrow[n \rightarrow \infty]{\text{m.s.}} X.$$

Then

$$\lim_{n \rightarrow \infty} \mathbb{E}[X_n] = \mathbb{E}[X]. \tag{2.19}$$

The s.p.'s $\{X(t) : t \in \mathcal{T}\}$ that we consider in the L_2 space are called second order s.p.'s (2-s.p.'s), and they satisfy that for each $t \in \mathcal{T}$, r.v.'s $X(t)$ are 2-r.v.'s, that is, $\mathbb{E}[(X(t))^2] < +\infty$ for each $t \in \mathcal{T}$.

If $\{X(t) : t \in \mathcal{T}\}$ is a 2-s.p., then its expectation $\mathbb{E}[X(t)]$ is always finite on \mathcal{T} . In fact, for each $t \in \mathcal{T}$, $X(t)$ is a 2-r.v., then by applying Cauchy-Schwarz inequality (2.12) with $X = X(t)$ and $Y = 1$, one gets

$$\mathbb{E}[|X(t)|] \leq \mathbb{E}[(X(t))^2] < +\infty.$$

Then by Jensen's inequality (2.14): $|\mathbb{E}[X(t)]| \leq \mathbb{E}[|X(t)|] < +\infty$ for each $t \in \mathcal{T}$. Therefore, we can consider a new s.p. $\{\tilde{X}(t) = X(t) - \mathbb{E}[X(t)] : t \in \mathcal{T}\}$ which is also a 2-s.p. and, moreover, it satisfies $\mathbb{E}[\tilde{X}(t)] = 0$. Hence, there is no loss of generality in assuming that the mean of a 2-s.p. is zero. In this case, we say that the s.p. is centered. As an extension of the covariance concept for r.v.'s, for each 2-s.p., we can associate it the following deterministic function in two-variables

$$\text{Cov}[X(t), X(s)] = \mathbb{E}[X(t)X(s)] - \mathbb{E}[X(t)]\mathbb{E}[X(s)], \quad t, s \in \mathcal{T}.$$

Note that as a consequence of Cauchy-Schwarz inequality (2.12), this function always exists if $X(t)$ is a 2-s.p. If we set $t = s$, then the covariance function becomes the variance function

$$\text{Var}[X(t)] = \text{Cov}[X(t), X(t)] = \mathbb{E}[(X(t))^2] - (\mathbb{E}[X(t)])^2, \quad t, s \in \mathcal{T}.$$

Since we have pointed out previously, we can assume without loss of generality that $\mathbb{E}[X(t)] = 0$, then the covariance function becomes

$$\Gamma_X(t, s) = \mathbb{E}[X(t)X(s)], \quad t, s \in \mathcal{T},$$

which is called the correlation function. This deterministic function characterizes every 2-s.p. in the sense that $\{X(t) : t \in \mathcal{T}\}$ is a 2-s.p. if, and only if, its correlation function $\Gamma_X(t, s)$ exists and is finite on $\mathcal{T} \times \mathcal{T}$. From above exposition, we realize that the correlation function or, equivalently, the covariance function determines the second-order properties of a 2-s.p. That is, second-order properties are those properties that are derivable from its covariance (or equivalently correlation) function. It is easy to see, therefore, that there is inherently a close relationship between linear operations (defined by membership in L_2) and second-order properties. Often, questions of second-order properties are most lucidly answered in the framework of the Hilbert space theory.

A more specific terminology for $\Gamma_X(t, s)$ is the auto-correlation function. It is usually denoted by $\Gamma_{X,X}(t, s)$. It is intended to distinguishing it from the cross-correlation function, defined by

$$\Gamma_{X,Y}(t, s) = \mathbb{E}[X(t)Y(s)], \quad t, s \in \mathcal{T}, \quad (2.20)$$

where the involved r.v.'s belong to two different s.p.'s, say, $\{X(t) : t \in \mathcal{T}\}$ and $\{Y(t) : t \in \mathcal{T}\}$.

The concept of m.s. convergence can be extended from a sequence of 2-r.v.'s to a 2-s.p. $\{X(t) : t \in \mathcal{T}\}$ where t lies over a finite interval \mathcal{T} . As in the analysis of ordinary functions, this extension leads to the notion of continuity and, in this case, continuity in the m.s. sense. Before introducing this concept as well as m.s. differentiation and m.s. integration, we need some previous results.

Proposition 2.2. ([61, p.7]) *Let X, Y be 2-r.v.'s and $\{X_n : n \geq 0\}$ be a sequence of 2-r.v.'s such that*

- i) $X_n \xrightarrow[n \rightarrow \infty]{m.s.} X$,
- ii) X_n is independent of Y for each $n \geq 0$,

then, X and Y are also independent r.v.'s.

Proof. Let $\phi_{X,Y}$ be the joint characteristic function of two r.v.'s X and Y . To prove that X and Y are independent, it is equivalent to show that

$$\phi_{X,Y}(u, v) = \phi_X(u)\phi_Y(v), \quad u, v \in \mathbb{R}.$$

Since $X_n \xrightarrow[n \rightarrow \infty]{m.s.} X$, one follows that $X_n \xrightarrow[n \rightarrow \infty]{d} X$, and it entails

$$\phi_{X_n}(u) \xrightarrow[n \rightarrow \infty]{} \phi_X(u), \quad u \in \mathbb{R}. \tag{2.21}$$

On the other hand, since $X_n \xrightarrow[n \rightarrow \infty]{d} X$ and $Y \xrightarrow[n \rightarrow \infty]{d} Y$, one gets

$$\phi_{X_n,Y}(u, v) \xrightarrow[n \rightarrow \infty]{} \phi_{X,Y}(u, v), \quad u, v \in \mathbb{R}. \tag{2.22}$$

By ii) and (2.21) and (2.22), we find

$$\phi_{X_n,Y}(u, v) = \phi_{X_n}(u)\phi_Y(v) \xrightarrow[n \rightarrow \infty]{} \phi_X(u)\phi_Y(v), \quad u, v \in \mathbb{R}. \tag{2.23}$$

By (2.22) and (2.23), we obtain

$$\phi_{X,Y}(u, v) = \phi_X(u)\phi_Y(v), \quad u, v \in \mathbb{R}.$$

Thus, X and Y are independent. \square

Lemma 2.2. ([56, p.88]) Let $\{X_n: n \geq 0\}$ and $\{Y_n: n \geq 0\}$ be two sequences of 2-r.v.'s m.s. convergent to the 2-r.v.'s X and Y , respectively, i.e.,

$$X_n \xrightarrow[n \rightarrow \infty]{m.s.} X, \quad Y_n \xrightarrow[n \rightarrow \infty]{m.s.} Y.$$

Then

$$\lim_{n \rightarrow \infty} E[X_n Y_n] = E[XY].$$

If we take $Y_n = X_n$, above Lemma 2.2 yields $E[(X_n)^2] \xrightarrow[n \rightarrow \infty]{} E[X^2]$. As a consequence, by applying (2.19) and (2.7), one follows that $\text{Var}[X_n] \xrightarrow[n \rightarrow \infty]{} \text{Var}[X]$. Summarizing,

Lemma 2.3. Let $\{X_n: n \geq 0\}$ a sequence of 2-r.v. m.s. convergent to the 2-r.v.'s X , then

$$\lim_{n \rightarrow \infty} E[X_n] = E[X], \quad \lim_{n \rightarrow \infty} \text{Var}[X_n] = \text{Var}[X].$$

The following results provides a necessary and sufficient condition for m.s. convergence in terms of correlation function.

Theorem 2.1. ([56, p.89]) Let $\{X_n(t): t \in \mathcal{T}, n \geq 0\}$ be a sequence of 2-s.p.'s. $\{X_n(t): t \in \mathcal{T}\}$ is m.s. convergent to a 2-s.p. $\{X(t): t \in \mathcal{T}\}$ as $n \rightarrow n_0$, if, and only if, the deterministic functions $E[X_n(t)X_{n'}(t)]$ converge to a finite function on \mathcal{T} as $n, n' \rightarrow n_0$ in any manner whatever. Then

$$\Gamma_{X_n}(t, s) \xrightarrow[n \rightarrow n_0]{} \Gamma_X(t, s),$$

on $\mathcal{T} \times \mathcal{T}$.

2.2.1 Mean square continuity

A 2-s.p. $\{X(t): t \in \mathcal{T}\}$ is said to be m.s continuous at $t \in \mathcal{T}$ if

$$\text{l.i.m.}_{\tau \rightarrow 0} X(t + \tau) = X(t),$$

for $t + \tau \in \mathcal{T}$, or

$$\lim_{\tau \rightarrow 0} \|X(t + \tau) - X(t)\|_2 = 0.$$

Example 2.2. Let $\{X_n : n \geq 1\}$ be a sequence of r.v.'s in L_2 and let $t \in \mathcal{T}$, \mathcal{T} being an interval of \mathbb{R} , then the 2-s.p. $\{n_0 X_{n_0} t^{n_0-1} : t \in \mathcal{T}\}$ for each $n_0 \in \mathbb{N}$ is m.s. continuous for each $t \in \mathcal{T}$. In fact,

$$\|n_0 X_{n_0} (t + \tau)^{n_0-1} - n_0 X_{n_0} t^{n_0-1}\|_2 = n_0 |(t + \tau)^{n_0-1} - t^{n_0-1}| \|X_{n_0}\|_2 \xrightarrow{\tau \rightarrow 0} 0,$$

because of $\|X_{n_0}\|_2 < +\infty$ as $X_{n_0} \in L_2$ for each $n_0 \in \mathbb{N}$ and the continuity of the deterministic function $f(t) = t^{n_0-1}$ with respect to t .

The following result provides a necessary and sufficient condition for m.s. continuity of a 2-s.p. in terms of its correlation function which is a deterministic function.

Theorem 2.2. A 2-s.p. $\{X(t) : t \in \mathcal{T}\}$ is m.s. continuous at t if, only if, its correlation function $\Gamma_X(t, s)$ is continuous at (t, t) .

If $\{X(t) : t \in \mathcal{T}\}$ is m.s. continuous at every $t \in \mathcal{T}$, then we say that $\{X(t) : t \in \mathcal{T}\}$ is m.s. continuous on the set \mathcal{T} . From Theorem 2.2, an analogous characterization of the m.s. continuity on a set \mathcal{T} can be established in terms of the continuity of the correlation function on $\mathcal{T} \times \mathcal{T}$. It is important to bear in mind that, in the discussion of m.s. continuity of $\Gamma_X(t, s)$, it is the joint continuity of the pair (t, s) that is of interest. Theorem 2.2 justifies in part the statement made at the beginning of this section: *the development of m.s. calculus follows in broad outline that of ordinary calculus.*

In practice, many times is very difficult to check that a s.p. is continuous in some probability sense, say, m.s., in probability or almost-surely, for instance. There is an useful concept, namely almost-surely sample continuity (or path continuity), which will be introduced in chapter 6, that it allows us to establish important results in the m.s. sense like the chain rule for s.p.'s (see Theorem 6.2). In chapter 6, we will state the celebrated Kolmogorov theorem that provides practical sufficient conditions for guaranteing that the path continuity of a s.p.

In closing, let us remark that m.s. continuity does not imply continuity properties at the sample function level, see [56, p.92] and [51, p.337].

2.2.2 Mean square differentiation

The concept of m.s. differentiation follows naturally from that of m.s. continuity. A 2-s.p. $\{X(t): t \in \mathcal{T}\}$ has a m.s. derivative $\dot{X}(t)$ at $t \in \mathcal{T}$ if

$$\text{l.i.m.}_{\tau \rightarrow 0} \frac{X(t + \tau) - X(t)}{\tau} = \dot{X}(t),$$

for $t + \tau \in \mathcal{T}$, or

$$\lim_{\tau \rightarrow 0} \left\| \frac{X(t + \tau) - X(t)}{\tau} - \dot{X}(t) \right\|_2 = 0.$$

Higher order derivatives are defined analogously.

Example 2.3. Consider the 2-s.p. $X(t) = At^n$, where A is a 2-r.v. and $t \in \mathcal{T}$. The m.s. derivative of $X(t)$ is given by $\dot{X}(t) = Ant^{n-1}$. In fact,

$$\begin{aligned} \lim_{\tau \rightarrow 0} \left\| \frac{X(t + \tau) - X(t)}{\tau} - \dot{X}(t) \right\|_2 &= \lim_{\tau \rightarrow 0} \left\| \frac{A(t + \tau)^n - At^n}{\tau} - Ant^{n-1} \right\|_2 \\ &= \|A\|_2 \lim_{\tau \rightarrow 0} \left| \frac{(t + \tau)^n - t^n}{\tau} - nt^{n-1} \right| = 0, \end{aligned}$$

where in the last step we have applied that $\|A\|_2 < \infty$ (because $A \in L_2$) and the deterministic differentiability of the function $f(t) = t^n$. In the same way it is straightforward to check that $\ddot{X}(t) = An(n-1)t^{n-2}$.

Next we state a characterization of m.s. differentiation of a 2-s.p. in terms of its correlation function.

Theorem 2.3. A 2-s.p. $\{X(t): t \in \mathcal{T}\}$ is m.s. differentiable at t if, only if, the second generalized derivative

$$\lim_{\tau, \tau' \rightarrow 0} \frac{\Gamma_X(t + \tau, s + \tau') - \Gamma_X(t + \tau, s) - \Gamma_X(t, s + \tau') + \Gamma_X(t, s)}{\tau \tau'},$$

exists at (t, t) and is finite.

The concept of m.s. differentiation can be extended directly on a set \mathcal{T} . M.s. differentiation has some analogous properties like those that its deterministic counterpart concept does. For instance, the m.s. derivative of a 2-s.p., if it exists, is unique; m.s. differentiation implies m.s continuity; m.s. differentiation is linear; etc. However, differentiation rules can not be directly extended from the deterministic framework to m.s. sense. As an illustrative example, we stress that, in general, the classical differentiation rule of the product is not true, without additional hypotheses, when factors are 2-s.p.'s., [62].

A nice m.s. differentiation property that links classical and stochastic differentiation is given by the following relationship

$$\mathbb{E}\left[X^{(n)}(t)\right] = \frac{d^n}{dt^n}\mathbb{E}[X(t)], \quad (2.24)$$

whenever $X(t)$ is n -times m.s. differentiable. Note that in (2.24), $\frac{d^n}{dt^n}$ denotes the derivative of order n in the deterministic sense, whereas $X^{(n)}(t)$ means the m.s. derivative of order n of s.p. $X(t)$. Generalizations of (2.24) can be found in [56, p.98]

$$\Gamma_{X^{(n)}, X^{(m)}}(t, s) = \mathbb{E}\left[X^{(n)}(t)X^{(m)}(s)\right] = \frac{\partial^{n+m}\Gamma_X(t, s)}{\partial t^n \partial s^m}, \quad (2.25)$$

whenever $X(t)$ is h -times m.s. differentiable, where $h = \max(n, m)$. In (2.25), $\Gamma_{X^{(n)}, X^{(m)}}(t, s)$ is the cross-correlation function of the s.p.'s $\{X^{(n)}(t) : t \in \mathcal{T}\}$ and $\{X^{(m)}(t) : t \in \mathcal{T}\}$ defined by (2.20).

The concept of m.s. differentiability leads to the notion of m.s. analyticity. A 2-s.p. $\{X(t) : t \in \mathcal{T}\}$ is said to be m.s. analytic on \mathcal{T} if it can be expanded in the m.s. convergent series

$$X(t) = \sum_{n=0}^{\infty} \frac{X^{(n)}(t_0)}{n!} (t - t_0)^n, \quad t, t_0 \in \mathcal{T}. \quad (2.26)$$

Again, this concept can be characterized in terms of the correlation function of the s.p. $X(t)$

Theorem 2.4. *A 2-s.p. $\{X(t) : t \in \mathcal{T}\}$ is m.s. analytic on \mathcal{T} if, only if, its correlation function $\Gamma_X(t, s)$ is analytic at (t, t) for every $t \in \mathcal{T}$, and then $\Gamma_X(t, s)$ is analytic on $\mathcal{T} \times \mathcal{T}$.*

Example 2.4. Let us consider the s.p. defined by $\{B(t) = B_1 t + B_0 : t \geq 0\}$ where B_0 and B_1 are standard Gaussian r.v.'s, i.e., $B_0, B_1 \sim N(\mu = 0; \sigma = 1)$ such that B_0 and B_1 are correlated: $E[B_0 B_1] = \gamma \in [-1, 1]$. Note that $E[B(t)] = 0$ since $E[B_0] = E[B_1] = 0$ and

$$\Gamma_B(t, t) = E \left[(B(t))^2 \right] = t^2 + 2\gamma t + 1 \geq 0.$$

Then by Theorem 2.4, $B(t)$ is m.s. analytic on the real line.

Next, we state a m.s. differentiation theorem of series, that may be found in [15]:

Theorem 2.5. Assume that for $n \geq 1$, the process $\{X_n(r) : r \in I\}$ satisfies

- i) $X_n(r)$ is m.s. differentiable and $\dot{X}_n(r)$ is m.s. continuous,
- ii) $X(r) = \sum_{n \geq 1} X_n(r)$ is m.s. convergent,
- iii) $\sum_{n \geq 1} \dot{X}_n(r)$ is m.s. uniform convergent in a neighborhood of each $r \in I$.

Then, for each $r \in I$, $X(r)$ is m.s. differentiable and

$$\dot{X}(r) = \sum_{n \geq 1} \dot{X}_n(r).$$

An important fact is that the 2-norm in L_2 does not provide a Banach algebra structure, i.e., it is not sub-multiplicative. Indeed, property

$$\|XY\|_2 \leq \|X\|_2 \|Y\|_2, \tag{2.27}$$

does not hold, in general, as it is easy to check.

Example 2.5. In fact, let $X = Y = U^{1/2}$ be two identical r.v.'s such that U is a Uniform r.v. on $[0, 1]$, then

$$\|XY\|_2 = (E[U^2])^{1/2} = \sqrt{3}/3 \geq 1/2 = E[U] = \|X\|_2 \|Y\|_2.$$

More generally, let Z be a non-constant positive 2-r.v. and let us take $X = Y = Z^{1/2}$, then

$$(\|XY\|_2)^2 - (\|X\|_2)^2 (\|Y\|_2)^2 = E[Z^2] - (E[Z])^2 = \text{Var}[Z] > 0$$

Note that, if X and Y are independent r.v.'s, then inequality (2.27) holds true, since in this case inequality (2.27) becomes an equality.

2.2.3 Mean square and mean fourth calculus

In chapters 5-7, we will be interested in constructing power series solution of random differential equations where uncertainty enters into the equation in a multiplicative way. The rigorous proof that such random series converge in the m.s. sense requires to use the following basic property

$$AX_n \xrightarrow[n \rightarrow \infty]{\text{m.s.}} AX. \tag{2.28}$$

This property holds true if $A \in L_2$, $\{X_n : n \geq 0\}$ is a sequence of 2-r.v.'s such that $X_n \xrightarrow[n \rightarrow \infty]{\text{m.s.}} X$ and A, X_n are independent r.v.'s for each n . However, independence hypothesis cannot be assumed in many practical cases like those that we have mentioned previously and will appear in chapters 5-7. Next example shows that property (2.28) does not hold, in general. We want to underline that it has been constructed from another one provided by courtesy of professor C.A. Braumann.

Example 2.6. Let $\mathcal{A} = \sigma(A)$ be the σ -algebra generated by the r.v.

$$A = \begin{cases} 0 & \text{w.p. } 1 - \frac{1}{n}, \\ 1 & \text{w.p. } \frac{1}{n}, \end{cases}$$

and let us consider the following sequence of r.v.'s defined through r.v. A

$$X_n = \begin{cases} 0 & \text{w.p. } 1 - \frac{1}{n}, \\ n^{A/2} & \text{w.p. } \frac{1}{n}. \end{cases}$$

In advance, let us assume that $X = 0$ w.p. 1. Note that by conditional expectation properties one gets

$$\mathbb{E} \left[(X_n - X)^2 | \mathcal{A} \right] = \mathbb{E} \left[(X_n - X)^2 | A \right] = n^A \times \frac{1}{n} + 0 \times \left(1 - \frac{1}{n} \right) = n^{A-1},$$

hence

$$\begin{aligned} \mathbb{E} \left[(X_n - X)^2 \right] &= \mathbb{E} \left[\mathbb{E} \left[(X_n - X)^2 | \mathcal{A} \right] \right] \\ &= n^{0-1} \times \left(1 - \frac{1}{n} \right) + n^{1-1} \times \frac{1}{n} \\ &= \frac{2}{n} - \frac{1}{n^2} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

So $X_n \xrightarrow[n \rightarrow \infty]{m.s.} 0 = X$. On the other hand,

$$\begin{aligned} (AX_n)^2 &= \begin{cases} 0 & \text{w.p. } 1 - \frac{1}{n}, \\ A^2 n^A & \text{w.p. } \frac{1}{n}, \end{cases} \\ \mathbb{E} \left[(AX_n - 0)^2 \right] &= \mathbb{E} \left[\mathbb{E} \left[(AX_n)^2 | \mathcal{A} \right] \right] = \mathbb{E} \left[A^2 n^A \right] \\ &= 0^2 n^0 \times \left(1 - \frac{1}{n} \right) + 1^2 \times n^1 \times \frac{1}{n} = 1 \neq 0. \end{aligned}$$

So, in this case we do not have property (2.28). In order to establish it, we will need to impose additional hypotheses.

This motivates the introduction of r.v.'s X such that $\mathbb{E}[X^4] < +\infty$ which will be denoted by 4-r.v.'s. By (2.5) a 4-r.v. is a 2-r.v. The set L_4 of all the 4-r.v.'s endowed with the norm

$$\|X\|_4 = \sqrt[4]{\mathbb{E}[X^4]}, \quad (2.29)$$

is a Banach space (see [3, p.9]). In the same way that 2-norm, the 4-norm $\|\cdot\|_4$ is not sub-multiplicative as it is easy to check by a counterexample similar to Example 2.5. A s.p. $\{X(t) : t \in \mathcal{T}\}$, where $\mathbb{E}[(X(t))^4] < +\infty$ for all $t \in \mathcal{T}$, will be called a 4-s.p. A sequence of 4-r.v.'s $\{X_n : n \geq 0\}$ is said to be mean fourth (m.f.) convergent to a 4-r.v. X if

$$\lim_{n \rightarrow \infty} \|X_n - X\|_4 = 0. \quad (2.30)$$

This type of convergence will be represented by $X_n \xrightarrow[n \rightarrow \infty]{m.f.} X$. The following lemma establishes the link between m.s. and m.f. convergence.

Lemma 2.4. *Let $\{X_n : n \geq 0\}$ be a sequence of 4-r.v.'s and suppose that $X_n \xrightarrow[n \rightarrow \infty]{m.f.} X$. Then $X_n \xrightarrow[n \rightarrow \infty]{m.s.} X$.*

Proof. Using the Schwarz inequality (2.12), one gets

$$(\|X_n - X\|_2)^2 = \mathbb{E}[1 \times (X_n - X)^2] \leq 1 \times \left(\mathbb{E}[(X_n - X)^4] \right)^{1/2} = (\|X_n - X\|_4)^2.$$

Since $\|X_n - X\|_4 \xrightarrow[n \rightarrow \infty]{} 0$ (because X_n is m.f. convergent to X), it immediately follows that $\|X_n - X\|_2 \xrightarrow[n \rightarrow \infty]{} 0$ and therefore $X_n \xrightarrow[n \rightarrow +\infty]{m.s.} X$. \square

Now, we can give sufficient conditions in order to property (2.28) holds true without assuming hypotheses based on independence.

Lemma 2.5. *Let A be a 2-r.v. and $\{X_n : n \geq 0\}$ a sequence of 4-r.v.'s such that $X_n \xrightarrow[n \rightarrow \infty]{m.f.} X$. Then $AX_n \xrightarrow[n \rightarrow \infty]{m.s.} AX$.*

Proof. By the definition of the norm $\|\cdot\|_2$ (see 2.18), we have

$$(\|A(X_n - X)\|_2)^2 = \mathbb{E}[A^2(X_n - X)^2]. \tag{2.31}$$

On the other hand, the hypothesis $X_n \xrightarrow[n \rightarrow \infty]{m.f.} X$ implies by definition $(X_n - X)^2 \xrightarrow[n \rightarrow \infty]{m.s.} 0$ and, as clearly $A^2 \xrightarrow[n \rightarrow \infty]{m.s.} A^2$, then from (2.31) and Lemma 2.2 one obtains $\|A(X_n - X)\|_2 \xrightarrow[n \rightarrow \infty]{} 0$ and hence $AX_n \xrightarrow[n \rightarrow \infty]{m.s.} AX$. \square

In closing this section, we underline that an extensive treatment of m.f. calculus in connection with m.s. calculus is provided in the paper [62].

2.2.4 Random variables of class p

Since we have already underlined, one of the most important problems to be faced in this piece is the lack of sub-multiplicativity of the 2-norm as well as 4-norm. This drawback becomes especially challenging when uncertainty enters into the random differential equation in a multiplicative way through a r.v.. (see chapters 5-7). Now, we introduce the class of r.v.'s to be play this role. First, we need to remember the following

Definition 2.1. ([22, p.335]) Let $\alpha_n = \{\alpha_n : n \geq 0\}$, $\beta_n = \{\beta_n : n \geq 0\}$ be two sequences of real numbers. The sequence α_n is said to be of order $\mathcal{O}(\beta_n)$ as $n \rightarrow \infty$, and this will be denoted by $\alpha_n = \mathcal{O}(\beta_n)$, if there is a constant $M > 0$ and a positive integer n_0 such that

$$|\alpha_n| \leq M |\beta_n|, \quad \forall n \geq n_0. \quad (2.32)$$

Equivalently, we say that $\alpha_n = \mathcal{O}(\beta_n)$ if $\left| \frac{\alpha_n}{\beta_n} \right|$ is bounded for $n \geq n_0$. Note that if the following limit

$$\lim_{n \rightarrow \infty} \left| \frac{\alpha_n}{\beta_n} \right|$$

exist and is finite, then $\alpha_n = \mathcal{O}(\beta_n)$.

Definition 2.2. Let A be a r.v. with finite absolute moments (with respect the origin). We say that A is of class p if there exists $p \geq 0$ such that

$$\|A^{n+1}\| = \mathcal{O}(n^p) \|A^n\|. \quad (2.33)$$

Taking $\alpha_n = \|A^{n+1}\| / \|A^n\|$ and $\beta_n = n^p$ in Definition 2.1, from (2.32) and a recurrence argument it is easy to see that condition (2.33) yields

$$\begin{aligned} \|A^{n+1}\| &\leq M^{n-n_0+1} (n(n-1)\cdots n_0)^p \|A^{n_0}\| \\ &= \frac{M^{1-n_0}}{((n_0-1)!)^p} M^n (n!)^p \|A^{n_0}\|, \quad p \geq 0, \quad \forall n \geq n_0. \end{aligned}$$

Since A has finite absolute moments, let $M_1 = \|A^{n_0}\| < +\infty$, then last condition writes as

$$\|A^n\| \leq H M^{n-1} ((n-1)!)^p, \quad p \geq 0, \quad \forall n > n_0, \quad (2.34)$$

where $H = M^{1-n_0} M_1 / ((n_0 - 1)!)^p$ and M are positive constants.

Now, let us show that the set of this type of r.v.'s is not empty. In fact, we are going to check that several relevant families of r.v.'s used extensively in probability, such that, Uniform, Exponential, Beta, Gamma or Gaussian, belong to this class.

Example 2.7. Let A be a Beta r.v. of parameters $\alpha > 0$, $\beta > 0$, i.e., $A \sim \text{Be}(\alpha; \beta)$. In this case A satisfies condition (2.33) for the 4-norm with $p = 0$. In fact

$$\frac{\|A^{n+1}\|_4}{\|A^n\|_4} = \sqrt[4]{\frac{\prod_{j=0}^3 (\alpha + 4n + j)}{\prod_{j=0}^3 (\alpha + \beta + 4n + j)}} = \mathcal{O}(1), \quad (2.35)$$

where we have used that the moments with respect to the origin of a Beta r.v. are given by (see Table 2.2)

$$E[A^k] = \frac{(\alpha + k - 1)!(\alpha + \beta - 1)!}{(\alpha + \beta + k - 1)!(\alpha - 1)!}, \quad k = 0, 1, 2, \dots$$

Since a Uniform r.v. on the interval $[0, 1]$ corresponds to a Beta r.v. of parameters $\alpha = \beta = 1$, this type of r.v.'s belong to r.v.'s of class p with $p = 0$.

Example 2.8. Let A be a Gamma r.v. of parameters $r > 0$, $a > 0$, i.e., $A \sim \text{Ga}(r; a)$. In this case A satisfies condition (2.33) for the 4-norm with $p = 1$. In fact

$$\frac{\|A^{n+1}\|_4}{\|A^n\|_4} = \sqrt[4]{\frac{(r + 4n)(r + 4n + 1)(r + 4n + 2)(r + 4n + 3)}{a^4}} = \mathcal{O}(n), \quad (2.36)$$

where we have used that the moments with respect to the origin of a Gamma r.v. are given by (see Table 2.2)

$$E[A^k] = \frac{r(r + 1) \cdots (r + k - 1)}{a^k}, \quad k = 0, 1, 2, \dots$$

By taking $a = \lambda$ and $r = 1$, a Gamma r.v. $A \sim \text{Ga}(r; a)$ becomes an Exponential r.v. $A \sim \text{Exp}(\lambda = a)$. As a consequence, Exponential r.v.'s also satisfies condition (2.33) for $p = 1$. In this case, from (2.36) one follows

$$\frac{\|A^{n+1}\|_4}{\|A^n\|_4} = \sqrt[4]{\frac{(4n+1)(4n+2)(4n+3)(4n+4)}{\lambda^4}} = \mathcal{O}(n). \quad (2.37)$$

Example 2.9. Let A be a Gaussian r.v. of parameters $\mu = 0$, $\sigma^2 > 0$, i.e., $A \sim \text{N}(0; \sigma^2)$. In this case A satisfies condition (2.33) for the 4-norm with $p = 1/2$. In fact,

$$\frac{\|A^{n+1}\|_4}{\|A^n\|_4} = \sigma \sqrt[4]{(4n+3)(4n+1)} = \mathcal{O}(n^{1/2}), \quad (2.38)$$

since

$$\mathbb{E}[A^{4k}] = \frac{\sigma^{4k} (4k)!}{2^{2k} (2k)!}. \quad (2.39)$$

Chapter 3

Mean square analytic solutions of coupled linear difference models with random initial condition and source term

This chapter deals with the construction of random discrete solutions of coupled linear difference equations incorporating uncertainty into both, the initial condition and the source term. First, sufficient conditions in order to guarantee the mean square stability of the discrete solution stochastic process are provided, then the main statistical functions, such as mean and covariance of the discrete solution stochastic process are given. Finally, illustrative examples of potential interest in applications like compartmental models are shown.

3.1 Introduction

Over the last few years, systems modeling has focused on producing more useful and realistic tools through the introduction of random factors inherent in the real world. This is achieved taking into account the complex nature of phenomena being studied and the possibility of measuring device errors. In the solution of these problems, the incorporation of random differential equations have been particularly fruitful. As a result, a significant number of recently published papers have been devoted to presenting mathematical models based on differential equations incorporating uncertainty into their formulation. Most of these contributions have focused on introducing uncertainty into these models by means of white noise, which is a Gaussian and stationary stochastic process. These types of equations driven by white noise have to be interpreted mathematically as Itô equations. In the literature these equations are referred to stochastic differential equations [38, 47]. Whereas, the term random differential equations is reserved for those differential equations that introduce randomness through another kind of stochastic processes [56]. Recent interesting contributions in both, random and stochastic differential equation approaches, can be found in [26] and [24]. Despite evident applicability to deterministic framework, the corresponding stochastic or random discrete models still require attention because research efforts have been mainly focused on the development of continuous models.

In this chapter, we consider random coupled difference systems of the form

$$\vec{X}_{n+1} = \mathbf{A}\vec{X}_n + \vec{B}_n, \quad n = 1, 2, \dots, \quad (3.1)$$

where \vec{X}_0 is the initial condition. In (3.1), the coefficient \mathbf{A} is a nonzero deterministic matrix in $\mathbb{R}^{r \times r}$, the source term \vec{B}_n is a vectorial sequence of random variables of size $r \times 1$ and the initial condition \vec{X}_0 is a random vector of the same size.

For the scalar case, model (3.1) may represent the problem where X_n denotes the population of infected people due to some infectious disease after n periods of time, and B_n denotes the migration. The vector case may represent several population subdivisions corresponding to distinct reactions to the illness. Model (3.1) also describes the uncertain amount of a drug X_n reaching the target organ of the body after a deterministic dosage has been administered in a different part of the body. Here, X_0 denotes the initial uncertain

amount of drug reaching the target organ at the beginning of the study. B_n denotes external sources, apart from dosage coming from nutritional sources, physical conditions and other factors. Coefficient A represents the rate of drug elimination between consecutive periods due to organ physiological behavior. In this chapter, as a first case, we will assume that A is a deterministic value, although in future studies we will consider that could be a random variable.

Random linear matrix difference equations of type (3.1) are of great importance in analyzing systems involving dynamic states as compartmental systems, widespread in Medicine as well as Biology [2, 40, 5]. These models have been successfully applied to studying the evolution of drug concentration in different organs of the human body, [31, 32]. The formulation of these type of systems is based on mass and energy balance considerations of organ compartments which interchange substance in accordance to inter-compartmental flow laws. However, the above references are only devoted to studying models of type (3.1) as well as their continuous counterparts in the deterministic framework, despite initial concentrations of drugs in the organs as well as inter-compartmental flows being seldom known in a deterministic way in practice due to measure errors and inherent complexity. Here, we present a more realistic approach, where in a first stage, we consider that the initial condition and the source term are modeled by means of random variables and discrete stochastic processes, respectively.

The aim of this chapter is to construct random solutions of problem (3.1) and its organization is as follows. Some deterministic results related to matrix spectral calculus are summarized in Section 3.2. This section also contains some stochastic definitions and results related to mean square convergence of vector random variables and vector discrete stochastic process that will play an important role in the following sections. Section 3.3 deals with the construction of the discrete process solution of (3.1), as well as providing sufficient conditions for its convergence in the mean square sense. In Section 3.4, we provide sufficient conditions under which the solution is mean square bounded. Section 3.5 is devoted to computing the main statistical functions such as the mean and the covariance matrix of the discrete solution process. Several illustrative examples of interest in applications are shown in section 3.6. Section 3.7 is devoted to emphasize the main conclusions of this chapter.

3.2 Preliminaries

In chapter 2, we have introduced the main probability and statistic notations, definitions and results that we will be used throughout this dissertation. However, for convenience, these concepts only were exposed in the scalar case. Current and next chapter are devoted to study random matrix difference and differential problems, respectively, hence the corresponding extension of some of those concepts and results introduced in chapter 2 are now required. Reader should notice that following presentation are only a straightforward extension to vector and matrix framework of those concepts already known in the L_2 -space of second order random variables (2-r.v.'s) in the scalar case and, hence for second order stochastic processes (2-s.p.'s).

Given a positive integer r , a second order random vector of size r is a vector $\vec{X} = (X^1, \dots, X^r)^T$ whose entries X^i lie in L_2 for $1 \leq i \leq r$ (here and in advance, the superscript T denotes vector or matrix transpose). The set L_2^r of all these vectors with the norm

$$\|\vec{X}\|_r = \max_{1 \leq i \leq r} \|X^i\|_2, \quad (3.2)$$

provides a Banach structure to L_2^r . Note that norm $\|\cdot\|_2$ is already defined by (2.18). We say that $\{\vec{X}_n : n \geq 0\}$ is a sequence of second order random vectors in L_2^r , if $\vec{X}_n \in L_2^r$ for each $n \geq 0$ integer.

The expectation of $\vec{X}_n = (X_n^1, \dots, X_n^r)^T$ is the deterministic vector

$$\mathbb{E}[\vec{X}_n] = (\mathbb{E}[X_n^1], \dots, \mathbb{E}[X_n^r])^T.$$

Note that

$$\mathbb{E}\left[(\vec{X}_n)^T\right] = \left(\mathbb{E}[\vec{X}_n]\right)^T,$$

holds true.

For each $n, m \geq 0$ integers, the covariance matrix of $\{\vec{X}_n : n \geq 0\} \in L_2^r$ is defined by the following square matrix of size r

$$\mathbf{\Lambda}_{\vec{X}_n, \vec{X}_m} = \mathbb{E}\left[\left(\vec{X}_n - \mathbb{E}[\vec{X}_n]\right)\left(\vec{X}_m - \mathbb{E}[\vec{X}_m]\right)^T\right] = (v_{nm}^{ij})_{r \times r}, \quad (3.3)$$

where its (i, j) -entry is given by

$$\begin{aligned} v_{nm}^{ij} &= \text{E} \left[(X_n^i - \text{E} [X_n^i]) (X_m^j - \text{E} [X_m^j]) \right] \\ &= \text{E} [X_n^i X_m^j] - \text{E} [X_n^i] \text{E} [X_m^j], \quad 1 \leq i, j \leq r, \quad n, m \geq 0. \end{aligned}$$

Note that v_{nn}^{ii} , denoted by $\text{Var} [X_n^i]$, is the variance of the r.v. X_n^i , $1 \leq i \leq r$. In the particular case that $r = 1$ and $\vec{X}_n = X$, (3.3) is the variance of the scalar 2-r.v. X , and it will be denoted by $\text{Var} [X]$. Note that it coincides with (2.7).

Given two sequences of random vectors in the space L_2^r , say, $\{\vec{X}_n : n \geq 0\} = \{(X_n^1, \dots, X_n^r)^T : n \geq 0\}$ and $\{\vec{Y}_m : m \geq 0\} = \{(Y_m^1, \dots, Y_m^r)^T : m \geq 0\}$, one defines their cross-covariance matrix by

$$\mathbf{\Lambda}_{\vec{X}_n, \vec{Y}_m} = \text{E} \left[(\vec{X}_n - \text{E} [\vec{X}_n]) (\vec{Y}_m - \text{E} [\vec{Y}_m])^T \right] = (\nu_{nm}^{ij})_{r \times r}, \quad (3.4)$$

where

$$\begin{aligned} \nu_{nm}^{ij} &= \text{E} \left[(X_n^i - \text{E} [X_n^i]) (Y_m^j - \text{E} [Y_m^j]) \right] \\ &= \text{E} [X_n^i Y_m^j] - \text{E} [X_n^i] \text{E} [Y_m^j], \quad 1 \leq i, j \leq r, \quad n, m \geq 0. \end{aligned}$$

In the specific case that $r = 1$, $X_n = X$ and $Y_m = Y$, expression (3.4) is called the covariance between the 2-r.v.'s X , Y , and, in accordance with (2.10), in advance, it will be denoted by $\text{Cov} [X, Y]$. Note that, it can be expanded as $\text{Cov} [X, Y] = \text{E} [X Y] - \text{E} [X] \text{E} [Y]$. From definitions above, one deduces that covariance matrix of \vec{X}_n is a particular case of cross-covariance matrix of vectors \vec{X}_n and $\vec{Y}_m = \vec{X}_m$. For $n, m \geq 0$ fixed integers and two sequences of random vectors of second order, say, $\{\vec{X}_n = (X_n^1, \dots, X_n^r)^T : n \geq 0\}$ and $\{\vec{Y}_m = (Y_m^1, \dots, Y_m^r)^T : m \geq 0\}$ such that $\{X_n^i\}_{i=1}^r$ and $\{Y_m^i\}_{i=1}^r$ are pairwise independent r.v.'s, then $\mathbf{\Lambda}_{\vec{X}_n, \vec{Y}_m}$ is the null matrix of size $r \times r$. In the specific case where $\vec{X}_n = \vec{Y}_n$, for $n \geq 0$ a fixed integer, being $\{X_n^i\}_{i=1}^r$ pairwise independent r.v.'s, $\mathbf{\Lambda}_{\vec{X}_n, \vec{X}_n}$ is the diagonal matrix $\text{Diag} (\text{Var} [X_n^i])$, $1 \leq i \leq r$. Also, note that the following property holds true:

$$\mathbf{\Lambda}_{\vec{Y}_m, \vec{X}_n} = \left(\mathbf{\Lambda}_{\vec{X}_n, \vec{Y}_m} \right)^T. \quad (3.5)$$

Finally, we recall a property which will be used extensively later. It provides the covariance matrix of two linear combination of random vectors in terms of the covariance matrices of each pair of the random vectors:

$$\mathbf{\Lambda}_{\vec{X}, \vec{Y}} = \sum_{j=1}^n \sum_{k=1}^m a_j b_k \mathbf{\Lambda}_{\vec{X}_j, \vec{Y}_k}, \quad \vec{X} = \sum_{j=1}^n a_j \vec{X}_j, \quad \vec{Y} = \sum_{k=1}^m b_k \vec{Y}_k, \quad (3.6)$$

where $\{a_j : 1 \leq j \leq n\}$ and $\{b_k : 1 \leq k \leq m\}$ are real numbers. This property extends relationship (2.11) to the matrix framework. In accordance with Definition (3.4), $\mathbf{\Lambda}_{\vec{X}, \vec{Y}}$ given by (3.6) is sometimes referred as to cross-covariance matrix of vectors \vec{X} and \vec{Y} .

Remark 3.1. *Given a sequence of random vectors $\{\vec{X}_n : n \geq 0\} \in L_2^r$, one can assume without loss in generality that they are centered at the origin with respect to its mean, that is, $E[\vec{X}_n] = \vec{0}$, for each $n \geq 0$. Otherwise, if $E[\vec{X}_n] = \vec{\mu}_n$, it is enough to consider $\vec{Y}_n = \vec{X}_n - \vec{\mu}_n$, then $E[\vec{Y}_n] = \vec{0}$. By assuming that $E[\vec{X}_n] = \vec{0}$, matrix (3.3) becomes*

$$\mathbf{\Gamma}_{\vec{X}_n, \vec{X}_n} = E \left[\vec{X}_n \left(\vec{X}_n \right)^T \right],$$

called correlation matrix.

We say that a sequence of second order random vectors $\{\vec{X}_n : n \geq 0\}$ is mean square (m.s.) convergent to $\vec{X} \in L_2^r$ if

$$\lim_{n \rightarrow \infty} \left\| \vec{X}_n - \vec{X} \right\|_r = 0, \quad (3.7)$$

and it will be denoted by $\text{l.i.m.}_{n \rightarrow \infty} \vec{X}_n = \vec{X}$ or $\vec{X}_n \xrightarrow[n \rightarrow \infty]{\text{m.s.}} \vec{X}$. From (3.2) it is evident that m.s. convergence in L_2^r is equivalent to the componentwise m.s. convergence.

From the corresponding m.s. properties for its components, see [56, p.88], if $\{\vec{X}_n : n \geq 0\}$ is a sequence of random vectors in L_2^r m.s. convergent to \vec{X} , then

$$E[\vec{X}_n] \xrightarrow[n \rightarrow \infty]{} E[\vec{X}], \quad \mathbf{\Lambda}_{\vec{X}_n, \vec{X}_n} \xrightarrow[n \rightarrow \infty]{} \mathbf{\Lambda}_{\vec{X}, \vec{X}}. \quad (3.8)$$

Note that this result is an extension of Lemma 2.3.

Given a matrix $\mathbf{A} = (a_{ij})$ in $\mathbb{R}^{r \times r}$, we denote by $\|\mathbf{A}\|_\infty$ the norm defined as [33, p.57]

$$\|\mathbf{A}\|_\infty = \max_{1 \leq i \leq r} \sum_{j=1}^r |a_{ij}|. \quad (3.9)$$

This matrix norm satisfies the submultiplicative property, that is, the following inequality

$$\|\mathbf{AB}\|_\infty \leq \|\mathbf{A}\|_\infty \|\mathbf{B}\|_\infty, \quad \mathbf{A}, \mathbf{B} \in \mathbb{R}^{r \times r},$$

holds true. In particular, one gets

$$\|\mathbf{A}^n\|_\infty \leq (\|\mathbf{A}\|_\infty)^n, \quad \forall n \geq 0.$$

Note that all above properties referred to matrix norm $\|\cdot\|_\infty$ are valid for a more general class of norms, called p -norms, defined as

$$\|\mathbf{M}\|_p = \sup_{\vec{v} \neq \vec{0}} \frac{\|\mathbf{M}\vec{v}\|_p}{\|\vec{v}\|_p}, \quad (3.10)$$

where $\mathbf{M} \in \mathbb{R}^{r \times r}$ and, for a vector $\vec{v} \in \mathbb{R}^r$ its p -norm is defined by

$$\|\vec{v}\|_p = (|v_1|^p + \dots + |v_r|^p)^{1/p}, \quad p \geq 1,$$

whereas

$$\|\vec{v}\|_\infty = \max\{|v_1|, \dots, |v_r|\}.$$

The norm (3.9) is the specific case of (3.10) with $p = +\infty$, (see, [33, p.52,55]).

Now we establish a result that will be crucial in the following:

Lemma 3.1. *Let \mathbf{A} be a matrix in $\mathbb{R}^{r \times r}$ and $\vec{X} \in L_2^r$, then*

$$\left\| \mathbf{A}\vec{X} \right\|_r \leq \|\mathbf{A}\|_\infty \left\| \vec{X} \right\|_r.$$

Proof. Given $\vec{X} = (X_n^1, \dots, X_n^r)^T$ by (3.2) and (3.9) it follows that

$$\begin{aligned} \|\mathbf{A}\vec{X}\|_r &= \max_{1 \leq i \leq r} \left\| \sum_{k=1}^r a_{ik} X^k \right\|_2 \leq \max_{1 \leq i \leq r} \sum_{k=1}^r \|a_{ik} X^k\|_2 \\ &= \max_{1 \leq i \leq r} \sum_{k=1}^r |a_{ik}| \|X^k\|_2 \leq \left(\max_{1 \leq i \leq r} \sum_{k=1}^r |a_{ik}| \right) \|\vec{X}\|_r \\ &= \|\mathbf{A}\|_\infty \|\vec{X}\|_r. \quad \square \end{aligned}$$

As a consequence of this result, it is easy to check that for \mathbf{A}, \mathbf{B} matrices in $\mathbb{R}^{r \times r}$ and $\{\vec{X}_n : n \geq 0\}, \{\vec{Y}_n : n \geq 0\}$ sequences of second order random vectors m.s. convergent, one gets

$$\text{l.i.m.}_{n \rightarrow \infty} (\mathbf{A} \vec{X}_n + \mathbf{B} \vec{Y}_n) = \mathbf{A} \text{l.i.m.}_{n \rightarrow \infty} \vec{X}_n + \mathbf{B} \text{l.i.m.}_{n \rightarrow \infty} \vec{Y}_n.$$

For $\mathbf{A} = (a_{ij})$ in $\mathbb{R}^{r \times r}$, one denotes by $\rho(\mathbf{A})$ the spectral radius defined as

$$\rho(\mathbf{A}) = \max \{|\lambda_1|, |\lambda_2|, \dots, |\lambda_r|\},$$

where $\lambda_1, \dots, \lambda_r$ are the eigenvalues of matrix \mathbf{A} . The set $\sigma(\mathbf{A}) = \{\lambda_i : 1 \leq i \leq r\}$ is called the spectrum of \mathbf{A} . For every matrix p -norm $\|\mathbf{A}\|_p$, in particular for $\|\mathbf{A}\|_\infty$, one gets the following inequality, see [50, chap.1]

$$\rho(\mathbf{A}) \leq \|\mathbf{A}\|_p, \quad p \geq 1. \quad (3.11)$$

A matrix $\mathbf{A} = (a_{ij})$ in $\mathbb{R}^{r \times r}$ is said to be of class \mathfrak{M} if for every eigenvalue λ such that $|\lambda| = \rho(\mathbf{A})$ every Jordan block associated with λ is of size 1×1 . An useful result concerning with powers of a matrix is given by the following

Proposition 3.1. ([50, p.25]) *Let $\mathbf{A} = (a_{ij})$ in $\mathbb{R}^{r \times r}$. Then $\lim_{n \rightarrow \infty} \mathbf{A}^n = \mathbf{0}$ if, and only if, $\rho(\mathbf{A}) < 1$. Moreover, $\|\mathbf{A}^n\|_p$ is bounded as $n \rightarrow \infty$ if, and only if, $\rho(\mathbf{A}) < 1$ or, $\rho(\mathbf{A}) = 1$ and \mathbf{A} is of class \mathfrak{M} .*

In particular, by (3.11) and Proposition 3.1 with $p = +\infty$ if

$$\|\mathbf{A}\|_\infty < 1, \quad (3.12)$$

then

$$\lim_{n \rightarrow \infty} \mathbf{A}^n = \mathbf{0}.$$

In the next section, we will require an useful deterministic result referred in literature by Kronecker's lemma:

Lemma 3.2. (*[48, p.233]*) *Let us suppose that the deterministic scalar series $\sum_{n \geq 0} \beta_n$ converges and let $\{\alpha_n : n \geq 0\}$ be an increasing sequence of real numbers that tends to $+\infty$, that is, $\alpha_n \uparrow +\infty$, then*

$$\lim_{n \rightarrow \infty} \frac{1}{\alpha_n} \sum_{j=0}^n \alpha_j \beta_j = 0.$$

The following result provides sufficient conditions for the m.s. convergence of a series of 2-r.v.'s

Proposition 3.2. (*[21, p.108]*) *Let $\{Y_n : n \geq 1\}$ be mutually independent 2-r.v.'s, with means $E[Y_n] = \mu_n$, and variances $Var[Y_n] = \sigma_n^2$, respectively, such that*

$$\sum_{n=1}^{\infty} \mu_n \text{ converges, } \sum_{n=1}^{\infty} \sigma_n^2 < +\infty.$$

Then, series $\sum_{n=1}^{\infty} Y_n$ m.s. converges.

3.3 Mean square solution of random matrix linear difference equations

In this section, we consider random linear models based on matrix difference equations of the form (3.1) where, we assume that the source term \vec{B}_n

as well as the initial condition \vec{X}_0 are in L_2^r . Recursively one gets

$$\begin{aligned}
 \vec{X}_{n+1} &= \mathbf{A}\vec{X}_n + \vec{B}_n \\
 &= \mathbf{A} \left(\mathbf{A}\vec{X}_{n-1} + \vec{B}_{n-1} \right) + \vec{B}_n \\
 &= \mathbf{A}^2 \vec{X}_{n-1} + \mathbf{A}\vec{B}_{n-1} + \vec{B}_n \\
 &= \mathbf{A}^2 \left(\mathbf{A}\vec{X}_{n-2} + \vec{B}_{n-2} \right) + \mathbf{A}\vec{B}_{n-1} + \vec{B}_n \\
 &= \mathbf{A}^3 \vec{X}_{n-2} + \mathbf{A}^2 \vec{B}_{n-2} + \mathbf{A}\vec{B}_{n-1} + \vec{B}_n \\
 &\vdots \\
 &= \mathbf{A}^{n+1} \vec{X}_0 + \sum_{j=0}^n \mathbf{A}^{n-j} \vec{B}_j,
 \end{aligned}$$

that is,

$$\vec{X}_n = \mathbf{A}^n \vec{X}_0 + \sum_{j=0}^{n-1} \mathbf{A}^{n-j-1} \vec{B}_j, \quad n = 0, 1, \dots, \quad (3.13)$$

for a given initial condition \vec{X}_0 .

Let us study conditions under which the discrete solution process (3.13) is m.s. convergent. Firstly, note that taking m.s. limits in (3.13) and using the fact that the m.s. limit of the sum of two vectors is the sum of their m.s. limits, one gets

$$\text{l.i.m.}_{n \rightarrow \infty} \vec{X}_n = \text{l.i.m.}_{n \rightarrow \infty} \mathbf{A}^n \vec{X}_0 + \text{l.i.m.}_{n \rightarrow \infty} \sum_{j=0}^{n-1} \mathbf{A}^{n-j-1} \vec{B}_j, \quad (3.14)$$

and considering the change of index $k = n - j$ in the above sum, last expression can be written as

$$\text{l.i.m.}_{n \rightarrow \infty} \vec{X}_n = \text{l.i.m.}_{n \rightarrow \infty} \mathbf{A}^n \vec{X}_0 + \text{l.i.m.}_{n \rightarrow \infty} \sum_{k=1}^n \mathbf{A}^{k-1} \vec{B}_{n-k}. \quad (3.15)$$

By Lemma 3.1 and Proposition 3.1, under hypothesis

$$\rho(\mathbf{A}) < 1, \quad (3.16)$$

the first m.s. limit in the right-hand side of (3.15) is the null random vector of size r

$$0 \leq \left\| \mathbf{A}^n \vec{X}_0 \right\|_r \leq \|\mathbf{A}^n\|_\infty \left\| \vec{X}_0 \right\|_r \xrightarrow{n \rightarrow \infty} 0,$$

since

$$\|\mathbf{A}^n\|_\infty \xrightarrow{n \rightarrow \infty} 0, \tag{3.17}$$

and $\left\| \vec{X}_0 \right\|_r < +\infty$. Moreover, from (3.17) one can assure the existence of a positive constant $L > 0$ such that

$$\|\mathbf{A}^n\|_\infty < L, \quad \forall n \geq 0. \tag{3.18}$$

On the other hand, by Lemma 3.1 and (3.18) one gets

$$0 \leq \sum_{k=1}^n \left\| \mathbf{A}^{k-1} \vec{B}_{n-k} \right\|_r \leq \sum_{k=1}^n \|\mathbf{A}^{k-1}\|_\infty \left\| \vec{B}_{n-k} \right\|_r \leq L \sum_{k=1}^n \left\| \vec{B}_{n-k} \right\|_r,$$

then, under hypothesis

$$\sum_{n=0}^{\infty} \left\| \vec{B}_n \right\|_r = \sum_{n=0}^{\infty} \max_{1 \leq i \leq r} \left(\left(\mathbb{E} \left[(B_n^i)^2 \right] \right)^{1/2} \right) < +\infty, \tag{3.19}$$

the existence of the second m.s. limit appearing in the right-hand side of (3.15) is guaranteed. Summarizing the following result has been established:

Theorem 3.1. *The vector discrete solution s.p. \vec{X}_n of the problem (3.1) is given by (3.13). Moreover, under hypotheses (3.16) and (3.19) \vec{X}_n is m.s. convergent.*

Remark 3.2. *From a practical point of view it is important to point out that in the scalar case, i.e., $r = 1$, sufficient conditions (3.16) and (3.19) for m.s. convergence of the solution s.p. (3.13) of the problem (3.1) can be written as follows:*

$$|a| < 1, \quad \sum_{n=0}^{\infty} \left(\mathbb{E} \left[(B_n)^2 \right] \right)^{1/2} < +\infty. \tag{3.20}$$

Remark 3.3. *An easier but important case is when $\vec{B}_n = \vec{B}$, i.e., the source term is a constant 2-s.p. (or equivalently, it is a second order random vector),*

then it permits an analogous treatment as in the deterministic framework. The solution of problem (3.1) is given by

$$\vec{X}_n = \mathbf{A}^n \vec{X}_0 + (\mathbf{I} - \mathbf{A}^n) (\mathbf{I} - \mathbf{A})^{-1} \vec{B},$$

whenever 1 does not belong to the spectrum of matrix \mathbf{A} , i.e., $\mathbf{I} - \mathbf{A}$ is nonsingular. Moreover, under hypothesis $\rho(\mathbf{A}) < 1$, by Proposition 3.1 and Lemma 3.1, one gets

$$\text{l.i.m.}_{n \rightarrow \infty} \vec{X}_n = (\mathbf{I} - \mathbf{A})^{-1} \vec{B}.$$

Theorem 3.1 provides sufficient conditions for m.s. convergence of (3.13) but, it does not require the m.s. limit value. Now, we provide sufficient conditions under which discrete s.p. (3.13) converges at the null random vector of size r . In fact, let us assume that hypotheses (3.12) and (3.19) hold true, then by (3.11) with $p = +\infty$ and Proposition 3.1, the first m.s. limit appearing in the right-hand side of (3.14) goes to the null random vector of size r . In order to prove that the second m.s. limit appearing in the right-hand side in (3.14) is also the null random vector, note that

$$\begin{aligned} 0 &\leq \left\| \sum_{j=0}^{n-1} \mathbf{A}^{n-j-1} \vec{B}_j \right\|_r &\leq \sum_{j=0}^{n-1} \left\| \mathbf{A}^{n-j-1} \vec{B}_j \right\|_r \\ &\leq \sum_{j=0}^{n-1} \|\mathbf{A}^{n-j-1}\|_\infty \left\| \vec{B}_j \right\|_r &\leq \sum_{j=0}^{n-1} (\|\mathbf{A}\|_\infty)^{n-j-1} \left\| \vec{B}_j \right\|_r \\ &\leq (\|\mathbf{A}\|_\infty)^{n-1} \sum_{j=0}^{n-1} \frac{\left\| \vec{B}_j \right\|_r}{\|\mathbf{A}\|_\infty^j} = \frac{(\|\mathbf{A}\|_\infty)^{-1}}{\left(\frac{1}{\|\mathbf{A}\|_\infty}\right)^n} \sum_{j=0}^{n-1} (\|\mathbf{A}\|_\infty)^{-j} \left\| \vec{B}_j \right\|_r \xrightarrow{n \rightarrow \infty} 0, \end{aligned}$$

where in the last step we have applied the Kronecker's lemma 3.2 for

$$\alpha_n = \left(\frac{1}{\|\mathbf{A}\|_\infty} \right)^n \uparrow +\infty, \quad \beta_n = \left\| \vec{B}_n \right\|_r,$$

as we are assuming that conditions (3.12) and (3.19) hold true. Note that $\|\mathbf{A}\|_\infty \neq 0$ since $\mathbf{A} \neq \mathbf{0}$. Summarizing the following result has been established:

Theorem 3.2. *Under hypotheses (3.12) and (3.19), the vector discrete solution s.p. \vec{X}_n of the problem (3.1) given by (3.13) is m.s. convergent to the null random vector of size r .*

In the two next remarks, we give different situations where whether above hypothesis (3.16) or (3.19) does not hold, but m.s. convergence of \vec{X}_n is established.

Remark 3.4. *Let us consider model (3.1) with $\mathbf{A} = \mathbf{I}$, being \mathbf{I} the identity matrix of size r , then by (3.14) one gets*

$$\text{l.i.m.}_{n \rightarrow \infty} \vec{X}_n = \vec{X}_0 + \text{l.i.m.}_{n \rightarrow \infty} \sum_{j=0}^{n-1} \vec{B}_j. \quad (3.21)$$

Although hypothesis (3.16) does not hold because $\rho(\mathbf{A}) = 1$, under condition (3.19) one can assure the m.s. convergence of \vec{X}_n . Now, we show an example where this situation arises. Let us consider model (3.1) in the scalar framework, that is, $r = 1$, with $\mathbf{A} = 1$, an arbitrary 2-r.v. X_0 as initial condition and the source term given by the following sequence of scalar 2-r.v.'s

$$B_n = (-1)^n \frac{C_n}{n+1}, \quad n = 1, 2, \dots,$$

being $\{C_n : n \geq 0\}$ r.v.'s mutually independent and identically distributed with means μ_C and variance σ_C^2 . To establish the existence of the m.s. limit (3.21), it is enough to check the m.s. convergence of the series with general term B_n . This follows from Proposition 3.2 taking into account that by Leibnitz's test one gets

$$\sum_{n=0}^{\infty} \text{E}[B_n] = \sum_{n=0}^{\infty} (-1)^n \frac{\text{E}[C_n]}{n+1} = \mu_C \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} \text{ converges,}$$

and

$$\sum_{n=0}^{\infty} \text{Var}[B_n] = \sum_{n=0}^{\infty} \frac{\text{Var}[C_n]}{(n+1)^2} = \sigma_C^2 \sum_{n=0}^{\infty} \frac{1}{(n+1)^2} = \frac{\sigma_C^2 \pi^2}{6}.$$

Remark 3.5. *When \mathbf{A} is a matrix of class \mathfrak{M} such that $\rho(\mathbf{A}) = 1$, the initial condition is the null random vector of size r and hypothesis (3.19) holds true, by (3.14) and Proposition 3.1, the discrete solution s.p. \vec{X}_n is m.s. convergent.*

3.4 Mean square boundedness of solutions

In this section is addressed to establish conditions on coefficients \mathbf{A} and \vec{B}_n in order to guarantee that solution of (3.1) is not m.s. unstable, in the sense that it remains m.s. bounded or may be m.s. convergent, when n goes to infinity.

For this, let us take a scalar linear difference equation that serves as an upper to (3.1), that is, a scalar equation whose solution y_n is always greater than the r -norm of the solution \vec{X}_n of (3.1). In the following, this difference equation will be called a majorant equation of (3.1).

Then, by imposing the well-known boundedness conditions on y_n , it will be assured the corresponding property for \vec{X}_n .

Let us take norms on (3.1), then by applying triangular inequality and Lemma 3.1 one gets

$$\left\| \vec{X}_{n+1} \right\|_r \leq \|\mathbf{A}\|_\infty \left\| \vec{X}_n \right\|_r + \left\| \vec{B}_n \right\|_r, \quad (3.22)$$

now denoting by $x_n = \left\| \vec{X}_n \right\|_r$, one obtains the following majorant deterministic scalar difference equation

$$x_{n+1} = \|\mathbf{A}\|_\infty x_n + \left\| \vec{B}_n \right\|_r, \quad n = 0, 1, \dots, \quad (3.23)$$

with initial condition $x_0 = \left\| \vec{X}_0 \right\|_r$. In fact, let us prove by induction principle that

$$\left\| \vec{X}_n \right\|_r \leq x_n, \quad n = 0, 1, \dots$$

The case $n = 0$ is evident by definition of the initial condition. Let us assume by induction hypothesis that

$$\left\| \vec{X}_k \right\|_r \leq x_k, \quad \forall k = 0, 1, \dots, n, \quad (3.24)$$

then for $k = n + 1$, from (3.22)-(3.24) one gets

$$\left\| \vec{X}_{n+1} \right\|_r \leq \|\mathbf{A}\|_\infty \left\| \vec{X}_n \right\|_r + \left\| \vec{B}_n \right\|_r \leq \|\mathbf{A}\|_\infty x_n + \left\| \vec{B}_n \right\|_r = x_{n+1}.$$

Moreover, if

$$\left\{ \left\| \vec{B}_n : n \geq 0 \right\| \right\} \text{ is m.s. uniformly bounded : } \left\| \vec{B}_n \right\|_r < M_B, \quad n \geq 0, \quad (3.25)$$

from (3.23) one obtains

$$x_{n+1} = \|\mathbf{A}\|_\infty x_n + \left\| \vec{B}_n \right\|_r \leq \|\mathbf{A}\|_\infty x_n + M_B, \quad n = 0, 1, \dots$$

Let \hat{x}_n be the solution of the scalar difference equation

$$\hat{x}_{n+1} = \|\mathbf{A}\|_\infty \hat{x}_n + M_B, \quad (3.26)$$

with $\hat{x}_0 = x_0 = \left\| \vec{X}_0 \right\|_r$. Again, it is easy to prove by principle of induction that

$$x_n \leq \hat{x}_n, \quad \forall n \geq 0.$$

From the deterministic theory, it is well-known that the solution of the scalar difference equation (3.26) is given by

$$\hat{x}_n = \frac{M_B}{1 - \|\mathbf{A}\|_\infty} + \left(\left\| \vec{X}_0 \right\|_r - \frac{M_B}{1 - \|\mathbf{A}\|_\infty} \right) (\|\mathbf{A}\|_\infty)^n, \quad \forall n \geq 0, \quad (3.27)$$

whenever $\|\mathbf{A}\|_\infty \neq 1$. Now, under condition (3.12), one can assure that $(\|\mathbf{A}\|_\infty)^n < 1$ for each $n \geq 1$ and, if initial condition satisfies

$$\left\| \vec{X}_0 \right\|_r > \frac{M_B}{1 - \|\mathbf{A}\|_\infty},$$

one gets

$$\hat{x}_n \leq \frac{M_B}{1 - \|\mathbf{A}\|_\infty} + \left(\left\| \vec{X}_0 \right\|_r - \frac{M_B}{1 - \|\mathbf{A}\|_\infty} \right) = \left\| \vec{X}_0 \right\|_r. \quad (3.28)$$

On the other hand, if initial condition satisfies the reverse inequality

$$\left\| \vec{X}_0 \right\|_r \leq \frac{M_B}{1 - \|\mathbf{A}\|_\infty},$$

from (3.27) one gets

$$\hat{x}_n \leq \frac{M_B}{1 - \|\mathbf{A}\|_\infty}. \quad (3.29)$$

Therefore from (3.28) and (3.29) one obtains the following bound for \hat{x}_n

$$\hat{x}_n \leq \max \left(\left\| \vec{X}_0 \right\|_r, \frac{M_B}{1 - \|\mathbf{A}\|_\infty} \right).$$

Summarizing the next result has been established:

Theorem 3.3. Under hypotheses (3.12) and (3.25), the vector discrete solution s.p. \vec{X}_n of the problem (3.1) is m.s. bounded as follows

$$\|\vec{X}_n\|_r \leq \max\left(\|\vec{X}_0\|_r, \frac{M_B}{1 - \|\mathbf{A}\|_\infty}\right), \quad (3.30)$$

where M_B is a positive constant satisfying condition (3.25).

Remark 3.6. It is important to point out that in the usual case which the vector source term $\{\vec{B}_n : n \geq 0\}$ is scalar (i.e., $r = 1$), say, $\{B_n : n \geq 0\}$ is a sequence of 2-r.v.'s with zero mean and common variance, condition (3.25) holds true. This is a typical hypothesis in the time series framework ([37]). Note that, it remains true in the corresponding vector scenario.

Remark 3.7. Let us study the particular case of (3.1) where \vec{X}_0 is a deterministic vector and the uncertainty of the source term can be expressed as follows

$$\vec{B}_n = \mathbf{E}[\vec{B}_n] + \vec{D}_n, \quad (3.31)$$

being that noise and expectation terms satisfy,

$$\|\vec{D}_n\|_r \leq M_D, \quad \forall n \geq 0,$$

and

$$\max\{|\mathbf{E}[B_n^i]| : 1 \leq i \leq r\} \leq M_E, \quad \forall n \geq 0,$$

respectively. Then one gets

$$\|\vec{B}_n\|_r \leq M_D + M_E := M_F, \quad \forall n \geq 0.$$

Let us consider the associated equation to (3.1) without noise

$$\vec{Y}_{n+1} = \mathbf{A}\vec{Y}_n + \mathbf{E}[\vec{B}_n], \quad (3.32)$$

and deterministic initial condition $\vec{Y}_0 = \vec{X}_0$. Note that, this equation is just what one should obtain taking expectations in (3.1) in the case where noise given by (3.31) satisfies: $\mathbf{E}[\vec{D}_n] = \vec{0}$, being $\vec{Y}_n = \mathbf{E}[\vec{X}_n]$.

In order to compare solutions of random and deterministic problems (3.1) and (3.32), respectively, let us consider the random difference vector: $\vec{Z}_n = \vec{X}_n - \vec{Y}_n$ which satisfies the random difference equation

$$\vec{Z}_{n+1} = \mathbf{A}\vec{Z}_n + \vec{D}_n.$$

From (3.30) and considering that $\vec{Z}_0 = \vec{0}$, one gets

$$\|\vec{Z}_n\|_r = \|\vec{X}_n - \vec{Y}_n\|_r \leq \frac{M_D}{1 - \|\mathbf{A}\|_\infty}.$$

This indicates that deterministic solution approximates random solution when noise-term norm is near to zero.

The next result provides sufficient conditions for bounding the solution of equation (3.1). It follows directly from expression (3.14) and Proposition 3.1.

Proposition 3.3. *If \mathbf{A} is of class \mathfrak{M} such that $\rho(\mathbf{A}) = 1$ and the partial sums of source term \vec{B}_n are m.s. bounded, for every initial condition $\vec{X}_0 \in L_2^r$, the discrete solution s.p. \vec{X}_n is m.s. bounded.*

Note that hypotheses of the above proposition are disjoint of conditions (3.12) and (3.25) imposed in Theorem 3.3.

Remark 3.8. *Both Theorem 3.3 and Proposition 3.3 provide sufficient conditions for m.s. boundedness of \vec{X}_n which involve the m.s. boundedness of the source term \vec{B}_n or of the m.s. boundedness of its partial sums. The m.s. boundedness of \vec{B}_n is also a necessary condition for the m.s. boundedness of \vec{X}_n . In fact, if \vec{X}_n is m.s. bounded, from Lemma 3.1, $\mathbf{A}\vec{X}_n$ is also m.s. bounded and, directly from equation (3.1), one deduces that \vec{B}_n are indeed m.s. bounded. As a result, if we check previously that \vec{B}_n are m.s. unbounded, then the solution \vec{X}_n will also be m.s. unbounded.*

3.5 Statistical functions of the mean square solution

Now, we focus on computing the main statistical functions of the m.s. solution (3.13) of equation of (3.1), namely its mean and its variance. Let

us take directly the expectation operator in (3.1), then one gets the following deterministic vector difference equation

$$\mathbf{E} \left[\vec{X}_{n+1} \right] = \mathbf{A} \mathbf{E} \left[\vec{X}_n \right] + \mathbf{E} \left[\vec{B}_n \right], \quad n = 0, 1, \dots,$$

with the initial condition $\mathbf{E} \left[\vec{X}_0 \right]$, whose solution is given by

$$\mathbf{E} \left[\vec{X}_n \right] = \mathbf{A}^n \mathbf{E} \left[\vec{X}_0 \right] + \sum_{j=0}^{n-1} \mathbf{A}^{n-j-1} \mathbf{E} \left[\vec{B}_j \right], \quad n = 0, 1, \dots \quad (3.33)$$

Note that, one gets the same relationship directly if we take the expectation operator in (3.13).

The computation of covariance matrix of random vector \vec{X}_n requires some additional effort. In fact, by using (3.6) one gets

$$\begin{aligned} \mathbf{\Lambda}_{\vec{X}_n, \vec{X}_m} &= \mathbf{E} \left[\left(\vec{X}_n - \mathbf{E} \left[\vec{X}_n \right] \right) \left(\vec{X}_m - \mathbf{E} \left[\vec{X}_m \right] \right)^T \right] \\ &= \mathbf{E} \left[\left(\left\{ \mathbf{A}^n \vec{X}_0 + \sum_{j=0}^{n-1} \mathbf{A}^{n-j-1} \vec{B}_j \right\} \right. \right. \\ &\quad \left. \left. - \left\{ \mathbf{A}^n \mathbf{E} \left[\vec{X}_0 \right] + \sum_{j=0}^{n-1} \mathbf{A}^{n-j-1} \mathbf{E} \left[\vec{B}_j \right] \right\} \right) \right. \\ &\quad \times \left(\left\{ \mathbf{A}^m \vec{X}_0 + \sum_{k=0}^{m-1} \mathbf{A}^{m-k-1} \vec{B}_k \right\} \right. \\ &\quad \left. \left. - \left\{ \mathbf{A}^m \mathbf{E} \left[\vec{X}_0 \right] + \sum_{k=0}^{m-1} \mathbf{A}^{m-k-1} \mathbf{E} \left[\vec{B}_k \right] \right\} \right)^T \right] \end{aligned}$$

$$\begin{aligned}
 &= \mathbf{E} \left[\left(\mathbf{A}^n (\vec{X}_0 - \mathbf{E} [\vec{X}_0]) + \sum_{j=0}^{n-1} \mathbf{A}^{n-j-1} (\vec{B}_j - \mathbf{E} [\vec{B}_j]) \right) \right. \\
 &\quad \times \left. \left(\mathbf{A}^m (\vec{X}_0 - \mathbf{E} [\vec{X}_0]) + \sum_{k=0}^{m-1} \mathbf{A}^{m-k-1} (\vec{B}_k - \mathbf{E} [\vec{B}_k]) \right)^T \right] \\
 &= \mathbf{E} \left[\left\{ \mathbf{A}^n (\vec{X}_0 - \mathbf{E} [\vec{X}_0]) + \sum_{j=0}^{n-1} \mathbf{A}^{n-j-1} (\vec{B}_j - \mathbf{E} [\vec{B}_j]) \right\} \right. \\
 &\quad \times \left. \left\{ (\vec{X}_0 - \mathbf{E} [\vec{X}_0])^T (\mathbf{A}^m)^T + \sum_{k=0}^{m-1} (\vec{B}_k - \mathbf{E} [\vec{B}_k])^T (\mathbf{A}^{m-k-1})^T \right\} \right] \\
 &= \mathbf{A}^n \mathbf{\Lambda}_{\vec{X}_0, \vec{X}_0} (\mathbf{A}^m)^T + \sum_{j=0}^{n-1} \sum_{k=0}^{m-1} \mathbf{A}^{n-j-1} \mathbf{\Lambda}_{\vec{B}_j, \vec{B}_k} (\mathbf{A}^{m-k-1})^T \\
 &\quad + \mathbf{A}^n \left(\sum_{k=0}^{m-1} \mathbf{\Lambda}_{\vec{X}_0, \vec{B}_k} (\mathbf{A}^{m-k-1})^T \right) + \left(\sum_{j=0}^{n-1} \mathbf{A}^{n-j-1} \mathbf{\Lambda}_{\vec{B}_j, \vec{X}_0} \right) (\mathbf{A}^m)^T.
 \end{aligned}$$

Summing up,

$$\begin{aligned}
 \mathbf{\Lambda}_{\vec{X}_n, \vec{X}_m} &= \mathbf{A}^n \mathbf{\Lambda}_{\vec{X}_0, \vec{X}_0} (\mathbf{A}^m)^T \\
 &\quad + \sum_{j=0}^{n-1} \sum_{k=0}^{m-1} \mathbf{A}^{n-j-1} \mathbf{\Lambda}_{\vec{B}_j, \vec{B}_k} (\mathbf{A}^{m-k-1})^T \\
 &\quad + \mathbf{A}^n \left(\sum_{k=0}^{m-1} \mathbf{\Lambda}_{\vec{X}_0, \vec{B}_k} (\mathbf{A}^{m-k-1})^T \right) \\
 &\quad + \left(\sum_{j=0}^{n-1} \mathbf{A}^{n-j-1} \mathbf{\Lambda}_{\vec{B}_j, \vec{X}_0} \right) (\mathbf{A}^m)^T.
 \end{aligned} \tag{3.34}$$

This general expression can be simplified if \mathbf{A} is invertible

$$\begin{aligned}
 \mathbf{\Lambda}_{\vec{X}_n, \vec{X}_m} &= \mathbf{A}^n \mathbf{\Lambda}_{\vec{X}_0, \vec{X}_0} (\mathbf{A}^m)^T \\
 &+ \mathbf{A}^n \left(\sum_{k=0}^{m-1} \mathbf{\Lambda}_{\vec{X}_0, \vec{B}_k} \left(\mathbf{A}^{-(k+1)} \right)^T \right) (\mathbf{A}^m)^T \\
 &+ \mathbf{A}^n \left(\sum_{j=0}^{n-1} \mathbf{A}^{-(j+1)} \mathbf{\Lambda}_{\vec{B}_j, \vec{X}_0} \right) (\mathbf{A}^m)^T \\
 &+ \mathbf{A}^n \left(\sum_{j=0}^{n-1} \sum_{k=0}^{m-1} \mathbf{A}^{-(j+1)} \mathbf{\Lambda}_{\vec{B}_j, \vec{B}_k} \left(\mathbf{A}^{-(k+1)} \right)^T \right) (\mathbf{A}^m)^T \quad (3.35) \\
 &= \mathbf{A}^n \left\{ \mathbf{\Lambda}_{\vec{X}_0, \vec{X}_0} + \sum_{j=0}^{n-1} \mathbf{A}^{-(j+1)} \mathbf{\Lambda}_{\vec{B}_j, \vec{X}_0} \right. \\
 &+ \sum_{j=0}^{n-1} \sum_{k=0}^{m-1} \mathbf{A}^{-(j+1)} \mathbf{\Lambda}_{\vec{B}_j, \vec{B}_k} \left(\mathbf{A}^{-(k+1)} \right)^T \\
 &\left. + \sum_{k=0}^{m-1} \mathbf{\Lambda}_{\vec{X}_0, \vec{B}_k} \left(\mathbf{A}^{-(k+1)} \right)^T \right\} (\mathbf{A}^m)^T.
 \end{aligned}$$

Some specific cases interesting in applications are treated in the following remarks.

Remark 3.9. When the random vector initial condition $\vec{X}_0 = (X_0^1, \dots, X_0^r)^T$ and the source term $\vec{B}_n = (B_n^1, \dots, B_n^r)^T$ are such that X_0^i is independent of B_n^j for each pair (i, j) with $1 \leq i, j \leq r$, $n \geq 0$ integer and \mathbf{A} is invertible, then

$$\mathbf{\Lambda}_{\vec{X}_0, \vec{B}_k} = \mathbf{0}_{r \times r} = \mathbf{\Lambda}_{\vec{B}_j, \vec{X}_0},$$

and the covariance matrix given by (3.35) can be simplified as follows

$$\mathbf{\Lambda}_{\vec{X}_n, \vec{X}_m} = \mathbf{A}^n \left\{ \mathbf{\Lambda}_{\vec{X}_0, \vec{X}_0} + \sum_{j=0}^{n-1} \sum_{k=0}^{m-1} \mathbf{A}^{-(j+1)} \mathbf{\Lambda}_{\vec{B}_j, \vec{B}_k} \left(\mathbf{A}^{-(k+1)} \right)^T \right\} (\mathbf{A}^m)^T.$$

Additional hypotheses allows us to simplify this expression. For instance, if \mathbf{A} is a symmetric matrix, last expression reads

$$\Lambda_{\vec{X}_n, \vec{X}_m} = \mathbf{A}^n \left\{ \Lambda_{\vec{X}_0, \vec{X}_0} + \sum_{j=0}^{n-1} \sum_{k=0}^{m-1} \mathbf{A}^{-(j+1)} \Lambda_{\vec{B}_j, \vec{B}_k} \mathbf{A}^{-(k+1)} \right\} \mathbf{A}^m,$$

since powers of \mathbf{A} as well as its inverse matrix are also symmetric matrices. Other interesting situation occurs when \mathbf{A} is l_0 -nilpotent, that is, there exists an integer l_0 such that $\mathbf{A}^l = \mathbf{0}$ for all $l \geq l_0$. Then

$$\Lambda_{\vec{X}_n, \vec{X}_m} = \mathbf{0}_{r \times r}, \quad \forall n, m : \min(n, m) \geq l_0.$$

Remark 3.10. Expectation and covariance discrete functions given by (3.33) and (3.34), respectively, take the following forms in the scalar case, that is, when $r = 1$:

$$E[X_n] = a^n E[X_0] + \sum_{j=0}^{n-1} a^{n-j-1} E[B_j], \quad n = 0, 1, \dots, \quad (3.36)$$

for a prefixed initial condition $E[X_0]$, and

$$\begin{aligned} \text{Cov}[X_n, X_m] &= a^{n+m} \left\{ \text{Var}[X_0] + \sum_{j=0}^{n-1} \sum_{k=0}^{m-1} a^{-(j+k+2)} \text{Cov}[B_j, B_k] \right. \\ &\quad \left. + \sum_{j=0}^{n-1} a^{-(j+1)} \text{Cov}[B_j, X_0] + \sum_{k=0}^{m-1} a^{-(k+1)} \text{Cov}[X_0, B_k] \right\}. \end{aligned}$$

Inasmuch as, $\text{Cov}[B_j, X_0] = \text{Cov}[X_0, B_j]$, in the particular case that $n = m$, one obtains the variance

$$\begin{aligned} \text{Var}[X_n] &= a^{2n} \left\{ \text{Var}[X_0] + \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} a^{-(j+k+2)} \text{Cov}[B_j, B_k] \right. \\ &\quad \left. + 2 \sum_{j=0}^{n-1} a^{-(j+1)} \text{Cov}[B_j, X_0] \right\}, \end{aligned}$$

which under usual hypothesis of pairwise independence between initial condition X_0 and source term $\{B_n : n \geq 0\}$ yields

$$\text{Var} [X_n] = a^{2n} \left\{ \text{Var} [X_0] + \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} a^{-(j+k+2)} \text{Cov} [B_j, B_k] \right\}. \quad (3.37)$$

3.6 Examples

Finally, we present illustrative applications of model (3.1). In these examples we will obtain the discrete solution process as well as its main statistical functions.

Although next examples are stated in a mathematical way, firstly, let us show that they are potentially useful in applications if we consider straightforward generalizations of others that appear in mathematical literature. For this goal, we have selected a simple, but illustrative example, that take place in long-term medical drug strategies [22, p.6]. Let us consider that a drug is administered once every four hours. Let D_n be the amount of the drug in the blood system at the n -th interval. The body eliminates a certain fraction p of the drug during each time interval. At the beginning, the amount administered is D_0 . In order to determine the evolution of D_n , we must create an equation to solve. Since the amount of drug in the patient's system at time $(n + 1)$ is equal to the amount at time n minus the fraction p that has been eliminated from the body, plus the new dosage D_0 , we arrive at the following scalar difference equation:

$$D_{n+1} = (1 - p)D_n + D_0, \quad n = 0, 1, 2, \dots$$

This model corresponds to equation (3.1) with $r = 1$, $\vec{X}_n = D_n$, $\mathbf{A} = 1 - p$, \vec{B}_n in this case is the constant value D_0 and $\vec{X}_0 = D_0$. However, note that now we are not considering uncertainty into this medical drug model. Of course, once D_n has been computed, a crucial question is to study the long-term behavior of D_n , i.e., to compute $\lim_{n \rightarrow \infty} D_n$.

Example 3.1. Let us consider the following discrete model

$$X_{n+1} = \frac{1}{2} X_n + \frac{1}{(n+1)^2} (2 + \cos((n+1)U)), \quad n = 0, 1, 2, \dots, \quad (3.38)$$

where U is a Uniform r.v., say, $U \sim \text{Un}([0, \pi])$ and X_0 is an Exponential 2-r.v. of parameter $\lambda = 1$, i.e., $X_0 \sim \text{Exp}(\lambda = 1)$, independent of r.v. U . From (3.13), the solution of problem (3.38) is given by

$$X_n = \left(\frac{1}{2}\right)^n X_0 + \sum_{j=0}^{n-1} \left(\frac{1}{2}\right)^{n-j-1} \frac{1}{(j+1)^2} (2 + \cos((j+1)U)), \quad n = 0, 1, 2, \dots \quad (3.39)$$

With regard to its asymptotic behavior, note that conditions (3.20)

$$|a| = \frac{1}{2} < 1, \quad \sum_{n=0}^{\infty} \left(\mathbb{E} \left[(B_n)^2 \right]\right)^{1/2} = \sum_{n=0}^{\infty} \left(\frac{9}{2(1+n)^4}\right)^{1/2} = \frac{\pi^2}{2\sqrt{2}} < +\infty,$$

hold true. Therefore the sequence X_n is m.s. convergent. Moreover, as condition (3.12) also holds true, by Theorem 3.2 one concludes that X_n is m.s. convergent to zero r.v.

In order to provide statistical information about the discrete solution s.p. (3.39), by considering (2.8), we compute the expectation of random source term

$$\mathbb{E}[B_n] = \frac{2}{(1+n)^2}, \quad n = 0, 1, \dots,$$

as well as its covariance

$$\begin{aligned} \text{Cov}[B_j, B_k] &= \mathbb{E}[B_j B_k] - \mathbb{E}[B_j] \mathbb{E}[B_k] \\ &= \frac{1}{(j+1)^2 (k+1)^2} \{ \mathbb{E}[\cos((j+1)U) \cos((k+1)U)] \\ &\quad + 2 \mathbb{E}[\cos((k+1)U)] + 2 \mathbb{E}[\cos((j+1)U)] \} \\ &= \begin{cases} 0 & \text{if } j \neq k \\ \frac{1}{2(j+1)^4} & \text{if } j = k, \end{cases} \end{aligned}$$

where we have used that $\mathbb{E}[\cos((k+1)U)] = 0$ and

$$\mathbb{E}[\cos((j+1)U) \cos((k+1)U)] = \begin{cases} 0 & \text{if } j \neq k, \\ \frac{1}{2} & \text{if } j = k, \end{cases} \quad j, k = 0, 1, \dots$$

Then by (3.36), (3.37) and taking into account that $E[X_0] = 1$ and $\text{Var}[X_0] = 1$ (see Table 2.1) one gets

$$E[X_n] = \left(\frac{1}{2}\right)^n \left(1 + 4 \sum_{j=0}^{n-1} \frac{2^j}{(1+j)^2}\right), \quad n = 0, 1, 2, \dots,$$

and

$$\text{Var}[X_n] = \left(\frac{1}{4}\right)^n \left(1 + 2 \sum_{j=0}^{n-1} \frac{4^j}{(1+j)^4}\right), \quad n = 0, 1, 2, \dots$$

Finally, as sequence $\{X_n : n \geq 0\}$ is m.s. convergent to zero, by (3.8) one gets

$$E[X_n] \xrightarrow{n \rightarrow \infty} 0, \quad \text{Var}[X_n] \xrightarrow{n \rightarrow \infty} 0,$$

as can be easily checked. Note that, the information about r.v.'s U and X_0 we have just considered has been their first and second order moments with respect to the origin, then any other statistical distributions for them could have been assumed.

Example 3.2. In this example, we study a matrix model of type (3.1). For instance, from a practical point of view, it can be interpreted as a typical three-compartment model in chemical kinetics or pharmacokinetics (see Figure 3.1). Let X_n^1 , X_n^2 and X_n^3 be the concentrations of a substance in the period n , with $n = 0, 1, 2, \dots$ at each compartment, respectively. We shall assume that the input term \vec{B}_n in (3.1) defining the entrance/exit flow for each compartment of the system is given by the following sequences

$$A_n = An \exp(-2n(n+1)), \quad B_n = B/(n+1)^2, \quad C_n = Cn/(n+1)^4,$$

respectively, where

$$A \sim \text{Be}(\alpha_A=2; \beta_A=1), \quad B \sim \text{Be}(\alpha_B=1; \beta_B=2), \quad C \sim \text{Be}(\alpha_C=1/50; \beta_C=1)$$

being A, B, C mutually independent r.v.'s.

Let us assume that there exists a concentration flow between compartments, as it is shown in Figure 3.1. These values are quoted on arrows connecting

compartments. This problem can be modeled by the random system (3.1) where

$$\mathbf{A} = \begin{bmatrix} 1 - \frac{1}{20} & \frac{1}{10} & \frac{1}{8} \\ \frac{1}{20} & 1 - \frac{1}{10} - \frac{2}{3} & 0 \\ 0 & \frac{2}{3} & 1 - \frac{1}{8} \end{bmatrix},$$

$$\vec{X}_0 = \vec{0}, \quad \vec{X}_n = \begin{bmatrix} X_n^1 \\ X_n^2 \\ X_n^3 \end{bmatrix}, \quad \vec{B}_n = \begin{bmatrix} -A n e^{-2(n+1)} \\ B (n+1)^{-2} \\ -C n (n+1)^{-4} \end{bmatrix}, \quad n = 0, 1, 2, \dots$$

The discrete solution s.p. of this problem is given by

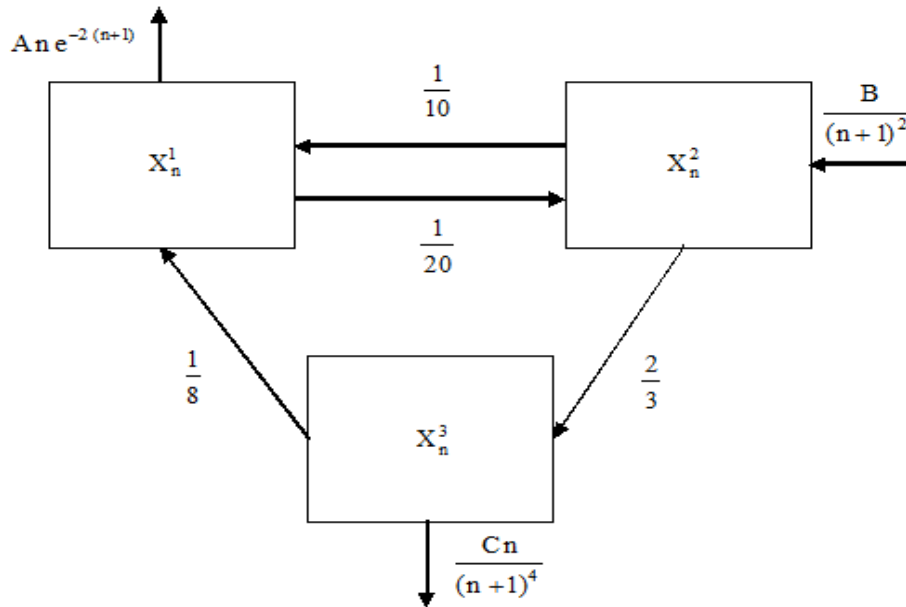


Figure 3.1: A three-compartment model in chemical kinetics. Note that the values on arrows connecting compartments indicate the rate of transfer (proportional to the concentration in the compartment). Whereas the values indicated on arrows coming from or going to the exterior of compartments indicate inputs or outputs to the system (they are not proportional to the concentration of the corresponding compartment)

$$\vec{X}_n = \sum_{j=0}^{n-1} \begin{bmatrix} 19/20 & 1/10 & 1/8 \\ 1/20 & 7/30 & 0 \\ 0 & 2/3 & 7/8 \end{bmatrix}^{n-j-1} \begin{bmatrix} -A j e^{-2(j+1)} \\ B (j+1)^{-2} \\ -C j (j+1)^{-4} \end{bmatrix}, \quad n = 0, 1, 2, \dots$$

Since $\sigma(\mathbf{A}) = \{1, 0.82288, 0.23545\}$, then $\rho(\mathbf{A}) = 1$ and \mathbf{A} is of class \mathfrak{M} . On the other hand, taking into account Table 2.2 it is easy to check that condition (3.19) holds true

$$\begin{aligned} \sum_{n=0}^{\infty} \|\vec{B}_n\|_{r=3} &= \sum_{n=0}^{\infty} \max \left\{ \frac{\sqrt{2}}{2} n e^{-2(n+1)}, \frac{\sqrt{6}}{6} \frac{1}{(n+1)^2}, \sqrt{\frac{1}{101}} \frac{n}{(n+1)^4} \right\} \\ &= \frac{\sqrt{6}}{6} \sum_{n=0}^{\infty} \frac{1}{(n+1)^2} = \frac{\sqrt{6} \pi^2}{36} < +\infty. \end{aligned}$$

Therefore by Remark 3.5 one can assure that \vec{X}_n is m.s. convergent. Since $E[A] = 2/3$, $E[B] = 1/3$, $E[C] = 1/51$ (see Table 2.1) and from (3.33), one computes the expectation of the solution

$$E[\vec{X}_n] = \sum_{j=0}^{n-1} \begin{bmatrix} 19/20 & 1/10 & 1/8 \\ 1/20 & 7/30 & 0 \\ 0 & 2/3 & 7/8 \end{bmatrix}^{n-j-1} \begin{bmatrix} -\frac{2}{3} j e^{-2(j+1)} \\ \frac{1}{3(j+1)^2} \\ -\frac{j}{51(j+1)^4} \end{bmatrix}, \quad n = 0, 1, 2, \dots$$

Figure 3.2 shows the average concentration in each compartment during the first twenty periods. In order to compute how variances change with time, first note that as $\vec{X}_0 = \vec{0}$, then $\Lambda_{\vec{X}_0, \vec{B}_k} = \Lambda_{\vec{B}_j, \vec{X}_0} = \Lambda_{\vec{X}_0, \vec{X}_0} = \mathbf{0}_{3 \times 3}$, being $\mathbf{0}_{3 \times 3}$ the square null matrix of size 3. On the other hand, we take advantage of mutually independence of r.v.'s involved in the model to write that

$$\Lambda_{\vec{B}_j, \vec{B}_k} = \begin{bmatrix} \frac{j k}{18} e^{-2(j+k+2)} & 0 & 0 \\ 0 & \frac{1}{18(j+1)^2(k+1)^2} & 0 \\ 0 & 0 & \frac{2500 j k}{262701 (j+1)^4(k+1)^4} \end{bmatrix},$$

then considering these partial computations in (3.35) and taking there $n = m$, diagonal elements of $\Lambda_{\vec{X}_n, \vec{X}_n}$ give us the variance of each concentration in period n . Table 3.1 shows the variance of the concentration at each compartment for different values of n .

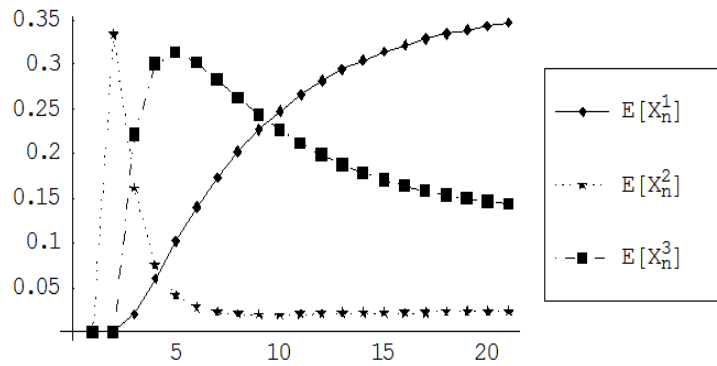


Figure 3.2: Plotting $E[\vec{X}_n]$ for $n = 0, 1, 2, \dots, 20$, for the Example 3.2

	$\text{Var}[X_n^1]$	$\text{Var}[X_n^2]$	$\text{Var}[X_n^3]$
n=0	0	0	0
n=1	0	0.0555556	0
n=2	0.0005742	0.0129784	0.0247657
n=3	0.0028833	0.0029106	0.0456958
n=4	0.0068947	0.0008996	0.0497679
n=5	0.0120381	0.0004232	0.0463320
n=10	0.0392776	0.0002286	0.0232788
n=15	0.0564256	0.0002699	0.0142004
n=20	0.0649329	0.0002956	0.0109922
n=25	0.0689968	0.0003074	0.0097842
n=30	0.0710117	0.0003125	0.0093097

Table 3.1: Variance of concentrations into each compartment at different periods for Example 3.2

3.7 **Conclusions**

In this chapter we have constructed a mean square solution stochastic process of random coupled linear difference equations incorporating uncertainty into both, the initial condition and the source term. It is a first stage in order to consider randomness in multiplicative coefficient in future studies. Also, we have provided sufficient conditions for guaranteeing mean square stability of the solution stochastic process. Furthermore, relevant statistical functions associated to the solution, like mean and covariance, were computed explicitly, in terms of the data. Finally, we have presented two illustrative examples of potential interest in applications.

The main results of this chapter have been accepted for publication in [11].

Chapter 4

Mean square analytic solution of coupled linear differential models with random initial condition and source term

This chapter deals with the construction of random power series solution of coupled linear differential equations containing uncertainty in both, initial condition and source term. Under appropriate hypotheses on the data, we establish that the random series solution constructed by a Fröbenius method is convergent in the mean square sense. A version of the so-called Mertens' theorem to multiply a couple of series, being one of them random and the other deterministic, as well as the so-called Cauchy's inequalities related to the coefficients of a mean square analytic stochastic process will be previously established. Also, the main statistical functions, such as the mean and the covariance, of the approximate solution stochastic process generated by truncation of the exact series solution are given. Finally, several illustrative examples are included.

4.1 Introduction

The aim of this chapter is to construct random power series solutions of coupled linear differential models of the form

$$\dot{\vec{X}}(t) = \mathbf{A}(t)\vec{X}(t) + \vec{B}(t), \quad |t| < c, \quad \vec{X}(0) = \vec{X}_0, \quad (4.1)$$

where the coefficient $\mathbf{A}(t)$ is a $\mathbb{R}^{r \times r}$ valued analytic function on $|t| < c$, the source term $\vec{B}(t)$ is a vector stochastic process of size $r \times 1$ which is mean square analytic on $|t| < c$ and the initial condition \vec{X}_0 is a random vector of the space L_2^r .

Random linear matrix differential equations of type (4.1) appear in the study of the response of a mass-spring linear oscillator to a random excitation [56, p.158]; in the determination of the effect on a earthbound structure of an earthquake-type disturbance [4, 57]; in the construction of compartmental models in pharmacokinetics [56, p.223] or in the formulation of epidemiological models [45]. The extension of the deterministic Malthus population model to the random framework leads in a natural way to problems of type (4.1). In fact, modeling population growth requires the estimation of some parameters such as the birth rate and the death rate. In practice, the determination of these parameters depends on different factors like sanitation and health care, wars, pollution, medicines, diet, psychological stress and anxiety. Other factors that should be considered are migration, living space restrictions, availability of food and water, climatology and epidemics. The huge complexity of these factors recommends to introduce them into mathematical models by means of random variables or stochastic process, instead of deterministic parameters or functions, respectively [56], [45], [8].

The study of random problems of type (4.1) provides a first stage to consider non-linear random models by means of linearization techniques [56, p.208]. As the exact solution process of (4.1) is not available in general, in [43, 16, 19] one develops random numerical methods for constructing approximate solutions of random differential systems. This approach can be considered as an alternative to the present work for approximating the solution stochastic process of (4.1).

This chapter is organized as follows. Section 4.2 contains some stochastic results that are not included in chapters 2 and 3, but that they will play

an important role in the following sections. Particularly, we provide a random matrix version of the Mertens's theorem related to product matrix series. Cauchy's inequalities for mean square analytic stochastic processes are also established in this section. Section 4.3 deals with the construction of a series solution of (4.1) by a random Fröbenius method, as well as we establish its convergence in the mean square sense under appropriate conditions on the data. This allows to construct an analytic-numerical series solution stochastic process of (4.1) by means of truncation. For this random finite series, the main statistical functions like mean and covariance matrix are computed in section 4.4. Several illustrative examples are given in section 4.5. In section 4.6 the main conclusions of this chapter are summarized.

4.2 Preliminaries

For the sake of clarity, this section contains some stochastic results that will be required in following sections but, for convenience, we do not have introduced previously in chapters 2 and 3. In this sense note that chapter 2 is devoted to introduce basic concepts and results for both discrete and continuous scalar s.p.'s, whereas section 3.2 of chapter 3 extends the corresponding part of chapter 2 to discrete vector s.p.'s since chapter 3 is devoted to compute m.s. analytic solutions of coupled linear difference models with random initial condition and source term. As the title of the current chapter announces, an extension of the corresponding concepts and results given in chapter 2 for continuous scalar s.p.'s to the vector framework is now required. This is the aim of the following exposition.

Throughout this chapter we will consider vector r.v.'s belonging to the Banach space $(L_2^r, \|\cdot\|_r)$ introduced previously in chapter 3 (see (3.2)). In this space the convergence of a sequence means convergence in the norm $\|\cdot\|_r$ according to (3.7). We say that $\{\vec{X}(t) : t \in \mathcal{T}\}$ is a second order vector stochastic process (2-v.s.p.), if $\vec{X}(t) \in L_2^r$ for each $t \in \mathcal{T}$, being \mathcal{T} the so-called space of times. The expectation function of $\vec{X}(t) = (X^1(t), \dots, X^r(t))^T$ is the deterministic vector function

$$\mathbf{E} [\vec{X}(t)] = (\mathbf{E} [X^1(t)], \dots, \mathbf{E} [X^r(t)])^T,$$

where T denotes the transpose vector operator. Note that by definition one gets

$$\mathbb{E} [(\vec{X}(t))^T] = \left(\mathbb{E} [\vec{X}(t)] \right)^T. \quad (4.2)$$

The covariance matrix function of $\{\vec{X}(t) : t \in \mathcal{T}\}$ is defined by the following square matrix of size $r \times r$

$$\mathbf{\Lambda}_{\vec{X}(t), \vec{X}(s)} = \mathbb{E} \left[\left(\vec{X}(t) - \mathbb{E} [\vec{X}(t)] \right) \left(\vec{X}(s) - \mathbb{E} [\vec{X}(s)] \right)^T \right] = (v^{ij}(t, s))_{r \times r} \quad (4.3)$$

where

$$\begin{aligned} v^{ij}(t, s) &= \mathbb{E} \left[(X^i(t) - \mathbb{E} [X^i(t)]) (X^j(s) - \mathbb{E} [X^j(s)]) \right] \\ &= \mathbb{E} [X^i(t)X^j(s)] - \mathbb{E} [X^i(t)] \mathbb{E} [X^j(s)], \quad 1 \leq i, j \leq r, t, s \in \mathcal{T}. \end{aligned} \quad (4.4)$$

Fixed $t \in \mathcal{T}$, note that $v^{ii}(t, t)$ denoted by $\text{Var} [X^i(t)]$, is the variance of the r.v. $X^i(t)$, $1 \leq i \leq r$. In the particular case that $r = 1$ and $\vec{X}(t) = X(t)$, (4.3)-(4.4) is the variance of the scalar 2-r.v. $X(t)$, and it will be denoted by $\text{Var} [X(t)]$, it coincides with (2.7).

Given two 2-v.s.p.'s in L_2^r , $\{\vec{X}(t) : t \in \mathcal{T}\} = \{(X^1(t), \dots, X^r(t))^T : t \in \mathcal{T}\}$, $\{\vec{Y}(t) : t \in \mathcal{T}\} = \{(Y^1(t), \dots, Y^r(t))^T : t \in \mathcal{T}\}$, their cross-covariance matrix function is defined by

$$\mathbf{\Lambda}_{\vec{X}(t), \vec{Y}(s)} = \mathbb{E} \left[\left(\vec{X}(t) - \mathbb{E} [\vec{X}(t)] \right) \left(\vec{Y}(s) - \mathbb{E} [\vec{Y}(s)] \right)^T \right] = (\nu^{ij}(t, s))_{r \times r}, \quad (4.5)$$

where

$$\begin{aligned} \nu^{ij}(t, s) &= \mathbb{E} \left[(X^i(t) - \mathbb{E} [X^i(t)]) (Y^j(s) - \mathbb{E} [Y^j(s)]) \right] \\ &= \mathbb{E} [X^i(t)Y^j(s)] - \mathbb{E} [X^i(t)] \mathbb{E} [Y^j(s)], \quad 1 \leq i, j \leq r, t, s \in \mathcal{T}. \end{aligned} \quad (4.6)$$

From definitions above, one deduces that covariance matrix of $\vec{X}(t)$ is a particular case of cross-covariance matrix of s.p.'s $\vec{X}(t)$ and $\vec{Y}(t) = \vec{X}(t)$. For $t \in \mathcal{T}$ fixed, and two 2-v.s.p.'s $\{\vec{X}(t) : t \in \mathcal{T}\} = \{(X^1(t), \dots, X^r(t))^T : t \in \mathcal{T}\}$, $\{\vec{Y}(t) : t \in \mathcal{T}\} = \{(Y^1(t), \dots, Y^r(t))^T : t \in \mathcal{T}\}$ such that 2-r.v.'s $\{X^i(t)\}_{i=1}^r$ and $\{Y^i(t)\}_{i=1}^r$ are pairwise independent, then $\mathbf{\Lambda}_{\vec{X}(t), \vec{Y}(t)}$ is the null matrix of size $r \times r$. In the specific case where $\vec{X}(t) = \vec{Y}(t)$, for $t \in \mathcal{T}$, being $\{X^i(t)\}_{i=1}^r$

pairwise independent r.v.'s, $\mathbf{\Lambda}_{\vec{X}(t), \vec{X}(t)}$ is the diagonal matrix $Diag(\text{Var}[X^i(t)])$, $1 \leq i \leq r$. Also, note that from property (4.2) and (4.5)-(4.6), the following property holds true:

$$\mathbf{\Lambda}_{\vec{Y}(t), \vec{X}(t)} = \left(\mathbf{\Lambda}_{\vec{X}(t), \vec{Y}(t)} \right)^T.$$

Since we pointed out at Remark 3.1 of chapter 3, there is not loss of generality in assuming that a 2-v.s.p. is centered at the origin, that is, $E[\vec{X}(t)] = \vec{0}$ for each $t \in \mathcal{T}$. In this case the covariance matrix given by (4.3) becomes

$$\mathbf{\Gamma}_{\vec{X}(t), \vec{X}(t)} = E \left[\vec{X}(t) \left(\vec{X}(t) \right)^T \right], \quad t \in \mathcal{T},$$

and it is referred to as correlation matrix function. In the following we will provide some results in terms of the covariance matrix function, but they could be stated through the correlation matrix function without any additional effort. The concept of m.s. continuity introduced in subsection 2.2.1 is easily extended to the vector framework. We say that the 2-v.s.p. $\{\vec{X}(t) : t \in \mathcal{T}\}$ is m.s. continuous on \mathcal{T} if

$$\lim_{\tau \rightarrow 0} \left\| \vec{X}(t + \tau) - \vec{X}(t) \right\|_r = 0,$$

for each $t \in \mathcal{T}$ with $t + \tau \in \mathcal{T}$.

Example 4.1. Let $\{\vec{X}_n : n \geq 0\}$ be a sequence in L_2^r and let $t \in \mathcal{T}$ being \mathcal{T} an interval of real line, then for each n_0 a nonnegative integer, the 2-v.s.p. given by $\{n_0 \vec{X}_{n_0} t^{n_0-1} : t \in \mathcal{T}\}$ is m.s. continuous for each $t \in \mathcal{T}$. In fact,

$$\left\| n_0 \vec{X}_{n_0} (t + \tau)^{n_0-1} - n_0 \vec{X}_{n_0} t^{n_0-1} \right\|_r = |n_0 ((t + \tau)^{n_0-1} - t^{n_0-1})| \left\| \vec{X}_{n_0} \right\|_r \xrightarrow{\tau \rightarrow 0} 0,$$

because $\left\| \vec{X}_{n_0} \right\|_r < \infty$ as $\vec{X}_{n_0} \in L_2^r$ for each n_0 and the continuity of the deterministic function $f(t) = t^{n_0-1}$ with respect to t .

Given $\{\vec{X}(t) : t \in \mathcal{T}\}$ a 2-v.s.p., we say that $\{\vec{X}(t) : t \in \mathcal{T}\}$ is m.s. differentiable on \mathcal{T} if

$$\lim_{\tau \rightarrow 0} \left\| \frac{\vec{X}(t + \tau) - \vec{X}(t)}{\tau} - \dot{\vec{X}}(t) \right\|_r = 0,$$

for all $t \in \mathcal{T}$ with $t + \tau \in \mathcal{T}$. In this case, $\{\dot{\vec{X}}(t) : t \in \mathcal{T}\}$ is called the m.s. derivative of $\{\vec{X}(t) : t \in \mathcal{T}\}$. This definition extends the corresponding introduced in section 2.2.2 for the scalar case.

Example 4.2. With the notation of Example 4.1, the 2-v.s.p. $\{\vec{X}_{n_0} t^{n_0} : t \in \mathcal{T}\}$ is m.s. differentiable on \mathcal{T} and its m.s. derivative is $\{n_0 \vec{X}_{n_0} t^{n_0-1} : t \in \mathcal{T}\}$. In fact

$$\left\| \frac{\vec{X}_{n_0}(t+\tau)^{n_0} - \vec{X}_{n_0} t^{n_0}}{\tau} - n_0 \vec{X}_{n_0} t^{n_0-1} \right\|_r = \left| \frac{(t+\tau)^{n_0} - t^{n_0}}{\tau} - n_0 t^{n_0-1} \right| \left\| \vec{X}_{n_0} \right\|_r \xrightarrow{\tau \rightarrow 0} 0$$

Following result is a direct extension to the vector framework of the m.s. differentiation theorem for series stated in Theorem 2.5 in chapter 2.

Proposition 4.1. Assume that $\{\vec{X}_n(t) : t \in \mathcal{T}\}$ is a 2-v.s.p. in L_2^r for each $n \in \mathbb{N}$, where \mathcal{T} is an interval, and satisfies:

- i) $\vec{X}(t) = \sum_{n \geq 0} \vec{X}_n(t)$ is m.s. convergent in L_2^r .
- ii) $\vec{X}_n(t)$ is m.s. differentiable and $\dot{\vec{X}}_n(t)$ is m.s. continuous for each $t \in \mathcal{T}$, $n \in \mathbb{N}$.
- iii) $\sum_{n \geq 0} \dot{\vec{X}}_n(t)$ is locally uniformly m.s. convergent in L_2^r .

Then for each $t \in \mathcal{T}$, $\{\vec{X}(t) : t \in \mathcal{T}\}$ is m.s. differentiable and its m.s. derivative is given by

$$\dot{\vec{X}}(t) = \sum_{n \geq 0} \dot{\vec{X}}_n(t).$$

Now we extend to the vector framework, the concept of m.s. analytic s.p. introduced previously in section 2.2.2. A 2-v.s.p. $\{\vec{X}(t) : t \in \mathcal{T}\}$ is said to be m.s. analytic on \mathcal{T} , if the series

$$\sum_{n \geq 0} \frac{\vec{X}^{(n)}(t_0)}{n!} (t - t_0)^n, \quad t, t_0 \in \mathcal{T},$$

is m.s. convergent to the 2-v.s.p. $\{\vec{X}(t) : t \in \mathcal{T}\}$. Here $\vec{X}^{(n)}(t_0)$ denotes the m.s. derivative of order n at $t = t_0$ of $\vec{X}(t)$. In connection with the m.s.

analyticity, we can extend the corresponding scalar characterization given in terms of the correlation function associated to the process (see [56, p.99] or Theorem 2.4) to the vector case by means of the correlation (or equivalently covariance) matrix function as follows:

Proposition 4.2. *A 2-v.s.p. $\{\vec{X}(t) : t \in \mathcal{T}\}$ is m.s. analytic on \mathcal{T} if, and only if, its covariance matrix function $\mathbf{\Lambda}_{\vec{X}(t), \vec{X}(t)}$ is analytic at point t for every $t \in \mathcal{T}$.*

Next we shall establish a random version of the Cauchy's inequalities in the m.s. sense. For that, let us consider a scalar m.s. analytic s.p. $\{B(t) : t \in \mathcal{T}\}$, i.e.,

$$B(t) = \sum_{n=0}^{\infty} B_n t^n, \quad B_n = \frac{B^{(n)}(0)}{n!}, \quad |t| < c, \quad (4.7)$$

(note that without loss of generality, it corresponds to (2.26) with $\mathcal{T} = (-c, c)$, $c > 0$ and $t_0 = 0$). By applying Theorem 2.4, m.s analyticity of s.p. $\{B(t) : t \in (-c, c), c > 0\}$ can be expressed in terms of the analyticity of the correlation function associated to $B(t)$ at the points (t, t) belonging to the diagonal of the square $\mathcal{T} \times \mathcal{T} = (-c, c) \times (-c, c)$, $c > 0$, and in this case, $\Gamma_B(t, s)$ is analytic on $\mathcal{T} \times \mathcal{T}$, that is:

$$\Gamma_B(t, s) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{\gamma_{nm}}{n! m!} t^n s^m, \quad |t|, |s| < c, \quad (4.8)$$

where

$$\gamma_{nm} = \left. \frac{\partial^{n+m} \Gamma_B(t, s)}{\partial t^n \partial s^m} \right|_{(t,s)=(0,0)}. \quad (4.9)$$

Since $\Gamma_B(t, s)$ is analytical on $\mathcal{T} \times \mathcal{T} = (-c, c) \times (-c, c)$, by the deterministic Cauchy's inequalities [20, p.222] one gets

$$\exists M_{\Gamma_B} > 0 : |\gamma_{nm}| \leq \frac{M_{\Gamma_B}}{\rho^{n+m}}, \quad 0 < \rho < c. \quad (4.10)$$

On the one hand, taking into account (4.8)-(4.10) and (2.25), one follows

$$|\gamma_{nm}| = \left| \mathbb{E} \left[B^{(n)}(0) B^{(m)}(0) \right] \right| \leq \frac{M_{\Gamma_B}}{\rho^{n+m}}, \quad 0 < \rho < c. \quad (4.11)$$

On the other hand, by the definition of the 2-norm, the expression of coefficients $B_n(0)$ given by (4.7) and expression (4.11) for $n = m$, yields

$$\begin{aligned} \|B_n\|_2 &= \sqrt{\mathbb{E}[B_n B_n]} = \sqrt{\mathbb{E}\left[\frac{B^{(n)}(0)}{n!} \frac{B^{(n)}(0)}{n!}\right]} = \frac{\sqrt{\mathbb{E}[B^{(n)}(0)B^{(n)}(0)]}}{n!} \\ &\leq \sqrt{\mathbb{E}[B^{(n)}(0)B^{(n)}(0)]} \leq \sqrt{\frac{M_{\Gamma_B}}{\rho^{2n}}} = \frac{(M_{\Gamma_B})^{1/2}}{\rho^n}. \end{aligned}$$

Taking into account the definition of $\|\cdot\|_r$ given by (3.2) and applying the previous development to each componentwise of \vec{B}_n one establishes the following

Theorem 4.1. *Let $\{\vec{B}(t) : t \in (-c, c), c > 0\}$ a m.s. analytic s.p. in L_2^r , i.e.,*

$$\vec{B}(t) = \sum_{n \geq 0} \vec{B}_n t^n, \quad |t| < c,$$

then there exists $M > 0$ such that

$$\|\vec{B}_n\|_r \leq \frac{M}{\rho^n}, \quad 0 < \rho < c, \quad n \geq 0.$$

The following lemma is a direct consequence of Lemma 3.1:

Lemma 4.1. *Let $\{\mathbf{M}_n : n \geq 0\}$ be a sequence of matrices in $\mathbb{R}^{r \times r}$ and let us assume that $\mathbf{M}_n \xrightarrow[n \rightarrow \infty]{} \mathbf{M}$ with respect to the norm $\|\cdot\|_\infty$ defined by (3.9).*

Then, for each $\vec{X} \in L_2^r$, one gets $\mathbf{M}_n \vec{X} \xrightarrow[n \rightarrow \infty]{m.s.} \mathbf{M} \vec{X} \in L_2^r$.

Proof. Note that by Lemma 3.1 one gets

$$\left\| \mathbf{M}_n \vec{X} - \mathbf{M} \vec{X} \right\|_r = \left\| (\mathbf{M}_n - \mathbf{M}) \vec{X} \right\|_r \leq \|\mathbf{M}_n - \mathbf{M}\|_\infty \left\| \vec{X} \right\|_r \xrightarrow[n \rightarrow \infty]{} 0. \quad \square$$

The following result is a vector random version of Mertens's theorem.

Proposition 4.3. *Let $\{\mathbf{A}_n : n \geq 0\}$ be a sequence of matrices in $\mathbb{R}^{r \times r}$ such that the series $\sum_{n \geq 0} \mathbf{A}_n$ is absolutely convergent to \mathbf{A} with respect to the norm*

$\|\cdot\|_\infty$. Let $\{\vec{X}_n : n \geq 0\}$ be a sequence in L_2^r such that $\sum_{n \geq 0} \vec{X}_n$ is m.s. convergent to $\vec{X} \in L_2^r$. If $\vec{C}_n = \sum_{k=0}^n \mathbf{A}_{n-k} \vec{X}_k$, then $\sum_{n \geq 0} \vec{C}_n$ is m.s. convergent to $\mathbf{A}\vec{X}$.

Proof. Let us introduce

$$\mathbf{A}'_n = \sum_{k=0}^n \mathbf{A}_k, \quad \vec{X}'_n = \sum_{k=0}^n \vec{X}_k, \quad \vec{C}'_n = \sum_{k=0}^n \vec{C}_k, \quad \vec{Y}'_n = \vec{X}'_n - \vec{X},$$

and note that we wish to prove that $\{\vec{C}'_n : n \geq 0\}$ is m.s. convergent to $\mathbf{A}\vec{X} \in L_2^r$. By definition of \vec{C}'_n we have

$$\begin{aligned} \vec{C}'_n &= \vec{C}_0 + \vec{C}_1 + \cdots + \vec{C}_n \\ &= \mathbf{A}_0 \vec{X}_0 + (\mathbf{A}_0 \vec{X}_1 + \mathbf{A}_1 \vec{X}_0) + \cdots + (\mathbf{A}_0 \vec{X}_n + \cdots + \mathbf{A}_n \vec{X}_0) \\ &= \mathbf{A}_0 (\vec{X}_0 + \cdots + \vec{X}_n) + \mathbf{A}_1 (\vec{X}_0 + \cdots + \vec{X}_{n-1}) + \cdots + \mathbf{A}_n \vec{X}_0 \\ &= \mathbf{A}_0 \vec{X}'_n + \mathbf{A}_1 \vec{X}'_{n-1} + \cdots + \mathbf{A}_n \vec{X}'_0 \\ &= \mathbf{A}_0 (\vec{Y}'_n + \vec{X}) + \mathbf{A}_1 (\vec{Y}'_{n-1} + \vec{X}) + \cdots + \mathbf{A}_n (\vec{Y}'_0 + \vec{X}) \\ &= (\mathbf{A}_0 + \mathbf{A}_1 + \cdots + \mathbf{A}_n) \vec{X} + (\mathbf{A}_0 \vec{Y}'_n + \mathbf{A}_1 \vec{Y}'_{n-1} + \cdots + \mathbf{A}_n \vec{Y}'_0) \\ &= \mathbf{A}'_n \vec{X} + (\mathbf{A}_0 \vec{Y}'_n + \mathbf{A}_1 \vec{Y}'_{n-1} + \cdots + \mathbf{A}_n \vec{Y}'_0). \end{aligned}$$

If we denote $\vec{Z}'_n = \mathbf{A}_0 \vec{Y}'_n + \mathbf{A}_1 \vec{Y}'_{n-1} + \cdots + \mathbf{A}_n \vec{Y}'_0$, then one gets

$$\vec{C}'_n = \mathbf{A}'_n \vec{X} + \vec{Z}'_n. \tag{4.12}$$

Note that by Lemma 4.1, $\{\mathbf{A}'_n \vec{X} : n \geq 0\}$ is m.s. convergent to $\mathbf{A}\vec{X}$. Taking into account (4.12), in order to prove the result, it is sufficient to show that $\{\vec{Z}'_n : n \geq 0\}$ is m.s. convergent to $\vec{0} \in L_2^r$. As $\{\vec{Y}'_n : n \geq 0\}$ is m.s. convergent to $\vec{0} \in L_2^r$, given $\epsilon > 0$, there exists $N_0 > 0$ such that

$$\|\vec{Y}'_n\|_r < \frac{\epsilon}{2\alpha}, \quad n \geq N_0, \tag{4.13}$$

being

$$\alpha = \sum_{n \geq 0} \|\mathbf{A}_n\|_\infty < +\infty. \quad (4.14)$$

Let $\beta_{N_0} > 0$ be defined by

$$\beta_{N_0} = \max \left\{ \left\| \vec{Y}'_0 \right\|_r, \dots, \left\| \vec{Y}'_{N_0} \right\|_r \right\}, \quad (4.15)$$

and since sequences $\{\mathbf{A}_n : n \geq 0\}, \{\mathbf{A}_{n-1} : n \geq 1\}, \dots, \{\mathbf{A}_{n-N_0} : n \geq N_0\}$ are convergent to the null matrix in $\mathbb{R}^{r \times r}$ with respect to the norm $\|\cdot\|_\infty$, we may also assume that

$$\|\mathbf{A}_{n-j}\|_\infty \leq \frac{\epsilon}{2(N_0 + 1)\beta_{N_0}}, \quad j = 0, 1, \dots, N_0. \quad (4.16)$$

Let $n > N_0$, then by Lemma 3.1, (4.13)-(4.16) one gets

$$\begin{aligned} \left\| \vec{Z}'_n \right\|_r &\leq \left\| \mathbf{A}_n \vec{Y}'_0 + \dots + \mathbf{A}_{n-N_0} \vec{Y}'_{N_0} \right\|_r + \left\| \mathbf{A}_{n-N_0-1} \vec{Y}'_{N_0+1} + \dots + \mathbf{A}_0 \vec{Y}'_n \right\|_r \\ &\leq \|\mathbf{A}_n\|_\infty \left\| \vec{Y}'_0 \right\|_r + \dots + \|\mathbf{A}_{n-N_0}\|_\infty \left\| \vec{Y}'_{N_0} \right\|_r \\ &\quad + \|\mathbf{A}_{n-N_0-1}\|_\infty \left\| \vec{Y}'_{N_0+1} \right\|_r + \dots + \|\mathbf{A}_0\|_\infty \left\| \vec{Y}'_n \right\|_r \\ &\leq \beta_{N_0}(N_0 + 1) \frac{\epsilon}{2(N_0 + 1)\beta_{N_0}} + (\|\mathbf{A}_{n-N_0-1}\|_\infty + \dots + \|\mathbf{A}_0\|_\infty) \frac{\epsilon}{2\alpha} \\ &\leq \frac{\epsilon}{2} + \alpha \frac{\epsilon}{2\alpha} = \epsilon. \quad \square \end{aligned}$$

4.3 Constructing a random power series solution

In the section we face the problem of constructing a solution s.p. of random differential system (4.1) by means of a random power series. We assume that coefficient $\mathbf{A}(t)$ is a matrix valued deterministic analytic function about $t = 0$, i.e.,

$$\mathbf{A}(t) = \sum_{n \geq 0} \mathbf{A}_n t^n, \quad |t| < c, \quad (4.17)$$

where $\mathbf{A}_n \in \mathbb{R}^{r \times r}$, for each $n \geq 0$, and $\vec{B}(t)$ is an analytic 2-v.s.p. about $t = 0$, i.e.,

$$\vec{B}(t) = \sum_{n \geq 0} \vec{B}_n t^n, \quad |t| < c, \quad (4.18)$$

where $\vec{B}_n \in L_2^r$ and, initial condition $\vec{X}_0 \in L_2^r$.

Let $F(\vec{X}, t) = \mathbf{A}(t)\vec{X} + \vec{B}(t)$, where $\vec{X} \in L_2^r$ and $|t| < c$. Then taking into account that $F(\vec{Y}, t) - F(\vec{X}, t) = \mathbf{A}(t)(\vec{Y} - \vec{X})$, by Lemma 3.1, it follows that

$$\left\| F(\vec{Y}, t) - F(\vec{X}, t) \right\|_r \leq \|\mathbf{A}(t)\|_\infty \left\| \vec{Y} - \vec{X} \right\|_r \leq k \left\| \vec{Y} - \vec{X} \right\|_r \quad (4.19)$$

where, from the analyticity of $\mathbf{A}(t)$, one gets

$$\tilde{a}_{ij} > \sup_{|t| < c} |a_{ij}(t)|, \quad k = \max_{1 \leq i \leq r} \sum_{j=1}^r \tilde{a}_{ij} < +\infty. \quad (4.20)$$

From Theorem 5.1.2. [56, p.118], (4.19)-(4.20) guarantees the existence of a unique m.s. solution s.p. of problem (4.1).

Let us seek a formal solution s.p. of i.v.p. (4.1) of the form

$$\vec{X}(t) = \sum_{n \geq 0} \vec{X}_n t^n, \quad (4.21)$$

where coefficients \vec{X}_n are second order random vectors of size r to be determined. Assuming that $\vec{X}(t)$ is termwise m.s. differentiable and applying Proposition 4.1, Proposition 4.3 as well as Example 4.2, it follows that

$$\dot{\vec{X}}(t) = \sum_{n \geq 0} (n+1) \vec{X}_{n+1} t^n,$$

$$\mathbf{A}(t)\vec{X}(t) = \left(\sum_{n \geq 0} \mathbf{A}_n t^n \right) \left(\sum_{n \geq 0} \vec{X}_n t^n \right) = \sum_{n \geq 0} \left(\sum_{k=0}^n \mathbf{A}_{n-k} \vec{X}_k \right) t^n,$$

and by imposing that $\vec{X}(t)$ given by (4.21) satisfies (4.1), one gets

$$\sum_{n \geq 0} (n+1) \vec{X}_{n+1} t^n = \sum_{n \geq 0} \left\{ \left(\sum_{k=0}^n \mathbf{A}_{n-k} \vec{X}_k \right) + \vec{B}_n \right\} t^n.$$

Henceforth, a candidate solution s.p. of the form (4.21) can be found by imposing that its coefficient \vec{X}_n satisfy the following relationship

$$(n+1)\vec{X}_{n+1} = \vec{B}_n + \sum_{k=0}^n \mathbf{A}_{n-k} \vec{X}_k, \quad (4.22)$$

with the initial condition $\vec{X}_0 \in L_2^r$, or equivalently

$$\vec{X}_{n+1} = \frac{1}{n+1} \left(\vec{B}_n + \sum_{k=0}^n \mathbf{A}_{n-k} \vec{X}_k \right), \quad n \geq 0. \quad (4.23)$$

Now, we will prove that the s.p. $\vec{X}(t)$ defined by (4.21) and (4.23) is m.s. convergent, termwise differentiable and then Proposition 4.1 is applicable. Let $\rho = |\hat{t}| < c$ be an arbitrary domain included into the domain $|t| < c$, firstly we shall show that

$$\sum_{n \geq 0} \left\| \vec{X}_n \right\|_r |t|^n < +\infty, \quad |t| < \rho.$$

On the one hand, since $\mathbf{A}(t)$ is analytic by the deterministic Cauchy inequalities (see [20, p.222]) there exists $M_1 > 0$ such that $\|\mathbf{A}_n\|_\infty \leq M_1/\rho^n$ for $0 < \rho < c$ and each $n \geq 0$. On the other hand, since $\vec{B}(t)$ is m.s. analytic, by Cauchy inequalities established in Theorem 4.1, there exists $M_2 > 0$ such that $\|\vec{B}_n\|_r \leq M_2/\rho^n$ for $0 < \rho < c$ and $n \geq 0$, then by taking $M = \max\{M_1, M_2\}$ one gets

$$\|\mathbf{A}_n\|_\infty \leq \frac{M}{\rho^n}, \quad \|\vec{B}_n\|_r \leq \frac{M}{\rho^n}, \quad n \geq 0. \quad (4.24)$$

From (4.22), taking into account (4.24) and Lemma 4.1, it follows that

$$\begin{aligned} (n+1) \left\| \vec{X}_{n+1} \right\|_r &\leq \left\| \vec{B}_n \right\|_r + \sum_{k=0}^n \|\mathbf{A}_{n-k}\|_\infty \left\| \vec{X}_k \right\|_r \\ &\leq \frac{M}{\rho^n} + \sum_{k=0}^n \frac{M}{\rho^{n-k}} \left\| \vec{X}_k \right\|_r \\ &= \frac{M}{\rho^n} \left(1 + \sum_{k=0}^n \rho^k \left\| \vec{X}_k \right\|_r \right). \end{aligned} \quad (4.25)$$

Let us introduce positive numbers y_n , such that

$$y_0 = \left\| \vec{X}_0 \right\|_r,$$

and inductively y_n is defined by the equation

$$(n+1)y_{n+1} = \frac{M}{\rho^n} \left(1 + \sum_{k=0}^n \rho^k y_k \right), \quad n \geq 0. \quad (4.26)$$

By definition of y_n and (4.25) one gets

$$\left\| \vec{X}_n \right\|_r \leq y_n, \quad n \geq 0.$$

From (4.26) it follows

$$y_{n+1} = \frac{M + n\rho^{-1}}{n+1} y_n,$$

henceforth if $|t| < \rho$ one gets

$$\lim_{n \rightarrow \infty} \frac{y_{n+1} |t|^{n+1}}{y_n |t|^n} = \lim_{n \rightarrow \infty} \frac{M\rho + n}{(n+1)\rho} |t| = \frac{|t|}{\rho} < 1. \quad (4.27)$$

Since $|t| < \rho < c$ is arbitrary, by (4.27) one deduces the $\|\cdot\|_r$ -absolute convergence in L_2^r of the series defined by (4.21) and (4.23) for all t such that $|t| < c$. Let us prove that $\vec{X}(t)$ defined by (4.21) and (4.23) is termwise m.s. differentiable at $t : 0 < |t| < c$. Let us take ρ such that $0 < |t| < \rho < c$, and taking into account that

$$\sum_{n \geq 0} \left\| \vec{X}_n \right\|_r \rho^n < +\infty,$$

given $\epsilon > 0$, there exists N_0 such that

$$\left\| \vec{X}_{n+1} \right\|_r \rho^{n+1} < \epsilon, \quad \forall n \geq N_0.$$

Note that for $n \geq N_0$ and t' such that $0 < |t'| < \rho$ one gets

$$\begin{aligned} \left\| (n+1)\vec{X}_{n+1} (t')^n \right\|_r &= (n+1) \left\| \vec{X}_{n+1} \right\|_r |t'|^n \\ &= (n+1) \left\| \vec{X}_{n+1} \right\|_r \rho^{n+1} \frac{1}{\rho} \left(\frac{|t'|}{\rho} \right)^n \\ &< \frac{n+1}{\rho} \epsilon \left(\frac{|t'|}{\rho} \right)^n = c_n, \end{aligned}$$

where

$$\lim_{n \rightarrow \infty} \frac{c_{n+1}}{c_n} = \lim_{n \rightarrow \infty} \frac{n+2}{n+1} \frac{|t'|}{\rho} = \frac{|t'|}{\rho} < 1.$$

Therefore, $\sum_{n \geq 0} (n+1) \vec{X}_{n+1} \tau^n$ is uniformly absolutely m.s convergent for any arbitrary $\tau : |\tau| \leq |t'|$. Example 4.2 shows that $\vec{X}_n t^n$ is m.s. differentiable whereas in Example 4.1 it was checked that $n \vec{X}_n t^{n-1}$ is m.s. continuous. Thus hypotheses of Proposition 4.1 are satisfied and series (4.21) and (4.23) is termwise m.s. differentiable on $|t| < c$.

Theorem 4.2. *Under hypotheses of analyticity of matrix coefficient $\mathbf{A}(t)$ and m.s. analyticity of source term $\vec{B}(t)$ on the common domain $|t| < c$, the random power series given by (4.21) and (4.23) m.s. converges to the exact solution s.p. of initial value problem (4.1).*

Remark 4.1. *If $\vec{B}(t)$ is the null vector s.p. and $\vec{X}_0 \in L_2^r$, then applying the above random Fröbenius method one deduces that the m.s. solution of (4.1) is given by $\vec{X}(t) = \exp(\mathbf{A}t) \vec{X}_0$, as it is well-known, see [56, chap.6]. However this result does not hold, in general, if coefficient $\mathbf{A}(t)$ is a matrix s.p. or a matrix r.v. without assuming additional conditions on $\mathbf{A}(t)$, see [18].*

Remark 4.2. *From a practical point of view it is important to point out that s.p.'s such that Brownian motion or white noise are not allowed to play the role of $\vec{B}(t)$ in problem (4.1) because of they are not m.s. differentiable anywhere, however interesting examples involving other s.p.'s are included in last section.*

4.4 Statistical functions of the mean square approximating solution

From a practical view-point, it is interesting to compute the expectation and the covariance matrix functions of the approximating solution s.p. since they give its basic statistical behavior. In order to compute the expectation of solution s.p. $\vec{X}(t)$ of problem (4.1) we have several ways. A direct approach is based on the following property (note that, it only is an extension to the vector framework of property (2.24) for $n = 1$)

$$\frac{d}{dt} \left(\mathbb{E} \left[\vec{X}(t) \right] \right) = \mathbb{E} \left[\dot{\vec{X}}(t) \right], \quad (4.28)$$

where the derivative of the left-hand side of (4.28) must be interpreted in the deterministic sense whereas, the derivative of the right-hand side must be considered in the m.s. sense. Then, taking expectations in (4.1), one gets the deterministic analytic system

$$\frac{d}{dt} \left(\mathbb{E}[\vec{X}(t)] \right) = \mathbf{A}(t) \mathbb{E}[\vec{X}(t)] + \mathbb{E}[\vec{B}(t)], \quad |t| < c, \quad \mathbb{E}[\vec{X}(0)] = \mathbb{E}[\vec{X}_0]. \quad (4.29)$$

Depending on the data $\mathbf{A}(t)$ and $\vec{B}(t)$, well-known deterministic methods are available in order to compute its exact solution $\mathbb{E}[\vec{X}(t)]$. Otherwise, accurate analytical-numerical methods have also been developed for constructing approximating solutions, see [42]. Other equivalent approach is exhibited in the following. First, we consider the approximating process generated by the truncation of the power series defined by (4.21) and (4.23)

$$\vec{X}_N(t) = \sum_{n=0}^N \vec{X}_n t^n, \quad (4.30)$$

then using linearity of the expectation operator one gets

$$\mathbb{E}[\vec{X}_N(t)] = \sum_{n=0}^N \mathbb{E}[\vec{X}_n] t^n, \quad (4.31)$$

where taking into account (4.23), series coefficients are given by

$$\mathbb{E}[\vec{X}_n] = \frac{1}{n} \left(\mathbb{E}[\vec{B}_{n-1}] + \sum_{k=0}^{n-1} \mathbf{A}_{n-k-1} \mathbb{E}[\vec{X}_k] \right), \quad n \geq 1, \quad (4.32)$$

being $\mathbb{E}[\vec{X}_0]$ and $\mathbb{E}[\vec{B}_n]$ for $n \geq 0$, data. In a similar way, from (4.23) one gets

$$\mathbb{E} \left[\left(\vec{X}_n \right)^T \right] = \frac{1}{n} \left(\mathbb{E} \left[\left(\vec{B}_{n-1} \right)^T \right] + \sum_{k=0}^{n-1} \mathbb{E} \left[\left(\vec{X}_k \right)^T \right] \left(\mathbf{A}_{n-k-1} \right)^T \right), \quad n \geq 1.$$

By Lemma 2.1, an important consequence of the m.s. convergence of series (4.21) and (4.23) established in Theorem 4.2 is that average approximation given by (4.31) and (4.32) tends as $N \rightarrow +\infty$ to the exact value $\mathbb{E}[\vec{X}(t)]$ for each t in the domain of convergence $|t| < c$. This is an important fact from a practical view-point since the exact solution of deterministic initial value

problem (4.29) is in practice exceptional.

Let us compute the covariance matrix of the approximating process $\vec{X}_N(t)$ defined by (4.30). Taking advantage of property (3.5), for this goal we will assume that

$$\mathbf{\Lambda}_{\vec{X}_0, \vec{X}_0}, \quad \mathbf{\Lambda}_{\vec{B}_i, \vec{X}_0} = \left(\mathbf{\Lambda}_{\vec{X}_0, \vec{B}_i} \right)^T, \quad \mathbf{\Lambda}_{\vec{B}_i, \vec{B}_j} = \left(\mathbf{\Lambda}_{\vec{B}_j, \vec{B}_i} \right)^T, \quad 0 \leq i, j \leq N,$$

are data. By considering (4.30), the definition of the covariance matrix of $\vec{X}_N(t)$ yields

$$\begin{aligned} \mathbf{\Lambda}_{\vec{X}_N(t), \vec{X}_N(t)} &= \mathbf{E} \left[\left(\vec{X}_N(t) - \mathbf{E} \left[\vec{X}_N(t) \right] \right) \left(\vec{X}_N(t) - \mathbf{E} \left[\vec{X}_N(t) \right] \right)^T \right] \\ &= \mathbf{E} \left[\vec{X}_N(t) \left(\vec{X}_N(t) \right)^T \right] - \mathbf{E} \left[\vec{X}_N(t) \right] \left(\mathbf{E} \left[\vec{X}_N(t) \right] \right)^T \\ &= \mathbf{E} \left[\left(\sum_{n=0}^N \vec{X}_n t^n \right) \left(\sum_{n=0}^N \left(\vec{X}_n \right)^T t^n \right) \right] \\ &\quad - \left(\sum_{n=0}^N \mathbf{E} \left[\vec{X}_n \right] t^n \right) \left(\sum_{n=0}^N \mathbf{E} \left[\left(\vec{X}_n \right)^T \right] t^n \right) \\ &= \sum_{i=0}^N \sum_{j=0}^N \mathbf{E} \left[\vec{X}_i \left(\vec{X}_j \right)^T \right] t^{i+j} \\ &\quad - \sum_{i=0}^N \sum_{j=0}^N \mathbf{E} \left[\vec{X}_i \right] \mathbf{E} \left[\left(\vec{X}_j \right)^T \right] t^{i+j} \\ &= \sum_{i=0}^N \sum_{j=0}^N \left(\mathbf{E} \left[\vec{X}_i \left(\vec{X}_j \right)^T \right] - \mathbf{E} \left[\vec{X}_i \right] \mathbf{E} \left[\left(\vec{X}_j \right)^T \right] \right) t^{i+j} \\ &= \sum_{i=0}^N \sum_{j=0}^N \mathbf{\Lambda}_{\vec{X}_i, \vec{X}_j} t^{i+j}, \end{aligned} \tag{4.33}$$

where linearity of expectation operator and property (4.2) have been applied. Note that expression (4.33) can be also derived as direct consequence of re-

relationship (3.6) for $a_i = t^i$ and $b_j = t^j$. Note that (4.33) can be developed as

$$\begin{aligned}
 \mathbf{\Lambda}_{\vec{X}_N(t), \vec{X}_N(t)} &= \mathbf{\Lambda}_{\vec{X}_0, \vec{X}_0} + \sum_{i=1}^N \mathbf{\Lambda}_{\vec{X}_i, \vec{X}_i} t^{2i} \\
 &+ \sum_{i=1}^N \left(\mathbf{\Lambda}_{\vec{X}_i, \vec{X}_0} + \left(\mathbf{\Lambda}_{\vec{X}_i, \vec{X}_0} \right)^T \right) t^i \\
 &+ \sum_{i=2}^N \sum_{j=1}^{i-1} \left(\mathbf{\Lambda}_{\vec{X}_i, \vec{X}_j} + \left(\mathbf{\Lambda}_{\vec{X}_i, \vec{X}_j} \right)^T \right) t^{i+j},
 \end{aligned} \tag{4.34}$$

where we have considered that for property (3.5), one gets $\mathbf{\Lambda}_{\vec{X}_j, \vec{X}_i} = \left(\mathbf{\Lambda}_{\vec{X}_i, \vec{X}_j} \right)^T$, what allow us to save computations.

Now we show that the involved computations in (4.34) can be handled in practice. From (4.23) and (4.32) one gets

$$\begin{aligned}
 \mathbf{\Lambda}_{\vec{X}_i, \vec{X}_0} &= \mathbf{E} \left[\left(\vec{X}_i - \mathbf{E}[\vec{X}_i] \right) \left(\vec{X}_0 - \mathbf{E}[\vec{X}_0] \right)^T \right] \\
 &= \mathbf{E} \left[\left(\frac{1}{i} \left\{ \vec{B}_{i-1} + \sum_{k=0}^{i-1} \mathbf{A}_{i-k-1} \vec{X}_k \right\} \right. \right. \\
 &\quad \left. \left. - \frac{1}{i} \left\{ \mathbf{E}[\vec{B}_{i-1}] + \sum_{k=0}^{i-1} \mathbf{A}_{i-k-1} \mathbf{E}[\vec{X}_k] \right\} \right) \left(\vec{X}_0 - \mathbf{E}[\vec{X}_0] \right)^T \right] \\
 &= \frac{1}{i} \mathbf{E} \left[\left(\left(\vec{B}_{i-1} - \mathbf{E}[\vec{B}_{i-1}] \right) + \sum_{k=0}^{i-1} \mathbf{A}_{i-k-1} \left(\vec{X}_k - \mathbf{E}[\vec{X}_k] \right) \right) \right. \\
 &\quad \left. \times \left(\vec{X}_0 - \mathbf{E}[\vec{X}_0] \right)^T \right] \\
 &= \frac{1}{i} \mathbf{\Lambda}_{\vec{B}_{i-1}, \vec{X}_0} + \frac{1}{i} \sum_{k=0}^{i-1} \mathbf{A}_{i-k-1} \mathbf{\Lambda}_{\vec{X}_k, \vec{X}_0}, \quad i \geq 1,
 \end{aligned}$$

thus, the terms $\mathbf{\Lambda}_{\vec{X}_i, \vec{X}_0}$ appearing in (4.34) can be computed from the data $\mathbf{\Lambda}_{\vec{X}_0, \vec{X}_0}$ and $\mathbf{\Lambda}_{\vec{B}_{i-1}, \vec{X}_0}$. In (4.34), only remains how to carry out the computa-

tion of matrices $\mathbf{\Lambda}_{\vec{X}_i, \vec{X}_j}$ for $i \geq 1, 1 \leq j \leq i$. For this goal, note that

$$\begin{aligned}
 \mathbf{\Lambda}_{\vec{X}_i, \vec{X}_j} &= \mathbb{E} \left[\left(\vec{X}_i - \mathbb{E}[\vec{X}_i] \right) \left(\vec{X}_j - \mathbb{E}[\vec{X}_j] \right)^T \right] \\
 &= \mathbb{E} \left[\left(\vec{X}_i - \mathbb{E}[\vec{X}_i] \right) \left(\frac{1}{j} \left\{ \left(\vec{B}_{j-1} - \mathbb{E}[\vec{B}_{j-1}] \right)^T \right. \right. \right. \\
 &\quad \left. \left. \left. + \sum_{l=0}^{j-1} \left(\vec{X}_l - \mathbb{E}[\vec{X}_l] \right)^T \mathbf{A}_{j-l-1}^T \right\} \right) \right] \tag{4.35} \\
 &= \frac{1}{j} \mathbf{\Lambda}_{\vec{X}_i, \vec{B}_{j-1}} + \frac{1}{j} \sum_{l=0}^{j-1} \mathbf{\Lambda}_{\vec{X}_i, \vec{X}_l} \mathbf{A}_{j-l-1}^T, \quad i \geq 1, 1 \leq j \leq i,
 \end{aligned}$$

then, one requires to know the matrices $\mathbf{\Lambda}_{\vec{X}_i, \vec{B}_{j-1}}$ and $\mathbf{\Lambda}_{\vec{X}_i, \vec{X}_l}$ for $0 \leq l \leq j-1$. Following the same procedure as the one used to obtain (4.35), one gets

$$\mathbf{\Lambda}_{\vec{X}_i, \vec{B}_{j-1}} = \frac{1}{i} \mathbf{\Lambda}_{\vec{B}_{i-1}, \vec{B}_{j-1}} + \frac{1}{i} \sum_{k=0}^{i-1} \mathbf{A}_{i-k-1} \mathbf{\Lambda}_{\vec{X}_k, \vec{B}_{j-1}}. \tag{4.36}$$

Thus, $\mathbf{\Lambda}_{\vec{X}_i, \vec{B}_{j-1}}$ one computes from the data $\mathbf{\Lambda}_{\vec{B}_{i-1}, \vec{B}_{j-1}}$ and $\mathbf{\Lambda}_{\vec{X}_0, \vec{B}_{j-1}}$. On the other hand, for $0 \leq l \leq j-1$ fixed, in order to compute $\mathbf{\Lambda}_{\vec{X}_i, \vec{X}_l}$ by (4.35) it is necessary to know $\mathbf{\Lambda}_{\vec{X}_i, \vec{B}_{l-1}}$ which is computable by (4.36) from the data $\mathbf{\Lambda}_{\vec{X}_0, \vec{B}_{l-1}}$ and $\mathbf{\Lambda}_{\vec{X}_i, \vec{X}_0}$.

4.5 Examples

In closing, we present illustrative applications of model (4.1) where the main statistical functions of the solution s.p. are obtained.

Example 4.3. *As a former test example, we will consider a particular case of i.v.p. (4.1) for which the exact mean of its solution s.p. is available. Hence it enables us to compare the goodness of the approximations obtained from our*

approach with respect to the exact ones. Let us consider model (4.1) where

$$\mathbf{A}(t) = \mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix}, \quad \vec{B}(t) = \begin{bmatrix} 0 \\ Ze^t \end{bmatrix}, \quad (4.37)$$

being Z an Exponential r.v. of parameter $\lambda = 0.5$, $Z \sim \text{Exp}(\lambda = 0.5)$, then $E[Z] = 2$ and $E[Z^2] = 8$ (see α_n , $n = 1, 2$ in Table 2.2 for an Exponential r.v. with $\lambda = 0.5$). Moreover, we are assuming through the second element of source term $\vec{B}(t)$ that incorporation of foreign prey individuals is allowed. Finally, let us suppose that initial conditions are 2-r.v.'s such that $E[X_1(0)] = 1$ and $E[X_2(0)] = 5$. In this case, coefficients \mathbf{A}_n of (4.17) are given by

$$\mathbf{A}_0 = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix}, \quad \mathbf{A}_n = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad n \geq 1.$$

Note that $\mathbf{A}(t)$ is analytic on the whole real line. The covariance matrix function of $\vec{B}(t)$ is

$$\mathbf{\Lambda}_{\vec{B}(t), \vec{B}(s)} = \begin{bmatrix} 0 & 0 \\ 0 & 4e^{t+s} \end{bmatrix},$$

which is analytic in the diagonal points (t, t) , then applying Proposition 4.2 one deduces that $\vec{B}(t)$ is m.s. analytic. Coefficients \vec{B}_n of (4.18) are given by

$$\vec{B}_n = \begin{bmatrix} 0 \\ Z \frac{1}{n!} \end{bmatrix}, \quad n \geq 0.$$

From Theorem 4.2, random power series given by (4.21) and (4.23) is m.s. convergent for all $t \in \mathbb{R}$ and as a result its average approximation (4.31) and (4.32) does.

Table 4.1 compares the expectation of the approximating s.p. $\vec{X}_N(t)$ for different orders of truncation N for $t \in [0, 2]$ with respect to the exact solution of i.v.p. (4.29), which is given by

$$E[\vec{X}(t)] = \begin{bmatrix} e^{-t} (1 + e^{2t} + 10t) / 2 \\ e^{-t} (9 + e^{2t} - 10t) / 2 \end{bmatrix}. \quad (4.38)$$

For that, we compute the componentwise absolute difference of vectors $E[\vec{X}(t)] - E[\vec{X}_N(t)]$, that will denoted by $|E[\vec{X}(t)] - E[\vec{X}_N(t)]|$. From Table 4.1 one observes that for t fixed, the approximations are better when N increases, and

for a fixed value of N , the approximations are worse when t departs from the origin, where the problem is proposed. Note that coefficients of $\mathbf{E}[\vec{X}_N(t)]$ defined by (4.32), in this case are given by

$$\mathbf{E}[\vec{X}_0] = \begin{bmatrix} 1 \\ 5 \end{bmatrix}, \quad (4.39)$$

$$\mathbf{E}[\vec{X}_n] = \frac{1}{n} \left\{ \begin{bmatrix} 0 \\ 2/(n-1)! \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix} \mathbf{E}[\vec{X}_{n-1}] \right\}, \quad n \geq 1. \quad (4.40)$$

Finally, we stress that in (4.37) an Exponential distribution for r.v. Z has been assumed, however in computations we have just used that $\mathbf{E}[Z] = 2$ and $\mathbf{E}[Z^2] = 8$. As a result, any other r.v. having the same first and second moments with respect to the origin could have been considered.

Example 4.4. Let us consider model (4.1) where

$$\mathbf{A}(t) = \begin{bmatrix} -e^{-t} & e^{-2t} \\ e^{-t} & -e^{-2t} \end{bmatrix}, \quad \vec{B}(t) = \begin{bmatrix} 0 \\ -B \end{bmatrix}, \quad \vec{X}_0 = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}, \quad (4.41)$$

being B an exponential r.v. with parameter $\lambda = 2$, i.e., $B \sim \text{Exp}(\lambda = 2)$ and, the initial conditions Z_1 and Z_2 are independent Gaussian r.v.'s with averages $\mu_1 = 0.5$, $\mu_2 = 1$ and variances $\sigma_1^2 = 1$, $\sigma_2^2 = 0.5$, respectively, i.e., $Z_1 \sim \text{N}(\mu_1 = 0.5; \sigma_1^2 = 1)$, $Z_2 \sim \text{N}(\mu_2 = 1; \sigma_2^2 = 0.5)$. We will also assume that r.v. B is independent of Z_1 and Z_2 . Note that coefficients \mathbf{A}_n of (4.17) are given by

$$\mathbf{A}_n = \frac{(-1)^n}{n!} \begin{bmatrix} -1 & 2^n \\ 1 & -2^n \end{bmatrix}, \quad n \geq 0,$$

whereas coefficients \vec{B}_n of (4.18) are given by

$$\vec{B}_0 = \begin{bmatrix} 0 \\ -B \end{bmatrix}, \quad \vec{B}_n = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad n \geq 1.$$

Table 4.2 compares the values of the mean obtained by using the truncation series method presented in this chapter for a truncation order of $N = 500$ (denoted by $\mathbf{E}[\vec{X}_{500}(t)]$) and Monte Carlo method with 10000 and 50000 simulations (denoted by $\tilde{\mu}_X^{10000}(t)$ and $\tilde{\mu}_X^{50000}(t)$), respectively, for different values of $t \in [0, 2]$.

Note that the diagonal elements of matrix $\mathbf{\Lambda}_{\vec{X}_N(t), \vec{X}_N(t)}$ given by (4.34) are the variance of each componentwise of the approximate solution $\vec{X}_N(t)$

t	$ E[\vec{X}(t)] - E[\vec{X}_5(t)] $	$ E[\vec{X}(t)] - E[\vec{X}_{10}(t)] $	$ E[\vec{X}(t)] - E[\vec{X}_{15}(t)] $
t=0	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$
t=0.25	$\begin{pmatrix} 0.00019 \\ 0.00022 \end{pmatrix}$	$\begin{pmatrix} 3.21521 \times 10^{-13} \\ 3.44169 \times 10^{-13} \end{pmatrix}$	$\begin{pmatrix} 2.978644 \times 10^{-20} \\ 6.889865 \times 10^{-20} \end{pmatrix}$
t=0.50	$\begin{pmatrix} 0.00593 \\ 0.00685 \end{pmatrix}$	$\begin{pmatrix} 6.43942 \times 10^{-10} \\ 6.89888 \times 10^{-13} \end{pmatrix}$	$\begin{pmatrix} 4.44089 \times 10^{-16} \\ 4.44089 \times 10^{-16} \end{pmatrix}$
t=0.75	$\begin{pmatrix} 0.04311 \\ 0.04963 \end{pmatrix}$	$\begin{pmatrix} 5.45273 \times 10^{-8} \\ 5.83769 \times 10^{-8} \end{pmatrix}$	$\begin{pmatrix} 3.64153 \times 10^{-14} \\ 3.86358 \times 10^{-14} \end{pmatrix}$
t=1.00	$\begin{pmatrix} 0.17415 \\ 0.19980 \end{pmatrix}$	$\begin{pmatrix} 1.26439 \times 10^{-6} \\ 1.35265 \times 10^{-6} \end{pmatrix}$	$\begin{pmatrix} 3.55005 \times 10^{-12} \\ 3.82716 \times 10^{-12} \end{pmatrix}$
t=1.25	$\begin{pmatrix} 0.51030 \\ 0.58320 \end{pmatrix}$	$\begin{pmatrix} 0.000014422 \\ 0.000015416 \end{pmatrix}$	$\begin{pmatrix} 1.24251 \times 10^{-10} \\ 1.33990 \times 10^{-10} \end{pmatrix}$
t=1.50	$\begin{pmatrix} 1.22120 \\ 1.38948 \end{pmatrix}$	$\begin{pmatrix} 0.000105024 \\ 0.000112173 \end{pmatrix}$	$\begin{pmatrix} 2.26358 \times 10^{-9} \\ 2.44218 \times 10^{-9} \end{pmatrix}$
t=1.75	$\begin{pmatrix} 2.54249 \\ 2.87833 \end{pmatrix}$	$\begin{pmatrix} 0.000561249 \\ 0.000598921 \end{pmatrix}$	$\begin{pmatrix} 2.62795 \times 10^{-8} \\ 2.83672 \times 10^{-8} \end{pmatrix}$
t=2.00	$\begin{pmatrix} 4.78222 \\ 5.38315 \end{pmatrix}$	$\begin{pmatrix} 0.002391560 \\ 0.00254969 \end{pmatrix}$	$\begin{pmatrix} 2.19401 \times 10^{-7} \\ 2.36951 \times 10^{-7} \end{pmatrix}$

Table 4.1: Comparison between exact expectation $E[\vec{X}(t)]$ given by (4.38) and its approximating values $E[\vec{X}_N(t)]$ given by (4.31), (4.39)-(4.40) for Example 4.3 for different values of the truncation order $N = 5, 10, 15$ at some points of the interval $[0, 2]$

t	$E[\vec{X}_{500}(t)]$	$\tilde{\mu}_X^{10000}(t)$	$\tilde{\mu}_X^{50000}(t)$
t=0	$\begin{pmatrix} 0.5 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 0.508549 \\ 0.99842 \end{pmatrix}$	$\begin{pmatrix} 0.502156 \\ 1.00535 \end{pmatrix}$
t=0.25	$\begin{pmatrix} 0.562088 \\ 0.820431 \end{pmatrix}$	$\begin{pmatrix} 0.566104 \\ 0.814969 \end{pmatrix}$	$\begin{pmatrix} 0.562045 \\ 0.820386 \end{pmatrix}$
t=0.50	$\begin{pmatrix} 0.55719 \\ 0.702523 \end{pmatrix}$	$\begin{pmatrix} 0.559087 \\ 0.696089 \end{pmatrix}$	$\begin{pmatrix} 0.556236 \\ 0.701119 \end{pmatrix}$
t=0.75	$\begin{pmatrix} 0.529724 \\ 0.603368 \end{pmatrix}$	$\begin{pmatrix} 0.532701 \\ 0.596578 \end{pmatrix}$	$\begin{pmatrix} 0.53055 \\ 0.601729 \end{pmatrix}$
t=1.00	$\begin{pmatrix} 0.501370 \\ 0.507439 \end{pmatrix}$	$\begin{pmatrix} 0.502799 \\ 0.500584 \end{pmatrix}$	$\begin{pmatrix} 0.501082 \\ 0.50612 \end{pmatrix}$
t=1.25	$\begin{pmatrix} 0.473928 \\ 0.40912 \end{pmatrix}$	$\begin{pmatrix} 0.475124 \\ 0.402362 \end{pmatrix}$	$\begin{pmatrix} 0.473689 \\ 0.408437 \end{pmatrix}$
t=1.50	$\begin{pmatrix} 0.449676 \\ 0.305809 \end{pmatrix}$	$\begin{pmatrix} 0.4515 \\ 0.30009 \end{pmatrix}$	$\begin{pmatrix} 0.450252 \\ 0.306797 \end{pmatrix}$
t=1.75	$\begin{pmatrix} 0.431126 \\ 0.201051 \end{pmatrix}$	$\begin{pmatrix} 0.432154 \\ 0.193539 \end{pmatrix}$	$\begin{pmatrix} 0.431034 \\ 0.20094 \end{pmatrix}$
t=2.00	$\begin{pmatrix} 0.415973 \\ 0.092112 \end{pmatrix}$	$\begin{pmatrix} 0.41669 \\ 0.0831062 \end{pmatrix}$	$\begin{pmatrix} 0.415657 \\ 0.0912403 \end{pmatrix}$

Table 4.2: Approximate values of the expectation for Example 4.4 by using the truncation series method ($E[\vec{X}_N(t)]$) with $N = 500$ and the Monte Carlo method ($\tilde{\mu}_X^m(t)$) with $m = 10000$ and $m = 50000$ simulations for different points t on the interval $[0, 2]$

obtained by the truncation of order N of the exact covariance matrix. Table 4.3 shows these approximating values at different points $t \in [0, 2]$ for $N = 500$ as well as the corresponding ones obtained by means of the Monte Carlo method with 10000 and 50000 simulations, denoted by $\text{Diag} \left(\mathbf{\Lambda}_{\vec{X}(t), \vec{X}(t)}^m \right)$, where m represents the number of simulations.

4.6 Conclusions

In this chapter we have constructed a m.s. power series solution of random coupled linear analytic differential equations incorporating uncertainty through both, the initial condition as well as the source/forcing term. For that, we have assumed that diffusion coefficient is a matrix valued analytic function and the forcing term is also a m.s. analytic in a common domain. The proposed method can be considered as an extension to the random framework of the well-know Fröbenius expansion for ordinary differential equations. By considering a truncation of the power series solution s.p., approximations of the mean and covariance matrix functions of the solution s.p. have been computed in terms of the data by means of certain recurrence relations. Even more, these approximations converge to the corresponding exact ones. Finally, we have presented several illustrative examples of potential interest in applications. The chapter must be considered as a first stage in order to consider randomness in multiplicative coefficient what constitutes the main aim of next chapter but in the scalar framework.

The main results of this chapter have been published in [9].

t	$\text{Diag} \left(\mathbf{\Lambda}_{\vec{X}_{500}(t), \vec{X}_{500}(t)} \right)$	$\text{Diag} \left(\mathbf{\Lambda}_{\vec{X}(t), \vec{X}(t)}^{10000} \right)$	$\text{Diag} \left(\mathbf{\Lambda}_{\vec{X}(t), \vec{X}(t)}^{50000} \right)$
t=0	$\begin{pmatrix} 1 \\ 0.5 \end{pmatrix}$	$\begin{pmatrix} 1.03296 \\ 0.505729 \end{pmatrix}$	$\begin{pmatrix} 1.00913 \\ 0.500876 \end{pmatrix}$
t=0.25	$\begin{pmatrix} 0.688811 \\ 0.399501 \end{pmatrix}$	$\begin{pmatrix} 0.704064 \\ 0.4032 \end{pmatrix}$	$\begin{pmatrix} 0.68839 \\ 0.399402 \end{pmatrix}$
t=0.50	$\begin{pmatrix} 0.540321 \\ 0.434211 \end{pmatrix}$	$\begin{pmatrix} 0.552287 \\ 0.440635 \end{pmatrix}$	$\begin{pmatrix} 0.540167 \\ 0.434097 \end{pmatrix}$
t=0.75	$\begin{pmatrix} 0.456362 \\ 0.520124 \end{pmatrix}$	$\begin{pmatrix} 0.463343 \\ 0.53217 \end{pmatrix}$	$\begin{pmatrix} 0.45317 \\ 0.520888 \end{pmatrix}$
t=1.00	$\begin{pmatrix} 0.397432 \\ 0.642328 \end{pmatrix}$	$\begin{pmatrix} 0.404134 \\ 0.658604 \end{pmatrix}$	$\begin{pmatrix} 0.395189 \\ 0.640848 \end{pmatrix}$
t=1.25	$\begin{pmatrix} 0.355297 \\ 0.791946 \end{pmatrix}$	$\begin{pmatrix} 0.361966 \\ 0.815316 \end{pmatrix}$	$\begin{pmatrix} 0.35387 \\ 0.789358 \end{pmatrix}$
t=1.50	$\begin{pmatrix} 0.329329 \\ 0.970143 \end{pmatrix}$	$\begin{pmatrix} 0.33081 \\ 1.00148 \end{pmatrix}$	$\begin{pmatrix} 0.323333 \\ 0.965544 \end{pmatrix}$
t=1.75	$\begin{pmatrix} 0.298298 \\ 1.183401 \end{pmatrix}$	$\begin{pmatrix} 0.307304 \\ 1.21727 \end{pmatrix}$	$\begin{pmatrix} 0.300296 \\ 1.16957 \end{pmatrix}$
t=2.00	$\begin{pmatrix} 0.292655 \\ 1.433712 \end{pmatrix}$	$\begin{pmatrix} 0.289352 \\ 1.46314 \end{pmatrix}$	$\begin{pmatrix} 0.282706 \\ 1.40184 \end{pmatrix}$

Table 4.3: Approximate values of the variance for Example 4.4 by using the truncation series method ($\text{Diag} \left(\mathbf{\Lambda}_{\vec{X}_N(t), \vec{X}_N(t)} \right)$) with $N = 500$ and the Monte Carlo method ($\text{Diag} \left(\mathbf{\Lambda}_{\vec{X}(t), \vec{X}(t)}^m \right)$) with $m = 10000$ and $m = 50000$ simulations for different points t on the interval $[0, 2]$

Chapter 5

Mean square power series solutions of first-order linear differential equations with multiplicative and additive randomness

This chapter is devoted to constructing random power series solutions of linear differential equations containing uncertainty through the diffusion coefficient, the forcing term as well as the initial condition. Under appropriate hypotheses on the data, we establish that the constructed random power series solution is mean square convergent in a certain interval about the origin whose length depends on the mean square norm of the random variable diffusion coefficient. Also, the main statistical functions of the approximating stochastic process solution generated by truncation of the exact power series solution are given. Finally, we apply the proposed technique to several illustrative examples.

5.1 Introduction

Ordinary differential equations provide a powerful deterministic description of a large class of important problems arising in different scientific areas. For an initial value problem

$$\frac{d}{dt}(x(t)) = f(x, t), \quad x(t_0) = x_0, \quad (5.1)$$

its solution represents the state of the system at time t_0 . In practice, even in the linear case, that is, when the right-hand side of (5.1) is given by $f(x, t) = ax + b$, the involved data a , b and x_0 need to be measured, therefore they will contain uncertainty. Moreover, many times the presence of randomness in the model also takes place not only due to measurement errors but the inherent complexity and difficulty of the considered physical phenomena. Then, it is proper to consider the above data as random variables, in advance denoted by A , B and X_0 , respectively, and the solution of (5.1) as a stochastic process, say $X(t)$, instead of deterministic magnitudes.

Since we have already underlined at chapter 1, there are two types of differential equation that contain stochastic influences. The first ones, the denominated stochastic differential equations introduce uncertainty through a stochastic process whose trajectories are not differentiable, namely, the white noise process or differential of the Brownian process. Since the pioneering works about the photoelectric phenomena due to A. Einstein based on Brownian motion at the beginning of the twenty century, scientists have paid great attention to these types of equations, (see for instance [8]). Their handling requires a special calculus called Itô Calculus whose rules differ considerably from the deterministic ones. In [3, 49], one provides an analytical treatment of these types of differential equation, whereas a numerical approach can be found in [38, 47]. On the other hand, in many practical situations may be more convenient to introduce chance into the model (5.1) through other stochastic processes apart from white noise (see [45], [56]). Then, one presents the so-called random differential equations. For these type of differential equation we can take advantage of the well-known deterministic methods of analysis. However additional difficulties appear especially when random differential equations contain uncertainty not only in the initial condition and the forcing term but besides in the coefficients (see [56], [29, p.66]). As a simple but clarifying

example, even in the linear case

$$\dot{X}(t) = A(t)X(t) + B(t), \quad X(t_0) = X_0, \quad (5.2)$$

(where the dot denotes the time derivative of the stochastic process $X(t)$ in the mean square sense), these difficulties arise not only in the determination of conditions under which the formal solution

$$X(t) = X_0 \exp\left(\int_{t_0}^t A(s)ds\right) + \int_{t_0}^t B(r) \exp\left(\int_r^t A(s)ds\right) dr,$$

is really a rigorous solution (see, for instance [18]) but also for computing, in a second step, its mean and variance functions. For this last goal, it deserves to be pointed out that in the case where $A(t) = A$ is a r.v., the mean of the solution stochastic process, $E[X(t)]$ does not satisfy, in general, the average differential equation (see [29, p.66])

$$\frac{d}{dt} (E[X(t)]) = E[A] E[X(t)] + E[B(t)].$$

As a consequence, the computation of this statistic information about the solution stochastic process as well another ones like variance becomes in many cases a very difficult task that only has been achieved in a few cases. For instance, in [56, p.220] one obtains the mean and the variance of $X(t)$ assuming that the initial condition is null and $A(t)$ and $B(t)$ are stationary and correlated Gaussian processes. Despite the relationship established in (5.2), being in general wrong, the so-called dishonest methods are based upon it as an alternative to get information about the first and greater moments of the solution stochastic process of (5.2) (see [46], [7] and [38, p.148]). The prior objective for random differential equations is the computation of the (exact) solution stochastic process, but as in the deterministic framework this task becomes unattainable in many cases. Otherwise it is necessary to provide reliable stochastic numerical methods for its approximation. In spite of these methods having been developed for stochastic differential equations (see [38]), apart from a few contributions (see for instance [35], [17]) these types of method are not available for random differential equations. In particular, to our knowledge the extension to random frameworks of the powerful Fröbenius method for constructing solution stochastic processes based on random power series are not considered in the literature apart from a paper [9], whose main ideas have been presented in the previous chapter 4. However, there we considered random matrix linear differential equations containing uncertainty only into the initial condition and the forcing term. Finally, it is worthwhile pointing

out that in the random framework it is of major interest to know the probability density function of the stochastic process solution for a given random differential equation. In [24], [44] one provides a general method to tackle this problem taking advantage of the powerful deterministic tools.

This chapter deals with problems of the form

$$\dot{X}(t) = AX(t) + B(t), \quad |t| < R, \quad X(0) = X_0, \quad (5.3)$$

where the diffusion coefficient A as well as the initial condition X_0 are random variables and the source term $B(t)$ is an analytic stochastic process in the interval $t \in]-R, R[$, $R > 0$. In advance, we will assume that A is a r.v. of class p with $0 \leq p \leq 1$ (see subsection 2.2.4). As it has been shown in Examples 2.7, 2.8 and 2.9 important statistical families of r.v.'s belong to this class. Note that model (5.3) extends in certain sense previous chapter since now we introduce randomness in a multiplicative way through coefficient A , although we will only consider the scalar framework. On the other hand, the study of random problems of type (5.3) provides a first stage to consider the case where diffusion coefficient A depends on time t , say $A(t)$, as well as it constitutes a basic tool to study nonlinear random models by means of linearization techniques (see, [56, chap.7-8]). Finally, we remark that with respect to applications, random linear differential equations of type (5.3) often appear in other areas like engineering, physics, epidemiology, pharmacokinetics, etc. as it can be checked in [56] and references therein.

This chapter is organized as follows. In section 5.2 we will consider the particular case of problem (5.3) where $B(t)$ is the null process and, we will construct a random power solution that allows us to define the exponential stochastic process by a random Fröbenius method. This simple scenario permits to show sharply some difficulties to be handled in section 5.3 where the random forcing term $B(t)$ will be considered. In this section, the construction of a power series solution of (5.3), as well as the proof of its convergence in the mean square sense under appropriate conditions is shown. This allows one to construct analytic-numerical solution series process by means of truncation whose main statistics properties (mean and covariance functions) are computed in section 5.4. Finally, several illustrative examples are given in section 5.5.

5.2 The case where $B(t)$ is null

This section is based on paper [63] which, for its part, constituted a starting point for article [10]. By using a Fröbenius method, we will construct a random power series solution of model

$$\dot{X}(t) = AX(t), \quad X(0) = X_0, \quad (5.4)$$

which represents a particular case of problem (5.3) with $B(t) = 0$. For that, we seek a formal solution of the form:

$$X(t) = \sum_{n \geq 0} X_n t^n, \quad (5.5)$$

where X_n are r.v.'s to be determined. By Example 2.3, the formal m.s. derivative of $X(t)$ is given by

$$\dot{X}(t) = \sum_{n \geq 0} (n+1)X_{n+1}t^n. \quad (5.6)$$

By imposing that (5.5)-(5.6) satisfy the random differential equation (5.4) it follows

$$\sum_{n \geq 0} (n+1)X_{n+1}t^n = \sum_{n \geq 0} AX_n t^n. \quad (5.7)$$

Note that in the last step we have assumed that the series (5.5) satisfies the hypotheses of Lemma 2.5, which justifies the commutation of the r.v. A and the infinite sum. Thus from (5.7) one gets

$$\sum_{n \geq 0} \{(n+1)X_{n+1} - AX_n\} t^n = 0.$$

Therefore a m.s. candidate solution s.p. of problem (5.4) can be obtained by imposing

$$X_{n+1} = \frac{1}{n+1}AX_n, \quad n \geq 0,$$

that allows to compute explicitly the coefficients X_n in terms of the data

$$X_n = \frac{1}{n!}A^n X_0, \quad n \geq 0.$$

Let us show that

$$X(t) = \sum_{n \geq 0} \frac{1}{n!}A^n X_0 t^n, \quad (5.8)$$

is not only a formal solution but a rigorous solution of problem (5.4). So first we need to justify the commutation performed in (5.7). Taking into account Lemma 2.5, it is sufficient to check that series (5.8) is m.f. convergent. For that, we will assume that A is a r.v. that satisfies condition (2.33) for the 4-norm with $0 \leq p \leq 1$ and it is independent of 4-r.v. X_0 , then by (2.34) (without loss of generality, we can assume that $n_0 = 0$) one gets

$$\begin{aligned} \sum_{n \geq 0} \frac{1}{n!} \|A^n X_0\|_4 |t|^n &= \sum_{n \geq 0} \frac{1}{n!} \|A^n\|_4 \|X_0\|_4 |t|^n \\ &\leq \sum_{n \geq 0} \frac{1}{n!} H M^{n-1} ((n-1)!)^p \|X_0\|_4 |t|^n, \end{aligned} \tag{5.9}$$

being H and M positive constants. Setting $\alpha_n^{(p)} = \frac{1}{n!} H M^{n-1} ((n-1)!)^p \|X_0\|_4 |t|^n$ and using the D'Alembert test one obtains

$$\lim_{n \rightarrow \infty} \frac{\alpha_{n+1}^{(p)}}{\alpha_n^{(p)}} = \lim_{n \rightarrow \infty} \frac{M n^p}{n+1} |t| = \begin{cases} M|t| & \text{if } p = 1, \\ 0 & \text{if } p < 1, \forall t. \end{cases} \tag{5.10}$$

From (5.9) and (5.10) we conclude that series (5.8) is absolutely m.f. convergent in the whole real line, if $p < 1$ and, in the domain $|t| < 1/M$ if $p = 1$. Observe that by Lemma 2.4, series (5.8) is also m.s. convergent into the corresponding domains above according to different cases for parameter p . Moreover, the argument above also justifies the uniform m.s. convergence of the series (5.8) as well as its formal derivative is given by (5.6) in any closed set contained into the domain of convergence. As a consequence of Theorem 2.5, the formal derivative performed in (5.6) is justified. Summarizing the following theorem has been established:

Theorem 5.1. *Consider the initial value problem given by (5.4), where A satisfies condition (2.33) for the 4-norm with $0 \leq p \leq 1$ and X_0 is a 4-r.v. independent of A . Then (5.8) is a m.s. solution of (5.4). Moreover, the domain of convergence of the random power series (5.8) depends on the value p associated to the r.v. A through the condition (2.33).*

From a practical point of view, it is interesting not only the obtention of the solution s.p. of the i.v.p. (5.4) but also providing its main statistical functions such as the mean and the variance. Note that this task cannot be performed

taking advantage of nice property (2.24) for $n = 1$ because for each t , A and $X(t)$ are not independent r.v.'s. Since $X(t)$ given by (5.8) is an infinite series we propose to proceed by truncation in order to get accurate approximations of the mean and the variance. Let us consider

$$X_N(t) = \sum_{n=0}^N \frac{1}{n!} A^n X_0 t^n, \quad (5.11)$$

then by applying the independence between A and X_0 (hence by Proposition 2.1 A^n and X_0 are also independent) one follows

$$\mathbb{E}[X_N(t)] = \mathbb{E}[X_0] \sum_{n=0}^N \frac{1}{n!} \mathbb{E}[A^n] t^n, \quad (5.12)$$

$$\begin{aligned} \mathbb{E}[(X_N(t))^2] &= \mathbb{E}[(X_0)^2] \left(\sum_{n=0}^N \frac{\mathbb{E}[A^{2n}] t^{2n}}{(n!)^2} \right. \\ &\quad \left. + 2 \sum_{n=1}^N \sum_{m=0}^{n-1} \frac{\mathbb{E}[A^{n+m}] t^{n+m}}{n!m!} \right). \end{aligned} \quad (5.13)$$

Inasmuch as we have already established the m.s. convergence of series (5.8), an important consequence of Lemma 2.3 is that (5.12) and (5.13) allow us to compute accurate approximations that converge to the exact mean and variance functions, respectively, as $N \rightarrow +\infty$.

Remark 5.1. *Since series $\sum_{n \geq 0} \frac{1}{n!} A^n t^n$ is m.s. convergent, by [18], it defines the exponential process e^{At} , consequently the m.s. solution given by (5.8) can be written as $X(t) = X_0 e^{At}$. The present approach links the corresponding one followed in [18] because permits to define the exponential s.p. as the m.s. solution of random differential model (5.4). However, it is important to point out that now, we have proposed a constructive method based on a random Fröbenius expansion whereas, in paper [18], first, a m.s. solution candidate of (5.4) was proposed and, m.s. differentiation exponential rules were required (see Theorem 8, [18]) in order to prove that it satisfies the initial value problem (5.4). Note that hypotheses in both approach are compatible.*

5.3 Solving the general case

In this section we will consider the random linear differential equation (5.3) where A is a r.v. satisfying condition (2.33) for the 4-norm with $0 \leq p \leq 1$, the initial condition X_0 is a 4-r.v. independent of A and $B(t)$ is an m.f. analytic s.p. about $t = 0$, i.e.,

$$B(t) = \sum_{n \geq 0} B_n t^n, \quad |t| < R, \quad (5.14)$$

where B_n are 4-r.v.'s independent of A for each n . Note that by Proposition 2.2 this implies that $B(t)$ is a r.v. independent of A for each t .

Let us seek a formal solution s.p. of problem (5.3) of the form

$$X(t) = \sum_{n \geq 0} X_n t^n, \quad (5.15)$$

where coefficients X_n are r.v.'s to be determined. Assuming that $X(t)$ is termwise m.s. differentiable and applying Example 2.3, it follows that

$$\dot{X}(t) = \sum_{n \geq 1} n X_n t^{n-1} = \sum_{n \geq 0} (n+1) X_{n+1} t^n. \quad (5.16)$$

Considering (5.16), we impose that $X(t)$ given by (5.15) satisfies (5.3), this yields

$$\sum_{n \geq 0} (n+1) X_{n+1} t^n = \sum_{n \geq 0} (A X_n + B_n) t^n. \quad (5.17)$$

Hence a candidate solution s.p. of model of (5.3) of the form (5.15) can be computed by imposing

$$X_{n+1} = \frac{1}{n+1} (A X_n + B_n), \quad n \geq 0, \quad (5.18)$$

with the initial condition $X_0 \in L_4$.

Now, we shall prove that the s.p. $X(t)$ defined by (5.15) and (5.18) is m.f. convergent (and then by Lemma 2.4 is also m.s. convergent) in a certain domain about the origin $t = 0$ where the random power series solution is

centered. Note that, by applying Lemma 2.5, it will justify the formal commutation between r.v. A and the infinite sum performed in (5.17). After that, we will show that the constructed series solution is termwise m.s. differentiable, by applying Proposition 2.5, which will justify the formal derivation considered in (5.16).

For the first goal, it is sufficient to show that

$$\sum_{n \geq 0} \|X_n\|_4 |t|^n < +\infty,$$

in a neighborhood of $t = 0$. It will be established by majoring this scalar series for another convergent one. Note that from recurrence (5.18) the coefficients X_n can be expressed as follows

$$X_n = \frac{1}{n!} A^n X_0 + \sum_{j=0}^{n-1} \frac{j!}{n!} A^{n-j-1} B_j, \quad n \geq 1. \quad (5.19)$$

Taking the 4-norm in previous expression and by applying independence between r.v.'s A and X_0 , as well as independence between r.v.'s A and B_n for each $n \geq 0$, one gets

$$\|X_n\|_4 \leq \frac{1}{n!} \|A^n\|_4 \hat{x}_0 + \sum_{j=0}^{n-1} \frac{j!}{n!} \|A^{n-j-1}\|_4 \|B_j\|_4 = \hat{x}_n, \quad n \geq 1, \quad (5.20)$$

where $\hat{x}_0 = \|X_0\|_4$.

In order to establish the convergence of the majorant deterministic series $\sum_{n \geq 0} \hat{x}_n |t|^n$, note that from expression (5.20) one deduces

$$\begin{aligned} \sum_{n \geq 0} \hat{x}_n |t|^n &= \|X_0\|_4 + \left(\sum_{n \geq 1} \frac{1}{n!} \|A^n\|_4 \|X_0\|_4 |t|^n \right) \\ &+ \sum_{n \geq 1} \left(\sum_{j=0}^{n-1} \frac{j!}{n!} \|A^{n-j-1}\|_4 \|B_j\|_4 \right) |t|^n. \end{aligned} \quad (5.21)$$

By applying that r.v. A satisfies condition (2.33) for the 4-norm with $0 \leq p \leq 1$ (and hence (2.34)) it is easy to check by following an analogous argument as we followed in (5.10) that first series of the right-hand side of (5.21) is convergent whenever $|t| < 1/M$ if $p = 1$, whereas if $0 \leq p < 1$, this series converges for each $t \in \mathbb{R}$.

In order to show the corresponding analysis for second series appearing in the right-hand side of (5.21), let us write it in the following way

$$\begin{aligned}
 \sum_{n \geq 1} \left(\sum_{j=0}^{n-1} \frac{j!}{n!} \|A^{n-j-1}\|_4 \|B_j\|_4 \right) |t|^n &= \left(\sum_{n \geq 1} \frac{1}{n!} \|A^{n-1}\|_4 |t|^n \right) \|B_0\|_4 \\
 &+ \sum_{n \geq 2} \left(\sum_{j=1}^{n-1} \frac{j!}{n!} \|A^{n-j-1}\|_4 \|B_j\|_4 \right) |t|^n \\
 &= \left(\sum_{n \geq 1} \frac{1}{n!} \|A^{n-1}\|_4 |t|^n \right) \|B_0\|_4 \quad (5.22) \\
 &+ \sum_{n \geq 3} \left(\sum_{j=1}^{n-2} \frac{j!}{n!} \|A^{n-j-1}\|_4 \|B_j\|_4 \right) |t|^n \\
 &+ \sum_{n \geq 2} \frac{1}{n} \|B_{n-1}\|_4 |t|^n.
 \end{aligned}$$

With respect to the convergence of the first numerical series of the right-hand side of (5.22), it is enough applying an analogous argument as in the previous one:

$$\frac{1}{n!} \|A^{n-1}\|_4 |t|^n \leq \frac{HM^{n-2}}{n!} ((n-2)!)^p |t|^n = \beta_n^{(p)},$$

then

$$\lim_{n \rightarrow \infty} \frac{\beta_{n+1}^{(p)}}{\beta_n^{(p)}} = \lim_{n \rightarrow \infty} \frac{M(n-1)^p}{n+1} |t| = \begin{cases} M|t| & \text{if } p = 1, \\ 0 & \text{if } p < 1, \forall t \in \mathbb{R}. \end{cases} \quad (5.23)$$

As a result, this series behaves as the previous one: it converges in the whole real line if $0 \leq p < 1$, whereas if $p = 1$ the convergence only takes place for

$|t| < 1/M$.

Regarding the convergence of the numerical scalar series appearing in the right-hand side of (5.22), by Cauchy inequalities (note that, we now require to apply an analogous result as Theorem 4.1 with $r = 1$ but in the m.f. framework that can be established following a similar reasoning as we considered there) one gets

$$\frac{1}{n} \|B_{n-1}\|_4 |t|^n \leq \frac{1}{n} \frac{\widehat{M}}{\rho^{n-1}} |t|^n = \gamma_n, \quad 0 < \rho < R,$$

where $R > 0$ defines the domain of convergence of the forcing term $B(t)$ and $\widehat{M} > 0$, henceforth

$$\lim_{n \rightarrow \infty} \frac{\gamma_{n+1}}{\gamma_n} = \lim_{n \rightarrow \infty} \frac{n}{n+1} \frac{|t|}{\rho} = \frac{|t|}{\rho} < 1,$$

then since ρ is an arbitrary number such that: $0 < \rho < R$, series is convergent in $|t| < R$.

Let us study the convergence of second series of the right-hand side of (5.22). Setting $M_1 = H\widehat{M}M^{-2}$ and following an analogous way as in the previous study one gets

$$\begin{aligned} \frac{j!}{n!} \|A^{n-j-1}\|_4 \|B_j\|_4 |t|^n &\leq \frac{j!}{n!} HM^{n-j-2} ((n-j-2)!)^p \frac{\widehat{M}}{\rho^j} |t|^n \\ &= M_1 M^{n-j} \left(\frac{j!}{n!}\right)^{1-p} \left(\frac{(n-2)!}{n!}\right)^p \\ &\times \left(\frac{(n-j-2)!j!}{(n-2)!}\right)^p \frac{|t|^n}{\rho^j} \tag{5.24} \\ &= M_1 \frac{M^{n-j}}{n^{1-p}(n-1)^{1-p}} \left(\frac{j!}{(n-2)!}\right)^{1-p} \\ &\times \frac{1}{n^p(n-1)^p} \frac{1}{\binom{n-2}{j}^p} \frac{|t|^n}{\rho^j}, \end{aligned}$$

where considering the following inequalities

$$0 < \frac{j!}{(n-2)!} \leq 1, \quad \binom{n-2}{j} \geq 1, \quad j = 1, \dots, n-2, \quad n \geq 3,$$

for each $0 \leq p \leq 1$, expression (5.24) can be bounded as

$$\begin{aligned} \frac{j!}{n!} \|A^{n-j-1}\|_4 \|B_j\|_4 |t|^n &\leq \frac{M_1 M^{n-j}}{n^{1-p} (n-1)^{1-p}} \frac{1}{n^p (n-1)^p} \frac{|t|^n}{\rho^j} \\ &= \frac{M_1 M^{n-j}}{n(n-1)} \frac{|t|^n}{\rho^j}. \end{aligned} \tag{5.25}$$

Note that given $\rho : 0 < \rho < R$, the positive constant M in (2.34) can be chosen such that $M\rho > 1$, therefore from (5.25) one obtains the majorant series

$$\begin{aligned} \sum_{n \geq 3} \left(\sum_{j=1}^{n-2} \frac{j!}{n!} \|A^{n-j-1}\|_4 \|B_j\|_4 \right) |t|^n &\leq M_1 \sum_{n \geq 3} \frac{1}{n(n-1)} \left(\sum_{j=1}^{n-2} (M\rho)^{-j} \right) (M|t|)^n \\ &\leq M_1 \sum_{n \geq 3} \frac{1}{n(n-1)} \left(\sum_{j=1}^{\infty} (M\rho)^{-j} \right) (M|t|)^n \\ &= \frac{M_1}{M\rho - 1} \sum_{n \geq 3} \frac{(M|t|)^n}{n(n-1)}. \end{aligned}$$

Note that under hypothesis $M\rho > 1$ one assures that the above involved geometric series is convergent. On the other hand, by the D'Alembert test, the last series is convergent whenever $M|t| < 1$ or equivalently for $t : |t| < 1/M$. From a practical view point the domain of convergence of series appearing in the right-hand side of (5.22) only can be assured in a neighborhood of the origin, $|t| < \min(1/M, R)$.

Previous exposition together with Examples 2.2 and 2.3 also justifies that series (5.15) whose coefficients are given by (5.18) is termwise m.s. differentiable. Summarizing, we have established the following

Theorem 5.2. *The problem (5.3), where A is a r.v. satisfying condition (2.33) for the 4-norm with $0 \leq p \leq 1$, the source term $B(t)$ is a m.f. analytic s.p. of the form (5.14), X_0 is 4-r.v. and the r.v. A is independent of X_0 and B_n for each $n \geq 0$, admits a random power series solution of the form (5.15), (5.18) which is m.s. convergent in a neighborhood of the origin.*

5.4 Computing the main statistical functions of the solution

Note that the solution s.p. of problem (5.3) has been computed by means an infinite power series given by (5.15), (5.18). In order to provide accurate approximations of the main statistical functions of the solution s.p. $X(t)$, we will truncate that series for a previously fixed initial condition, say X_0 , that is,

$$X_N(t) = X_0 + \sum_{n=1}^N X_n t^n. \quad (5.26)$$

As far as statistical behavior is concerned, we will be mainly interested in computing its mean and covariance functions.

For the first one, we will assume that $E[X_0]$, $E[B_n]$, $n = 0, 1, 2, \dots$ and $E[A^m]$, $m = 1, 2, \dots$, are data. Then by taking into account the linearity of the expectation operator, from (5.26) one gets

$$E[X_N(t)] = E[X_0] + \sum_{n=1}^N E[X_n] t^n, \quad (5.27)$$

where by applying independence between A and X_0 and, A and B_n for each $n \geq 0$ (which from a practical viewpoint is realistic), from (5.19) and Proposition 2.1, the coefficients $E[X_n]$ satisfy the following explicit relationship in terms of the data:

$$E[X_n] = \frac{1}{n!} E[A^n] E[X_0] + \sum_{j=0}^{n-1} \frac{j!}{n!} E[A^{n-j-1}] E[B_j], \quad n \geq 1. \quad (5.28)$$

Now we are interested in computing the covariance function of the approxi-

mating s.p. given by (5.26). From (2.11) one gets

$$\begin{aligned}
 \text{Cov} [X_N(t), X_N(s)] &= \text{Cov} \left[\sum_{n=0}^N X_n t^n, \sum_{m=0}^N X_m s^m \right] \\
 &= \sum_{n=0}^N \sum_{m=0}^N \text{Cov} [X_n, X_m] t^n s^m \\
 &= \text{Var} [X_0] + \sum_{n=1}^N \sum_{m=1}^N \text{Cov} [X_n, X_m] t^n s^m \\
 &\quad + \sum_{n=1}^N \text{Cov} [X_n, X_0] t^n + \sum_{m=1}^N \text{Cov} [X_0, X_m] s^m,
 \end{aligned} \tag{5.29}$$

where from (5.19) coefficients $\text{Cov} [X_0, X_m]$, $\text{Cov} [X_n, X_0]$ and $\text{Cov} [X_n, X_m]$, $n, m \geq 1$ can be expressed as follows

$$\begin{aligned}
 \text{Cov} [X_0, X_m] &= \text{Cov} \left[X_0, \frac{1}{m!} A^m X_0 + \sum_{k=0}^{m-1} \frac{k!}{m!} A^{m-k-1} B_k \right] \\
 &= \frac{1}{m!} \{ \text{Cov} [X_0, A^m X_0] \\
 &\quad + \sum_{k=0}^{m-1} k! \text{Cov} [X_0, A^{m-k-1} B_k] \}, \quad m \geq 1,
 \end{aligned} \tag{5.30}$$

$$\begin{aligned}
 \text{Cov} [X_n, X_0] &= \text{Cov} \left[\frac{1}{n!} A^n X_0 + \sum_{j=0}^{n-1} \frac{j!}{n!} A^{n-j-1} B_j, X_0 \right] \\
 &= \frac{1}{n!} \{ \text{Cov} [A^n X_0, X_0] \\
 &\quad + \sum_{j=0}^{n-1} j! \text{Cov} [A^{n-j-1} B_j, X_0] \}, \quad n \geq 1,
 \end{aligned} \tag{5.31}$$

$$\begin{aligned}
 \text{Cov} [X_n, X_m] &= \text{Cov} \left[\frac{1}{n!} A^n X_0 + \sum_{j=0}^{n-1} \frac{j!}{n!} A^{n-j-1} B_j, \right. \\
 &\quad \left. \frac{1}{m!} A^m X_0 + \sum_{k=0}^{m-1} \frac{k!}{m!} A^{m-k-1} B_k \right] \\
 &= \frac{1}{n! m!} \{ \text{Cov} [A^n X_0, A^m X_0] \\
 &\quad + \sum_{k=0}^{m-1} k! \text{Cov} [A^n X_0, A^{m-k-1} B_k] \\
 &\quad + \sum_{j=0}^{n-1} j! \text{Cov} [A^{n-j-1} B_j, A^m X_0] \\
 &\quad + \sum_{j=0}^{n-1} \sum_{k=0}^{m-1} j! k! \text{Cov} [A^{n-j-1} B_j, A^{m-k-1} B_k] \}, \quad n, m \geq 1.
 \end{aligned} \tag{5.32}$$

Although expressions above allow us to compute the covariance of the approximating s.p. whenever the involved joint statistical behavior of the data is known, under independence hypothesis they can be simplified.

Note that from a practical viewpoint, independence assumption between r.v.'s A , X_0 and B_n , for each $n \geq 0$, is realistic. So, in the following, we will assume this independence, henceforth the covariances appearing in (5.30)-(5.32) admit the following representation in terms of the statistical moments with respect to the origin and variance of each involved r.v.

First, let us compute the values of the following covariances: $\text{Cov}[X_0, A^m X_0]$, $\text{Cov}[A^n X_0, X_0]$ and $\text{Cov}[A^n X_0, A^m X_0]$, respectively,

$$\text{Cov} [X_0, A^m X_0] = \text{Var} [X_0] \text{E} [A^m], \quad m \geq 1, \tag{5.33}$$

$$\text{Cov} [A^n X_0, X_0] = \text{Var} [X_0] \text{E} [A^n], \quad n \geq 1, \tag{5.34}$$

and

$$\begin{aligned} \text{Cov} [A^n X_0, A^m X_0] &= \text{E} [A^{n+m}] \text{E} [(X_0)^2] \\ &- \text{E} [A^n] \text{E} [A^m] (\text{E} [X_0])^2, \quad n, m \geq 1. \end{aligned} \quad (5.35)$$

With respect to terms of the form $\text{Cov} [A^n X_0, A^{m-k-1} B_k]$, $n \geq 0$, one gets

$$\text{Cov} [X_0, A^{m-k-1} B_k] = 0, \quad m \geq 1, \quad (5.36)$$

$$\begin{aligned} \text{Cov} [A^n X_0, A^{m-k-1} B_k] &= (\text{E} [A^{n+m-k-1}] - \text{E} [A^n] \text{E} [A^{m-k-1}]) \\ &\times \text{E} [X_0] \text{E} [B_k], \quad n, m \geq 1. \end{aligned} \quad (5.37)$$

With respect to terms of the form $\text{Cov} [A^{n-j-1} B_j, A^m X_0]$, $m \geq 0$ one gets

$$\text{Cov} [A^{n-j-1} B_j, X_0] = 0, \quad n \geq 1, \quad (5.38)$$

$$\begin{aligned} \text{Cov} [A^{n-j-1} B_j, A^m X_0] &= (\text{E} [A^{n+m-j-1}] - \text{E} [A^{n-j-1}] \text{E} [A^m]) \\ &\times \text{E} [X_0] \text{E} [B_j], \quad n, m \geq 1. \end{aligned} \quad (5.39)$$

and finally

$$\begin{aligned} \text{Cov} [A^{n-j-1} B_j, A^{m-k-1} B_k] &= \text{E} [A^{n+m-j-k-2}] \text{E} [B_j B_k] \\ &- \text{E} [A^{n-j-1}] \text{E} [A^{m-k-1}] \\ &\times \text{E} [B_j] \text{E} [B_k], \quad n, m \geq 1. \end{aligned} \quad (5.40)$$

In particular, by considering (5.30)-(5.40), following computations show that

the variance function of (5.26) is given by

$$\begin{aligned}
 \text{Var} [X_N(t)] &= \text{Cov} [X_N(t), X_N(t)] \\
 &= \text{Var} [X_0] \left(1 + 2 \sum_{n=1}^N \frac{1}{n!} \text{E}[A^n] t^n \right) \\
 &+ \sum_{n=1}^N \sum_{m=1}^N \frac{1}{n! m!} \left\{ \text{E}[A^{n+m}] \text{E}[(X_0)^2] - \text{E}[A^n] \text{E}[A^m] (\text{E}[X_0])^2 \right. \\
 &+ \sum_{k=0}^{m-1} k! \left(\text{E}[A^{n+m-k-1}] - \text{E}[A^n] \text{E}[A^{m-k-1}] \right) \text{E}[X_0] \text{E}[B_k] \quad (5.41) \\
 &+ \sum_{j=0}^{n-1} j! \left(\text{E}[A^{n+m-j-1}] - \text{E}[A^{n-j-1}] \text{E}[A^m] \right) \text{E}[X_0] \text{E}[B_j] \\
 &+ \sum_{j=0}^{n-1} \sum_{k=0}^{m-1} j! k! \left(\text{E}[A^{n+m-j-k-2}] \text{E}[B_j B_k] \right. \\
 &\left. - \text{E}[A^{n-j-1}] \text{E}[A^{m-k-1}] \text{E}[B_j] \text{E}[B_k] \right) \left. \right\} t^{n+m}.
 \end{aligned}$$

5.5 Examples

There are many physical, biological or engineering problems which can lead to models of the form (5.4) or more generally (5.1). For instance, in population modeling, a basic but important population growth model is based upon the rate of change of population of certain species at time t , given by $\frac{dP(t)}{dt}$, is proportional to the population at that instant, given by P , that is, $\frac{dP(t)}{dt} = kP(t)$, where the parameter k is regarded as the difference between the birth and death population at time t . The birthrate and deathrate are influenced by many circumstances like climatology, water, epidemics, living space restrictions and availability of food. The factors above make more realistic to model the parameter k as a r.v. instead of a deterministic value, even

more the initial population P_0 should be also considered a r.v. because of in practice, it is not known with certainty, see [45]. Another physical situation that lead to model (5.4) can be found in thermodynamics. Indeed, in many physical situations the rate of change of temperature T of a body can be assumed proportional to the difference between the temperature of the body and the temperature α of the surrounding medium. If $\alpha = 0$, then $\frac{dT(t)}{dt} = kT(t)$, where the parameter k as well as the initial temperature T_0 should again be considered r.v.'s because of measure device errors. Several similar examples can be found in another related physical situations such that evaporation and diffusion problems. Now, in the next examples we do not focus on any specific model, although taking into account previous comments they could be applied to population or thermodynamics models, for instance.

Example 5.1. *Let us consider model (5.4) where A is a Uniform r.v. on the interval $[0, 1]$, i.e., $A \sim \text{Un}([0, 1])$ and the initial condition X_0 is a 4-r.v. such that $E[X_0] = 1$, $E[(X_0)^2] = 2$ and so $\text{Var}[X_0] = 1$. Since A is just a Beta r.v. of parameters $\alpha = 1$ and $\beta = 1$, by Example 2.7 it satisfies condition (2.33) for the 4-norm for $p = 0$. Note that both, A and X_0 are 4-r.v.'s. On the other hand, we know (see Table 2.2)*

$$E[A^n] = \frac{1}{n+1}. \quad (5.42)$$

In the following, we shall assume that X_0 and A are independent r.v.'s. From (5.12)-(5.13) and (2.7) (or equivalently from (5.41)), respectively, it is easy to compute the expectation and the variance of the approximating random power series (5.11) (see [18])

$$E[X_N(t)] = 1 + \sum_{n=1}^N \frac{t^n}{(n+1)!}, \quad (5.43)$$

$$\begin{aligned} \text{Var}[X_N(t)] &= 1 + 2 \sum_{n=1}^N \frac{t^n}{(n+1)!} \\ &+ \sum_{n=1}^N \sum_{m=1}^N \frac{1}{n!m!} \left(\frac{2}{n+m+1} - \frac{1}{(n+1)(m+1)} \right) t^{n+m}, \end{aligned} \quad (5.44)$$

whereas the exact expectation and variance functions are given by

$$E[X(t)] = \begin{cases} \frac{1}{t} (e^t - 1) & \text{if } t \neq 0, \\ 1 & \text{if } t = 0, \end{cases} \quad (5.45)$$

and

$$\text{Var}[X(t)] = \begin{cases} \frac{1}{t} (2e^t \sinh(t)) - \frac{1}{t^2} (e^t - 1)^2 & \text{if } t \neq 0, \\ 1 & \text{if } t = 0, \end{cases} \quad (5.46)$$

respectively. Note that, since the hypothesis of Theorem 5.1 holds, the random power series solution given by (5.8) is m.s. convergent for each t . As a consequence of Lemma 2.3, the approximate expectation as well as the variance of order N given by (5.43) and (5.44), converge as $N \rightarrow \infty$ to (5.45) and (5.46), respectively, for all t . This feature is shown in Figures 5.1 and 5.2, when we take $0 \leq t \leq 4$. In order to make Figures 5.1 and 5.2 clearer, we have considered truncations of order $N = 2, 4, 6$ because the associated truncation error on the interval $[0, 4]$ for $N = 20$ is at least of orders 10^{-9} and 10^{-7} , respectively.

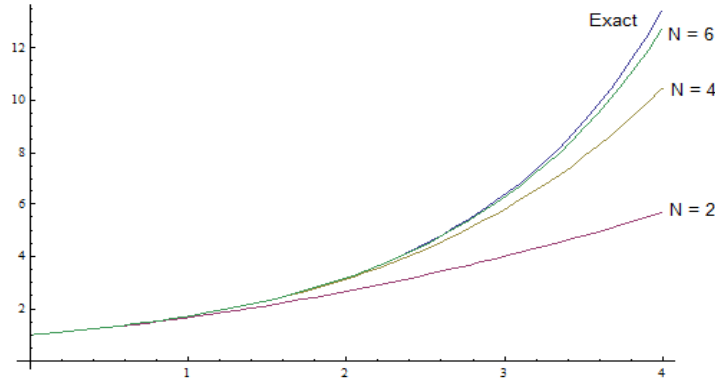


Figure 5.1: Exact and approximate mean for Example 5.1

Example 5.2. Let us consider a model of the form (5.4), where A is a standard Gaussian r.v. $A \sim N(0; 1)$ and X_0 is a 4-r.v. independent of A such that $E[X_0] = 1$ and $E[(X_0)^2] = 2$. Note that by Example (2.9), r.v. A satisfies

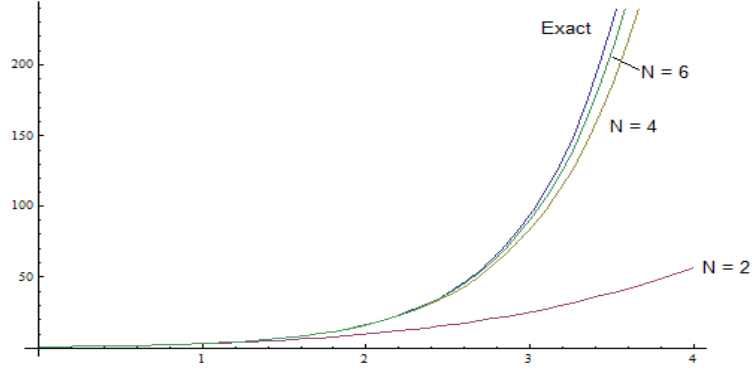


Figure 5.2: Exact and approximate variance for Example 5.1

condition (2.33) for the 4-norm for $p = 1/2$, so by Theorem 5.1 series (5.8) is m.s. convergent for all $t \in \mathbb{R}$. As a consequence of Lemma 2.2 we can get accurate approximations of the exact mean and variance from (5.12) and (5.13), respectively, by taking N large enough. Since

$$\mathbb{E} [A^{2k}] = \frac{(2k)!}{2^k k!},$$

by (5.12) and (5.13) it follows

$$\mathbb{E} [X_N(t)] = \sum_{n=0}^{\lfloor \frac{N}{2} \rfloor} \frac{t^{2n}}{2^n n!}, \quad (5.47)$$

where $\lfloor \cdot \rfloor$ denotes the integer part function, and

$$\begin{aligned} \mathbb{E} [(X_N)^2] &= 4 \sum_{r=1}^{\lfloor \frac{N}{2} \rfloor} \sum_{s=0}^{r-1} \frac{(2(r+s+1))! t^{2(r+s+1)}}{2^{r+s+1} (2s+1)! (2r+1)! (r+s+1)!} \\ &+ 4 \sum_{r=1}^{\lfloor \frac{N}{2} \rfloor} \sum_{s=0}^{r-1} \frac{(2(r+s))! t^{2(r+s)}}{2^{r+s} (2s)! (2r)! (r+s)!} \\ &+ 2 \sum_{n=0}^N \frac{(2n)! t^{2n}}{(n!)^3 2^n} \quad \text{if } N \text{ is odd,} \end{aligned} \quad (5.48)$$

and

$$\begin{aligned}
 E[(X_N)^2] &= 4 \sum_{r=1}^{\lfloor \frac{N-1}{2} \rfloor} \sum_{s=0}^{r-1} \frac{(2(r+s+1))! t^{2(r+s+1)}}{2^{r+s+1} (2s+1)! (2r+1)! (r+s+1)!} \\
 &+ 4 \sum_{r=1}^{\lfloor \frac{N}{2} \rfloor} \sum_{s=0}^{r-1} \frac{(2(r+s))! t^{2(r+s)}}{2^{r+s} (2s)! (2r)! (r+s)!} \\
 &+ 2 \sum_{n=0}^N \frac{(2n)! t^{2n}}{(n!)^3 2^n} \quad \text{if } N \text{ is even.}
 \end{aligned} \tag{5.49}$$

Table 5.1 shows the mean of the solution s.p. $X(t)$ given by (5.47) for different values of the truncation order $N = 10, 15$ and some select points on the interval $t \in [0, 2]$. The columns $\mu_X^{50000}(t)$ and $\mu_X^{100000}(t)$ contain the correspondent average approximations by using a Monte Carlo method for 50000 and 100000 simulations of the r.v.'s data A and X_0 , respectively. Since in this case from Remark 5.1, the exact solution s.p. is available, we compare the results with the theoretical ones, that can be obtained from $E[X(t)] = E[X_0]E[e^{At}] = e^{t^2/2}$. Note that, the convergence $E[X_N(t)] \xrightarrow{N \rightarrow \infty} E[X(t)]$ takes place in the whole real line, that is for each t , whereas for N fixed the approximations $E[X_N(t)]$ improves when t is near to the origin $t = 0$, where the random power series solution is centered. We observe that Monte Carlo method improves its approximations when the number of simulations increase from 50000 to 100000, as it is expected, but from a computational point of view, we realize that truncation method can be considered, in this particular case, better because it provides more accurate results by using very low values of the order of truncation N , and moreover it saves computational time.

Table 5.2 shows the correspondent results for the variance of $X(t)$ computed from (5.48)-(5.49) as well as by means of a Monte Carlo procedure. Here $\sigma_X^{2;m}(t)$ denotes the approximate variance obtained by applying Monte Carlo method for m simulations. Note that the exact variance is given by $\text{Var}[X(t)] = e^{t^2} (2e^{t^2} - 1) > 0$. Like in the case of computations for the mean, similar comments can be made for comparing both methods.

Example 5.3. Let us consider again a model of the form (5.4) where A is a Gamma r.v. of parameters $r > 0$, $a > 0$, i.e., $A \sim \text{Ga}(r; a)$, and the initial condition X_0 is a 4-r.v. independent of A such that $E[X_0] = 1$ and

t	$E[X_{10}(t)]$	$E[X_{15}(t)]$	$\mu_X^{50000}(t)$	$\mu_X^{100000}(t)$	$E[X(t)]$
0.00	1.0000	1.0000	1.01225	1.00299	1.0000
0.20	1.0202	1.0202	1.01759	1.02309	1.0202
0.40	1.08329	1.08329	1.07993	1.08598	1.08329
0.60	1.19722	1.19722	1.19302	1.19954	1.19722
0.80	1.37713	1.37713	1.37275	1.3786	1.37713
1.00	1.64870	1.64872	1.64661	1.64805	1.64872
1.20	2.05422	2.05443	2.0612	2.04814	2.05443
1.40	2.66303	2.66443	2.69619	2.64347	2.66446
1.60	3.58921	3.59643	3.6907	3.5375	3.59664
1.80	5.02075	5.05166	5.29397	4.89525	5.05309
2.00	7.26667	7.38095	7.96505	6.97603	7.38906

Table 5.1: Approximated and exact expectation for Example 5.2 with $A \sim N(0; 1)$

t	$\text{Var}[X_{10}(t)]$	$\text{Var}[X_{15}(t)]$	$\sigma_X^{2;50000}(t)$	$\sigma_X^{2;100000}(t)$	$\text{Var}[X(t)]$
0.00	1.0000	1.0000	0.995243	1.00269	1.0000
0.20	1.12576	1.12576	1.12245	1.12674	1.12576
0.40	1.58074	1.58074	1.59068	1.57785	1.58074
0.60	2.67553	2.67554	2.70583	2.66201	2.67554
0.80	5.29649	5.29680	5.51251	5.24664	5.29688
1.00	12.0513	12.0598	13.2413	11.8691	12.0598
1.20	31.2506	31.4068	37.305	30.6881	31.4079
1.40	91.5262	93.6678	121.849	91.4166	93.7016
1.60	297.066	320.892	451.471	319.45	321.735
1.80	1035.76	1261.56	1846.27	1332.75	1278.41
2.00	3741.12	5622.86	8118.9	6629.60	5907.32

Table 5.2: Approximated and exact variance for Example 5.2 with $A \sim N(0;1)$

$E[(X_0)^2] = 2$. By Example 2.8, A satisfies condition (2.33) for the 4-norm for $p = 1$. In this case, from (2.36) the value M in (2.34) can be chosen greater than $4/a$, hence by (5.10) we can assure that the random power series solution given by (5.8) m.s. converges for $|t| < a/4$. Observe that this domain increases as parameter a does, independently of parameter r . Consequently from (5.12)-(5.13), Lemma 2.3 and the moments with respect to the origin of a Gamma r.v. (see Table 2.2), the moments

$$E[X_N(t)] = \sum_{n=0}^N \frac{(r+n-1)!t^n}{n!(r-1)!a^n}, \quad (5.50)$$

$$E[(X_N(t))^2] = 2 \sum_{n=0}^N \frac{(r+2n-1)!t^{2n}}{a^{2n}(n!)^2(r-1)!} + 4 \sum_{n=1}^N \sum_{m=0}^{n-1} \frac{(r+n+m-1)!t^{n+m}}{n!m!(r-1)!a^{n+m}}, \quad (5.51)$$

also converge in the same domain as $N \rightarrow \infty$. It is worthwhile pointing out that in practice, the domain of convergence for the mean approximation given by (5.50) can be greater than $|t| < a/4$. This fact is expected from the own definition of the m.s. convergence. In this case, it is straightforward to check by D'Alembert test that series given by (5.50) converges for $|t| < a$. Table 5.3 illustrates this situation for the case where $A \sim \text{Ga}(r = 1; a = 2)$. In this case, the convergence for average approximation given by (5.50) one presents for $|t| < 2$.

Whereas Table 5.4 shows that variance approximation computed from (2.7), (5.51) and (5.50) (or equivalently from (5.41)) converges for $|t| < 1$. The computations of the exact mean $E[X(t)]$ and variance $\text{Var}[X(t)]$ in Tables 5.3 and 5.4 have been performed using formulae (2.8) and (2.7). In both tables, we underline the better results that again our approach provide with respect to the ones provided by a Monte Carlo procedure. Finally observe that as a consequence of this example, the exponential process e^{At} with $A \sim \text{Ga}(r; a)$ is only defined for $|t| < a/4$, hence the domain of definition for e^{At} in this case increases when parameter a does.

Example 5.4. Let us consider a random linear model based on (5.3), where the diffusion coefficient A is a Uniform r.v. on the interval $[0, 1]$, the forcing term $B(t)$, given by (5.14), is such that: $B_n = B/n!$, where B is a r.v. with zero mean, $E[B] = 0$, a constant variance $E[B^2] = \gamma \in (0, 1)$ and such that $E[B^4] = \nu$. The initial condition X_0 is a 4-r.v. such that $E[X_0] = 1/2$ and

t	$E[X_{30}(t)]$	$E[X_{40}(t)]$	$\mu_X^{50000}(t)$	$\mu_X^{100000}(t)$	$E[X(t)]$
0.00	1.0000	1.0000	0.994639	0.999233	1.0000
0.25	1.14286	1.14286	1.13788	1.14185	1.14286
0.50	1.3333	1.3333	1.3289	1.3323	1.3333
0.75	1.6000	1.6000	1.59522	1.59912	1.6000
1.00	2.0000	2.0000	1.98796	1.99703	2.0000
1.25	2.66667	2.66667	2.60862	2.64002	2.66667
1.50	3.99946	3.99997	3.67272	3.78619	4.000
1.75	7.87255	7.96647	5.66778	6.0675	8.000
2.00	31.000	41.000	9.76052	11.1415	divergent
2.25	300.19	992.791	18.8786	23.610	divergent
2.50	4034.97	37611.8	40.6549	56.843	divergent
2.75	51680.9	1.24847×10^6	95.5911	151.087	divergent
3.00	575251.0	3.3172×10^7	240.002	430.89	divergent

Table 5.3: Approximate and exact expectation for Example 5.3 with $A \sim \text{Ga}(r = 1; a = 2)$

t	$\text{Var}[X_{30}(t)]$	$\text{Var}[X_{40}(t)]$	$\sigma_X^{2;50000}(t)$	$\sigma_X^{2;100000}(t)$	$\text{Var}[X(t)]$
0.00	1.0000	1.0000	0.999724	0.9996	1.0000
0.25	1.3605	1.3605	1.36553	1.36139	1.3605
0.50	2.2222	2.2222	2.22092	2.22561	2.2222
0.75	5.4400	5.4400	4.85051	5.07959	5.4400
1.00	107.485	145.594	15.9162	18.9778	divergent
1.50	6.4017×10^{10}	1.8704×10^{14}	508.132	927.164	divergent
3.00	1.9486×10^{28}	5.8964×10^{37}	2.4406×10^8	1.5241×10^9	divergent

Table 5.4: Approximate and exact variance for Example 5.3 with $A \sim \text{Ga}(r = 1; a = 2)$

$E[(X_0)^2] = 1/2$. We will assume that A, X_0 and $B_n, n \geq 0$ are independent r.v.'s. As a consequence, $E[B(t)] = 0$ and applying the corresponding m.f. analyticity characterization (that can be stated in analogous way that Theorem 2.4), $B(t)$ is m.f. analytic in the whole real line since

$$\begin{aligned} \Gamma_B(t, t, t, t) &= E[(B(t))^4] = \sum_{n=0}^{\infty} \sum_{j=0}^n \sum_{l=0}^{n-j} \sum_{k=0}^{n-j-l} E[B_{n-j-l-k} B_k B_l B_j] t^n \\ &= \sum_{n=0}^{\infty} \sum_{j=0}^n \sum_{l=0}^{n-j} \sum_{k=0}^{n-j-l} \frac{E[B^4]}{(n-j-l)! k! l! j!} t^n = \nu \sum_{n=0}^{\infty} \sum_{j=0}^n \sum_{l=0}^{n-j} \sum_{k=0}^{n-j-l} \frac{t^n}{(n-j-l)! k! l! j!}. \end{aligned}$$

From (5.27)-(5.28) and (5.42), one gets

$$E[X_N(t)] = \frac{1}{2} \sum_{n=0}^N \frac{t^n}{(n+1)!}. \tag{5.52}$$

Table 5.5 collects the values of the expectation of the approximating process given by (5.52) for different orders of truncation, $N = 5, 10, 20$ at points $t_i = i/4$ with $i = 0, 1, 2, \dots, 12$ belong to the interval $[0, 3]$.

Table 5.5 shows that these values differ with respect to the corresponding ones obtained by the application of the dishonest method (see [46], [38]). Note

that taking the expectation operator in problem (5.3) and applying property (2.24) for $n = 1$, the mean function of the solution s.p. by the dishonest method is given by $y_d(t) = (1/2) \exp((1/2) t)$ which results in solving the initial value problem

$$\frac{d}{dt} (y_d(t)) = \frac{1}{2} y_d(t), \quad y_d(0) = \frac{1}{2}.$$

t	$E[X_5(t)]$	$E[X_{10}(t)]$	$E[X_{20}(t)]$	$y_d(t)$
$t = 0.00$	0.50000	0.50000	0.50000	0.50000
$t = 0.25$	0.56805	0.56805	0.56805	0.56657
$t = 0.50$	0.64872	0.64872	0.64872	0.64201
$t = 0.75$	0.74465	0.74467	0.74467	0.72750
$t = 1.00$	0.85903	0.85914	0.85914	0.82436
$t = 1.25$	0.99569	0.99614	0.99614	0.9341
$t = 1.50$	1.15918	1.16056	1.16056	1.0585
$t = 1.75$	1.35484	1.35846	1.35846	1.1994
$t = 2.00$	1.58889	1.59726	1.59726	1.35914
$t = 2.25$	1.86851	1.88615	1.88616	1.54011
$t = 2.50$	2.20193	2.23647	2.23651	1.74517
$t = 2.75$	2.59849	2.66221	2.66234	1.97754
$t = 3.00$	3.06875	3.18068	3.18092	2.24084

Table 5.5: Comparison between approximate expectations obtained from the proposed method and dishonest one for Example 5.4

For the computation of the variance of $X_N(t)$, we apply (5.41) taking into account that in this case $\text{Var}[X_0] = 1/4$, $E[A^n] = 1/(n + 1)$ and $E[B_j B_k] =$

$\frac{\gamma}{j!k!}$ then substituting these values into (5.41) one obtains

$$\begin{aligned} \text{Var} [X_N(t)] &= \frac{1}{4} \left(1 + 2 \sum_{n=1}^N \frac{t^n}{(n+1)!} \right) \\ &+ \sum_{n=1}^N \sum_{m=1}^N \frac{1}{n!m!} \left\{ \frac{1}{2(n+m+1)} - \frac{1}{4(n+1)(m+1)} \right. \\ &\left. + \gamma \sum_{j=0}^{n-1} \sum_{k=0}^{m-1} \frac{1}{n+m-j-k-1} \right\} t^{n+m}. \end{aligned}$$

As far as the variance is concerned, we consider several cases with respect to the value of the correlation parameter γ associated to the source term. Table 5.6 collects the variance for different orders of truncation ($N = 3, 5, 10$) on the interval $[0, 1]$ when $\gamma = 0.5$. It is worthwhile pointing out that results do not change for N greater than 10.

Table 5.7 shows the variance for different fixed values of time variable, $t = 0.25, 0.5, 0.75$, for several nonnegative values of correlation parameter, $\gamma = 0.1, \dots, 0.9, 1$ when the order of truncation is $N = 20$.

5.6 Conclusions

This chapter consider the construction of random power series solution of first-order linear random differential equation (5.3) where uncertainty is incorporated through, the multiplicative coefficient A , the forcing term $B(t)$ and the initial conditions X_0 . By assuming that A is a r.v. of certain class (for which, relevant r.v.'s, like, Beta, Gaussian or Gamma are included), $B(t)$ is an m.s. analytical s.p. (with coefficients are denoted by B_n) and the independence between A , B_n for each $n \geq 0$ and X_0 , we have established the m.s. convergence of the random power series on a neighborhood of the origin. In the special case where $B(t)$ is the null s.p., the study allows us to define the exponential s.p. $\exp(At)$. By means of the truncation of the power series solution s.p.,

$t \setminus \gamma = 0.5$	$\text{Var}[X_3(t)]$	$\text{Var}[X_5(t)]$	$\text{Var}[X_{10}(t)]$
$t = 0.00$	0.25000	0.25000	0.25000
$t = 0.10$	0.282796	0.282797	0.282797
$t = 0.20$	0.335439	0.335482	0.335482
$t = 0.30$	0.415761	0.416098	0.416099
$t = 0.40$	0.533817	0.535332	0.535341
$t = 0.50$	0.702281	0.707283	0.707329
$t = 0.60$	0.93689	0.950435	0.950618
$t = 0.70$	1.25691	1.28888	1.28947
$t = 0.80$	1.68567	1.75381	1.75547
$t = 0.90$	2.25109	2.38544	2.38961
$t = 1.00$	2.98627	3.23528	3.24489

Table 5.6: Approximate variance for $\gamma = 0.5$ at $t = 0, 0.1, 0.2, \dots, 0.9, 1$ for different values of the truncation order $N = 3, 5, 10$ in Example 5.4

approximations of the mean, variance and covariance functions of the solution s.p. are provided in terms of the data. These approximations converge to the corresponding exact functions. Finally, we have presented several illustrative examples where the mean and the variance are computed from our approach. These values are compared with respect to the approximations obtained by Monte Carlo and dishonest methods. Examples show that the approach here presented provide better approximations.

The main results of this chapter have been published in [10].

$\gamma \setminus N = 20$	$\text{Var}[X_{20}(0.25)]$	$\text{Var}[X_{20}(0.5)]$	$\text{Var}[X_{20}(0.75)]$
$\gamma = 0.1$	0.335172	0.492107	0.785679
$\gamma = 0.2$	0.344304	0.545913	0.965324
$\gamma = 0.3$	0.353436	0.599718	1.14497
$\gamma = 0.4$	0.362569	0.653524	1.32461
$\gamma = 0.5$	0.371701	0.707329	1.50426
$\gamma = 0.6$	0.380833	0.761135	1.6839
$\gamma = 0.7$	0.389965	0.81494	1.86355
$\gamma = 0.8$	0.399098	0.868746	2.04319
$\gamma = 0.9$	0.40823	0.922551	2.22284
$\gamma = 1.0$	0.417362	0.976357	2.40248

Table 5.7: Approximate variance for different values of γ , $N = 20$ at values $t = 0.25, 0.50, 0.75$ in Example 5.4

Chapter 6

Trigonometric stochastic processes solutions of linear second-order random differential equations

In this chapter we introduce some elementary trigonometric stochastic processes. These processes arise as mean square solutions of linear second-order random differential equations by using a random Fröbenius method. These random differential equations are the counterpart of the correspondent deterministic ones. After solving them, we will compute the main statistic functions of their solution and the results will be compared with respect to those obtained from other techniques, namely, the so-called dishonest method as well as Monte Carlo method. Apart from this, we will establish important operational properties of the trigonometric stochastic processes.

6.1 Introduction

Random mathematical models are powerful tools to capture the complexity and uncertainties of real problems. In recent papers [18], [62] several first order random differential models are solved using mean square calculus. It is well-known that managing deterministic differential models, the classical trigonometric functions $\sin(at)$ and $\cos(at)$ are solutions of the second-order ordinary differential equation $\frac{d^2x(t)}{dt^2} + a^2x(t) = 0$ where a^2 is a positive real number.

Following ideas of [18], [62], in this chapter we will provide mean square power series solutions of random differential equations of type

$$\ddot{X}(t) + A^2X(t) = 0, \tag{6.1}$$

where A^2 is a positive random variable satisfying certain additional properties addressed in a natural way to introduce trigonometric stochastic processes. Once random differential equation (6.1) has been solved, we will be interested in computing the main statistical functions of its solution stochastic process like the mean and variance. This goal will be achieved by means of the truncation of the power series solution previously provided. We will compare these results for the average and variance with respect to the correspondent provided for another techniques like Monte Carlo method [28] or the so-called dishonest method [46], [38]. The differences between these methods are illustrated with an example. Besides important operational properties of the introduced trigonometric stochastic processes will be included.

This brief chapter links with the previous one in the sense that it is also devoted to solve linear random differential equations by means a Fröbenius method.

This chapter is organized as follows: Section 6.3 develops a random Fröbenius method for solving differential equation (6.1). After that, we will introduce sinus and cosinus trigonometric stochastic processes as well as some of their more important algebraic relationships. Section 6.4 is addressed to compute approximations of the main statistical properties of the solution stochastic process of equation (6.1). An illustrative example is included in order to show that these approximations improve others obtained from alternative methods.

6.2 Preliminaries

In this section we present the main results that will be required throughout this chapter. We begin by introducing the concept of almost-surely sample continuous s.p. which is an useful concept in dealing with s.p.'s.

Definition 6.1. ([66, p.55]) *We say that a s.p. $\{X(t) : t \in \mathcal{T}\}$ defined on an interval \mathcal{T} is almost surely (a.s.) sample path continuous, i.e., $\{X(t) : t \in \mathcal{T}\}$ has continuous paths with probability one (w.p. 1) if*

$$\mathbb{P} \left[\bigcup_{t \in \mathcal{T}} \left\{ \omega \in \Omega : \lim_{\Delta t \rightarrow 0} |X(t + \Delta t)(\omega) - X(t)(\omega)| \neq 0 \right\} \right] = 0.$$

Then, for every $t \in \mathcal{T}$, the s.p. $\{X(t) : t \in \mathcal{T}\}$ is a.s. continuous, [66, p.55]. An important example of s.p. satisfying this property is the Brownian motion (see, [66, p.58]). The following criterium, namely Kolmogorov condition, is of great advantage in order to check that a s.p. is a.s. path-continuous because it only involves two-dimensional distributions. Moreover, it can usually be verified by direct computations.

Theorem 6.1 (Kolmogorov's criterium). ([66, p.57]) *Let $\{X(t) : t \in \mathcal{T}\}$ be a separable s.p. and let \mathcal{T} be a finite interval. If there exist strictly positive constants α , β and C such that*

$$\mathbb{E} [|X(t+h) - X(t)|^\alpha] \leq C h^{1+\beta},$$

then, $\{X(t) : t \in \mathcal{T}\}$ is a.s. sample continuous on \mathcal{T} .

Example 6.1. *Later we will need to check that the following s.p. $X(t) = At$, being A a 2-r.v., is a.s. sample continuous on every finite interval $\mathcal{T} = [a, b]$. In fact,*

$$\mathbb{E} [|X(t+h) - X(t)|^2] = \mathbb{E} [|Ah|^2] = |h|^2 \mathbb{E} [|A|^2] \leq C |h|^2,$$

where the positive constant C can be select such that: $C > \mathbb{E} [|A|^2]$ (since A is a 2-r.v). Then taking $\alpha = 2 > 0$, $\beta = 1 > 0$, by applying Kolmogorov's criterium one follows that the s.p. $X(t) = At$ is a.s. path continuous on \mathcal{T} .

Next, we state a m.s. chain rule for a C^1 -function of a m.f. differentiable s.p.'s.

Theorem 6.2 (Chain Rule). ([62]) *Let f be a real function with continuous derivative f' and $\{X(t) : t \in T = [a, b]\}$ be a 4-s.p. satisfying*

- i) $X(t)$ is m.f. differentiable.*
- ii) $X(t)$ is a.s. path continuous.*
- iii) There are $r > 4$ and $\delta > 0$ such that $\sup_{s \in [-\delta, \delta]} \mathbb{E} \left[\left| f'(x) \Big|_{x=X(t+s)} \right|^r \right] < +\infty$.*

Then, the 2-s.p. $f(X(t))$ is m.s. differentiable and the m.s. derivative is given by

$$\frac{df(X(t))}{dt} = f'(x) \Big|_{x=X(t)} \frac{dX(t)}{dt}.$$

Example 6.2. *Let us check that the m.s. derivative of s.p. $Z(t) = \cos(At)$, being A a 2-r.v., is given by $Z(t) = -A \sin(At)$. Let us apply the previous chain rule for $f(u) = \cos(u)$ which is a real function with continuous derivative. The s.p. $X(t) = At$ is m.f. differentiable (as it can be directly checked as we done in Example 2.3 but for m.s. differentiation); it is a.s. path continuous (see Example 6.1) and, condition iii) of Theorem 6.2 also holds*

$$\sup_{s \in [-\delta, \delta]} \mathbb{E} [|-\sin(A(t+s))|^r] \leq 1 < +\infty,$$

for any $r > 0$ (in particular taking any $r > 4$) since sinus function lies between -1 and 1 and the general expectation property

$$\text{If } x_1 \leq X(\omega) \leq x_2 \text{ then } x_1 \leq \mathbb{E}[X] \leq x_2, \quad \forall \omega \in \Omega.$$

In chapter 4 a random version of Mertens's theorem was established (see Proposition 4.3). It allows us to justify the product between an absolutely convergent series and a m.s. absolutely convergent series of r.v.'s. Later an extended version of Proposition 4.3 will be required in order to establish some relevant trigonometric identities for s.p.'s. This new version generalizes the previous one since it considers the case in which both series are random. For the demonstration of such result m.f. convergence of both involved series are required.

Proposition 6.1. *Let $\{A_n : n \geq 0\}$ be a sequence of 4-r.v.'s such that series $\sum_{n \geq 0} A_n$ is m.f. absolutely convergent to A . Let $\{X_n : n \geq 0\}$ be a sequence of 4-r.v.'s such that $\sum_{n \geq 0} X_n$ is m.f. convergent to $X \in L_4$. If $C_n = \sum_{k=0}^n A_{n-k} X_k$, then $\sum_{n \geq 0} C_n$ is m.s. convergent to AX .*

Proof. We will follow the proof presented at Proposition 4.3. Then, first we introduce the following notation

$$A'_n = \sum_{k=0}^n A_k, \quad X'_n = \sum_{k=0}^n X_k, \quad C'_n = \sum_{k=0}^n C_k, \quad Y'_n = X'_n - X.$$

We know that

$$C'_n = A'_n X + (A_0 Y'_n + A_1 Y'_{n-1} + \cdots + A_n Y'_0).$$

Now we denote $Z'_n = A_0 Y'_n + A_1 Y'_{n-1} + \cdots + A_n Y'_0$, then one gets

$$C'_n = A'_n X + Z'_n. \tag{6.2}$$

Since by hypotheses $\{A'_n : n \geq 0\}$ is m.f. convergent and X is a 4-r.v. (hence a 2-r.v.), by Lemma 2.5 one gets that $\{A'_n X : n \geq 0\}$ is m.s. convergent to AX . Taking into account (6.2), in order to prove the result, it is sufficient to show that $\{Z'_n : n \geq 0\}$ is m.s. convergent to $0 \in L_2$. As $\{Y'_n : n \geq 0\}$ is m.f. convergent to 0, by Lemma 2.4 it is also m.s. convergent to $0 \in L_2$, then given $\epsilon > 0$, there exists $N_0 > 0$ such that

$$\|Y'_n\|_4 < \frac{\epsilon}{2\alpha}, \quad n \geq N_0 \quad \text{or,} \quad \left(\| (Y'_n)^2 \|_2 \right)^{1/2} < \frac{\epsilon}{2\alpha}, \quad n \geq N_0, \tag{6.3}$$

being

$$\alpha = \sum_{n \geq 0} \|A_n\|_4 = \sum_{n \geq 0} \left(\| (A_n)^2 \|_2 \right)^{1/2} < +\infty.$$

Let $\beta_{N_0} > 0$ be defined by

$$\beta_{N_0} = \max \left\{ \left(\| (Y'_0)^2 \|_2 \right)^{1/2}, \dots, \left(\| (Y'_{N_0})^2 \|_2 \right)^{1/2} \right\},$$

and since sequences $\{A_n : n \geq 0\}, \{A_{n-1} : n \geq 1\}, \dots, \{A_{n-N_0} : n \geq N_0\}$ are m.f. convergent to the null r.v., we can also assume that

$$\|A_{n-j}\|_4 = \left(\| (A_{n-j})^2 \|_2 \right)^{1/2} \leq \frac{\epsilon}{2(N_0 + 1)\beta_{N_0}}, \quad j = 0, 1, \dots, N_0. \tag{6.4}$$

Let $n > N_0$, then by (6.3)-(6.4) one gets

$$\begin{aligned}
 \|Z'_n\|_2 &\leq \|A_n Y'_0 + \cdots + A_{n-N_0} Y'_{N_0}\|_2 \\
 &+ \|A_{n-N_0-1} Y'_{N_0+1} + \cdots + A_0 Y'_n\|_2 \\
 &\leq \left(\|A_n\|_2\right)^{1/2} \left(\|Y'_0\|_2\right)^{1/2} \\
 &+ \cdots \\
 &+ \left(\|A_{n-N_0}\|_2\right)^{1/2} \left(\|Y'_{N_0}\|_2\right)^{1/2} \\
 &+ \left(\|A_{n-N_0-1}\|_2\right)^{1/2} \left(\|Y'_{N_0+1}\|_2\right)^{1/2} \\
 &+ \cdots \\
 &+ \left(\|A_0\|_2\right)^{1/2} \left(\|Y'_n\|_2\right)^{1/2} \\
 &\leq \beta_{N_0} (N_0 + 1) \frac{\epsilon}{2(N_0 + 1)\beta_{N_0}} \\
 &+ \left(\left(\|A_{n-N_0-1}\|_2\right)^{1/2} + \cdots + \left(\|A_0\|_2\right)^{1/2}\right) \frac{\epsilon}{2\alpha} \\
 &\leq \frac{\epsilon}{2} + \alpha \frac{\epsilon}{2\alpha} = \epsilon. \quad \square
 \end{aligned}$$

6.3 Trigonometric processes solutions of random differential equations

Let us consider the random i.v.p. based on random differential equation (6.1), i.e.,

$$\ddot{X}(t) + A^2 X(t) = 0, \quad X(0) = Y_0, \quad \dot{X}(0) = Y_1, \quad (6.5)$$

where A^2 is a positive r.v. satisfying property (2.33) (note that this implies that $E[(A^2)^k] < +\infty$, for each $k = 0, 1, 2, \dots$, hence A^2 is in particular a 4-r.v.). Let us also assume that A^2 is independent of r.v.'s Y_0 and Y_1 .

We seek a formal solution of the form:

$$X(t) = \sum_{n \geq 0} X_n t^n, \quad (6.6)$$

where X_n are r.v.'s to be determined. By Example 2.3, the formal second m.s. derivative of $X(t)$ is given by

$$\ddot{X}(t) = \sum_{n \geq 0} (n+2)(n+1)X_{n+2}t^n. \quad (6.7)$$

By imposing that (6.6)-(6.7) satisfy the random differential equation (6.5) it follows

$$\ddot{X}(t) + A^2 X(t) = \sum_{n \geq 0} ((n+2)(n+1)X_{n+2} + A^2 X_n) t^n = 0. \quad (6.8)$$

Note that in the last step we have assumed that the series (6.6) satisfies the hypotheses of Lemma 2.5, which justifies the commutation of the r.v. A^2 and the infinite sum. Therefore a m.s. candidate solution process of problem (6.5) can be obtained by imposing

$$X_{n+2} = -\frac{A^2 X_n}{(n+2)(n+1)}, \quad n \geq 0,$$

that allows to compute explicitly the coefficients X_n in terms of the data

$$X_{2k} = \frac{(-1)^k (A^2)^k X_0}{(2k)!}, \quad X_{2k+1} = \frac{(-1)^k (A^2)^k X_1}{(2k+1)!}, \quad k = 0, 1, 2, \dots$$

Let us show that

$$X(t) = \sum_{k \geq 0} \frac{(-1)^k (A^2)^k X_0}{(2k)!} t^{2k} + \sum_{k \geq 0} \frac{(-1)^k (A^2)^k X_1}{(2k+1)!} t^{2k+1}, \quad (6.9)$$

where $X_0 = Y_0$ and $X_1 = Y_1$, is not only a formal solution but a rigorous solution of i.v.p. (6.5). So first we need to justify the commutation implicitly performed in (6.8). Taking into account Lemma 2.4, it is sufficient to check

that both series appearing in (6.9) are m.f. convergent. For that, we can assume in (2.34) that $n_0 = 0$ without loss of generality, then by considering that A^2 is a r.v. independent of $Y_0 = X_0$ as well as A^2 satisfies condition (2.33) for the 4-norm, for the first series one gets

$$\begin{aligned} \sum_{k \geq 0} \frac{\|(-1)^k (A^2)^k Y_0\|_4 |t|^{2k}}{(2k)!} &= \sum_{k \geq 0} \frac{\|(A^2)^k\|_4 \|Y_0\|_4 |t|^{2k}}{(2k)!} \\ &\leq \sum_{k \geq 0} \frac{HM^{k-1} ((k-1)!)^p \|Y_0\|_4 |t|^{2k}}{(2k)!}. \end{aligned} \tag{6.10}$$

Setting

$$\alpha_k^{(p)} = \frac{1}{(2k)!} HM^{k-1} ((k-1)!)^p \|Y_0\|_4 |t|^{2k},$$

and using the D'Alembert test one obtains

$$\lim_{k \rightarrow \infty} \frac{\alpha_{k+1}^{(p)}}{\alpha_k^{(p)}} = \lim_{k \rightarrow \infty} \frac{M |t|^2 k^p}{(2k+2)(2k+1)} = 0, \quad p < 2, \forall t. \tag{6.11}$$

The corresponding study for the second series appearing in (6.9) can be made in the same way and it is easy to check that the domain of m.f. convergence does not change. Then, from (6.10) and (6.11) we conclude that the series (6.9) is absolutely m.f. convergent in the whole real line, if $p < 2$. Therefore by Lemma 2.4 series given by (6.9) is m.s. convergent in the same domain. Moreover, the above argument also justifies the uniform m.s. convergence of the series given by (6.9) as well as its formal derivative given by (6.7) in any closed set contained into the domain of convergence. As a consequence of Theorem 2.5 the formal derivative performed in (6.7) is justified. Summarizing the following theorem has been established:

Theorem 6.3. *Consider the random i.v.p. (6.5), where A^2 satisfies condition (2.33) for the 4-norm and $0 \leq p < 2$ and Y_0 and Y_1 are 4-r.v. both independent of A^2 . Then (6.9) is a m.s. solution of (6.5) in the whole t -real line.*

Due to the m.s. convergence of the first series appearing in (6.9), for a r.v. A such that A^2 satisfies condition (2.33) for the 4-norm, we can define the cosinus trigonometric s.p. as follows

$$\cos(At) = \sum_{k \geq 0} \frac{(-1)^k (A^2)^k}{(2k)!} t^{2k}. \tag{6.12}$$

As the second series of (6.9) does not correspond exactly to the s.p. $\sin(At)$, we will proceed in the following way to define it. Let us denote $Z(t) = \cos(At)$ then by Example 6.2 one gets

$$-A \sin(At) = \dot{Z}(t) = \sum_{k \geq 1} \frac{(-1)^k (A^2)^k}{(2k-1)!} t^{2k-1},$$

then isolating $\sin(At)$ and considering Lemma 2.5 one gets

$$\sin(At) = \sum_{k \geq 1} \frac{(-1)^{k-1} (A^2)^{k-\frac{1}{2}}}{(2k-1)!} t^{2k-1} = \sum_{k \geq 0} \frac{(-1)^k (A^2)^{k+\frac{1}{2}}}{(2k+1)!} t^{2k+1}. \quad (6.13)$$

This allows to represent the solution process given by (6.9) as follows

$$X(t) = Y_0 \cos(At) + \frac{Y_1}{A} \sin(At). \quad (6.14)$$

On the other hand, if r.v. A^2 satisfies condition (2.33) for the 4-norm then we have established the m.f. absolute convergence of series (6.12) and (6.13) then by Proposition 6.1 one gets

$$(\cos(At))^2 + (\sin(At))^2 = 1, \quad \forall t \in \mathbb{R}.$$

By assuming that A^2 , B^2 and $(A \pm B)^2$ satisfy condition (2.33) again by Proposition 6.1 one follows

$$\begin{aligned} \cos((A \pm B)t) &= \cos(At) \cos(Bt) \mp \sin(At) \sin(Bt), \\ \sin((A \pm B)t) &= \sin(At) \cos(Bt) \pm \sin(Bt) \cos(At). \end{aligned}$$

6.4 Statistical properties of the solution process. An example

From a practical point of view, it is interesting not only the obtention of the solution s.p. of the i.v.p. (6.5) but also providing its main statistical functions such as the mean and the variance.

With respect to the computation of the expectation, this task cannot be performed solving the corresponding deterministic linear differential equation that the so-called dishonest approach handles by taking advantage of nice property (2.24) because for each t , A^2 and $X(t)$ are not independent r.v.'s (see, [46], [38]).

Since $X(t)$ given by (6.9) is an infinite series we will proceed by truncation:

$$X_N(t) = Y_0 \sum_{k=0}^N \frac{(-1)^k}{(2k)!} (A^2)^k t^{2k} + Y_1 \sum_{k=0}^N \frac{(-1)^k}{(2k+1)!} (A^2)^k t^{2k+1}.$$

In the following, we will assume that r.v. A^2 is independent of Y_0 and Y_1 , respectively, which is realistic in applications. Then

$$\begin{aligned} \mathbb{E}[X_N(t)] &= \mathbb{E}[Y_0] \sum_{k=0}^N \frac{(-1)^k}{(2k)!} \mathbb{E}[(A^2)^k] t^{2k} \\ &+ \mathbb{E}[Y_1] \sum_{k=0}^N \frac{(-1)^k}{(2k+1)!} \mathbb{E}[(A^2)^k] t^{2k+1}, \end{aligned} \tag{6.15}$$

and

$$\begin{aligned} &\mathbb{E}[(X_N(t))^2] \\ &= \sum_{k=0}^N \mathbb{E}[(A^2)^{2k}] \left(\frac{\mathbb{E}[(Y_0)^2] t^{4k}}{((2k)!)^2} + \frac{\mathbb{E}[(Y_1)^2] t^{4k+2}}{((2k+1)!)^2} \right) \\ &+ 2 \sum_{k=1}^N \sum_{j=0}^{k-1} (-1)^{k+j} \mathbb{E}[(A^2)^{k+j}] \left\{ \frac{\mathbb{E}[(Y_0)^2] t^{2(k+j)}}{(2k)!(2j)!} + \frac{\mathbb{E}[(Y_1)^2] t^{2(k+j+1)}}{(2k+1)!(2j+1)!} \right\} \\ &+ 2\mathbb{E}[Y_0 Y_1] \sum_{k=0}^N \sum_{j=0}^N (-1)^{k+j} \mathbb{E}[(A^2)^{k+j}] \frac{t^{2(k+j)+1}}{(2k)!(2j+1)!}. \end{aligned} \tag{6.16}$$

The following example illustrates the results presented previously.

Example 6.3. Consider the i.v.p. (6.5), where $A^2 \sim Be(\alpha = 2; \beta = 1)$, independent of 4-r.v.'s Y_0 and Y_1 which satisfy $\mathbb{E}[Y_0] = 1$, $\mathbb{E}[(Y_0)^2] = 2$,

t	$E[X_{10}(t)]$	$E[X_{15}(t)]$	$E^d[X(t)]$	$\mu_X^{50000}(t)$	$\mu_X^{100000}(t)$	$E[X(t)]$
0.00	1.0000	1.0000	1.0000	0.998141	1.00304	1.0000
0.20	1.18581	1.18581	1.18581	1.18454	1.18754	1.18581
0.40	1.34013	1.34013	1.34007	1.33954	1.34055	1.34013
0.60	1.459	1.459	1.45867	1.45916	1.46015	1.459
0.80	1.53952	1.53952	1.53846	1.54048	1.53751	1.53952
1.00	1.57993	1.57993	1.57731	1.58174	1.57691	1.57993
2.00	1.20122	1.20122	1.16022	1.20699	1.19638	1.20122
3.00	0.18816	0.18816	0.01168	0.19547	0.18783	0.18816
4.00	-0.73967	-0.73967	-1.14423	-0.73491	-0.73200	-0.73967
5.00	-0.998991	-0.998992	-1.57876	-1.0003	-0.98568	-0.998992
10.0	1.65839	-0.02348	0.860002	-0.01653	-0.03144	-0.02343

Table 6.1: Comparison between the exact expectation and its approximations obtained from series truncation, dishonest and Monte Carlo methods for Example 6.3 with $A \sim \text{Be}(\alpha = 2; \beta = 1)$

$E[Y_1] = 1$, $E[(Y_1)^2] = 3$ and $E[Y_0Y_1] = 0$, that is, Y_0 and Y_1 are uncorrelated r.v.'s. Note that by Example 2.7, A^2 satisfies condition (2.33) for the 4-norm with $p = 0$.

Table 6.1 shows the numerical results of the mean $E[X_N(t)]$ for different values of the truncation order N ($N = 10, 15$) and independent variable t . These values have been obtained from (6.15). Table 6.1 also includes approximations computed from Monte Carlo method taking 5×10^4 and 10^5 simulations (they have been denoted by $\mu_X^{50000}(t)$ and $\mu_X^{100000}(t)$, respectively). Moreover this table includes approximations provided by the so-called dishonest method ($E^d[X(t)]$). Exact values of the mean $E[X(t)]$ can be computed directly from expression (6.14) and they are included in the last column of Table 6.1.

By comparing these approximations for the expectation, we realize that

truncation method developed in this chapter provides values that coincide with the exact ones from $t = 0$ to $t = 10$ even taking low truncation order N .

Approximations provided by Monte Carlo method are also agree with exact values but for a fixed number of simulations they become worse as t departs from origin. On the other hand, dishonest method only provides acceptable approximations near $t = 0$, where the i.v.p. is posed.

Figure 6.1 shows the results for the standard deviation function. Note that the approximation of this function by using the truncation method can be performed from (6.15), (6.16) and (2.7) (in Figure 6.1 we have taken $N = 15$ as truncation order). These values coincide with the corresponding obtained directly by using (6.14), (2.8) and (2.7) in the interval $t \in [0, 10]$ whereas Monte Carlo method with 10^5 simulations provides worse approximations specially near $t = 0$.

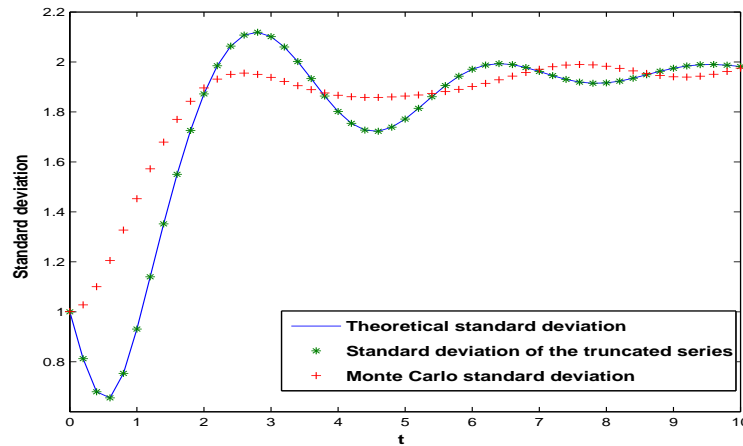


Figure 6.1: Comparison between the exact standard deviation and its approximations obtained from series truncation (with order of truncation $N = 15$) and Monte Carlo (with 10^6 simulations) methods for Example 6.3 with $A^2 \sim \text{Be}(\alpha = 2; \beta = 1)$

6.5 Conclusions

In this chapter we have constructed a m.s. power series solution of second-order linear random differential equation (6.5) incorporating uncertainty through both, the multiplicative coefficient A^2 and the initial conditions Y_0 and Y_1 . For that, we have assumed that diffusion coefficient A^2 is a positive r.v. satisfying certain quite general condition that many important positive r.v.'s hold like Beta or Gamma, for instance; for the initial conditions we have assumed that both are 4-r.v.'s independent of r.v. A^2 which a realistic assumption from a viewpoint. Under such hypotheses we have shown that the constructed random power series solution is m.s. convergent. As a consequence we have introduced some basic trigonometric s.p.'s, namely, sinus and cosinus, that are the correspondent counterpart of the deterministic ones. Moreover, important algebraic properties about these s.p.'s have been established. By considering a truncation of the power series solution s.p., approximations of the mean and standard deviation functions of the solution s.p. have been computed in terms of the data by means of explicit formulae. Finally, we have presented an illustrative example where the mean and the standard deviation are computed from our approach as well as Monte Carlo and dishonest method. This example shows that the technique here developed provide better approximations.

The main results of this chapter have been accepted for publication in [12].

Chapter 7

Mean square power series solutions of random Hermite differential equations

This chapter deals with the construction of random power series solutions of second order linear differential equations of Hermite containing uncertainty through its coefficient as well as the initial conditions. Under appropriate hypotheses on the data, we establish that the constructed random power series solution is mean square convergent. Conditions in order to obtain random polynomial solutions are provided. This will allow us to introduce the concept of random Hermite polynomial that constitutes the random counterpart of the classical concept of Hermite polynomial. Moreover, the main statistical functions of the approximating solution stochastic process generated by truncation of the exact power series solution are given. Finally, several illustrative examples for computing the mean and the variance of the solution stochastic process are included in order to compare the truncation series method with respect to another approach.

7.1 Introduction

In this paper, we develop a power series method to construct mean square convergent series solutions of the so-called random Hermite differential equation

$$\ddot{X}(t) - 2t\dot{X}(t) + AX(t) = 0, \quad -\infty < t < +\infty, \quad (7.1)$$

where the coefficient A is a random variable satisfying certain conditions to be specified later.

Hermite differential equation has a great interest in the development of applied mathematics, not only because it appears in a natural way solving partial differential equations of the mathematical physics by using the separation of variables technique, but also because its prevalent role in other areas of mathematics such as functional analysis or orthogonal polynomials theory.

An important difficulty to be overcome is the lack of sub-multiplicativity of the mean square norm (and as a consequence of the mean fourth norm) and the necessity of bounding products of random variables. Particularly interesting is the case where A is a discrete random variable taking a finite number of even integer values with certain probabilities. In such case, it appears mean square finite series solutions of problem (7.1) that will be referred as to random Hermite polynomials.

The chapter is organized as follows. Section 7.2 introduces the concepts of fundamental set of solution process for equation (7.1) as well as random polynomial. Section 7.3 deals with the proof of an important inequality relating the norm of the product of random variables with the product of the certain powers of the norm each random variable factor. This result manages satisfactorily the lack of submultiplicativity of the mean square and mean fourth norms. Section 7.4 deals with the construction of mean square convergent power series solution of (7.1) in the case where A is a random variable satisfying certain conditions related to the exponential growth of its absolute moments with respect to the origin. The case where A is a discrete random variable taking a finite number of even values leads to the concept of random Hermite polynomials whose definition and obtention is shown in section 7.5. Statistical functions of the truncated random power series solutions, and par-

ticularly of random Hermite polynomials, are studied in section 7.6. In section 7.7 some illustrative examples are presented showing the discrepancies with respect to the so-called dishonest and Monte Carlo methods. In last section 7.8, we summing up the main results of this chapter.

7.2 Preliminaries

We start this section by introducing some definitions that will be required later. These concepts are the corresponding extensions to the random framework of their deterministic counterparts.

Definition 7.1. *Let A_1 and A_2 be r.v.'s, and let $X_1(t)$ and $X_2(t)$ be two solutions of the second-order random differential equation*

$$\ddot{X}(t) + A_1\dot{X}(t) + A_2X(t) = 0, \quad -\infty < t < +\infty. \quad (7.2)$$

We say that $\{X_1(t), X_2(t)\}$ is a fundamental set of solution s.p.'s of (7.2) in $-\infty < t < +\infty$, if any solution $X(t)$ of (7.2) admits a unique representation of the form

$$X(t) = C_1X_1(t) + C_2X_2(t), \quad t \in (-\infty, +\infty), \quad (7.3)$$

where C_1 and C_2 are r.v.'s uniquely determined by $X(t)$.

Definition 7.2. *Let $\mathcal{S} = \{X_1(t), X_2(t)\}$ be two solutions of (7.2), the s.p. given by*

$$W_{\mathcal{S}}(t) = X_1(t)\dot{X}_2(t) - X_2(t)\dot{X}_1(t),$$

is called the wronskian s.p. of \mathcal{S} .

The following result provides sufficient conditions for a pair of solutions of (7.2) defines a fundamental set. We omit the proof because it follows in broad outline the same steps as in the deterministic case.

Proposition 7.1. *If $\mathcal{S} = \{X_1(t), X_2(t)\}$ is a pair of solution s.p.'s of the random differential equation (7.2) in $-\infty < t < +\infty$ and there exists $t_0 \in$*

$(-\infty, +\infty)$ such that $W_{\mathcal{S}}(t_0) \neq 0$, then $\mathcal{S} = \{X_1(t), X_2(t)\}$ is a fundamental set of solution s.p.'s of (7.2).

As we have underlined in the first section, we are interested in determining solution s.p.'s of differential equation (7.1) by means of random power series which under certain conditions become random polynomials. Now we introduce both concepts.

Definition 7.3. *Given a family of r.v.'s $\{X_k : k \geq 0\}$ and $t \in \mathcal{T}$, $\sum_{k \geq 0} X_k t^k$ is called a random power series. If $P[X_k = 0] = 1$ for all $k > m$, that is, $P[\{\omega \in \Omega : X_k(\omega) = 0, \forall k > m\}] = 1$, then $\sum_{k=0}^m X_k t^k$ is said to be a random polynomial in t of degree m .*

7.3 A crucial inequality

In dealing with the norm of a product of r.v.'s, an important fact in this dissertation, is that the 2-norm $\|\cdot\|_2$ given by (2.18) does not provide a Banach algebra structure to L_2 , and as a consequence the 4-norm $\|\cdot\|_4$ given by (2.29) neither is sub-multiplicative, that is, property $\|XY\| \leq \|X\| \|Y\|$ does not hold for neither norms (see Example 2.5). This situation difficult our next target. Indeed, in the next section we will obtain the solution of (7.1) by means of a random power series which general term is the product of r.v.'s. In order to establish the absolute m.f. convergence of this series (thus, it m.s. convergence), we will require to bound the 4-norm of a product of r.v.'s in terms of the 4-norm of each factor. Next we address our work to provide some results to overcome this difficulty.

Next proposition contains the main result in this section and it constitutes a generalization of (2.12).

Proposition 7.2. *Let $\{X_i\}_{i=1}^n$, $n \geq 2$ be r.v.'s such that $E[(X_i)^{2^n}] < \infty$,*

$i = 1, 2, \dots, n$, then

$$\mathbb{E} \left[\left| \prod_{i=1}^n X_i \right| \right] \leq \left(\prod_{i=1}^n \mathbb{E} \left[(X_i)^{2^{n-1}} \right] \right)^{1/2^{n-1}}, \quad n \geq 2. \quad (7.4)$$

Proof. It follows by induction on the number n of the involved r.v.'s. For $n = 2$, (7.4) is exactly the Schwarz's inequality given by (2.12). Note that since by hypothesis $\mathbb{E} \left[(X_i)^4 \right] < \infty$ for $i = 1, 2$, then by (2.5) one gets $\mathbb{E} \left[(X_i)^2 \right] < \infty$, $i = 1, 2$. For $n = 3$, by applying twice (2.12) one obtains

$$\begin{aligned} \mathbb{E} [|X_1 X_2 X_3|] &\leq \left(\mathbb{E} \left[(X_1 X_2)^2 \right] \right)^{1/2} \left(\mathbb{E} \left[(X_3)^2 \right] \right)^{1/2} \\ &= \left(\mathbb{E} \left[(X_1)^2 (X_2)^2 \right] \right)^{1/2} \left(\mathbb{E} \left[(X_3)^2 \right] \right)^{1/2} \\ &\leq \left(\mathbb{E} \left[(X_1)^4 \right] \right)^{1/4} \left(\mathbb{E} \left[(X_2)^4 \right] \right)^{1/4} \left(\mathbb{E} \left[(X_3)^2 \right] \right)^{1/2} \\ &\leq \left(\mathbb{E} \left[(X_1)^4 \right] \right)^{1/4} \left(\mathbb{E} \left[(X_2)^4 \right] \right)^{1/4} \left(\mathbb{E} \left[(X_3)^4 \right] \right)^{1/4}, \end{aligned}$$

where in the last inequality we have applied (2.13) for $X = X_3$ and $m = 2 \leq 4 = n$. Now, we assume by induction hypothesis that (7.4) holds and we shall establish that

$$\mathbb{E} \left[\left| \prod_{i=1}^{n+1} X_i \right| \right] \leq \left(\prod_{i=1}^{n+1} \mathbb{E} \left[(X_i)^{2^n} \right] \right)^{1/2^n}.$$

Indeed, by the induction hypothesis one gets

$$\begin{aligned} \mathbb{E} \left[\left| \prod_{i=1}^{n+1} X_i \right| \right] &= \mathbb{E} [| (X_1) (X_2) \cdots (X_{n-1}) (X_n X_{n+1}) |] \\ &\leq \left(\mathbb{E} \left[(X_1)^{2^{n-1}} \right] \right)^{1/2^{n-1}} \left(\mathbb{E} \left[(X_2)^{2^{n-1}} \right] \right)^{1/2^{n-1}} \cdots \\ &\quad \times \left(\mathbb{E} \left[(X_{n-1})^{2^{n-1}} \right] \right)^{1/2^{n-1}} \left(\mathbb{E} \left[(X_n)^{2^{n-1}} (X_{n+1})^{2^{n-1}} \right] \right)^{1/2^{n-1}}, \end{aligned}$$

next we apply the Liapounov's inequality (2.13) to the $n - 1$ first factors and, the Schwarz's inequality (2.12) to the last factor of the right-hand side and

one obtains

$$\begin{aligned} \mathbb{E} \left[\left| \prod_{i=1}^{n+1} X_i \right| \right] &\leq \left(\mathbb{E} \left[(X_1)^{2^n} \right] \right)^{1/2^n} \left(\mathbb{E} \left[(X_2)^{2^n} \right] \right)^{1/2^n} \cdots \left(\mathbb{E} \left[(X_{n-1})^{2^n} \right] \right)^{1/2^n} \\ &\quad \times \left(\mathbb{E} \left[(X_n)^{2^n} \right] \right)^{1/2^n} \left(\mathbb{E} \left[(X_{n+1})^{2^n} \right] \right)^{1/2^n} \\ &= \left(\prod_{i=1}^{n+1} \mathbb{E} \left[(X_i)^{2^n} \right] \right)^{1/2^n} . \quad \square \end{aligned}$$

Thus the result is established.

Now we rewrite the inequality (7.4) in terms of the 4-norm given by (2.29). For this, we will assume that $\mathbb{E} \left[(Y_i)^{2^{n+1}} \right] < +\infty$, $i = 1, 2, \dots, n$ and we apply Proposition 7.2 to the r.v.'s $X_i = (Y_i)^4$, $i = 1, 2, \dots, n$, thus one obtains

$$\begin{aligned} \mathbb{E} \left[\left(\prod_{i=1}^n Y_i \right)^4 \right] &\leq \left(\prod_{i=1}^n \mathbb{E} \left[\left((Y_i)^4 \right)^{2^{n-1}} \right] \right)^{1/2^{n-1}} \\ &= \prod_{i=1}^n \left(\mathbb{E} \left[\left((Y_i)^{2^{n-1}} \right)^4 \right] \right)^{1/2^{n-1}} , \end{aligned}$$

taking fourth roots in this expression it follows

$$\begin{aligned} \left(\mathbb{E} \left[\left(\prod_{i=1}^n Y_i \right)^4 \right] \right)^{1/4} &\leq \left(\prod_{i=1}^n \mathbb{E} \left[\left((Y_i)^{2^{n-1}} \right)^4 \right] \right)^{1/2^{n+1}} \\ &= \prod_{i=1}^n \left(\left(\mathbb{E} \left[\left((Y_i)^{2^{n-1}} \right)^4 \right] \right)^{1/4} \right)^{1/2^{n-1}} \end{aligned}$$

i.e., considering the 4-norm definition one gets

$$\left\| \prod_{i=1}^n Y_i \right\|_4 \leq \prod_{i=1}^n \left(\left\| (Y_i)^{2^{n-1}} \right\|_4 \right)^{1/2^{n-1}} , \quad \mathbb{E} \left[(Y_i)^{2^{n+1}} \right] < +\infty, \quad i = 1, \dots, n. \quad (7.5)$$

7.4 Solving random Hermite differential equations

This section is addressed to obtain the solution of the random differential equation (7.1) by means of a random power series as well as establishing its m.s. convergence. As we will see later in this section, inequality (7.5) will play a crucial role in order to prove this type of convergence. In the following we will assume that the absolute moments with respect to the origin of r.v. A appearing into (7.1) increase at the most exponentially, i.e, there exist a nonnegative integer n_0 and positive constants H and M such that

$$E [|A|^n] \leq H M^n < +\infty, \quad \forall n \geq n_0. \quad (7.6)$$

Let us seek a formal solution s.p. of problem (7.1) of the form

$$X(t) = \sum_{n \geq 0} X_n t^n, \quad (7.7)$$

where coefficients X_n are 2-r.v.'s. Assuming that $X(t)$ is termwise m.s. differentiable and applying Example 2.3, it follows that

$$\dot{X}(t) = \sum_{n \geq 1} n X_n t^{n-1}, \quad t \dot{X}(t) = \sum_{n \geq 1} n X_n t^n, \quad (7.8)$$

$$\ddot{X}(t) = \sum_{n \geq 2} n(n-1) X_n t^{n-2} = 2X_2 + \sum_{n \geq 1} (n+2)(n+1) X_{n+2} t^n. \quad (7.9)$$

By imposing that (7.7)-(7.9) satisfy (7.1), one gets

$$AX_0 + 2X_2 + \sum_{n \geq 1} ((n+2)(n+1)X_{n+2} + (A-2n)X_n) t^n = 0. \quad (7.10)$$

Therefore a candidate m.s. solution process of problem (7.1) can be obtained by imposing

$$\left. \begin{aligned} AX_0 + 2X_2 &= 0, \\ (n+2)(n+1)X_{n+2} + (A-2n)X_n &= 0, \quad n \geq 1, \end{aligned} \right\}$$

i.e.,

$$X_{n+2} = \frac{(2n-A)}{(n+2)(n+1)} X_n, \quad n \geq 0. \quad (7.11)$$

By a recursive reasoning, these coefficients X_n can be computed as follows

$$X_{2k+2} = \frac{(-A)(4-A)(8-A)\cdots(4k-A)}{(2k+2)!} X_0, \quad k \geq 0, \quad (7.12)$$

$$X_{2k+3} = \frac{(2-A)(6-A)(10-A)\cdots(4k+2-A)}{(2k+3)!} X_1, \quad k \geq 0. \quad (7.13)$$

Note that, all above equalities must be understanding with probability 1. Therefore taking into account relationships (7.12)-(7.13), the s.p. (7.7) can be expressed in terms of the data

$$X(t) = X_0 X_1(t) + X_1 X_2(t), \quad (7.14)$$

where

$$\begin{aligned} X_1(t) &= \left(1 + \sum_{k \geq 0} \frac{t^{2k+2}}{(2k+2)!} \prod_{j=0}^k (4j-A) \right), \\ X_2(t) &= \left(t + \sum_{k \geq 0} \frac{t^{2k+3}}{(2k+3)!} \prod_{j=0}^k (4j+2-A) \right). \end{aligned} \quad (7.15)$$

Note that in (7.10) we have formally commuted r.v. A and a random infinite sum, then property (2.28) has been applied implicitly. In order to legitimate this commutation, hypotheses of Lemma 2.5 must be checked. Therefore, we address our work to establish the m.f. convergence of the two random series given by (7.15), and since $(L_4, \|\cdot\|_4)$ is a Banach space, that is equivalent to prove that both series are absolutely convergent in the 4-norm. Thus, for each $t \in (-\infty, +\infty)$, we consider the numerical series associated to the first series in (7.15) given by

$$\sum_{k \geq 0} \frac{|t|^{2k+2}}{(2k+2)!} \left\| \prod_{j=0}^k (4j-A) \right\|_4, \quad (7.16)$$

and note that by hypothesis $E[|A|^n] < +\infty, \forall n \geq 0$, then applying (7.5) one gets

$$\left\| \prod_{j=0}^k (4j-A) \right\|_4 \leq \prod_{j=0}^k \left(\left\| (4j-A)^{2^k} \right\|_4 \right)^{\frac{1}{2^k}}.$$

Now we can bound each factor of the above right-hand side by considering the so-called c_s -inequality (2.15). Indeed, let k be a fixed nonnegative integer and,

for each $j = 0, 1, \dots, k$ let us consider the following bound of the j -th factor

$$\begin{aligned} \left(\left\| (4j - A)^{2^k} \right\|_4 \right)^{\frac{1}{2^k}} &= \left(\mathbb{E} \left[|4j - A|^{2^{k+2}} \right] \right)^{\frac{1}{2^{k+2}}} \\ &\leq \left(2^{2^{k+2}-1} \left((4j)^{2^{k+2}} + \mathbb{E} \left[|A|^{2^{k+2}} \right] \right) \right)^{\frac{1}{2^{k+2}}} \quad (7.17) \\ &\leq \left(2^{2^{k+2}-1} \left((4k)^{2^{k+2}} + \mathbb{E} \left[|A|^{2^{k+2}} \right] \right) \right)^{\frac{1}{2^{k+2}}}. \end{aligned}$$

Under hypothesis (7.6), we can assure the existence of a integer, say k_0 , such that: $(4k)^{2^{k+2}} \geq H M^{2^{k+2}}$ for each $k \geq k_0$, so from (7.17) one gets

$$\begin{aligned} \left(\left\| (4j - A)^{2^k} \right\|_4 \right)^{\frac{1}{2^k}} &\leq \left(2^{2^{k+2}-1} \left((4k)^{2^{k+2}} + \mathbb{E} \left[|A|^{2^{k+2}} \right] \right) \right)^{\frac{1}{2^{k+2}}} \\ &\leq \left(2^{2^{k+2}} \left((4k)^{2^{k+2}} \right) \right)^{\frac{1}{2^{k+2}}} = 8k, \quad \forall k \geq k_0. \end{aligned} \quad (7.18)$$

As we are only interested in convergence, we can assume without loss of generality that $k_0 = 0$. Then series (7.16) can be majorized by series

$$\sum_{k \geq 0} \frac{|t|^{2^{k+2}}}{(2k+2)!} (8k)^{k+1},$$

which is convergent for all t as it can be directly checked by D'Alembert test. Therefore, we have proven that numerical series given by (7.16) is convergent, thus the first random series of (7.15) is m.f. convergent, so also by Lemma 2.4 m.s. convergent. Following an analogous procedure, it is easy to establish the m.s. convergence of the second series in (7.15) for all t . Note that the above reasoning shows that both solution series $X_1(t)$ and $X_2(t)$ given by (7.15) are m.s. uniformly convergent, therefore taking into account Examples 2.2 and 2.3 and Theorem 2.5, the formal differentiation considered in (7.8)-(7.9) is justified. On the other hand, taking $t_0 = 0$ and considering that $X_1(0) = 1$, $\dot{X}_1(0) = 0$, $X_2(0) = 0$ and $\dot{X}_2(0) = 1$, one gets that $W_S(0) = 1 \neq 0$, then by Proposition 7.1 and (7.3), the solution of random differential equation (7.1) with random initial conditions $X(0) = Y_0$ and $\dot{X}(0) = Y_1$ is given by

$$X(t) = Y_0 X_1(t) + Y_1 X_2(t), \quad t \in (-\infty, +\infty), \quad (7.19)$$

where $X_1(t)$ and $X_2(t)$ are defined in (7.15). Summarizing the following result has been established

Theorem 7.1. *The random differential equation (7.1) with initial conditions $X(0) = Y_0$ and $\dot{X}(0) = Y_1$, where A is a r.v. satisfying condition (7.6), admits a random power series solution of the form (7.19) where $X_1(t)$ and $X_2(t)$ are given by (7.15). Moreover the solution is m.s. convergent for each $t \in (-\infty, +\infty)$.*

From (7.18), one deduces that previous exposition also remains true by assuming the following growth condition about the absolute moments of r.v. A

$$\exists k_1 : \mathbb{E} \left[|A|^{2^{k+2}} \right] \leq (4k)^{2^{k+2}}, \quad \forall k \geq k_1, \quad (7.20)$$

which is less restrictive than (7.6). However, from a practical viewpoint previous condition is not easy to check because of the lack of explicit expressions for the absolute moments with respect to the origin of relevant r.v.'s such as Binomial, Poisson, etc. about which we are mainly interested in order to find polynomial solutions of random differential equation (7.1). In order to overcome this difficulty, from now on, we will deal with r.v.'s A taking values on any bounded interval in the set \mathbb{R} of the real numbers, i.e., such that $a_1 \leq A(\omega) \leq a_2$, for each $\omega \in \Omega$. Note that this class of r.v.'s satisfy condition (7.6). Indeed, let us assume that A is a continuous r.v. with density function $f_A(a)$, then taking $H = \max(|a_1|, |a_2|)$, one gets

$$\mathbb{E} [|A|^n] = \int_{a_1}^{a_2} |a|^n f_A(a) da \leq H^n.$$

Note that, by substituting the integral for a sum, previous conclusion remains true if A is a discrete r.v. We emphasize that important r.v.'s such as Binomial, Hypergeometric, Uniform or Beta take values on bounded sets of the real line. Otherwise, we can use the truncation method (see [48]) for dealing with unbounded r.v.'s such as Poisson, Gaussian or Exponential (see Example 7.2 below).

7.5 Introducing the random Hermite polynomials

Next we address to motivate the introduction of random Hermite polynomials as well as some aspects related to them that arise in a natural way when

we extend this deterministic concept to the stochastic framework. First, note that taking into account (7.11)-(7.15), one deduces that if there exists $n \geq 0$ such that $P[A = 2n] = 1$, that is, A is a (degenerate) discrete r.v. whose total probability mass is concentrated at the point $2n$, then the random differential equation (7.1) has a random polynomial solution. Since every even integer number takes the form $4k$ or $4k + 2$, $k \geq 0$, if there exists $k \geq 0$ such that $P[A = 4k] = 1$ (or $P[A = 4k + 2] = 1$), then $X_1(t)$ (or $X_2(t)$) given by (7.15) generates a (random) polynomial solution of degree $2k$ (or $2k + 1$). These random (degenerate) solutions can be interpreted as corresponding Hermite polynomials that one presents in the deterministic framework. However in the random scenario there are richer situations that deserve to be considered.

Indeed, in the case that A is a continuous r.v. since $P[A = 2n] = 0$ for every integer $n \geq 0$, then with probability 1, one can conclude that there are not random polynomial solutions of differential equation (7.1). Whereas if A is a discrete r.v. that only takes different even values (not concentrated in just one even value), then there will exist with probability 1, random polynomial solutions (see later, Example 7.3). This case generalizes the concept of Hermite polynomial solution from the deterministic framework (moreover, note that this situation contains the previous case where A was a degenerated r.v. concentrated just at a point).

For the case that A is a discrete r.v. whose values lie in a set containing even numbers, random differential equation (7.1) does not have any random polynomial solution, but it admits some sample representations which are (deterministic) polynomials. Then, considering the solution s.p. as a family of trajectories, we can assign the probability p_{pol} that the random power series given by (7.19) and (7.15) has random polynomial sample solutions. Note that this series becomes a random polynomial if, and only if, $p_{pol} = 1$. Thus if $0 \leq p_{pol} < 1$, then $X(t)$ given by (7.19) and (7.15) produces sample solutions with probability p_{pol} of random differential equation (7.1) with initial conditions $Y_0(\omega)$, $Y_1(\omega)$, but not a random solution. As an illustrative example, let us assume that A is Binomial r.v. of parameters n and p , $A \sim \text{Bi}(n; p)$. Table 7.1 shows the probability p_{pol} for different values of the parameters n and p . As in general, $\text{Bi}(n; p) = \text{Bi}(n; 1 - p)$, we only consider values $0 < p \leq 0.5$. One observes that for p fixed, these values decrease up to 0.5 as n increases. The considerations above motivate the following result:

Corollary 7.1. *Let us consider the random differential equation (7.1) with*

n	$p_{pol}, p = 0.1$	$p_{pol}, p = 0.2$	$p_{pol}, p = 0.3$	$p_{pol}, p = 0.4$	$p_{pol}, p = 0.5$
2	0.82	0.68	0.58	0.52	0.5
4	0.7048	0.5648	0.5128	0.5008	0.5
6	0.631072	0.523328	0.502048	0.500032	0.5
8	0.583886	0.508398	0.500328	0.500001	0.5
10	0.553687	0.503023	0.500052	0.5	0.5
100	0.5	0.5	0.5	0.5	0.5

Table 7.1: Probabilities of generating random polynomial (sample) solutions when A is a Binomial r.v. of parameters n and p

initial conditions $X(0) = Y_0$ and $\dot{X}(0) = Y_1$, where the (discrete) r.v. A takes only a finite number of even integer values, that is, $P[A = 2m_j] = p_j > 0$, $1 \leq j \leq n$ with $\sum_{j=1}^n p_j = 1$. Then this i.v.p. has a random polynomial solution $H_{m_n}(t)$ of the degree m_n .

Definition 7.4. Let m_n be a positive integer. The m_n -th random Hermite polynomial or the random Hermite polynomial of degree m_n , is the random m.s. solution defined through $X_1(t)$ and $X_2(t)$ of problem (7.1) with initial conditions $X(0) = Y_0$ and $\dot{X}(0) = Y_1$ given by (7.15), being A the discrete r.v. taking the finite number of even integer values $2m_j$, with probabilities $P[A = 2m_j] = p_j > 0$, $1 \leq j \leq n$ with $\sum_{j=1}^n p_j = 1$.

With respect to Definition 7.4 and keeping this notation, it is important to point out that under conditions of Corollary 7.1, a random Hermite polynomial solution can be interpreted as a collection of deterministic Hermite polynomials, which, for each $j : 1 \leq j \leq n$, have a probability p_j of sampling. The degree m_n of the random Hermite polynomial $H_{m_n}(t)$ is the greatest of all degrees corresponding to each (deterministic) Hermite polynomials, but it is not necessary that its sample associated probability p_n be also greater than p_j for all $j : 1 \leq j \leq n$, (see Example 7.3).

Remark 7.1. Note that the situation where A is the discrete r.v. taking all

the even integer values with $P[A = 2j] = 2^{-(j+1)}$, $j = 0, 1, 2, \dots$, then for each j one obtains a sample polynomial solution, but since j lies in the positive integer numbers, the degree of the random polynomial solution cannot be defined according with Definition 7.4.

7.6 Statistical functions of the mean square random power series solution

This section deals with the computation of the main statistical functions of the m.s. solution of (7.1) given by (7.19) and (7.15) such that the mean and variance in terms of the data $E[Y_0]$, $E[Y_1]$, $E[Y_0Y_1]$, $E[(Y_0)^2]$, $E[(Y_1)^2]$ as well as certain moments related to algebraic transformations of the random coefficient A that will be specified later.

First at all, let us take the expectation operator in the random differential equation (7.1). Then by applying property (2.24) one gets

$$\frac{d^2}{dt^2}(\mu_X(t)) - 2t \frac{d}{dt}(\mu_X(t)) + E[AX(t)] = 0. \quad (7.21)$$

Note that, (7.21) is not a suitable equation for computing $\mu_X(t)$ because the term $E[AX(t)]$ cannot be factorized as $A\mu_X(t)$, in general. Nevertheless some methods, like the so-called dishonest method (see, [46], [38, p.148]), accept the above factorization as an alternative to handle the problem of computing the mean of the solution process. Since we shall see right away through examples, our approach avoid the above approximations and it allows us to provide reliable values for the mean and the variance into quite general situations.

In practice, as it occurs in the deterministic framework, it will be unfeasible the computation of the mean through the infinite series given by (7.14)-(7.15).

Then, we will consider the truncation of order N

$$\begin{aligned}
 X_N(t) &= X_0 \left(1 + \sum_{k=0}^N \frac{t^{2k+2}}{(2k+2)!} \prod_{j=0}^k (4j - A) \right) \\
 &+ X_1 \left(t + \sum_{k=0}^N \frac{t^{2k+3}}{(2k+3)!} \prod_{j=0}^k (4j + 2 - A) \right).
 \end{aligned} \tag{7.22}$$

Henceforth, we will assume that r.v. A is independent of initial conditions $X(0) = Y_0$ and $\dot{X}(0) = Y_1$ (note that from an applied point of view, this hypothesis is realistic), then taking the expectation operator in (7.22) one gets

$$\begin{aligned}
 \mu_{X_N}(t) &= E[Y_0] \left(1 + \sum_{k=0}^N \frac{t^{2k+2}}{(2k+2)!} E \left[\prod_{j=0}^k (4j - A) \right] \right) \\
 &+ E[Y_1] \left(t + \sum_{k=0}^N \frac{t^{2k+3}}{(2k+3)!} E \left[\prod_{j=0}^k (4j + 2 - A) \right] \right).
 \end{aligned} \tag{7.23}$$

Depending whether r.v. A is discrete, with probability mass function $p_A(a)$, or continuous, with probability density function $f_A(a)$, the expectation terms involved in (7.23) can be computed as follows

$$E \left[\prod_{j=0}^k (4j - A) \right] = \begin{cases} \sum_{a:p_A(a)>0} \prod_{j=0}^k (4j - a) p_A(a), \\ \int_{-\infty}^{\infty} \prod_{j=0}^k (4j - a) f_A(a) da, \end{cases}$$

and

$$E \left[\prod_{j=0}^k (4j + 2 - A) \right] = \begin{cases} \sum_{a:p_A(a)>0} \prod_{j=0}^k (4j + 2 - a) p_A(a), \\ \int_{-\infty}^{\infty} \prod_{j=0}^k (4j + 2 - a) f_A(a) da. \end{cases}$$

Now taking into account the expression (2.7) for computing the variance of the truncated solution process, now we only require to calculate $E \left[(X_N(t))^2 \right]$. In

order to save executing time in the computer, it is convenient considering the following relationship

$$E \left[(X_N(t))^2 \right] = \sum_{n=0}^N E \left[(X_n)^2 \right] t^{2n} + 2 \sum_{n=1}^N \sum_{m=0}^{n-1} E [X_n X_m] t^{n+m}.$$

With the usual convection $\prod_{i=u}^v f(i) = 1$ if $v < u$, the terms involved in the two previous sums can be computed as follows:

$$E [X_n X_m] = \frac{1}{n! m!} \begin{cases} E \left[(Y_0)^2 \right] E \left[P_1 \left(\frac{n-2}{2} \right) P_1 \left(\frac{m-2}{2} \right) \right] & \text{if } \begin{matrix} n=0, 2, 4, \dots, \\ m=0, 2, 4, \dots, \end{matrix} \\ E [Y_0 Y_1] E \left[P_1 \left(\frac{n-2}{2} \right) P_2 \left(\frac{m-3}{2} \right) \right] & \text{if } \begin{matrix} n=0, 2, 4, \dots, \\ m=1, 3, 5, \dots, \end{matrix} \\ E [Y_0 Y_1] E \left[P_2 \left(\frac{m-3}{2} \right) P_1 \left(\frac{n-2}{2} \right) \right] & \text{if } \begin{matrix} n=1, 3, 5, \dots, \\ m=0, 2, 4, \dots, \end{matrix} \\ E \left[(Y_1)^2 \right] E \left[P_2 \left(\frac{n-3}{2} \right) P_2 \left(\frac{m-3}{2} \right) \right] & \text{if } \begin{matrix} n=1, 3, 5, \dots, \\ m=1, 3, 5, \dots, \end{matrix} \end{cases} \quad (7.24)$$

where we have denoted

$$P_1(k) = \prod_{j=0}^k (4j - A), \quad P_2(k) = \prod_{j=0}^k (4j + 2 - A),$$

and the expectations appearing in (7.24) can be computed as follows

$$E [P_1(k) P_1(l)] = \begin{cases} \sum_{a: p_A(a) > 0} \prod_{j=0}^k (4j - a) \prod_{j=0}^l (4j - a) p_A(a), \\ \int_{-\infty}^{+\infty} \prod_{j=0}^k (4j - a) \prod_{j=0}^l (4j - a) f_A(a) da, \end{cases}$$

$$E [P_1(k) P_2(l)] = \begin{cases} \sum_{a: p_A(a) > 0} \prod_{j=0}^k (4j - a) \prod_{j=0}^l (4j + 2 - a) p_A(a), \\ \int_{-\infty}^{+\infty} \prod_{j=0}^k (4j - a) \prod_{j=0}^l (4j + 2 - a) f_A(a) da, \end{cases}$$

$$\begin{aligned}
 E [P_2(k)P_1(l)] &= \begin{cases} \sum_{a:p_A(a)>0} \prod_{j=0}^k (4j+2-a) \prod_{j=0}^l (4j-a) p_A(a), \\ \int_{-\infty}^{+\infty} \prod_{j=0}^k (4j+2-a) \prod_{j=0}^l (4j-a) f_A(a) da, \end{cases} \\
 E [P_2(k)P_2(l)] &= \begin{cases} \sum_{a:p_A(a)>0} \prod_{j=0}^k (4j+2-a) \prod_{j=0}^l (4j+2-a) p_A(a), \\ \int_{-\infty}^{+\infty} \prod_{j=0}^k (4j+2-a) \prod_{j=0}^l (4j+2-a) f_A(a) da. \end{cases}
 \end{aligned}$$

Taking into account Lemma 2.3 as well as the m.s. convergence of random series given by (7.19) and (7.15), the convergence of the mean and the variance of the truncated solution (7.22) to the corresponding exact values are warranted under the hypotheses of Theorem 7.1.

7.7 Examples

In this section we provide several illustrative examples. The results obtained to approximate the mean and the variance by means of the series method presented in this paper are compared with respect to the corresponding ones provided by the dishonest and Monte Carlo approach.

Example 7.1. *Let us consider a random differential equation of the form (7.1) where A is a Beta r.v. with parameters $\alpha = 2$ and $\beta = 3$, i.e., $A \sim \text{Be}(\alpha = 2; \beta = 3)$ and the initial conditions Y_0 and Y_1 are (positive) correlated r.v.'s such that $E[Y_0Y_1] = 0.3$ and, $E[Y_0] = 1$, $E[(Y_0)^2] = 2$, $E[Y_1] = 2$, $E[(Y_1)^2] = 5$. Note that r.v. A satisfies conditions of Theorem 7.1 since it takes values on a bounded interval. Then the m.s. solution of model (7.1) with initial conditions Y_0 and Y_1 is given by (7.19) and (7.15). Table 7.2 shows the expectation of the truncated solution s.p. for different values of the*

truncation order N (denoted by $\mu_{X_N}(t)$) at different values of the time parameter t as well as the corresponding values obtaining by the dishonest ($\mu_X^d(t)$) and Monte Carlo methods ($\tilde{\mu}_X^m(t)$) by using m simulations. One observes

t	$\mu_{X_5}(t)$	$\mu_{X_{10}}(t)$	$\mu_{X_{20}}(t)$	$\mu_X^d(t)$	$\tilde{\mu}_X^m(t)$ $m = 5 \times 10^4$	$\tilde{\mu}_X^m(t)$ $m = 10^5$
0.00	1.00000	1.00000	1.00000	1.00000	0.99826	0.99995
0.25	1.49575	1.49575	1.49575	1.49574	1.49502	1.49611
0.50	2.01775	2.01775	2.01775	2.01761	2.01814	2.01855
0.75	2.63184	2.63184	2.63184	2.63099	2.63362	2.63315
1.00	3.44890	3.44891	3.44891	3.44545	3.45263	3.45088
1.25	4.69234	4.69267	4.69267	4.68111	4.69942	4.69553
1.50	6.86741	6.87373	6.87373	6.83843	6.88590	6.87801
1.75	11.2081	11.2854	11.2855	11.1806	11.30880	11.29240
2.00	20.8513	21.5482	21.5514	21.2341	21.60060	21.56350

Table 7.2: Comparison of the mean for different methods in Example 7.1

that for values of t near of the origin (where the initial conditions are established and the series solution s.p. is centered), the approximations obtained by the method proposed in this chapter coincide for different truncation orders of the series solution. In fact, the approximations are accurate for $N = 5$ about $t = 0$. These values differ from the corresponding ones obtained by the dishonest method. Regarding approximations obtained by means of the Monte Carlo method it is worthwhile pointing out that they improve as the number m of simulations increases, and in general, they provide better approximations than those obtained by the dishonest method. Table 7.3 compares the values of variance for the truncation method with respect to the Monte Carlo method. In order to show that accurate approximations of the variance require greater values of N , in Table 7.3 we have considered values of N that differ from those we have taken in Table 7.2. Finally, note that as A is a continuous r.v., in this case the initial value problem does not have random (Hermite) polynomial solutions.

t	$\text{Var}[X_{10}(t)]$	$\text{Var}[X_{20}(t)]$	$\text{Var}[X_{50}(t)]$	$\widetilde{\text{Var}}_X^m(t)$ $m = 5 \times 10^4$	$\widetilde{\text{Var}}_X^m(t)$ $m = 10^5$
0.00	1.00000	1.00000	1.00000	1.002	0.997809
0.25	1.03943	1.03943	1.03943	1.04321	1.03741
0.50	1.18425	1.18425	1.18425	1.18964	1.18298
0.75	1.54404	1.54432	1.54432	1.55086	1.54517
1.00	2.44517	2.45458	2.45458	2.4609	2.461183
1.25	4.84512	5.01194	5.01197	5.01346	5.035022
1.50	11.2801	13.3658	13.3675	13.3465	13.4437
1.75	24.477	46.2562	46.3373	46.2146	46.6174
2.00	105.199	206.118	209.017	208.359	210.273

Table 7.3: Comparison of the variance for different methods in Example 7.1

Example 7.2. In this example we take advantage of the so-called truncation method (see [48]) to deal with a r.v. A that neither satisfy condition (7.6) nor condition (7.20). Let us consider model (7.1) where A is a Gaussian r.v., $A \sim N(\mu = 5; \sigma = 1)$, and the initial conditions Y_0 and Y_1 are uncorrelated r.v.'s such that $E[Y_0] = 1$, $E[(Y_0)^2] = 2$, $E[Y_1] = 2$, $E[(Y_1)^2] = 5$. Note that r.v. A has unbounded codomain. In order to overcome this difficulty, we will consider the truncation of this r.v. on the interval $[\mu - 3\sigma, \mu + 3\sigma] = [2, 8]$ that will contain all the values of A with probability 0.997. The probability density function associated to the new censored r.v., say B , is

$$f_B(b) = \frac{\exp\left(-\frac{1}{2}(b-5)^2\right)}{\int_2^8 \exp\left(-\frac{1}{2}(x-5)^2\right) dx}, \quad 2 \leq b \leq 8.$$

In this way, B satisfies hypotheses of Theorem 7.1 since it takes values on a bounded interval. Table 7.4 compares approximations of the expectation of the solution s.p. computed by the truncation series, dishonest and Monte Carlo methods. For the truncation method, Table 7.4 only shows results for $N = 5, 10$ because for greater values the obtained results do not change.

t	$\mu_{X_5}(t)$	$\mu_{X_{10}}(t)$	$\mu_X^d(t)$	$\tilde{\mu}_X^m(t)$ $m = 5 \times 10^4$	$\tilde{\mu}_X^m(t)$ $m = 10^5$
0.00	1.00000	1.00000	1.00000	0.996945	0.99917
0.25	1.32907	1.32907	1.32889	1.3267	1.3286
0.50	1.26473	1.26473	1.26175	1.26405	1.26507
0.75	0.7451	0.745101	0.729061	0.747267	0.746903
1.00	-0.271569	-0.271568	-0.324665	-0.265094	-0.267331
1.25	-1.80638	-1.80635	-1.93991	-1.7937	-1.79838
1.50	-3.85904	-3.85876	-4.13681	-3.83768	-3.84559
1.75	-6.41039	-6.4087	-6.90081	-6.37705	-6.38944
2.00	-9.43986	-9.43401	-10.1448	-9.39179	-9.4106

Table 7.4: Comparison of the mean for different methods in Example 7.2

Table 7.5 shows approximations for the variance obtained by means of the truncation method as well as Monte Carlo method. Analogous comments we made in Example 7.1 can be done again.

Example 7.3. Let us consider model (7.1) where A is a discrete r.v. whose values and associated probabilities are specified in the two first columns of Table 7.6. In this case, the random differential equation (7.1) with initial conditions $X(0) = Y_0$ and $\dot{X}(0) = Y_1$ has random (Hermite) polynomial solutions of degree 9. For instance, with probability 1/8 one will present as (sample) solution the polynomial $1 - 2t^2$ generated from the term $X_1(t)$ (given by (7.15)) that defines the general solution (7.14). Apart from a constant, this polynomial coincides with the deterministic Hermite polynomial, given by $4t^2 - 2$. In fact, it is enough taking as initial condition Y_0 a r.v. concentrated in the point -2 , i.e., $P[Y_0 = -2] = 1$ for obtaining the Hermite (sample) polynomial of degree 2, $4t^2 - 2$. In the same way, with probability 1/4 we will obtain as solution the polynomial $t - 2/3t^3$ which is generated from the term $X_2(t)$ given by (7.15). For an initial condition Y_1 such that $P[Y_1 = -12] = 1$ one obtain just the deterministic Hermite polynomial of degree 3, $8t^3 - 12t$. Following an analogous

t	$\text{Var}[X_{10}(t)]$	$\text{Var}[X_{20}(t)]$	$\text{Var}[X_{50}(t)]$	$\widetilde{\text{Var}}_X^m(t)$ $m = 5 \times 10^4$	$\widetilde{\text{Var}}_X^m(t)$ $m = 10^5$
0.00	1.00000	1.00000	1.00000	0.99837	0.998648
0.25	0.774331	0.774331	0.774331	0.77285	0.774595
0.50	0.37752	0.37752	0.37752	0.379991	0.380405
0.75	0.54182	0.541811	0.541811	0.550826	0.546854
1.00	2.10416	2.10396	2.10396	2.11649	2.10742
1.25	5.48791	5.48669	5.48669	5.49385	5.48397
1.50	10.4335	10.4463	10.4463	10.454	10.4573
1.75	17.6092	18.0106	18.0117	18.1235	18.1765
2.00	38.0585	43.7945	43.8404	44.5363	44.7127

Table 7.5: Comparison of the variance for different methods in Example 7.2

procedure, we fill the Table 7.6. We stress the degree of the random polynomial solution is 9 and, the sample polynomial solution associated to this degree is 1/6, whereas the sample polynomial solution of degree 5 has the greater probability of occurrence, 1/3.

7.8 Conclusions

In this chapter we have constructed a m.s. power series solution of the so-called random Hermite differential equation incorporating uncertainty through both, the initial conditions and the coefficient A (see, expression (7.1)). For that, we have introduced previously the concept of random polynomial and fundamental set of solutions stochastic processes of a second-order random linear differential equation. In order to show that the constructed random power series solution converges in the m.s. sense, we have required to manage the norm of a product of r.v.'s. For this task, the establishment of inequal-

A	$P[A]$	initial condition	Hermite sample polynomial
4	1/8	$Y_0 = -2$	$-2 + 4t^2$
6	1/4	$Y_1 = -12$	$-12t + 8t^3$
10	1/3	$Y_1 = 120$	$120t - 160t^3 + 32t^5$
12	1/8	$Y_0 = -120$	$-120 + 720t^2 - 480t^4 + 64t^6$
18	1/6	$Y_1 = 30240$	$30240t - 80640t^3 + 48384t^5 - 9216t^7 + 512t^9$

Table 7.6: Hermite sample polynomials and its associated probabilities in Example 7.3

ity (7.5) was crucial. An exponential growth condition involving the absolute moments with respect to the origin of r.v. A has been assumed in order to demonstrate the m.s. convergence of the random power series solution. An important fact related to the obtained results is that every r.v. A with bounded codomain, such as Binomial, Hypergeometric, Uniform, Beta, satisfies condition above; otherwise accurate approximations of r.v. A can be achieved by using the so-called truncation method. As a consequence, relevant r.v.'s such as Gaussian, Gamma or Poisson can also play the role of r.v. A . We have also determined conditions under which the random power series solution is finite and generates polynomial solutions referred as to random Hermite polynomial solutions. By considering a truncation of the power series solution s.p., approximations of the mean and covariance matrix functions of the solution s.p. have been computed explicitly in terms of the data. Finally, we have shown several examples that illustrate the different concepts and theoretical results previously presented.

References

- [1] A. Adomian (1994). Solving Frontier Problems of Physics: The Decomposition Method, Kluwer, Boston.
- [2] D.H. Anderson (1983). Compartmental Modeling and Tracer Kinetics, Lecture Notes in Biomathematics, vol. 50, Springer-Verlag, Berlin.
- [3] L. Arnold (1974). Stochastic Differential Equations. Theory and Applications, John Wiley, New York.
- [4] R. Barron, A.G. Ayala (2002). El método de yuxtaposición de dominios en la solución numérica de ecuaciones diferenciales estocásticas, (in Spanish), Proc. Métodos Numéricos en Ingeniería y Ciencias Aplicadas, (E. Oñate et all eds.) CIMNE, Barcelona, 267–276.
- [5] D.S. Bernstein, D.C. Hyland (1993). Compartmental modeling and second-moment analysis of state space systems, SIAM J. Matrix Anal. Appl. 14, 880–901.
- [6] F. Black, M. Scholes (1973). The pricing of options and corporate liabilities, J. Polit. Econ., 81, 637–659.
- [7] W.E. Boyce (1968). Probabilistic Methods in Applied Mathematics I, Ed. A.T. Bharucha-Reid, Academic Press, New York.
- [8] C.A. Braumann (1999). Variable effort fishing models in random environments, Math. Biosci. 156, 1–19.

References

- [9] G. Calbo, J.C. Cortés, L. Jódar (2008). Random analytic solution of coupled differential models with uncertain initial condition and source term, *Comput. Math. Applic.* 56, 785–798.
- [10] G. Calbo, J.C. Cortés, L. Jódar (2010). Mean square power series solution of random linear differential equations, *Comput. Math. Applic.* 59, 559–572.
- [11] G. Calbo, J.C. Cortés, L. Jódar (2010). Random matrix difference models arising in long-term medical drug strategies, *Appl. Math. Comput.* 217, 2149–2161.
- [12] G. Calbo, J.C. Cortés, L. Jódar, L. Villafuerte (2010). Analytic stochastic processes solutions of second-order random differential equations, *Appl. Math. Lett.* 23, 1421–1424.
- [13] R. Cameron, W. Martin (1947). The orthogonal development of nonlinear functionals in series of Fourier– Hermite functionals, *Ann. Math.* 48, 385–392.
- [14] V. Chellaboina, W.M. Haddad, J.M. Bailey, J. Ramakrishnan (2002). On the absence of oscillations in compartmental dynamics systems, *Proc. IEEE Conference on Decision and Control*, Nevada, 1663–1668.
- [15] J.C. Cortés, P. Sevilla-Peris, L. Jódar (2005). Analytic-numerical approximating processes of diffusion equation with data uncertainty, *Comput. Math. Appl.* 49, 1255–1266.
- [16] J.C. Cortés, L. Jódar, L. Villafuerte (2006). A random Euler method for solving differential equations with uncertainties, *Progress in Industrial Mathematics E.C.M.I. 2006 (European Conference on Mathematics for Industry)*, Vol.12, Springer-Verlag, 944–948.
- [17] J.C. Cortés, L. Jódar, L. Villafuerte (2007). Mean square numerical solution of random differential equations: facts and possibilities, *Comput. Math. Appl.* 53, 1098–1106.
- [18] J.C. Cortés, L. Jódar, L. Villafuerte (2009). Random linear-quadratic mathematical models: computing explicit solutions and applications, *Math. Comput. Simulat.* 79, 2076–2090.

- [19] J.C. Cortés, L. Jódar, L. Villafuerte, R.J. Villanueva (2010). Mean square convergent numerical methods for nonlinear random differential equations, *Trans. Comput. Syst. (serie LNCS: Special issue on Computations of Stochastic Nonlinear Systems)*, Springer-Verlag, 1–21.
- [20] J. Dieudonné (1960). *Foundations of Modern Analysis*, Academic Press, New York.
- [21] J.L. Doob (1953). *Stochastic Processes*, John Wiley and Sons Inc. Publications in Statistics, New York.
- [22] S. Elaydi (2005). *An Introduction to Difference Equations*, Third. Ed., UTM Collection, Springer, New York.
- [23] M. El-Tawil (2003). The application of WHEP technique on stochastic partial differential equations, *Internat. J. Differ. Equat. and Applic.* 7, 325–337.
- [24] M. El-Tawil (2005). The approximate solutions of some stochastic differential equations using transformations, *Appl. Math. Computat.* 164, 167–178.
- [25] M. El-Tawil (2008). The Homotopy Wiener-Hermite Expansion and Perturbation Technique (WHEP), *Transact. on Computat. Sci. I. LNCS 4750* (Springer), 159–180.
- [26] M. El-Tawil, W. El-Tahan, A. Hussein (2005). A proposed technique of SFEM on solving ordinary random differential equation, *Appl. Math. Computat.* 161, 35–47.
- [27] M. El-Tawil, G. Mahmoud (2005). The solvability of parametrically forced oscillators using WHEP technique, *Mech. and Mech. Engin.* 3, 181–188.
- [28] G.S. Fishman (1996). *Monte Carlo: Concepts, Algorithms and Applications*, Springer-Verlag, New York.
- [29] T.C. Gard (1988). *Introduction to Stochastic Differential Equations*, Marcel Dekker, New York.
- [30] R. Ghanem, P. Spanos (1991). *Stochastic Finite Elements: Spectral Approach*, Springer, New York.
- [31] M. Gibaldi, D. Perrier (1975). *Pharmacokinetics*, Marcel Dekker, New York.

References

- [32] K. Godfrey (1975). *Compartmental Models and Their Applications*, Marcel Dekker, New York.
- [33] G. Golub, C.F. Van Loan (1996). *Matrix Computations*, Third Ed., The Johns Hopkins University Press, Baltimore MD.
- [34] G.R. Grimmett, D. Stirzaker (2000). *D.R. Stirzaker, Probability and Random Processes*, Clarendon Press, Oxford.
- [35] L. Grüne, P.E. Kloeden (2001). Pathwise approximation of random differential equations, *BIT* 41, 711–721.
- [36] W.M. Haddad, V. Chellaboina, E. August (2001). Stability and dissipativity for discrete-time nonnegative and compartmental dynamical system, *Proc. IEEE Conference on Decision and Control*, Florida, 4236–4241.
- [37] J.D. Hamilton (1994). *Time Series Analysis*, Princeton University Press, New Jersey.
- [38] D. Henderson, P. Plaschko (2006). *Stochastic Differential Equations in Science and Engineering*, World Scientific, Singapore.
- [39] N. Hofmann, P. Mathé (1997). On quasi-Monte Carlo methods of stochastic differential equations, *Math. Comput. Simulat.* 38, 63–68.
- [40] J.A. Jacquez (1985). *Compartmental Analysis in Biology and Medicine*, 2nd. ed., University of Michigan Press, Michigan.
- [41] A. Jahedi, G. Ahmadi (1983). Application of Wiener-Hermite expansion to nonstationary random vibration of a Duffing oscillator, *Trans. ASME* 50, 436–442.
- [42] L. Jódar, J.C. Cortés, J.L. Morera (1996). Construction and computation of variable coefficient differential problems, *Comput. Math. Applic.* 32, 41–50.
- [43] L. Jódar, J.C. Cortés, L. Villafuerte (2007). Reliable numerical solutions of random differential systems, *Proc. (in CD) 4th Internat. Congress, 2nd National Congress Numerical Methods in Engineering and Applied Sciences*, (M.C. Suárez et al eds.) U.M.S.N.H., México.
- [44] S. Kadry (2007). On the generalization of probabilistic transformation method, *Appl. Math. Computat.* 190, 1284–1289.

-
- [45] B. Kegan, R.W. West (2005). Modeling the simple epidemic with deterministic differential equations and random initial conditions, *Math. Biosci.* 195, 179–193.
- [46] J.B. Keller (1963). Stochastic equations and wave propagation in random media, *Procc. Symp. Appl. Math. of A.M.S. Providence Rhode Island* 16, 145–170.
- [47] E. Kloeden, E. Platen (1992). *Numerical Solution of Stochastic Differential Equations*, Springer, Berlin.
- [48] M. Loève (1963). *Probability Theory*, Van Nostrand–Reinhold, New Jersey.
- [49] B. Øksendal (1998). *Stochastic Differential Equations. An Introduction with Applications*, 5th. ed., Springer-Verlag, Berlin Heidelberg.
- [50] J.M. Ortega (1990). *Numerical Analysis. A Second Course*, SIAM, Philadelphia.
- [51] A. Papoulis (1991). *Probability, Random Variables and Stochastic Processes*, McGraw Hill, New York.
- [52] P.A. Ruymgaart, T.T. Soong (1971). A sample theory of ordinary stochastic differential equations, *Hawaii Internat. Conf. System Sci.*, 2nd, Honolulu, Hawaii, 603-606. Western Periodiclas Co., Honolulu.
- [53] P.A. Ruymgaart, T.T. Soong (1971). A sample treatment of Langevin-type stochastic differential equation, *J. Math. Anal. Appl.* 34, 325–338.
- [54] D. Stanescu, B. Chen-Charpentier, B.J. Jensen, P.J.S. Colberg (2009). Random coefficient differential models of growth of anaerobic photosynthetic bacteria, *Electr. Transact. Numer. Anal.* 34, 44–58.
- [55] T.T. Soong, L.J. Skidmore (1966). Unified guidance analysis in design of space trajectories, *. Spacecr. Rockets* 3, 98-103.
- [56] T.T. Soong (1973). *Random Differential Equations in Science and Engineering*, Academic Press, New York.
- [57] T.T. Soong, M. Grigoriu (1993). *Random Vibration of Mechanical and Structural Systems*, Prentice Hall, México.
- [58] J.P. Spanos, W. Iwan (1978). The existence and uniqueness of solution generated by equivalent linearization, *Int. J. Nonlinear Mech.* 13, 71–78.
-

References

- [59] J.L. Strand (1968). Stochastic Ordinary Differential Equations, Ph.D. Thesis, Univ. of California, Berkeley, California.
- [60] J.L. Strand (1973). Random ordinary differential equations, *J. Diff. Equat.* 7, 538–553.
- [61] L. Villafuerte (2007). Numerical and Analytical Mean Square Solutions for Random Differential Models , Ph.D. Thesis, Univ. Politécnica de Valencia, Valencia, Spain.
- [62] L. Villafuerte, C.A. Braumann, J.C. Cortés, L. Jódar (2010). Random differential operational calculus: Theory and applications, *Comput. Math. Appl.* 59, 115–125.
- [63] L. Villafuerte, J.C. Cortés, L. Jódar (2010). Constructing power series solutions for random differential models, *Proc. (in CD) 5th Internat. Congress Numerical Methods in Engineering and Applied Sciences*, (M.C. Suárez et all eds.) U.M.S.N.H., México.
- [64] W. Wagner (1988). MonteCarlo evaluation of functionals of solutions of stochastic differential equations. Variance reduction and numerical methods, *Stoch. Anal. Appl.* 6, 447–468.
- [65] R.M. Wilcox, R. Bellmann (1970). On the approximate moment equations of a nonlinear stochastic differential equation, *J. Math. Anal. Appl.* 29, 384–391.
- [66] E. Wong, B. Hajek (1985). *Stochastic Processes in Engineering System*, Springer Verlag, New York.

List of Tables

2.1	Some relevant r.v.'s and their mean and variance	32
2.2	Some relevant r.v.'s and their moments with respect to the origin	32
3.1	Variance of concentrations into each compartment at different periods for Example 3.2	81
4.1	Comparison between exact expectation $E[\vec{X}(t)]$ given by (4.38) and its approximating values $E[\vec{X}_N(t)]$ given by (4.31), (4.39)-(4.40) for Example 4.3 for different values of the truncation order $N = 5, 10, 15$ at some points of the interval $[0, 2]$	103
4.2	Approximate values of the expectation for Example 4.4 by using the truncation series method ($E[\vec{X}_N(t)]$) with $N = 500$ and the Monte Carlo method ($\tilde{\mu}_X^m(t)$) with $m = 10000$ and $m = 50000$ simulations for different points t on the interval $[0, 2]$	104
4.3	Approximate values of the variance for Example 4.4 by using the truncation series method ($\text{Diag}(\mathbf{\Lambda}_{\vec{X}_N(t), \vec{X}_N(t)})$) with $N = 500$ and the Monte Carlo method ($\text{Diag}(\mathbf{\Lambda}_{\vec{X}(t), \vec{X}(t)}^m)$) with $m = 10000$ and $m = 50000$ simulations for different points t on the interval $[0, 2]$	106
5.1	Approximated and exact expectation for Example 5.2 with $A \sim N(0; 1)$	128

List of Tables

5.2	Approximated and exact variance for Example 5.2 with $A \sim N(0; 1)$	129
5.3	Approximate and exact expectation for Example 5.3 with $A \sim \text{Ga}(r = 1; a = 2)$	131
5.4	Approximate and exact variance for Example 5.3 with $A \sim \text{Ga}(r = 1; a = 2)$	132
5.5	Comparison between approximate expectations obtained from the proposed method and dishonest one for Example 5.4	133
5.6	Approximate variance for $\gamma = 0.5$ at $t = 0, 0.1, 0.2, \dots, 0.9, 1$ for different values of the truncation order $N = 3, 5, 10$ in Example 5.4	135
5.7	Approximate variance for different values of γ , $N = 20$ at values $t = 0.25, 0.50, 0.75$ in Example 5.4	136
6.1	Comparison between the exact expectation and its approximations obtained from series truncation, dishonest and Monte Carlo methods for Example 6.3 with $A \sim \text{Be}(\alpha = 2; \beta = 1)$. . .	147
7.1	Probabilities of generating random polynomial (sample) solutions when A is a Binomial r.v. of parameters n and p	162
7.2	Comparison of the mean for different methods in Example 7.1 .	167
7.3	Comparison of the variance for different methods in Example 7.1	168
7.4	Comparison of the mean for different methods in Example 7.2 .	169
7.5	Comparison of the variance for different methods in Example 7.2	170
7.6	Hermite sample polynomials and its associated probabilities in Example 7.3	171

List of Figures

3.1	A three-compartment model in chemical kinetics. Note that the values on arrows connecting compartments indicate the rate of transfer (proportional to the concentration in the compartment). Whereas the values indicated on arrows coming from or going to the exterior of compartments indicate inputs or outputs to the system (they are not proportional to the concentration of the corresponding compartment)	79
3.2	Plotting $E[\vec{X}_n]$ for $n = 0, 1, 2, \dots, 20$, for the Example 3.2 . . .	81
5.1	Exact and approximate mean for Example 5.1	125
5.2	Exact and approximate variance for Example 5.1	126
6.1	Comparison between the exact standard deviation and its approximations obtained from series truncation (with order of truncation $N = 15$) and Monte Carlo (with 10^6 simulations) methods for Example 6.3 with $A^2 \sim \text{Be}(\alpha = 2; \beta = 1)$	148

