Stochastic Analysis with Applications to Dynamical Systems

by

Themistoklis P. Sapsis

Diploma Thesis

Supervisor: G. A. Athanassoulis, Professor NTUA

Athens
September 2005
Preface

This thesis deals with the analysis of multidimensional nonlinear dynamical systems subject to general external stochastic excitation. The method used to address this problem is based on the Characteristic Functional. As it is known, the characteristic functional includes the full probability structure of a stochastic process. Hence the knowledge of the characteristic functional induces the knowledge of every statistical quantity related with the stochastic process. For Stochastic Dynamical Systems, the joint Characteristic Functional of the response and the excitation is governed by linear Functional Differential Equations, i.e. differential equations with Volterra derivatives. This kind of equations appeared for the first time at the paper “Statistical Hydromechanics and Functional Calculus” by E. Hopf (1952) in the context of the analysis of Turbulence. Even though such an equation contains the full probabilistic structure of the system response cannot be a straightforward method of solution of the problem. Apparently no practical methods have been found for the efficiently solution of this type of equations.

The present work deals with two very important issues, concerning Functional Differential Equations that describe Stochastic Dynamical Systems. First, we show that apparently all known equations partially describing the probability structure of the response under specific assumptions (such as Fokker-Planck Equation, Liouville Equation, Moment Equations, etc) can be derived directly through the Functional Differential Equation. Additionally, we derive some new (to the best our knowledge) Partial Differential Equations governing the characteristic functions of the response for systems under general stochastic excitation. In a second level, a new method is proposed for the solution of Functional Differential Equations. This method is related with the concepts of Kernel Probability Measures and Kernel Characteristic Functionals in infinite dimensional spaces, which are defined and discussed extensively. In this way we propose a method for treating the Functional Differential Equations at the infinite dimensional level, without any reduction to PDE’s. A simplification of the proposed method is implemented and a numerical example is presented.

This thesis is divided into five chapters.

In Chapter 1 we briefly present the basic background needed for our study. We recall some basic elements of set theory as well as the notions of probability measure and probability space. In the last section of the chapter we review the concept of random variable and give a summary of some results concerning their characteristic properties.

The first part of Chapter 2 presents the concept of stochastic process as well as an extensive discussion concerning the Kolmogorov Theorem. This discussion will be very useful for Chapter 3 where we will define the stochastic process using the measure theoretic approach. The second part of chapter 2 deals with some classifications of the stochastic processes. Stochastic processes are categorized with respect to their memory, their regularity and their ergodic properties.

Chapter 3 deals with the stochastic processes through the measure theoretic approach. In the first two sections we extensively discuss the properties of a probability measure defined on a Banach space as well as its finite dimensional produced measures. In Sections 3.3-3.5 the characteristic functional is defined and studied. Especially we study the connection of characteristic functional with more conventional statistical quantities describing the stochastic processes, such as moments,
characteristic functions etc. Finally, in the three last sections of the chapter, we study specific cases of probability measures/characteristic functionals, such as the Gaussian and the Poisson. Special attention has been given to the characteristic functional introduced by V. Tatarskii, which induces as special (or asymptotic) cases almost every known characteristic functional.

Before we proceed to the general study of Stochastic Dynamical Systems, we present the topic of mean-square calculus. Although classical, it’s very important for a complete description of Stochastic Differential Equations. Among the topics we examine, of special interest is the presentation for the Integral of stochastic processes where we distinguish the case of integration of a stochastic process with respect to another independent stochastic process and the case of integration over Martingales with non-anticipative dependence. At the end of the chapter we present some results and criteria for the sample path properties of stochastic process.

Finally in the beginning of Chapter 5 the notion of a stochastic differential equation is defined and some general theorems concerning the existence and uniqueness of its solutions are proved. Moreover, necessary conditions for bounded of all moments of the probability measure describing the system response are given. Section 5.2 deals with the formulation of functional differential equations for the joint characteristic functional describing (jointly) the probability structure of the excitation and the response. The formulation concerns stochastic systems described by nonlinear ODE’s with polynomial nonlinearities. We distinguish between the case of a m.s.-continuous stochastic excitation and an orthogonal-increment excitation. Before the general results we present two specific systems, the Duffing oscillator and the van der Poll oscillator. In the next section 5.3 we prove how the general functional differential equation describing the probability structure of a system can produce a family of partial differential equations for the characteristic functions of various orders. Special cases are the Fokker-Planck equation, as well as the Liouville equation. Of great importance are the partial differential equations proved at Section 5.3.4 that describe the probability structure of the system for the case of m.s.-continuous excitation. In the last Section 5.5 we introduce the notion of the kernel characteristic functional that generalizes the corresponding idea of kernel density functions in infinite dimensional spaces. We prove that the Gaussian characteristic functional has the kernel property, thus being an efficient tool for the study of FDEs. The proof is based on existent theorems for extreme values of stochastic processes. In Section 5.4.4 we discuss how the kernel characteristic functionals can be efficiently used for the solution of functional differential equations. The basic idea is the localization in the physical phase space domain (probability measure) expressed in terms of the characteristic functional. It should be noted that the Gaussian measures are of great importance in this direction since most (if not all) of the existing analytical results concerning infinite dimensional integrals pertain Gaussian measures. In this way we derive a set of partial differential equations that governs the functional parameters of the kernel characteristic functionals. Finally, in Section 5.4.5, we a simplified set of equations is derived in the basis of specific assumptions concerning the probability interchange between probability kernels. The latter equations are applied for the numerical study of a simple nonlinear ODE.
Acknowledgments

Acknowledgments are owed to everyone who helped and assisted me, each one on his/her own way, during the elaboration of this thesis. However, let me take this opportunity to thank some people individually.

First, I would like to express my deep gratitude to the supervisor of this work, Prof. Gerassimos A. Athanassoulis, for his guidance, encouragement and support throughout the whole duration of my studies at the school of Naval Architecture and Marine Engineering of NTUA. I would like to thank him for bringing me in touch, through his inspiring teaching, with various mathematical problems and especially the examined one. This thesis could never have been written or even conceived without all the stimulating discussions that we had the last four years.

The valuable contribution of Mr. Panagiotis Gavriliadis, PhD candidate, through the critical discussions as well as of Mr. Agissilaos Athanassoulis, PhD candidate, through his stimulating comments that led to a number of improvements, are greatly appreciated.

Finally I have to thank my family Panagiotis, Euaggelia and Pantelis Sapsis as well as Kiriaki Kofiani, student of NTUA. This thesis could never have been taken the current form without their continuous and unconditional love and support. This thesis is dedicated to them with immense gratitude.

Th. Sapsis

September 2005
Contents

Preface iii

Acknowledgments v

Contents vii

List of Symbols xi

Chapter 1: BACKGROUND CONCEPTS FROM PROBABILITY AND ANALYSIS

1.1. Elements of set theory 2
    1.1.1. Set operations 2
    1.1.2. Borel Fields 4

1.2. Axioms of Probability – Definition of Probability Measure 5

1.3. Random Variables 7
    1.3.1. Distribution Functions and Density Functions 7
    1.3.2. Moments 9
    1.3.3. Characteristic Functions 10
    1.3.4. Gaussian Random Variables and the Central Limit Theory 13
    1.3.5. Convergence of a Sequence of Random Variables 16
        1.3.5.a. Mean-Square Convergence 16
        1.3.5.b. Convergence in Probability 17
        1.3.5.c. Almost Sure Convergence 17
        1.3.5.d. Convergence in Law or Distribution 17
        1.3.5.e. Relationships Between the Four Modes of Convergence 17

1.4. References 18
## Chapter 2: STOCHASTIC PROCESSES. DEFINITION AND CLASSIFICATION

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1. Basic Concepts of Stochastic Processes</td>
<td>20</td>
</tr>
<tr>
<td>2.1.1. Definition of a Continuous-parameter Stochastic Process</td>
<td>20</td>
</tr>
<tr>
<td>2.1.2. Moments</td>
<td>23</td>
</tr>
<tr>
<td>2.1.3. Characteristic Functions</td>
<td>24</td>
</tr>
<tr>
<td>2.2. Classification of Stochastic Process</td>
<td>26</td>
</tr>
<tr>
<td>2.2.1. Classification of Stochastic Process based upon Memory</td>
<td>26</td>
</tr>
<tr>
<td>2.2.1.a. Purely Stochastic Process - Generalized Stochastic Process</td>
<td>26</td>
</tr>
<tr>
<td>2.2.1.b. Markov Process</td>
<td>29</td>
</tr>
<tr>
<td>2.2.1.c. Independent Increment Process</td>
<td>31</td>
</tr>
<tr>
<td>2.2.1.d. Martingales</td>
<td>38</td>
</tr>
<tr>
<td>2.2.2. Classification of Stochastic Process based upon Regularity</td>
<td>41</td>
</tr>
<tr>
<td>2.2.2.a. Stationary and Wide-Sense Stationary Processes</td>
<td>41</td>
</tr>
<tr>
<td>2.2.2.b. Spectral Densities and Correlation Functions</td>
<td>42</td>
</tr>
<tr>
<td>2.2.3. Classification of Stochastic Process based upon Ergodicity</td>
<td>45</td>
</tr>
<tr>
<td>2.2.4. Gaussian Process</td>
<td>47</td>
</tr>
<tr>
<td>2.3. References</td>
<td>49</td>
</tr>
</tbody>
</table>

## Chapter 3: PROBABILITY MEASURES AND THEIR CHARACTERISTIC FUNCTIONAL IN HILBERT (FUNCTION) SPACES

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1. Probability Measures in Hilbert Spaces and their finite-dimensional relatives</td>
<td>52</td>
</tr>
<tr>
<td>3.1.1. Stochastic Processes as Random variables defined on Polish Spaces</td>
<td>52</td>
</tr>
<tr>
<td>3.1.2. Cylindric Sets</td>
<td>53</td>
</tr>
<tr>
<td>3.1.3. Restriction of a probability measure over a Polish space, to finite-dimensional subspaces</td>
<td>54</td>
</tr>
<tr>
<td>3.1.4. Conditions for the existence of probability measures over a Polish space</td>
<td>55</td>
</tr>
<tr>
<td>3.2. Cylinder functions and their Integral</td>
<td>59</td>
</tr>
<tr>
<td>3.3. Characteristic Functional of the Probability Measure</td>
<td>61</td>
</tr>
<tr>
<td>3.4. Mean and Covariance Operator for Probability Measures</td>
<td>69</td>
</tr>
<tr>
<td>3.5. Moments from the Characteristic Functional</td>
<td>72</td>
</tr>
<tr>
<td>3.6. Characteristic Functions from the Characteristic Functional</td>
<td>75</td>
</tr>
<tr>
<td>3.7. Gaussian Measure</td>
<td>77</td>
</tr>
</tbody>
</table>
3.8. Tatarskii Characteristic Functional 80
3.9. Characteristic Functionals derived from Tatarskii Functional 85
3.10. References 90

Chapter 4: STOCHASTIC CALCULUS – PRINCIPLES AND RESULTS

4.1. Mean-Square Calculus of 2nd-Order Process (smooth and non-smooth) 94
   4.1.1. Preliminaries 94
   4.1.2. Mean-Square Convergence 94
   4.1.3. Mean-Square Continuity 96
   4.1.4. Mean-Square Differentiation 97
   4.1.5. Mean-Square Integration I – Integration of a Stochastic Process with respect to another (Independent) Process 100
   4.1.6. Mean-Square Integration II – Integration over Martingale Stochastic Process with non-anticipative dependence 103
      4.1.6.a A formula for changing variables (Ito Formula) 106
      4.1.6.b Stochastic Differentials 107

4.2. Analytical Properties of Sample Functions 110
   4.2.1. Sample Function Integration 110
   4.2.2. Sample Function Continuity 111
   4.2.3. Sample Function Differentiation 113

4.3. References 116

Chapter 5: PROBABILISTIC ANALYSIS OF THE RESPONSES OF DYNAMICAL SYSTEMS UNDER STOCHASTIC EXCITATION

5.1. Stochastic Differential Equations – General Theory 119
   5.1.1. General Problems of the Theory of Stochastic Differential Equations 119
   5.1.2. The Existence and Uniqueness Theorems of Stochastic Differential Equations 124
   5.1.3. Bounds on the Moments of Solutions of Stochastic Differential Equations 128
   5.1.4. Continuous dependence on a parameter of solutions of stochastic equations 129
5.2. Hopf-type Functional Differential Equations (FDE) for the Characteristic Functional of Dynamical Systems

5.2.1. Duffing oscillator
5.2.2. Van Der Poll Oscillator
5.2.3. Non-Autonomous Dynamical Systems under Stochastic Excitation I – Case of m.s. Continuous Excitation
5.2.4. Non-Autonomous Dynamical Systems under Stochastic Excitation II – Case of Independent Increment Excitation

5.3. Reduction of the Hopf-type FDE to PDEs for (joint) characteristic functions or (joint) probability densities

5.3.1. Systems with Random Initial Conditions
5.3.2. Moment Equations from the FDE
5.3.3. Excitation with Independent Increments - Fokker-Planck Equation
5.3.4. Partial Differential Equation for (N+M)-dimensional joint (excitation-response) characteristic functions in the case of general m.s. Continuous Excitation

5.4. Kernel representations for the Characteristic Functional

5.4.1. Exact Formula for Integrals of Special Functionals
5.4.2. Integrals of Variations and of Derivatives of Functionals
5.4.3. Kernel Characteristic Functionals
5.4.4. Superposition of Kernel Characteristic Functionals for the numerical solution of FDEs
5.4.5. Simplification of the set of FDE’s governing the Kernel Characteristic Functionals

5.5. References

Subjects for Future Research

Appendix I : Volterra Derivatives of Gaussian Characteristic Functional

Appendix II : Optimization algorithm

Appendix III : Application on Ship Stability and Capsizing on Random Sea

Index
**List of Symbols**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Omega$</td>
<td>Space of all possible events of a random experiment</td>
</tr>
<tr>
<td>$\mathcal{B}(\Omega)$</td>
<td>$\sigma$ – field of subsets of $\Omega$</td>
</tr>
<tr>
<td>$\mathcal{P}$</td>
<td>Probability measure on the $\mathcal{B}(\Omega)$</td>
</tr>
<tr>
<td>$\mathcal{P}_X$</td>
<td>Induced probability measure of the random variable/stochastic process $X$</td>
</tr>
<tr>
<td>$F_X$</td>
<td>Probability distribution function of the random variable $X(\beta)$</td>
</tr>
<tr>
<td>$f_X$</td>
<td>Probability density function of the random variable $X(\beta)$</td>
</tr>
<tr>
<td>$E^\beta(\cdot)$</td>
<td>Mean value operator</td>
</tr>
<tr>
<td>$\phi_X$</td>
<td>Characteristic function of the random variable $X(\beta)$</td>
</tr>
<tr>
<td>$\mathcal{N}[m,\sigma]$</td>
<td>Gaussian random variable with mean $m$ and variance $\sigma$</td>
</tr>
<tr>
<td>$I \subseteq \mathbb{R}$</td>
<td>Time interval</td>
</tr>
<tr>
<td>$\Pi_H(\cdot)$</td>
<td>Linear projection operator to the subspace $H$</td>
</tr>
<tr>
<td>$D(I)$</td>
<td>Space of real valued functions on $I \subseteq \mathbb{R}$</td>
</tr>
<tr>
<td>$\mathcal{Z}$</td>
<td>Space of random variables/stochastic processes with finite second moment</td>
</tr>
<tr>
<td>$C^\infty_c(I) \subset D(I)$</td>
<td>Space of infinite differentiable functions with compact support</td>
</tr>
<tr>
<td>$\mathcal{B}_\sigma(\Omega)$</td>
<td>current of $\sigma$ – algebras</td>
</tr>
<tr>
<td>$S_\rho$</td>
<td>Sphere with radius $\rho$</td>
</tr>
<tr>
<td>$\mathcal{F}_X$</td>
<td>Characteristic functional of the probability measure $\mathcal{P}_X$</td>
</tr>
<tr>
<td>$\langle \cdot, \cdot \rangle$</td>
<td>Dual product</td>
</tr>
<tr>
<td>$\mathcal{S}$</td>
<td>Space of nuclear operators</td>
</tr>
<tr>
<td>$m_{\rho}$</td>
<td>Mean value element of the probability measure $\mathcal{P}$</td>
</tr>
<tr>
<td>$C_{\rho}$</td>
<td>Correlation operator of the probability measure $\mathcal{P}$</td>
</tr>
<tr>
<td>$\mathcal{C}^2$</td>
<td>Space of twice differentiable functionals</td>
</tr>
<tr>
<td>$\delta \mathcal{F}(x)[z]$</td>
<td>Frechet derivative of the functional $\mathcal{F}$ at the direction $z$</td>
</tr>
<tr>
<td>$\frac{\delta \mathcal{F}}{\delta x(t)}$</td>
<td>Volterra derivative of the functional $\mathcal{F}(\cdot)$ at the point $t$</td>
</tr>
</tbody>
</table>
Chapter 1

Background Concepts from Probability

The mathematical theory of probability and the basic concepts of random variables form the theoretical background of the present work. The reader is assumed to be familiar with these concepts, so only some basic definitions and results are summarized in this chapter mainly for establishing a consistent system of notation, and for later reference.

In Section 1.1 we recall some basic elements of set theory and give the definition of a Borel field that will be used extensively in what follows.

Section 1.2 deals with the notion of the probability measure. We present the three axioms of probability and discuss the completeness of the mathematical description of a random experiment.

Finally in the last section we give the definition of a general random variable and is given the various probability density functions and characteristic functions. All analytical means used for describing its probability structure are defined and briefly discussed.

For more detailed discussions the reader is referred to the literature ATHANASSOULIS, G.A., (Stochastic Modeling and Forecasting of Ship Systems) and FRISTEDT, B. & GRAY. L. (A Modern Approach to Probability Theory).
1.1. Elements of Set Theory

Events and combinations of events occupy a central place in probability theory. The mathematics of events is closely tied to the theory of sets, and this section constitutes a summary of some elements of this theory.

A set is a collection of arbitrary objects. These objects are called elements of the set and they can be of any kind with any specified properties. We may consider, for example, a set of numbers, a set of points or a set of functions.

A set containing no elements is called an empty set and . We distinguish between sets containing a finite number of elements and those having an infinite number. They are called, respectively, finite sets and infinite sets. An infinite set is called enumerable or countable if all its elements can be arranged in such a way that there is a one-to-one correspondence between them and all positive integers; thus, a set containing all positive integers 1,2,... is a simple example of an enumerable set. A nonenumerable or noncountable set is one where the above mentioned one-to-one correspondence cannot be established. A simple example of nonenumerable set is the set of all points on a straight line segment.

If every element of a set \( A \) is also an element of a set \( B \), the set \( A \) is called a subset of \( B \) and this is represented symbolically by

\[
A \subset B
\]

It is clear that an empty set is a subset of any set. In the case where both \( A \subset B \) and \( B \subset A \) hold, the set \( A \) is then equal to \( B \), and we write

\[
A = B
\]

We now give meaning to a particular set we shall called space. In our development, we shall consider only sets which are subsets of a fixed (nonempty) set. This "largest" set containing all elements of all the sets under consideration is called the space, and it is denoted by the symbol \( \Omega \). The class of all the sets in \( \Omega \) is called the space of sets in \( \Omega \). It will be denoted by \( \Omega (\Omega) \).

Consider a subset \( A \) in \( \Omega \). The set of all elements in \( \Omega \) which are not elements of \( A \) is called the complement of \( A \), and we denote it by \( A' \). The following relations clearly hold:

\[
\Omega' = \emptyset, \quad \emptyset' = \Omega, \quad (A')' = A
\]

1.1.1. Set Operations

Let us now consider operations of sets \( A,B,C, \ldots \) which are subsets of space \( \Omega \). We are primarily concerned with addition, subtraction, and multiplication of these sets.

The union or sum of \( A \) and \( B \), denoted by \( A \cup B \), is the set of all elements belonging to \( A \) or \( B \) or both.

The intersection or product of \( A \) and \( B \), written as \( A \cap B \) is the set of all elements which are common to \( A \) and \( B \).
If \( A \cap B = \emptyset \), the sets \( A \) and \( B \) contain no common elements, and we call \( A \) and \( B \) mutually exclusive or disjoint. The symbol " + " shall be reserved to denote the union of two disjoint sets.

The definitions of the union and the intersection can be directly generalized to those involving an arbitrary number (finite or countably infinite) of sets. Thus, the set

\[
A_1 \cup A_2 \cup \ldots \cup A_n = \bigcup_{j=1}^{n} A_j
\]

stands for the set of all elements belonging to one or more of the sets \( A_j, j = 1, 2, \ldots, n \). The intersection

\[
A_1 \cap A_2 \cap \ldots \cap A_n = \bigcap_{j=1}^{n} A_j
\]

is the set of all elements common to all \( A_j, j = 1, 2, \ldots, n \). The sets \( A_j, j = 1, 2, \ldots, n \) are called mutually exclusive if

\[
A_i \cap A_j \quad \text{for every } i, j = 1, 2, \ldots, n
\]  

(3)

It is easy to verify that the union and the intersection operations of sets are associative, commutative, and distributive, that is

\[
(A \cup B) \cup C = A \cup (B \cup C) = A \cup B \cup C
\]  

(4)

\[
A \cup B = B \cup A
\]  

(5)

\[
(A \cap B) \cap C = A \cap (B \cap C) = A \cap B \cap C
\]  

(6)

\[
A \cap B = B \cap A
\]  

(7)

\[
A \cap (B \cup C) = (A \cap B) \cup (A \cap C)
\]  

(8)

Clearly, we also have

\[
A \cup A = A, \quad A \cup \emptyset = A, \quad A \cap \emptyset = \emptyset
\]

\[
A \cup \Omega = \Omega, \quad A \cap \Omega = A, \quad A \cap A^c = \emptyset
\]

Moreover, the following useful relations hold

\[
A \cup (B \cap C) = (A \cup B) \cap (A \cup C)
\]  

(9)

\[
(A \cap B)^c = A^c \cup B^c
\]  

(10)

\[
A \cup B = A \cup A^c \cap B
\]  

(11)

The last two relations are referred to as De Morgan's law.
Finally, we define the *difference* between two sets. The difference $A - B$ is the set of all elements belonging to $A$ but not to $B$. From the definition, we have the simple relations

$A - \emptyset = A$, \hspace{1cm} $\Omega - A = A'$, \hspace{1cm} $A - B = A \cap B'$

### 1.1.2. Borel Fields

**Definition 1.1**: Given a space $\Omega$, a *Borel field*, or a $\sigma$–*field* $\mathcal{B}$ of subsets of $\Omega$ is a class of, in general noncountable, subsets $A_j, j = 1, 2, \ldots$, having the following properties:

1. $\Omega \in \mathcal{B}$
2. If $A \in \mathcal{B}$ then $A' \in \mathcal{B}$
3. If $A_j \in \mathcal{B}, j = 1, 2, \ldots$, then $\bigcup_{j=1}^{\infty} A_j \in \mathcal{B}$. \[\blacksquare\]

The first two properties imply that

$\emptyset - \Omega' \in \mathcal{B}$ \hspace{1cm} (12)

and, with the aid of De Morgan's law, the second and the third properties lead to

$\bigcap_{j=1}^{\infty} A_j = \left( \bigcup_{j=1}^{\infty} A_j' \right)' \in \mathcal{B}$ \hspace{1cm} (13)

A Borel field is thus a class of sets, including the empty set $\emptyset$ and the space $\Omega$, which is closed under all countable unions and intersections of its sets. It is clear, of course, that a class of all subsets of $\Omega$ is a Borel field. However, in the development of the basic concepts of probability, this particular Borel field is too large and impractical. We in general consider the *smallest* class of subsets of $\Omega$ which is a Borel field and it contains all sets and elements under consideration.
1.2. Axioms of Probability – Definition of Probability Measure

The basic concepts of probability theory are centered around the idea of a random experiment $E$, whose outcomes are events. Given a random experiment $E$ the collection of all possible events is called a sample space, whose elements are the simple events. Observable events enter the sample space as its subsets. The definitions of events and sample space provide a framework within which the analysis of events can be performed, and all definitions and relations between events in probability theory can be described by sets and set operations in the theory of sets. Consider the space $\Omega$ of elements $\omega$ and with subsets $A, B, \ldots$

In our subsequent discussion of probability theory, we shall assume that both $\Omega$ and $\emptyset$ are observable events. We also require that the collection of all observable events associated with a random experiment $E$ is a Borel field $\mathcal{B}$ which implies that all events formed through countable unions and intersections of observable events are observable, and they are contained in $\mathcal{B}$.

We now introduce the notion of a probability function. Given a random experiment $E$, a finite number $\mathcal{P}(A)$ is assigned to every event $A$ in the Borel field $\mathcal{B}$ of all observable events. The number $\mathcal{P}(A)$ is a set function and is assumed to be defined for all sets in $\mathcal{U}$. We thus have the following

Definition 2.1 : Let the space $\Omega$ of elements $\omega$ and the Borel field $\mathcal{B}(\Omega)$ generated by the open subsets of $\Omega$. The set function $\mathcal{P} : \mathcal{B}(\Omega) \rightarrow [0,1]$ will be a probability function or probability measure iff the following axioms hold

1. $\mathcal{P}(A) \geq 0$ for every $A \in \mathcal{B}(\Omega)$
2. $\mathcal{P}(\Omega) = 1$
3. for any countable collection of mutually disjoint sets $A_1, A_2, \ldots$, $\mathcal{P}\left(\bigcup_j A_j\right) = \sum_j \mathcal{P}(A_j)$

Hence, Axioms 1-3 define a countably additive and nonnegative set function $\mathcal{P}(A), A \in \mathcal{B}(\Omega)$. The following properties associated with the probability function can be easily deduced from the axioms stated above

$\mathcal{P}(\emptyset) = 0$ (1)

If $A \subseteq B$, then

$\mathcal{P}(A) \leq \mathcal{P}(B)$ (2)

We also see that in general,

$\mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(B) - \mathcal{P}(A \cap B)$ (3)
A mathematical description of a random experiment is now complete. As we have seen above, it consists of three fundamental constituents: A sample space $\Omega$, a Borel field $\mathcal{B}(\Omega)$ of observable events, and the probability function $\mathcal{P} : \mathcal{B}(\Omega) \rightarrow [0,1]$. These three quantities constitute a probability space associated with a random experiment, and it is denoted by $(\Omega, \mathcal{B}, \mathcal{P})$. 
1.3. Random Variables

Consider a random experiment $E$ whose outcomes $\beta$ are elements of $\Omega$ in the underlying probability space $(\Omega, \mathcal{B}, \mathcal{P})$. These outcomes are in general represented by real numbers, functions, or objects of various kinds. In order to construct a model for a random variable, we assume that for all experiments it is possible to assign a real number, function, or other object of various kinds $X(\omega)$ for each $\omega$ following a certain set of rules. The "number" $X(\omega)$ is really a \textit{point function} defined over the domain of the basic probability space. We have the following

\textbf{Definition 3.1} : Let $(\Omega, \mathcal{B}, \mathcal{P})$ be a probability space, $X: \Omega \rightarrow \mathbb{R}$ a \textit{point function} and $\mathcal{B}_X$ a Borel field over $\mathbb{R}$. The mapping $X$ is a random variable, iff is $(\mathcal{B}, \mathcal{B}_X)$-measurable.

In the above definition the probability space $(\Omega, \mathcal{B}, \mathcal{P})$ is modelling a stochastic experiment and the set $\mathcal{B}$ is modelling the quantities, at those we focus our attention. The mapping $X: \Omega \rightarrow \mathbb{R}$ is modelling the procedure of taking the quantity $X(\omega)$. The above value is a realization of the random variable.

\textbf{Definition 3.2} : Let $(\Omega, \mathcal{B}, \mathcal{P})$ be a probability space, $(\mathbb{R}, \mathcal{B}_\mathbb{R})$ a measurable space and $X: \Omega \rightarrow \mathbb{R}$ a random variable. On the Borel field $\mathcal{B}_\mathbb{R}$ of the measurable space $(\mathbb{R}, \mathcal{B}_\mathbb{R})$ we define the \textit{induced probability measure}

\[ \mathcal{P}_X(B) = \mathcal{P}(X^{-1}(B)) \quad \text{for every } B \in \mathcal{B}_\mathbb{R} \]  

(1)

The produced probability space $(\mathbb{R}, \mathcal{B}_\mathbb{R}, \mathcal{P}_X)$, will be called as $X$-\textit{induced probability space}.

1.3.1. Distribution Functions and Density Functions

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

\[ F_X(x) = \mathcal{P}_X(X(\beta) \leq x) \]  

(2)

is defined as the \textit{distribution function} of $X(\beta)$. In the above, the subscript $X$ identifies the r.v. The distribution function always exists. By definition, it is a nonnegative, continuous to the right, and nondecreasing function of the real variable $x$. Moreover, we have

\[ F_X(-\infty) = 0 \quad \text{and} \quad F_X(\infty) = 1 \]  

(3)
A r.v. \( X(\beta) \) is called a continuous r.v. if its associated distribution function is continuous and differentiable almost everywhere. It is a discrete r.v. when the distribution function assumes the form of a staircase with a finite or countably infinite jumps. For a continuous r.v. \( X(\beta) \), the derivative

\[
f_X(x) = \frac{dF_X(x)}{dx}
\]

in this case exists and is called the density function of the r.v. \( X \). It has the properties,

\[
f_X(x) \geq 0
\]

\[
\int_a^b f_X(x)dx = F_X(b) - F_X(a)
\]

\[
\int_{-\infty}^{\infty} f_X(x)dx = 1
\]

On the other hand, the density function of a discrete r.v. or of a r.v. of the mixed type does not exist in the ordinary sense. However, it can be constructed with the aid of the Dirac delta function. Consider the case where a r.v. \( X \) takes on only discrete values \( x_1, x_2, \ldots, x_n \). A consistent definition of its density function is

\[
f_X(x) = \sum_{j=1}^{n} p_j \delta(x - x_j)
\]

where

\[
p_j = \mathbb{P}_X(X(\beta) = x_j)
\]

This definition of \( f_X(x) \) is consistent in the sense that it has all the properties indicated by equations (5)-(7), and the distribution function \( F_X(x) \) is recoverable from equation (4) by integration.

The definition of the distribution function and the density function can be readily extended to the case of many random variables. Thus, the joint distribution function of a sequence of \( n \) r.v.'s, \( \{X_n(\beta)\} \), is defined by

\[
F_{x_1,x_2,\ldots,x_n}(x_1, x_2, \ldots, x_n) = \mathbb{P}_{X_1,X_2,\ldots,X_n}(X_1 \leq x_1 \cap X_2 \leq x_2 \cap \ldots \cap X_n \leq x_n)
\]

The corresponding joint density function is

\[
f_{x_1,x_2,\ldots,x_n}(x_1, x_2, \ldots, x_n) = \frac{\partial^n F_{x_1,x_2,\ldots,x_n}(x_1, x_2, \ldots, x_n)}{\partial x_1 \partial x_2 \ldots \partial x_n}
\]

if the indicated partial derivative exists. The finite sequence \( \{X_n(\beta)\} \) may be regarded as the components of an \( n \)-dimensional random vector \( X(\beta) \). The distribution function and the density function of \( X(\beta) \) are identical with (10) and (11), but they may be written in the more
1.3. RANDOM VARIABLES

compact forms $F_X(x)$ and $f_X(x)$, where $x$ is a vector with components $x_1, x_2, \ldots, x_n$. In the sequel, the existence of the density functions shall always be assumed.

1.3.2. Moments

Some of the most useful information concerning a random variable is revealed by its moments, particularly those of the first and second order. The nth moment of a r.v. $X(\beta)$ is defined by

$$E^\beta \left( (X(\beta))^n \right) = \begin{cases} 
\sum_{i} x_i^n p_i & \text{for discrete r.v.} \\
\int_{-\infty}^{+\infty} x^n f_X(x) \, dx & \text{for continuous r.v.}
\end{cases}$$

(12)

if $\int_{-\infty}^{+\infty} x^n f_X(x) \, dx$ is finite. The first moment $E^\beta \left( X(\beta) \right)$ gives the statistical average of the r.v. $X(\beta)$. We will usually denote it by $m_X$.

The central moments of a r.v. $X(\beta)$ are the moments of $X(\beta)$ with respect to its mean. Hence, the nth central moment of $X(\beta)$, is defined as

$$E^\beta \left( (X(\beta) - m_X)^n \right) = \begin{cases} 
\sum_{i} (x_i - m_X)^n p_i & \text{for discrete r.v.} \\
\int_{-\infty}^{+\infty} (x - m_X)^n f_X(x) \, dx & \text{for continuous r.v.}
\end{cases}$$

(13)

The moments of two or more random variables are defined in a similar fashion. The joint moments of the r.v.’s $X(\beta)$ and $Y(\beta)$, are defined by

$$E^\beta \left( (X(\beta))^n \cdot (Y(\beta))^m \right) = \begin{cases} 
\sum_{i,j} x_i^n y_j^m p_{ij} & \text{for discrete r.v.} \\
\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^n \cdot y^m \cdot f_{XY}(x, y) \, dx \, dy & \text{for continuous r.v.}
\end{cases}$$

(14)

The central joint moments of the r.v.’s $X(\beta)$ and $Y(\beta)$, are defined by

$$E^\beta \left( (X(\beta) - m_X)^n \cdot (Y(\beta) - m_Y)^m \right) = \begin{cases} 
\sum_{i,j} (x_i - m_X)^n (y_j - m_Y)^m p_{ij} & \text{for disc. r.v.} \\
\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - m_X)^n \cdot (y - m_Y)^m f_{XY}(x, y) \, dx \, dy & \text{for cont. r.v.}
\end{cases}$$

(15)

when they exist.

For $n = m = 1$ we have the covariance function

$$R_{XY} = E^\beta \left( X(\beta) \cdot Y(\beta) \right)$$

(16)
and the correlation function
\[ C_{XY} = E^{\beta} \left( (X(\beta) - m_X)(Y(\beta) - m_Y) \right) \]  
\hspace{1cm} (17)

Obviously, we have
\[ C_{XY} = R_{XY} - m_X m_Y \]  
\hspace{1cm} (18)

The r.v.’s \( X(\beta) \) and \( Y(\beta) \), are said to be uncorrelated if
\[ C_{XY} = 0 \]  
\hspace{1cm} (19)

which implies
\[ E^{\beta} (X(\beta)Y(\beta)) = E^{\beta} (X(\beta)) E^{\beta} (Y(\beta)) \]  
\hspace{1cm} (20)

In the particular case when
\[ E^{\beta} (X(\beta)Y(\beta)) = 0 \]  
\hspace{1cm} (21)

the r.v.’s are called orthogonal.

It is clearly seen from Eq. (18) that two independent r.v.’s with finite second moments are uncorrelated. We point out that the converse of this statement is not necessarily true.

In closing this subsection, we state two inequalities involving the moments of random variables which will be used extensively in the latter chapters.

**Tchebycheff Inequality.**
\[ \Pr \left( \left| X(\beta) \right| \geq k \right) \leq \frac{E^{\beta} \left( \left| X(\beta) \right|^n \right)}{k^n} \]  
\hspace{1cm} (22)

**Hölder Inequality.**
\[ E^{\beta} \left( \left| X(\beta) \cdot Y(\beta) \right| \right) \leq \left[ E^{\beta} \left( \left| X(\beta) \right|^n \right) \right]^{\frac{1}{n}} \cdot \left[ E^{\beta} \left( \left| Y(\beta) \right|^m \right) \right]^{\frac{1}{m}} \]  
\hspace{1cm} (23)

where,
\[ n, m > 1, \quad \text{with} \quad \frac{1}{n} + \frac{1}{m} = 1 \]

The proof of the above inequalities can be found at KOLMOGOROV, A.N. & FOMIN, S.V, (Introductory Real Analysis).

### 1.3.3. Characteristic Functions

**Definition 3.3:** Let \( X(\beta) \) be a r.v. with density function \( f_X(x) \). The characteristic function of \( X(\beta) \), \( \phi_X(u) \) is defined by
\[ \phi_X(u) = E^{\beta} \left( e^{iuX} \right) = \begin{cases} \sum_j e^{iu_j} p_j & \text{for discrete r.v.} \\ \int_{-\infty}^{\infty} e^{iuX} f_X(x) \, dx & \text{for continuous r.v.} \end{cases} \]  
\hspace{1cm} (24)
It is seen that the characteristic function \( \phi_X(u) \) and the density function \( f_X(x) \) form a Fourier transform pair. Since a density function is absolute integrable, its associated characteristic function always exists. Furthermore, it follows from the theory of Fourier transforms that the density function is uniquely determined in terms of the characteristic function by

\[
    f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux} \phi_X(u) du \tag{25}
\]

Eq. (25) points out one of the many uses of a characteristic function. In many physical problems, it is often more convenient to determine the density function of a random variable by first determining its characteristic function and then performing the Fourier transform as indicated by equation (25).

Another important property of the characteristic function is its simple relation with the moments. The MacLaurin series of \( \phi_X(u) \) gives

\[
    \phi_X(u) = \phi_X(0) + \phi_X'(u)u + \phi_X''(u)\frac{u^2}{2} + \ldots \tag{26}
\]

Now, we see that

\[
    \phi_X(0) = \int_{-\infty}^{\infty} f_X(x) dx = 1
\]
\[
    \phi_X'(0) = i \int_{-\infty}^{\infty} x f_X(x) dx = i E_\beta(\beta) X(\beta)
\]
\[
    \vdots
\]
\[
    \phi_X^{(n)}(0) = i^n \int_{-\infty}^{\infty} x^n f_X(x) dx = i^n E_\beta(\beta)^n X(\beta)^n \tag{27}
\]

The joint characteristic function of many r.v.’s \( \{X_i(\beta)\} \) is defined by

\[
    \phi_{X_1, X_2, \ldots, X_n}(u_1, u_2, \ldots, u_n) = E\left\{ e^{i\sum_{j=1}^{n} u_j x_j} \right\} = \\
    = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) dx_1 \cdots dx_n \tag{28}
\]

Analogous to the one-random-variable case, the joint density function \( f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) \) is uniquely determined in terms of \( \phi_{X_1, X_2, \ldots, X_n}(u_1, u_2, \ldots, u_n) \) by the \( n \)-dimensional Fourier transform

\[
    f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) = \\
    = \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{i\sum_{j=1}^{n} u_j x_j} \phi_{X_1, X_2, \ldots, X_n}(u_1, u_2, \ldots, u_n) du_1 du_2 \cdots du_n \tag{29}
\]

The moments, if they exist, are related to \( \phi_{X_1, X_2, \ldots, X_n}(u_1, u_2, \ldots, u_n) \) as
\[ E^\beta \left( (X_1(\beta))^n \cdots (X_n(\beta))^n \right) = \frac{1}{i^{m_1+\cdots+m_n}} \sum_{\alpha} \frac{\partial^{m_1+\cdots+m_n}}{\partial u_1^{m_1} \cdots \partial u_n^{m_n}} \phi_{X_1,\ldots,X_n}(0,0,\ldots,0) \]  

(30)

Since a density function is absolute integrable, its associated characteristic function always exists. Here are some of the basic properties of the characteristic function:

1) \( \phi(0,0,\ldots,0) = 1 \)

2) \( \phi(-u) = \overline{\phi(u)} \)

3) \( \phi(u) \) is uniformly continuous for each \( u \in \mathbb{R}^n \)

4) \( \phi(u) \) is a positive–definite function

Concerning the characterization of characteristic functions, there is the following

**Theorem 3.4 [Bochner]**: The function \( g : \mathbb{R}^n \to \mathbb{R} \) satisfies the above conditions if and only if there exists an \( n \)-dimensional random vector \( X(\beta) \), whose characteristic function is \( g(u) \).

**Proof**: The “if” assertion is immediate, since, if \( g \) is a characteristic function with density function \( G \), then, letting \( u \) and \( v \) range over an arbitrary but finite set in \( \mathbb{R} \),

\[
\sum_{u,v} g(u-v)h(u)\overline{h(v)} = \int \left| \sum_{u,v} e^{i(u-v)x}h(u)\overline{h(v)} \right|^2 dG(x) = \int \sum_u e^{iux}h(u)^2 dG(x) \geq 0
\]

Conversely, let \( g \) on \( \mathbb{R} \) be nonnegative-definite and continuous. It coincides on \( \mathbb{R} \) with a characteristic function. If it does on the set \( S_r \) (dense in \( \mathbb{R} \)) of all rationals of the form of the form \( k/2^n \), \( k = 0, \pm 1, \pm 2, \ldots \), \( n = 1, 2, \ldots \). For every integer \( n \), let \( S_n \) be the corresponding subset of all rationals of the form \( k/2^n \) so that \( S_n \uparrow S_r \). Since \( g \) is nonnegative-definite on \( \mathbb{R} \), it is nonnegative-definite on every \( S_n \). Therefore, by Herglotz Lemma (see LOÈVE, Probability Theory) there exists characteristic functions such that \( g(k/2^n) = f_n(k/2^n) \) whatever be \( k \) and \( n \). Since \( S_n \uparrow S_r \), it follows that \( f_n \to g \) on \( S_r \). Let \( 0 \leq \theta \), \( \theta_n \leq 1 \), so that, by Herglotz Lemma,

\[
1 - \text{Re} f_n \left( \theta/2^n \right) = \int_{-\pi}^{\pi} 1 - \cos(\theta x) dF_n(2^nx) \leq \int_{-\pi}^{\pi} 1 - \cos x dF_n(2^nx) = 1 - \text{Re} f(1/2^n)
\]

Therefore, by the elementary inequality \( |a+b|^2 \leq 2|a|^2 + 2|b|^2 \) and the increments inequality, for fixed \( h = (k_n + \theta_n)/2^n \),
\[ |1 - f_n(h)|^2 \leq 2 |1 - f_n(k_n / 2^n)|^2 + 4 \left( 1 - \text{Re} f_n(\theta_n / 2^n) \right)^2 \leq 2 |1 - g(k_n / 2^n)| + 4 \left( 1 - \text{Re} g(1 / 2^n) \right) \]

Since \( g \) is continuous at the origin, it follows that the sequence \( f_n \) is equicontinuous. Hence, by Ascoli’s Theorem, it contains a subsequence converging to continuous function \( f \) so that \( g = f \) on \( S \) and hence on \( R \). Since by the continuity theorem \( f \) is a characteristic function, the proof is complete.

The “only if” assertion can be proved directly, and this direct proof will extend to a more general case: For every \( T > 0 \) and \( x \in R \)

\[ p_T(x) = \frac{1}{T} \int_0^T \int_0^T g(u-v)e^{-(u-v)x}dudv \geq 0 \]

since, \( g \) on \( R \) being nonnegative-definite and continuous, the integral can be written as a limit of nonnegative Riemann sums. Let \( u = v + t \), integrate first with respect to \( u \) and set \( g_T(t) = \left( 1 - \frac{|t|}{T} \right) g(t) \) or 0 according as \( |t| \leq T \) or \( |t| \geq T \). The above relation becomes

\[ p_T(x) = \frac{1}{T} \int g_T(t)e^{-ixt}dt \geq 0 \]

Now multiply both sides by \( \frac{1}{2\pi} \left( 1 - \frac{|x|}{X} \right) e^{ixu} \) and integrate with respect to \( x \) on \((-X, X)\). The relation becomes

\[ \frac{1}{2\pi} \int_{-X}^{X} \left( 1 - \frac{|x|}{X} \right) p_T(x)e^{ixu} = \frac{1}{2\pi} \int \frac{\sin \left( \frac{1}{2} X(t-u) \right)}{\frac{1}{4} X(t-u)^2} g_T(t)dt \]

The l.h.s. is a characteristic function (since its integrand is a product of \( e^{ixu} \) by a nonnegative function) and the r.h.s. converges to \( g_T(u) \) as \( X \to \infty \). Therefore, \( g_T \) is the limit of a sequence of characteristic functions. Since it is continuous at the origin, the continuity theorem applies and \( g_T \) is a characteristic function. Since \( g_T \to g \) as \( T \to \infty \), the same theorem applies, and the assertion is proved. ■

An extensive discussion for the characteristic functions and their properties can be found at ATHANASSOULIS, G.A., (Stochastic Modeling and Forecasting of Ship Systems) and LOÈVE, (Probability Theory, p.198-213).

1.3.4 Gaussian Random Variables and the Central Limit Theory

The most important distribution of random variables which we encounter in physical problems is the Gaussian or normal distribution.
Definition 3.5: A r.v. \( X (\beta) \) is Gaussian or normal if its density function \( f_X (x) \) has the form

\[
f_X (x) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{(x-m)^2}{2\sigma^2} \right\}
\]

(31)

Its distribution function is the integral

\[
F_X (x) = \int_{-\infty}^{x} \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{(t-m)^2}{2\sigma^2} \right\} dt
\]

(32)

It is seen that a Gaussian r.v. is completely characterized by the two parameters \( m \) and \( \sigma \); they are, as can be easily shown, the mean and the standard deviation of \( X (\beta) \), respectively. Usually we denote it as \( \mathcal{N}[m, \sigma] \).

Alternatively, a Gaussian r.v. can be defined in terms of its characteristic function. It is of the form

\[
\phi_X (u) = \exp \left\{ ium - \frac{1}{2} \sigma^2 u^2 \right\}
\]

(33)

As indicated by equation (30), the moments of \( X (\beta) \) can be easily derived from equation (33) by successively differentiating \( \phi(u) \) with respect to \( u \) and setting \( u = 0 \). The results, expressed in terms of its central moments, are

\[
E^\beta \left( (X (\beta) - m_X)^n \right) = \begin{cases} 0, & n \text{ odd} \\ 1 \cdot 3 \cdot \ldots \cdot (n-1) \cdot \sigma^n, & n \text{ even} \end{cases}
\]

(34)

In addition to having these simple properties, another reason for the popularity of a Gaussian r.v. in physical applications stems from the central limit theorem stated below. Rather than giving the theorem in general terms, it serves our purpose quite well by stating a more restricted version due to Lindeberg.

Theorem 3.6 [Central Limit Theorem]: Let \( \{X_n (\beta)\} \) be a sequence of mutually independent and identically distributed r.v.'s with means \( m \) and variances \( \sigma^2 \). Let

\[
X (\beta) = \sum_{j=1}^{n} X_j (\beta)
\]

(35)

and let the normalized r.v. \( Y (\beta) \) be defined as

\[
Y (\beta) = \frac{X (\beta) - nm}{\sigma \sqrt{n}}
\]

(36)

Then the distribution function of \( Y (\beta) \), \( F_Y (y) \), converges to the zero mean, unit variance Gaussian distribution as \( n \to \infty \) for every fixed \( y \).

Proof: We first remark that the r.v. \( Y (\beta) \) is simply the "normed" r.v. \( X (\beta) \) with mean zero and variance one. In terms of the characteristic functions \( \phi_X (u) \) of the r.v.'s \( X_j (\beta) \) the characteristic function \( \phi_y (u) \) of \( Y (\beta) \) has the form
\[ \phi_Y(u) = E^\beta(e^{iuY}) = \left[ \exp\left(-ium/\sigma n^{1/2}\right)\phi_X\left(u/\sigma n^{1/2}\right) \right] \]  

(37)

In view of the expansion (26), we can write

\[
\phi_Y(u) = \left[ \exp\left(-ium/\sigma n^{1/2}\right)\left[1 + iu\sigma n^{1/2} + \frac{(\sigma^2 + m^2)}{2}\left(\frac{iu}{\sigma n^{1/2}}\right)^2 + \ldots \right] \right]^n = \\
\left[1 - u^2/2n + o\left(\frac{u^2}{n}\right)\right]^n \to \exp\left(-\frac{u^2}{2}\right) 
\]

(38)

as \( n \to \infty \). In the last step we have used the elementary identity in calculus,

\[ \lim_{n \to \infty} (1 + c/n)^n = e^c \]

for any real \( c \). Equation (38) shows that \( \phi_Y(u) \) approaches the characteristic function of a zero mean and unit variance Gaussian distribution in the limit. The proof is thus complete. ■

The central limit theorem thus indicates that often a physical phenomenon has a Gaussian distribution when it represents the sum of a large number of small independent random effects. This indeed is a model or at least an idealization of many physical phenomena we encounter. Consider a sequence \( \{X_n(\beta)\} \) of \( n \) r.v.'s. They are said to be jointly Gaussian if its associated joint density function has the form

\[
f_X(x) = f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) = \\
= (2\pi)^{-n/2} |A|^{-1/2} \exp\left[-\frac{1}{2}(x - \mu)^\tau A^{-1}(x - \mu)\right] 
\]

(39)

where

\[ m^\tau = [m_1, m_2, \ldots, m_n] = [E^\beta(X_1(\beta)), E^\beta(X_2(\beta)), \ldots, E^\beta(X_n(\beta))] \]

(40)

and \( A = [\mu_{ij}] \) is the \( n \times n \) covariance matrix of \( X(\beta) \) with

\[ \mu_{ij} = E^\beta\left[\left(X_j(\beta) - m_j\right)\left(X_j(\beta) - m_j\right)\right] 
\]

(41)

Again, we see that a joint Gaussian distribution is completely characterized by the first- and second-order joint moments.

Parallel to our discussion of a single Gaussian r.v., an alternative definition of a sequence of jointly Gaussian distributed random variables is that it has the joint characteristic function

\[
\phi_X(u) = \exp\left\{im^\tau u - \frac{1}{2}u^\tau A u\right\} 
\]

(42)

This definition is sometimes preferable since it avoids the difficulties when the covariance matrix \( A \) becomes singular. The joint moments of \( X(\beta) \) can be obtained by differentiating the joint characteristic function \( \phi_X(u) \) and setting \( u = 0 \).
It is clear that, since the joint moments of the first- and second-order completely specify the joint Gaussian distribution, these moments also determine the joint moments of orders higher than two. We can show that, if the r.v.'s \( \{X_n(\beta)\} \) have zero means, all odd order moments of these r.v.'s vanish and, for \( n \) even

\[
E^\beta (X_1 X_2 \cdots X_n) = \sum_{m_1, m_2, \ldots, m_n} E^\beta (X_{m_1} X_{m_2}) \cdot E^\beta (X_{m_2} X_{m_3}) \cdots E^\beta (X_{m_n-1} X_{m_n})
\]  

The sum above is taken over all possible combinations of \( n/2 \) pairs of \( n \) r.v.'s. The number of terms in the summation is \( 1 \cdot 3 \cdot 5 \cdot \ldots \cdot (n-3) \cdot (n-1) \).

1.3.5. Convergence of a Sequence of Random Variables

In this section we recall the different types of convergence of a sequence of random variables. These results will be used in the following chapters.

1.3.5.a. Mean-Square Convergence

Let \( \{X_n\}, n \in \mathbb{N} \) be a sequence of second-order random variables. The sequence \( \{X_n\}, n \in \mathbb{N} \) converges in mean-square to a second-order random variable \( X \) as \( n \to \infty \) iff

\[
\lim_{n \to \infty} E^\beta \left( \left( X_n - X \right)^2 \right)^{1/2} = 0
\]  

Preposition 3.7 : The sequence \( \{X_n\}, n \in \mathbb{N} \), of second-order random variables converges in mean-square to a second-order random variable \( X \) such that \( E^\beta (X^2) = x \), \( x > 0 \) if and only if

\[
\lim_{n, n' \to \infty} E^\beta (X_n \cdot X_{n'}) = x
\]

Proof : Assume that eq. (45) is true. Then,

\[
E^\beta \left( (X_n - X_n')^2 \right) = E^\beta (X_n \cdot X_n) + E^\beta (X_n' \cdot X_{n'}) - 2 \cdot E^\beta (X_n \cdot X_{n'}) \to 0
\]

as \( m, m' \to \infty \). Therefore \( \{X_n\}, n \in \mathbb{N} \) is a Cauchy sequence. But since the space of second-order random variables is complete (See LOÈVE, Probability Theory I), any Cauchy sequence of second-order random variables convergences to a second-order random variable. Hence the sequence \( \{X_n\}, n \in \mathbb{N} \) convergences and we have \( \lim_{n, n' \to \infty} E^\beta (X_n \cdot X_{n'}) = x \).

Conversely, if the sequence \( \{X_n\}, n \in \mathbb{N} \) converges to a second-order random variable, then we have eq. (45) because any convergent sequence is a Cauchy sequence and therefore

\[
E^\beta \left( (X_n - X_{n'})^2 \right) \to 0 \text{ as } n, n' \to \infty.
\]

\[\blacksquare\]
1.3.5.b. Convergence in Probability

A sequence \( \{X_n\}, n \in \mathbb{N} \) of random variables defined on a probability space \((\Omega, \mathcal{B}, \mathcal{P})\), converges in probability, to a random variable \(X\) as \(n \to \infty\), iff for every \(\varepsilon > 0\),
\[
\lim_{n \to \infty} \mathcal{P} \left( \left| X_n(\beta) - X(\beta) \right| \geq \varepsilon \right) = 0 \tag{46}
\]
Convergence in probability is also called stochastic convergence or convergence in measure. Note that \(\left| X_n - X \right| \geq \varepsilon\) denotes the subset of \(\mathcal{B}\) such that \(\{\omega \in \Omega : \left| X_n(\beta) - X(\beta) \right| \geq \varepsilon\}\).

1.3.5.c. Almost Sure Convergence

A sequence \( \{X_n\}, n \in \mathbb{N} \) of random variables defined on a probability space \((\Omega, \mathcal{B}, \mathcal{P})\), converges almost surely (a.s.), to a random variable \(X\) as \(n \to \infty\), iff
\[
\mathcal{P} \left( \lim_{n \to \infty} X_n(\beta) = X(\beta) \right) = 0 \tag{47}
\]
This means that subset \(\Omega_0 = \{\beta \in \Omega : \lim_{n \to \infty} X_n(\beta) = X(\beta)\}\) is \(\mathcal{P}\) — negligible : \(\mathcal{P}(\Omega_0) = 0\).

We can also write eq. (47) as
\[
\mathcal{P} \left( \lim_{n \to \infty} X_n(\beta) = X(\beta) \right) = 1 \tag{48}
\]
Convergence \(\mathcal{P}\) — a.s. is also called convergence almost everywhere, or almost certain convergence, or convergence with probability one.

1.3.5.d. Convergence in Law or Distribution

A sequence \( \{X_n\}, n \in \mathbb{N} \) of random variables with probability distributions \(F_{X_n}\), converges in law (or in probability distribution) to a random variable \(X\) with probability distribution \(F_X\) if the sequence of bounded positive measures \(\{F_{X_n}\}\) converges weakly to \(F_X\), i.e. if for all real continuous functions \(f\) on \(\mathbb{R}\) such that \(f(x) \to 0\) as \(|x| \to \infty\), we have
\[
\lim_{n \to \infty} \int_{\mathbb{R}} f(x) F_{X_n}(dx) = \int_{\mathbb{R}} f(x) F_X(dx) \tag{49}
\]

1.3.5.e. Relationships Between the Four Modes of Convergence

Almost Sure Convergence

\[\rightarrow\] Convergence in Probability

Mean-Square Convergence

\[\rightarrow\] Convergence in Law
1.4. References

Chapter 2

Stochastic Processes – Definition and Classification

Random variables or random vectors 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. This parameter can be either time or spatial variable, depending on the application. Typical examples of the first case are ODEs with inherent randomness or random excitation and these cases will mainly concern us in this thesis.

In the present chapter we will present the basic definitions and theorems for the description of the above mathematical objects as well as the fundamental theorems that ensure their existence. Moreover we will make a basic classification of the stochastic processes in relation with fundamental properties such as memory and ergodicity.

More specifically in Section 2.2.1 the stochastic processes are classified according to the manner in which their present state is dependent upon their past history. We thus have the purely random processes without any memory. In order to obtain a mathematic satisfactory theory of these processes we introduce the concept of generalized stochastic processes. The next step in this kind of classification is the Markov processes that can be described completed form the second order quantities. Moreover, we emphasize the case of independent increment processes and martingale stochastic processes that play important role to the theory of stochastic dynamical systems.

In the next Section 2.2.2 the concept of stationarity is presented. We formulate the basic definitions as well as all the necessary results for the direction of stochastic dynamical systems. Stationary stochastic processes are very useful for the description of linear systems since, under suitable assumptions, we can formulate the whole problem in terms of the spectral densities.

Section 2.2.3 deals with the important issue of ergodicity, fundamental property for dynamical systems, that relates statistical or ensemble averages of a stationary stochastic process to time averages of its individual sample functions. The interchangeability of ensemble and time averages has considerable appeal in practice since we can compute statistical averages of a stochastic process in terms of a long single observation of one sample function.

Finally in the last Section 2.2.4 we present the Gaussian stochastic process with its basic properties.
2.1. Basic Concepts of Stochastic Processes

The foregoing description of a stochastic process is based upon the axiomatic approach in probability theory, originally due to Kolmogorov. Let us call each realization of a given random experiment a sample function or a member function. This description suggests that a s.p. \(X(t; \beta), t \in I \subseteq \mathbb{R}\) is a family of sample functions of the variable \(t\), all defined on the same underlying probability space \((\Omega, \mathcal{B}, P)\). Intuitively, this sample theoretic approach to the concept of a stochastic process is physically plausible since a stochastic process describing a physical phenomenon generally represents a family of individual realizations. To each realization there corresponds a definite deterministic function of \(t\). However, a difficult point arises. A stochastic process defined this way is specified by the probability of realizations of various sample functions. This leads to the construction of probability measures on a functional space and it generally requires advanced mathematics in measure theory and functional analysis. We will consider this approach extensively in Chapter 3. On the other hand, Definition 1.1 presents a simpler approach to the concept of a stochastic process which is sufficient for the subsequent development.

2.1.1. Definition of a Continuous-parameter Stochastic Process

At a fixed \(t\), a s.p. \(X(t; \beta), t \in I \subseteq \mathbb{R}\), is a random variable. Hence, another characterization of a stochastic process is to regard it as a family of random variables, say, \(X(t_1; \beta), X(t_2; \beta), \ldots\) depending upon a parameter \(t \in I \subseteq \mathbb{R}\). The totality of all the random variables defines the s.p. \(X(t; \beta)\). For discrete-parameter stochastic processes, this set of random variables is finite or countably infinite. For continuous-parameter processes, the number is noncountably infinite. Thus, we see that the mathematical description of a stochastic process is considerably more complicated than that of a random variable. It, in fact, is equivalent to the mathematical description of an infinite and generally noncountable number of random variables.

We recall from Section 1.3 of Chapter 1 that a finite sequence of random variables is completely specified by its joint distribution or joint density functions. It thus follows from the remarks above that a s.p. \(X(t; \beta), t \in I \subseteq \mathbb{R}\), is defined by a family of joint distribution or density functions, taking into account that we are now dealing with infinitely many random variables.

Let us now make this definition more precise. The index set \(I\) usually is the real line or an interval of the real line. But the above definition can be extended to the case of an index set with \(n\)-dimensions.

**Definition 1.1**: Let a family of r.v.’s \(X(t_1; \beta), t_i \in I\) such that for every finite set \(\{t_1, t_2, \ldots, t_n\}\) the corresponding set of r.v.’s \(X_1(\beta) = X(t_1; \beta), X_2(\beta) = X(t_2; \beta), \ldots, X_n(\beta) = X(t_n; \beta)\), has a well defined joint probability distribution function.
2.1. DEFINITION OF STOCHASTIC PROCESSES

\[ F_{t_1, t_2, \ldots, t_n}(x_1, x_2, \ldots, x_n) = \mathcal{P}\left( X(t_1, \beta) < x_1 \cap \ldots \cap X(t_n, \beta) < x_n \right) \]  

(1)

that satisfies the following conditions (Kolmogorov compatibility conditions):

a) the symmetry condition: if \( \{i_1, i_2, \ldots, i_n\} \) is a permutation of numbers 1, 2, \ldots, \( n \) then for arbitrary \( n \geq 1 \)

\[ F_{t_{i_1}, t_{i_2}, \ldots, t_{i_n}}(x_{i_1}, x_{i_2}, \ldots, x_{i_n}) = F_{t_{i_1}, t_{i_2}, \ldots, t_{i_n}}(x_1, x_2, \ldots, x_n) \]  

(2)

b) the consistency condition: for \( m < n \) and arbitrary \( t_{m+1}, \ldots, t_n \in I \)

\[ F_{t_{i_1}, \ldots, t_{i_m}, t_{m+1}, \ldots, t_n}(x_1, x_2, \ldots, x_m, \infty, \ldots, \infty) = F_{t_{i_1}, t_{i_2}, \ldots, t_{i_n}}(x_1, x_2, \ldots, x_n) \]  

(3)

Then this family of joint distribution functions defines a s.p. \( X(t; \beta), t \in I \).

In the special case where \( F_{t_1, t_2, \ldots, t_n}(x_1, x_2, \ldots, x_n) \) is differentiable we can alternative use the corresponding probability density function defined by

\[ f_{t_1, t_2, \ldots, t_n}(x_1, x_2, \ldots, x_n) = \frac{\partial^n F_{t_1, t_2, \ldots, t_n}(x_1, x_2, \ldots, x_n)}{\partial x_1 \partial x_2 \cdots \partial x_n} \]  

(4)

This family of probability density functions satisfies the following conditions:

a) the positiveness condition

\[ f_{t_1, t_2, \ldots, t_n}(x_1, x_2, \ldots, x_n) \geq 0 \]  

(5)

b) the normalization property

\[ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{t_1, t_2, \ldots, t_n}(x_1, x_2, \ldots, x_n) \cdot dx_1 \cdot dx_2 \cdots dx_n = 1 \]  

(6)

Based upon the definition above, another convenient and often useful method of specifying a stochastic process is to characterize it in terms of an analytic formula as a function of containing random variables as parameters. Let \( A_1(\beta), A_2(\beta), \ldots, A_n(\beta) \) be a set of random variables. A s.p. \( X(t; \beta) \) characterized this way has the general form

\[ X(t; \beta) = g(t; A_1, A_2, \ldots, A_n) \]  

(7)

where the functional form of \( g \) is given.

Example 1.2 [Pierson – Longuet Higgins Model] : Representation (7) can be applied for the description of random wind generated, water waves at a specific point. In this case we can determine experimentally the spectral distribution, i.e., the amplitude that corresponds to every frequency. Hence, the stochasticity is introduced from the phase of every harmonic component. More specifically we use the following representation for the random field at a specific point

\[ X(t; \beta) = \sum_{i=1}^{N} A_i \cos(\omega_i t + \varphi_i(\beta)) \]  

(8)
where $(A_i, \omega_i), i = 1, \ldots, N$ are given and $\varphi_i(\beta), i = 1, \ldots, N$ are independent random variables, uniformly distributed at $[0, 2\pi]$. We must note that representation (8) is consistent with the theory of linear water waves. Moreover it can be easily proved that the above stochastic process is stationary and ergodic. Additionally for large number of harmonic components $N$ its probability distribution functions follow the Gaussian law.

As we mention at the beginning a stochastic process can also specified by the probability of realizations of various sample functions. For the following definitions we assume that $\mathcal{H}$ is a measurable, metric space.

**Definition 1.3**: Let $(\Omega, \mathcal{B}(\Omega), \mathcal{P})$ be a probability space, $X : \Omega \to \mathcal{H}$ a point function from the space of elementary events $\Omega$ to a metric space $\mathcal{H}$ and $\mathcal{B}(\mathcal{H})$ a Borel field over $\mathcal{H}$. The mapping $X$ is a stochastic process, iff is $(\mathcal{B}(\Omega), \mathcal{B}(\mathcal{H}))$-measurable.

Together with the definition of the stochastic process we have, as in the case of random variables the induced probability measure.

**Definition 1.4**: Let $(\Omega, \mathcal{B}(\Omega), \mathcal{P})$ be a probability space, $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$ a measurable space and $X : \Omega \to \mathcal{H}$ a stochastic process. On the Borel field $\mathcal{B}(\mathcal{H})$ of the measurable space $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$ we define the induced probability measure

$$\mathcal{P}^X (B) = \mathcal{P}^X \left[ X^{-1}(B) \right], \quad \text{for every } B \in \mathcal{B}(\mathcal{H})$$

(9)

The produced probability space $(\mathcal{H}, \mathcal{B}(\mathcal{H}), \mathcal{P}^X)$, will be called as $X$-induced probability space.

To connect the above probability measure defined on the infinite dimensional space $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$ with the joint probability distribution functions we have the following definition.

**Definition 1.5**: Let $(\Omega, \mathcal{B}(\Omega), \mathcal{P})$ be a probability space, $X : \Omega \to \mathcal{H}$ a stochastic process and $(\mathcal{H}, \mathcal{B}(\mathcal{H}), \mathcal{P}^X)$ the induced probability space. Let $\Pi_N$ be a projection of $\mathcal{H}$ with $\dim \Pi_N [\mathcal{H}] = N$. For every $\Pi_N$ we introduce the following probability measure defined on $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$,

$$\mathcal{P}^{X \Pi_N} (B) = \mathcal{P}^X \left[ \Pi_N^{-1}[B] \right] = \mathcal{P}^X \left[ \{x \in \mathcal{H} : \Pi_N [x] \in B \} \right], \quad B \in \mathcal{B}(\mathbb{R}^N)$$

(10)

**Example 1.6**: For the case of $\mathcal{H} = C(I)$ being the space of real, continuous functions of a real variable $t \in I$, a very important measure is for $\Pi_J [x(t)] = (x(t_1), x(t_2), \ldots, x(t_n))$ for every $J = \{t_1, t_2, \ldots, t_n\} \subset I$. For this projection operator we have

$$\mathcal{P}^{X \Pi_J} (B) = \mathcal{P}_{C(I)} \left[ \{x(\cdot) \in C(I) : \Pi_J [x(\cdot)] \in B \} \right], \quad B \in \mathcal{B}(\mathbb{R}^N)$$

(11)
2.1. DEFINITION OF STOCHASTIC PROCESSES

Now, let's restrict the family of $B \in \mathcal{B} \left( \mathbb{R}^N \right)$ to those of the form $B = (-\infty, x_1) \times (-\infty, x_2) \times \ldots \times (-\infty, x_N)$. It is easy to see that we get the joint probability distributions

$$F_{x_1, x_2, \ldots, x_n} (x_1, x_2, \ldots, x_n) = \mathcal{P}_{C(I)} \left( \left\{ x(\cdot) \in C(I) ; \prod_I [x(\cdot)] \in B \right\} \right)$$

where, $B = (-\infty, x_1) \times (-\infty, x_2) \times \ldots \times (-\infty, x_N)$.

The above approach will be discussed comprehensively in Chapter 3. To connect the Definition 1.3 with the first Definition 1.1 we can think of $\mathcal{H}$ as the space of real function of a real variable $D(I)$. We then have the following

**Theorem 1.7 [Kolmogorov Theorem]** : For every given family of finite-dimensional distributions satisfying conditions a), b), there exists a probability space $(\Omega, \mathcal{B}, \mathcal{P})$ and a stochastic process $X(t; \beta)$, $t \in I$ defined on it that possesses the given distribution as finite-dimensional distributions.


In what follows we assume the existence of a probability space $(\Omega, \mathcal{B}(\Omega), \mathcal{P})$. The stochastic process $X : \Omega \rightarrow C(I)$ is defined on it and the induced probability space is $(C(I), \mathcal{B}(C(I)), \mathcal{P}_{C(I)})$.

2.1.2. Moments

As in the case of random variables, some of the most important properties of a s.p. are characterized by its moments, particularly those of the first and the second order. In the sequel, the existence of density functions shall be assumed.

The moments of the probability density function $f_t(x)$, are defined as the expectation value $E^\beta \left( (X(t; \beta))^n \right)$. We have

$$R^n_x (t) \equiv E^\beta \left( (X(t; \beta))^n \right) = \int_{-\infty}^{+\infty} x^n f_t(x) dx$$

As special cases we have the first order moment

$$R^1_x (t) = m_x (t) = E^\beta \left( X(t; \beta) \right) = \int_{-\infty}^{+\infty} x f_t(x) dx$$

which is the mean value of $X(t; \beta)$ at time $t$. Additionally the second order moment

$$R^2_x (t) = m_x (t) = E^\beta \left( (X(t; \beta))^2 \right) = \int_{-\infty}^{+\infty} x^2 f_t(x) dx$$
is the variance of $X(t; \beta)$ at time $t$. Analogously we can define $n$–point moments. We have

$$R_{X_1,\ldots,X_n}(t_1,t_2,\ldots,t_m) \equiv E^\beta \left( X(t_1; \beta)^{n_1} \cdot X(t_2; \beta)^{n_2} \cdot \ldots \cdot X(t_m; \beta)^{n_m} \right) =$$

$$= \int_{-\infty}^{+\infty} x_1^{n_1} \cdot x_2^{n_2} \cdot \ldots \cdot x_m^{n_m} f_{t_1,t_2,\ldots,t_m}(x_1,x_2,\ldots,x_m) dx_1 \ldots dx_m$$

A special case is the autocorrelation function,

$$R_{XX}(t_1,t_2) \equiv E^\beta \left( X(t_1; \beta) \cdot X(t_2; \beta) \right)$$

One very important moment of a stochastic process $X(t; \beta)$ is the autocovariance function defined by

$$C_{XX}(t_1,t_2) \equiv E^\beta \left[ X(t_1; \beta) - m(t_1) \right] \left[ X(t_2; \beta) - m(t_2) \right]$$

The importance of the correlation functions or, equivalently, the covariance functions rests in part on the fact that their properties define the mean square properties of the stochastic processes associated with them. This will become clear in Chapter 4. In what follows, we note some of the important properties associated with these functions.

1. $R_{XX}(t_1,t_2) = R_{XX}(t_2,t_1)$ for every $t_1,t_2 \in I$  
   (18)

2. $R_{XX}^2(t_1,t_2) \leq R_{XX}(t_1,t_1) \cdot R_{XX}(t_2,t_2)$ for every $t_1,t_2 \in I$  
   (19)

3. $R_{XX}(t_1,t_2)$ is positive–definite on $I \times I$.  
   (20)

### 2.1.3. Characteristic Functions

Following the definition of a characteristic function associated with a sequence of random variables, the characteristic functions of a s.p. $X(t; \beta), t \in I$, are defined in an analogous way. There is, of course, a family of characteristic functions since $X(t; \beta)$ is characterized by a family of density functions. Again, we assume the existence of the density functions in what follows.

The characteristic function $\phi_h(u)$ is defined in the following manner for the one dimensional case

$$\phi_h(u) \equiv E^\beta(e^{iuX}) = \int_{-\infty}^{+\infty} e^{iuX} f_i(x) dx$$

That is, $\phi_h(u)$ is the Fourier transform of $f_i(x)$. By setting $u = 0$ we can note immediately that

$$\phi_h(0) = \int_{-\infty}^{+\infty} f_i(x) dx = 1$$
2.1. DEFINITION OF STOCHASTIC PROCESSES

The characteristic functional may be defined for \( n \) variables \( u_1, u_2, \ldots, u_n \) analogously. Here

\[
\phi_{t_1, t_2, \ldots, t_n} (u_1, u_2, \ldots, u_n) \equiv E^\beta \left( \exp \left[ i \sum_{k=1}^{n} X \left( t_k \cdot \beta \right) \cdot u_k \right] \right) =
\]

\[
= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \exp \left[ i \sum_{k=1}^{n} x_k \cdot u_k \right] \cdot f_{t_1, t_2, \ldots, t_n} (x_1, x_2, \ldots, x_n) \, dx_1 \cdots dx_n
\]

(23)

Here are some of the basic properties of the characteristic function of a stochastic process:

1) \( \phi_{t_1, t_2, \ldots, t_n} (u_1, u_2, \ldots, u_n) \leq \phi_{t_1, t_2, \ldots, t_n} (0, 0, \ldots, 0) = 1 \)

2) \( \phi_{t_1, t_2, \ldots, t_n} (u) = \phi_{t_1, t_2, \ldots, t_n} (-u) \)

3) \( \phi_{t_1, t_2, \ldots, t_n} (u_1, u_2, \ldots, 0) = \phi_{t_1, t_2, \ldots, t_n} (u_1, u_2, \ldots, u_{n-1}) \)

4) \( \phi_{t_1, t_2, \ldots, t_n} (u) \) is uniformly continuous for each \( u \in \mathbb{R}^n \)

5) \( \phi_{t_1, t_2, \ldots, t_n} (u) \) is a positive-definite function

The extension of \( \phi \) to infinitely many variables involves the concept of a functional and will be considered in Chapter 3. There is an intimate connection between the moments and derivatives of \( \phi \) with respect to \( u \). Following the analysis we made for random variables at Section 1.3 of Chapter 1 we will have the similar result

\[
R^{n_1, n_2, \ldots, n_n}_{X_1, X_2, \ldots, X_n} (t_1, t_2, \ldots, t_n) = \frac{1}{i^{n_1 + \ldots + n_n}} \frac{\partial^{n_1 + \ldots + n_n}}{\partial u_{n_1} \cdots \partial u_{n_n}} \phi_{t_1, t_2, \ldots, t_n} (0, 0, \ldots, 0).
\]

(24)
2.2. **Classification of Stochastic Process**

In what follows we assume the existence of a probability space \((\Omega, \mathcal{B}(\Omega), \mathbb{P})\). The stochastic process \(X: \Omega \rightarrow C(I)\) is defined on it and the induced probability space is \((C(I), \mathcal{B}(C(I)), \mathbb{P}_{\mathcal{C}(I)})\).

We have seen that a s.p. \(X(t; \beta), t \in I\), is defined by a system of probability distribution functions. This system consists of

\[
F_{t_1, \ldots, t_n}(x_1, x_2, \ldots, x_n) = \mathbb{P}\left( X(t_1, \beta) < x_1 \cap \ldots \cap X(t_n, \beta) < x_n \right)
\]

for every \(n\) and for every finite set \(\{t_1, t_2, \ldots, t_n\}\) in the fixed index set \(I\). In general, all probability distribution functions are needed to specify a stochastic process completely. It is not uncommon, however, that stochastic processes modelling physical phenomena are statistically simpler processes, simple in the sense that all the statistical information about a given process is contained in a relatively small number of probability distribution functions. The statistical complexity of a stochastic process is determined by the properties of its distribution functions and, in the analysis of stochastic processes, it is important to consider them in terms of these properties.

We consider in the remainder of this chapter classification of stochastic processes according to the properties of their distribution functions. Presented below are two types of classifications that are important to us in the analysis to follow, namely, classification based upon its **memory** and classification based upon the **statistical regularity** of a process over the index set \(I\). We note that these two classifications are not mutually exclusive and, in fact, many physically important processes possess some properties of each.

### 2.2.1. Classification Based Upon Memory

In this classification, a s.p. \(X(t; \beta), t \in I\), is classified according to the manner in which its present state is dependent upon its past history. It is centred around one of the most important classes of stochastic processes, that is, the Markov processes.

#### 2.2.1.a. Purely Stochastic Process – Generalized Stochastic Process

According to memory properties, the simplest s.p. is one without memory. They characterized by the following

**Definition 2.1:** A s.p. \(X(t; \beta), t \in I\) is **fully stochastic** when the random variables \(X(t_1; \beta)\) and \(X(t_2; \beta)\) are independent for every \(t_1 \neq t_2\).

From the above definition it is clear that the \(n^{th}\) order distribution function can be written as product of first order distributions.
\[ F_{x_1, x_2, \ldots, x_n}(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} F_x(x_i) \]  

(1)

The probability structure of the fully stochastic process is characterized from the first order distribution only. Although mathematically simple, a continuous-parameter purely stochastic process is not physically realizable because it implies absolute independence between the present and the past no matter how closely they are spaced. In physical situations we certainly expect \(X(t; \beta)\) to be dependent on \(X(t'; \beta)\) when \(t_1\) is sufficiently close to \(t_2\). In practical applications, therefore, purely stochastic processes with a continuous parameter are considered to be limiting processes. In order to obtain a mathematic satisfactory theory of the process just described one has to introduce the concept of generalized s.p.

Let \(C^\infty_c(I)\) be the space of all infinitely differentiable functions \(\varphi(t): I \rightarrow \mathbb{R}\) vanishing identically outside a finite closed interval; it is a space of test functions. Let \(\mathcal{Y}\) be a Hilbert space of random variables defined on \((\Omega, \mathcal{B}(\Omega), \mathcal{P})\) having finite second moment.

**Definition 2.2** : A generalized s.p. on \(I\) is a continuous linear mapping \(\Phi\) from \(C^\infty_c(I)\) to \(\mathcal{Y}\); the value of a generalized stochastic process \(\Phi\) at \(\varphi\) will be denoted by \(\{\varphi, \Phi\}\) or \(\Phi(\varphi)\).

**Example 2.3** : Let \(X(t; \beta), t \in I\) be a s.p. with finite second moment (a mapping from \(I\) to \(\mathcal{Y}\)). As we shall see in Chapter 4, for a very wide class of processes and for \(\varphi(t) \in C^\infty_c(I)\) there exists the following integral

\[ \{\varphi, \Phi_X\} = \Phi_X(\varphi) = \int_S \varphi(t) X(t; \beta) dt \]

(2)

where \(S\) is an arbitrary compact set in \(I\) (\(X(t; \beta)\) is a locally integrable stochastic process on \(I\)). The above mapping is linear and continuous. Therefore, every locally integrable s.p. on \(I\) determines a generalized s.p. in the form presented above.

One of the important advantages of a generalized stochastic process is the fact that its derivative always exists and is itself a generalized s.p.

**Definition 2.4** : The derivative \(\Phi'\) with respect to \(t\) of a generalized s.p. \(\Phi\) in \(C^\infty_c(I)\) is defined by means of the formula

\[ \{\varphi, \Phi'\} = \left\{ \frac{d\varphi}{dt}, \Phi \right\} \text{ for all } \varphi(t) \in C^\infty_c(I) \]

(3)

**Remark 2.5** : It can be verified that the mapping \(\varphi \rightarrow \left\{ \frac{d\varphi}{dt}, \Phi \right\}\) is a continuous linear mapping from \(C^\infty_c(T)\) to \(\mathcal{Y}\). The derivative \(\Phi^{(\alpha)}\) for any given \(\alpha \geq 1\) are given by

\[ \{\varphi, \Phi^{(\alpha)}\} = (-1)^\alpha \left\{ \frac{d^\alpha \varphi}{dt^\alpha}, \Phi \right\} \]

(4)

**Remark 2.6** : It is clear from Example 2.3 that the appropriate definition of \(\{\varphi, \Phi_X'\}\) is
\[
\left\{ \varphi, \Phi_X \right\} = -\int_s^t \varphi'(s) X(t; \beta) \, ds
\]  
(5)

From the above it follows that an ordinary process \( X(t; \beta) \) may also be regarded as always differentiable, provided its derivative be a generalized process instead of an ordinary process.  

From the standpoint of correlation theory, a generalized s.p. \( \Phi \) or \( \Phi(\varphi) \) is characterized by its mean
\[
m_\varphi(\varphi) = E^\beta(\Phi(\varphi))
\]  
(6)

and correlation operator
\[
C(\varphi, \psi) = E^\beta(\Phi(\varphi)\Phi(\psi)) \quad \varphi, \psi \in C_c^\infty(I)
\]  
(7)

which is assumed to be linear and continuous with respect to \( \varphi \) and \( \psi \).

A typical example of a generalized stochastic process is Gaussian white noise \( \xi \).

**Definition 2.7:** A Gaussian white noise \( \xi \) or \( \xi(\varphi) \) is a generalized s.p. such that there exist a Wiener process \( W(t; \beta) \ t \in I \) satisfying \( \xi(t; \beta) = W'(t; \beta) \), i.e.

\[
\left\{ \varphi, \xi \right\} = \left\{ \varphi, W' \right\} = -\left\{ \frac{d\varphi}{dt}, W \right\} \quad \text{for all } \varphi(t) \in C_c^\infty(I)
\]  
(8)

Note that for every \( \varphi(t) \in C_c^\infty(I) \), \( \left\{ \varphi, \xi \right\} \) is a Gaussian r.v., and that for every finite number of functions \( \varphi_1, \varphi_2, \ldots, \varphi_n \in C_c^\infty(I) \), the random variables \( \left\{ \varphi_i, \xi \right\} ; 1 \leq i \leq n \) are jointly Gaussian. This is the reason why \( \xi \) is called a Gaussian white noise.

**Theorem 2.8:** If \( \xi \) is a white Gaussian noise, then

\[
E^\beta(\xi(\varphi)) = 0 \quad \text{for all } \varphi(t) \in C_c^\infty(I)
\]  
(9)

and

\[
C_\xi(\varphi, \psi) = \int_T \varphi(t)\psi(t) \, dt \quad \varphi, \psi \in C_c^\infty(I)
\]  
(10)

**Proof:**

\[
E^\beta(\xi(\varphi)) = E^\beta(\left\{ \varphi, \xi \right\}) = E^\beta\left( \int_T -\frac{d\varphi}{dt} W(t) \, dt \right) = -E^\beta\left( \int_T -\frac{d\varphi}{dt} W(t) \, dt \right) = 0
\]

since \( E^\beta(W(t)) = 0 \) for all \( t \in I \) and \( \varphi \) is arbitrary. To show the second equality, let \( \varphi, \psi \in C_c^\infty(I) \). Then

\[
C_\xi(\varphi, \psi) = E^\beta(\left\{ \varphi, \xi \right\} \cdot \left\{ \psi, \xi \right\}) = E^\beta\left( \int_T -\frac{d\varphi}{dt} W(t_1) \, dt_1 \cdot \int_T -\frac{d\psi}{dt} W(t_2) \, dt_2 \right) =
\]

\[
= \int_T \int_T \frac{d\varphi}{dt_1} \frac{d\psi}{dt_2} E^\beta(W(t_1) \cdot W(t_2)) \, dt_1 dt_2 = \int_T \int_T \frac{d\varphi}{dt_1} \frac{d\psi}{dt_2} \min(t_1, t_2) \, dt_1 dt_2
\]

After integrating by parts we obtain the desired equality.  

\textbf{Remark 2.9:} It is worth noting that the formula \((10)\) can be put in the form
\[
C_\xi(\varphi, \psi) = \int_T \int_T \delta(t-s) \varphi(t) \psi(t) \, ds \, dt, \quad \varphi, \psi \in C_0^\infty(I)
\]
which implies that the correlation function of the Gaussian white noise is the generalized function \(\delta(t-s)\). ■

\subsection{2.1.2.b. Markov Stochastic Process}

The Markov s.p., is a subset of the wide class of s.p. which characterized only from the second order distribution function. More specifically, for the Markov s.p., the random variables defined for all the values of \(t > \tau\) are independent from the random variables defined for \(t < \tau\) if we know the value of the stochastic process at time \(\tau\). The above property is expressed at the following

\textbf{Definition 2.10}: A s.p. \(X(t; \beta), t \in I\) is Markov, if for every \(n\) and every \(\{t_1 < t_2 < \ldots < t_n\}\) in \(I\) we have
\[
F_{\pi_n}(x_{n-1}, \ldots, x_1, x_{n-1}, t_{n-1}) = F_{\pi_n}(x_{n-1}, t_{n-1})
\]
or equivalently,
\[
\mathcal{D}_\Pi(t_n)\left(\mathcal{D}_\Pi(t_{n-1}) \cdots \mathcal{D}_\Pi(t_1) \right) = B, \quad B \in \mathcal{B}(\mathbb{R})
\]
where \(\Pi(t_n)\) is the projection operator \(\Pi(t_n)\left[\mathcal{F}(\cdot)\right] = x(t_n)\) for every \(t_n\). (See Example 1.6. of Section 2.1) ■

\textbf{Remark 2.11}: The equivalent definition, presented above, can be used to generalize the concept of Markov s.p. to the case of a general Hilbert space, whose elements, are defined on some ordered set. ■

As we can easily see, the conditional distribution of \(X(t_n; \beta)\) for given values of \(X(t_1; \beta), X(t_2; \beta), \ldots, X(t_{n-1}; \beta)\) depends only at the most recent known value \(X(t_{n-1}; \beta)\).

The above definition implies that the basic characteristics of a Markov Process are described by a function
\[
F_M(x, t; y, s) = F_t(x | y, s)
\]
The function \(F_M(x, t; y, s)\) is called a \textit{transition probability}. It gives the probability of a transition of a process from a state \(y\) at time \(s\) to one of the states for which \(\{X(t; \beta) < x\}\) at time \(t\). We can easily see that for fixed \(t, s, y\) the transition probability \(F_M(x, t; y, s)\) is a probability distribution function.

\textbf{Proposition 2.12}: The probability structure of a Markov s.p. is completely characterized by the second order probability distribution functions.
Proof: To see that a Markov process is completely specified by its first and second distribution functions, we start with the general relation
\[
F_{t_{i+1},t_{i}}(x_{1},x_{2},...,x_{n}) = F_{t_{i}}(x_{i}) \prod_{j=1}^{n-1} F_{t_{i+1}}(x_{i+1} | x_{i},t_{j})
\]
Using the definition of the Markov s.p. we immediately see that for \( n \geq 3 \) and \( \{t_{i} < t_{2} < \ldots < t_{n}\} \)
\[
F_{t_{i},t_{i+1},...,t_{n}}(x_{1},x_{2},...,x_{n}) = F_{t_{i}}(x_{i}) \prod_{j=1}^{n-1} F_{t_{i+1}}(x_{i+1} | x_{i},t_{j})
\]
But,
\[
F_{t_{i+1}}(x_{i+1} | x_{i},t_{j}) = \frac{F_{t_{i+1},t_{i}}(x_{i+1},x_{i})}{F_{t_{i}}(x_{i})}
\]
Consequently,
\[
F_{t_{i},t_{i+1},...,t_{n}}(x_{1},x_{2},...,x_{n}) = \prod_{j=1}^{n-1} F_{t_{i+1}}(x_{i+1},x_{i}) \prod_{i=2}^{n-1} F_{t_{i}}(x_{i})
\]

**Proposition 2.13**: Let \( t_{1} < t_{2} < t_{3} \). For the transition probability density function the following Smoluchowski-Chapman-Kolmogorov (SCK) holds,
\[
f_{M}(x_{3},t_{3};x_{1},t_{1}) = \int_{-\infty}^{+\infty} f_{M}(x_{3},t_{3};x_{2},t_{2}) f_{M}(x_{2},t_{2};x_{1},t_{1}) dx_{2}
\]
or equivalently,
\[
F_{M}(x_{3},t_{3};x_{1},t_{1}) = \int_{-\infty}^{+\infty} F_{M}(x_{3},t_{3};x_{2},t_{2}) dF_{M}(x_{2},t_{2};x_{1},t_{1})
\]

**Proof**: Consider the general relation
\[
f_{t_{1},t_{2}}(x_{1},x_{2}) = \int_{-\infty}^{+\infty} f_{t_{1},t_{2},t_{3}}(x_{1},x_{2},x_{3}) dx_{2}
\]
Let \( t_{1} < t_{2} < t_{3} \). Then using the relation
\[
f_{t_{1},t_{2}}(x_{1},x_{2},x_{3}) = f(x_{1},t_{1} | x_{2},t_{2};x_{3},t_{3}) \cdot f(x_{2},t_{2} | x_{3},t_{3}) f_{t_{3}}(x_{3})
\]
and equations (15) we clearly have
\[
f_{t_{1},t_{2}}(x_{1},x_{2},x_{3}) = f(x_{1},t_{1} | x_{2},t_{2}) \cdot f(x_{2},t_{2} | x_{1},t_{1}) \cdot f_{t_{3}}(x_{3})
\]
Then making use of equation (21) we have the desired equation. \( \blacksquare \)

In the analysis of Markov Processes and in modeling of real random phenomena we frequently deal with so called stationary transition probabilities characterizing a homogeneous Markov s.p.
Definition 2.14: A Markov s.p. \( X(t; \beta) \), \( t \in I \) is said to be homogeneous (with respect to time) if for arbitrary \( s \in I \) and \( t \in I \), \( s < t \), the transition probability depends only on the difference \( \tau = t - s \), that is

\[
F_{\text{HM}}(x, t; y, s) = F_{\text{HM}}(x, \tau; y)
\]  

(22)

The function \( F_{\text{HM}}(x, \tau; y) \) determines the probability of transition from a state \( y \) at time \( s \) to one of the states for which \( \{X(t; \beta) < x\} \) at time \( t \). This probability does not depend on the position of \( s \), but only from the difference \( t - s \). In this case the Smoluchowski-Chapman-Kolmogorov equation takes the form

\[
F_{\text{HM}}(x_1, \tau_1 + \tau_2; x_3) = \int_{-\infty}^{+\infty} F_{\text{HM}}(x_1, \tau_1; x_2) dF_{\text{HM}}(x_2, \tau_2; x_3)
\]  

(23)

The SCK equation has important implications in the theory of Markov s.p.; among others, as it can be shown (SOBCZYK, K., Stochastic Differential Equations. p. 31), it implies the existence of a semi-group of operators associated with a homogeneous Markov s.p.

Preposition 2.15: Let the homogeneous Markov s.p. \( X(t; \beta) \), \( t \in I \) with transition probability \( F_{\text{HM}}(x, \tau; y) \). For every bounded, measurable, function \( f \) defined on the state space \( R \subset \mathbb{R} \) (the set of possible values of the s.p.) of the process we define the operator \( T_t \) acting on the functions \( f \) as

\[
T_t[f(*)] = \int_R f(x) dF_{\text{HM}}(x, t; y)
\]  

(24)

then,

a) Operators \( T_t \) are linear, positive and continuous.

b) Operators \( T_t \) form a semi-group, i.e.

\[
T_{t+s} = T_t \circ T_s = T_s \circ T_t, \quad s, t > 0.
\]  

(25)

Proof: For the proof of a) we refer to (SOBCZYK, K., Stochastic Differential Equations. p. 31). Part b) can be proved using the form (23) of Smoluchowski-Chapman-Kolmogorov equation.

2.1.3.c. Processes with Independent Increments

Mathematical models of numerous real phenomena can be described by stochastic processes whose increments are independent random variables.

Definition 2.16: A s.p. \( X(t; \beta) \), \( t \in I \) is said to be an independent–increment process, if for every \( n \) and every \( \{t_1 < t_2 < \ldots < t_n\} \) in \( I \), the random variables,

\[
X(t_1; \beta), X(t_2; \beta) - X(t_1; \beta), \ldots, X(t_n; \beta) - X(t_{n-1}; \beta)
\]

are independent.
The definition above implies that all finite-dimensional distributions of the process \( X(t; \beta) \) are completely determined by the distribution of the random variable \( X(t; \beta) \) for each \( t \) and the distributions of the increments \( X(t_2; \beta) - X(t_1; \beta) \) for all possible values of \( t_1, t_2 \), such that \( t_1 < t_2 \). However, as we shall show at the next proposition, the distributions of the increments \( X(t_2; \beta) - X(t_1; \beta) \) can be expressed in terms of the distributions of \( X(t; \beta) \). So, we have the following,

**Remark 2.17:** For the special case of an independent increment process, where \( X(0; \beta) = 0 \) almost surely we will have

\[
X(t; \beta) = (X(t; \beta) - X(s; \beta)) + (X(s; \beta) - X(0; \beta))
\]

where \( (X(t; \beta) - X(s; \beta)) \) and \( (X(s; \beta) - X(0; \beta)) \) are independent random variables. Consequently stochastic process \( X(t; \beta) \) verifies the Markov property. ■

**Proposition 2.18:** Let the independent–increment s.p. \( X(t; \beta), t \in I \). Then all the finite-dimensional distributions of the process \( X(t; \beta) \) can be expressed in terms of the one dimensional distribution of the random variable \( X(t; \beta) \) for every fixed \( t \).

**Proof:** We will prove the above assertion in two steps. First we have for the general finite dimensional characteristic function

\[
\phi_{h_1, \ldots, h_n}(u_1, u_2, \ldots, u_n) = E^{\beta} \left[ \exp \left( \sum_{k=1}^{n} X(t_k; \beta) \cdot u_k \right) \right] =
\]

\[
= E^{\beta} \left[ \exp \left( i \sum_{k=1}^{n} u_k X(t_k; \beta) + i \sum_{i=1}^{n} u_i \times [X(t_2; \beta) - X(t_1; \beta)] + \ldots + i \sum_{i=n}^{n} u_i \times [X(t_n; \beta) - X(t_{n-1}; \beta)] \right) \right] =
\]

\[
= \phi_{h_1}(u_1) \cdot \phi_{h_2}(u_2) \cdots \phi_{h_n}(u_n) \tag{26}
\]

Let us now take two arbitrary time instants \( t_1 < t_2 \). It is clear that \( X(t_2; \beta) = X(t_1; \beta) + [X(t_2; \beta) - X(t_1; \beta)] \), i.e. \( X(t_2; \beta) \) is a sum of two independent random variables. Hence the following relation holds for the characteristic functions,

\[
\phi_{h_1, h_2}(u) = \frac{\phi_{h_1}(u)}{\phi_{h_2}(u)} \tag{27}
\]

for each \( t_1 < t_2 \). Using the eq.(26), we have

\[
\phi_{h_1, h_2, \ldots, h_n}(u_1, u_2, \ldots, u_n) = \phi_{h_1}(u_1 + u_2 + \ldots + u_n) \cdot \frac{\phi_{h_2}(u_2 + \ldots + u_n)}{\phi_{h_2}(u_2 + \ldots + u_n)} \cdots \frac{\phi_{h_n}(u_n)}{\phi_{h_n}(u_n)} \tag{28}
\]

**Remark 2.19:** It should be noticed that the one dimensional distribution of the independent–increment s.p. cannot be given arbitrary. The one dimensional characteristic function has to be such that \( \phi_{h_1, h_2}(u) \) given by (27) is also a characteristic function. ■

An important class of independent–increment s.p. is formed by the processes with orthogonal increments. The Riemann–Stieljes integrals with respect to process with orthogonal increments
are particularly important in the study of stationary stochastic processes. Moreover, as we shall prove at the sequel every stochastic process has orthogonal increments if and only if is a martingale process (Preposition 2.39).

**Definition 2.20:** A s.p. \( X(t; \beta), t \in I \) is said to be an orthogonal increment process, if for every \( t < s \leq u < v \) in \( I \),
\[
E^\beta \left[ \left( X(t; \beta) - X(s; \beta) \right)^2 \right] < \infty
\]
and
\[
E^\beta \left[ \left( X(t; \beta) - X(s; \beta) \right) \left( X(u; \beta) - X(v; \beta) \right) \right] = 0
\]
(29)

**Remark 2.21:** From the above definition we can easily see that an independent–increment s.p. that fulfills eq. (29) and has constant mean value, i.e. \( m_x(t) = \text{const.} \) is also an orthogonal increment process.

We now define the function
\[
F(t; t_0) = \begin{cases} 
E^\beta \left[ \left( X(t; \beta) - X(t_0; \beta) \right)^2 \right], & t \geq t_0 \\
-E^\beta \left[ \left( X(t; \beta) - X(t_0; \beta) \right)^2 \right], & t < t_0
\end{cases}
\]
(31)

Clearly, we have \( F(t; t_0) = -F(t_0; t) \) for every \( t, t_0 \in I \).

**Theorem 2.22:** Let the orthogonal increment s.p. \( X(t; \beta), t \in I \). Then for the variance of the increment \( X(t_1; \beta) - X(t_2; \beta) \) for every \( t_1 < t_2 \) with \( t_1, t_2 \in I \) the following relation holds
\[
F(t; t_1) = F(t; t_2) + E^\beta \left[ \left( X(t_1; \beta) - X(t_2; \beta) \right)^2 \right], & t \in I
\]
(32)

**Proof:** Without any loss of generality we can assume \( t_1 < t_2 \). Under this assumption we have three cases. Indeed, consider the new constant \( t_1 \). This constant can be choose such that \( t_1 < t_0 < t \) or \( t_0 < t_1 < t \). We examine the first case. (For the sequel we will denote \( F(t; t_1) \) as \( F_1(t) \))

Case I):

\[
F_1(t) = -E^\beta \left[ \left( X(t; \beta) - X(t_1; \beta) \right)^2 \right] =
\]
\[
= -E^\beta \left[ \left( X(t; \beta) - X(t_1; \beta) + X(t_2; \beta) - X(t_2; \beta) \right)^2 \right] =
\]
\[
= -E^\beta \left[ \left( X(t; \beta) - X(t_2; \beta) \right)^2 \right] - E^\beta \left[ \left( X(t_1; \beta) - X(t_2; \beta) \right)^2 \right] -
\]
\[
-2E^\beta \left[ \left( X(t; \beta) - X(t_2; \beta) \right) \left( X(t_2; \beta) - X(t_1; \beta) \right) \right]
\]
\[
= F_2(t) - E^\beta \left[ \left( X(t_2; \beta) - X(t_1; \beta) \right)^2 \right] + 2E^\beta \left[ \left( X(t_2; \beta) - X(t_1; \beta) \right)^2 \right]
\]
\[
= F_2(t) + E^\beta \left[ \left( X(t_2; \beta) - X(t_1; \beta) \right)^2 \right]
\]
Case II):

\[
F_i(t) = E^g \left( \left[ X(t; \beta) - X(t_i; \beta) \right]^2 \right) = \\
= E^g \left( \left[ X(t; \beta) - X(t_1; \beta) + X(t_2; \beta) - X(t_2; \beta) \right]^2 \right) = \\
= E^g \left( \left[ X(t; \beta) - X(t_2; \beta) \right]^2 \right) E^g \left( \left[ X(t_2; \beta) - X(t_1; \beta) \right]^2 \right) + \\
+ 2E^g \left( [X(t; \beta) - X(t_2; \beta)][X(t_2; \beta) - X(t_1; \beta)] \right) \\
= E^g \left( \left[ X(t; \beta) - X(t_2; \beta) \right]^2 \right) E^g \left( \left[ X(t_2; \beta) - X(t_1; \beta) \right]^2 \right) + \\
- 2E^g \left( [X(t; \beta) - X(t_2; \beta)][X(t; \beta) - X(t_2; \beta)] \right) \\
= -E^g \left( \left[ X(t; \beta) - X(t_2; \beta) \right]^2 \right) + E^g \left( \left[ X(t_2; \beta) - X(t_1; \beta) \right]^2 \right) \\
= F_i(t) + E^g \left( \left[ X(t_2; \beta) - X(t_1; \beta) \right]^2 \right)
\]

Case III):

\[
F_i(t) = E^g \left( \left[ X(t; \beta) - X(t_i; \beta) \right]^2 \right) = \\
= E^g \left( \left[ X(t; \beta) - X(t_1; \beta) + X(t_2; \beta) - X(t_2; \beta) \right]^2 \right) = \\
= E^g \left( \left[ X(t; \beta) - X(t_2; \beta) \right]^2 \right) E^g \left( \left[ X(t_2; \beta) - X(t_1; \beta) \right]^2 \right) + \\
+ 2E^g \left( [X(t; \beta) - X(t_2; \beta)][X(t_2; \beta) - X(t_1; \beta)] \right) \\
= F_i(t) + E^g \left( \left[ X(t_2; \beta) - X(t_1; \beta) \right]^2 \right)
\]

We thus have complete the proof. ■

Hence, we see that the change of the parameter \( t_0 \) change the function \( F(t; t_0) \) only by a positive constant. For the sequel we will denote \( F(t; t_0) \) as \( F(t) \) and we will call it as the characteristic of \( X(t; \beta) \). From the above theorem we can easily see that the variance of the increment \( X(t_i; \beta) - X(t_2; \beta) \) can be described by

\[
E^g \left( \left[ X(t_i; \beta) - X(t_2; \beta) \right]^2 \right) = F(t_2) - F(t_1), \quad t_i < t_2 \in I
\] (33)

Thus, for every orthogonal increment process we have
Preposition 2.23 [Levy property]: Let the s.p. \( X(t; \beta) \) \( t \in I \) with orthogonal increments and \( a = t_1 < t_2 < \ldots < t_n = b \) an arbitrary partition of the interval \([a, b]\) and \( \Delta = \max(t_{i+1} - t_i) \). Then,

\[
\lim_{n \to \infty} \sum_{i=0}^{n-1} \left[ X(t_{i+1}; \beta) - X(t_i; \beta) \right]^2 = F(b) - F(a) \quad \square
\]

Now we are going to describe some specific cases of s.p.’s described above. The most important examples of independent–increment s.p.’s are the Wiener process and the Poisson process. For a complete study of processes with independent increments we refer to PUGACHEV, V.S. & SINITSYN, (Stochastic Differential Systems).

\( \quad \)

a) \textbf{Wiener Process}

\textbf{Definition 2.24}: A s.p. \( W(t; \beta) \) \( t \in I \) is called a \textit{Wiener} process or \textit{Brownian motion} process if:

a) \( \mathcal{D}_x^{\infty}(t_{(0-0)}(\{0\}) = \mathcal{D}_x^{\infty}(\{x(\cdot) \in C(I) : x(0) \in \{0\}\}) = 1 \)

b) for every \( n \) and every \( \{t_1 < t_2 < \ldots < t_n\} \) in \( T \), the increments \( W(t_2; \beta) - W(t_1; \beta), \ldots, W(t_n; \beta) - W(t_{n-1}; \beta) \) are independent

c) for arbitrary \( t \) and \( h > 0 \) the increment \( W(t + h; \beta) - W(t; \beta) \) has a Gaussian distribution with

\[
E^\beta(W(t + h; \beta) - W(t; \beta)) = \mu h \\
E^\beta\left([W(t + h; \beta) - W(t; \beta)]^2\right) = \sigma^2 h
\]

where \( \mu \) a real constant, called drift coefficient and \( \sigma^2 \) a positive constant, called variance. \( \square \)

Without restriction one can assume that \( \mu = 0 \) and \( \sigma^2 = 1 \). In what follows we consider this case (standard Wiener process).

\textbf{Preposition 2.25}: The standard Wiener s.p. \( W(t; \beta) \) \( t \in I \) has the following properties:

a) \( W(t; \beta) \) is a centred, second–order, Gaussian s.p., i.e.

\[
E^\beta(W(t; \beta)) = 0 \quad \text{and} \quad E^\beta\left([W(t; \beta)]^2\right) < \infty
\]

This process is not stationary and its covariance function is written as

\[
C_{ww}(t_1, t_2) = E^\beta(W(t_1; \beta)W(t_2; \beta)) = \min(t_1, t_2), \quad t_1, t_2 \in I
\]

b) \( W(t; \beta) \) is a Markov s.p. whose system of transition probabilities is homogeneous and for \( \tau > 0 \), is written as

\[
F_{\alpha\beta}(x, \tau; z) = \frac{1}{\sqrt{2\pi \tau}} \exp\left(-\frac{1}{2} \frac{(y-x)^2}{\tau}\right) dy
\]

(37)

c) \( W(t; \beta) \) is mean–square continuous on \( I \).

d) The joint probability density function has the form
\[ f^{W}_{t_1,t_2,\ldots,t_n}(x_1, x_2, \ldots, x_n) = \exp \left\{ -\sum_{i=1}^{n} \frac{1}{2} \left( \frac{x_i - x_{i-1}}{t_i - t_{i-1}} \right)^2 \right\} \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \] (38)

**Proof**: We only give some elements of the proof herein. For more details we refer to SOIZE, C. (Mathematical Methods in Signal Analysis). For the proof of point a) we can write for \( t > 0 \),

\[ W(t; \beta) = (W(t; \beta) - W(0; \beta)) + W(0; \beta) \]

We then use Definition 2.24. Point b) can be proved by direct application of Remark 2.17 and using the Gaussian structure proved at point a). The proof of point c) is obtained by calculating the mean square norm of the difference, i.e. by direct application of m.s.-continuous definition. For point d) we notice that the vector \( (W(t_1; \beta), \ldots, W(t_n; \beta)) \) is a non-degenerate linear transformation of the Gaussian vector \( (W(t_1; \beta) - W(t_0; \beta), \ldots, W(t_n; \beta) - W(t_{n-1}; \beta)) \). For the last vector the density is given by

\[ f^{\mathcal{N}}_{t_1,t_2,\ldots,t_n}(x_1, x_2, \ldots, x_n) = \exp \left\{ -\sum_{i=1}^{n} \frac{1}{2} \cdot \frac{x_i^2}{t_i - t_{i-1}} \right\} \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \] (39)

Hence after some calculus we get the desired result. ■

The fundamental properties of the realizations of a Wiener s.p. are summarized in the following

**Theorem 2.26**: Let the Wiener s.p. \( W(t; \beta) \), \( t \in I \). Almost all realizations (sample functions) of \( W(t; \beta) \), are continuous, nowhere differentiable and have unbounded variation in every finite interval.

**Proof**: The continuity follows from the Kolmogorov criterion which will be formulated in Section 2/Chapter 4 (see Example 2.8/Chapter 4). The non–differentiability for fixed \( t \) can be made clear as follows: the distribution of the difference quotient

\[ \frac{1}{h} [W(t + h; \beta) - W(t; \beta)] \]

is the Gaussian distribution \( \mathcal{N}(0, \frac{1}{h}) \) which diverges as \( h \to 0 \); therefore this quotient can not converge with positive probability to a finite random variable in any probabilistic sense. The last assertion of the theorem follows from Preposition 2.23 and some additional considerations connected with almost sure convergence. ■

**b) Poisson process**

**Definition 2.27.** A s.p. \( N(t; \beta) \), \( t \in I \) is called a non–homogeneous Poisson process if

a) the s.p. \( N(t; \beta) \) has independent increments

b) \( \mathcal{D}^{N}_{I(t=0)}(\{0\}) = \mathcal{D}^{N}_{D(t)}(\{N(\bullet) \in D(I) : N(0) \in \{0\}\}) = 1 \)
2.2. CLASSIFICATION OF STOCHASTIC PROCESSES

c) for arbitrary $t_1, t_2$ the difference $N(\alpha; \beta) - N(\beta; \beta)$ have a Poisson distribution, that is there exist a function $\lambda(t)$ that for each integer number $k > 0$

$$\mathcal{D}_N^{\alpha}(t, \beta) = \mathcal{D}_N^{\beta}(t, \beta) \left\{ \{N(\cdot) \in D(I) : N(t_1) - N(t_2) \in \{k\}\} \right\} = \frac{[\lambda(t) - \lambda(t_2)]^k}{k!} e^{[\lambda(t) - \lambda(t_1)]}$$

If $\lambda(t) = \nu t$, $\nu > 0$ then the process $N(t; \beta)$ is called a homogeneous Poisson process. ■

It is clear that the processes $N(t; \beta)$ for each $t \in I$ are random variables with Poisson distribution

$$\mathcal{D}_N^{\alpha}(t, \beta) = \mathcal{D}_N^{\beta}(t, \beta) \left\{ \{N(\cdot) \in D(I) : N(t) \in \{k\}\} \right\} = \frac{[\lambda(t)]^k}{k!} e^{-\lambda(t)} \tag{40}$$

**Preposition 2.28** : Let $N(t; \beta)$ be a Poisson process with mean value $\lambda(t)$. Then $N(t; \beta)$ is a Markov process with transition probability

$$\mathcal{D}_N^{\alpha}(t, \beta) = \mathcal{D}_N^{\beta}(t, \beta) \left\{ \{N(t) - N(s) \in \{k\}\} \right\} = \frac{[\lambda(t) - \lambda(s)]^k}{k!} e^{-(\lambda(t) - \lambda(s))} \tag{41}$$

For the case when $\lambda(t) = t$ we have a homogeneous Poisson process and the homogenous transition probability has the form

$$\mathcal{D}_N^{\alpha}(t, \beta) = \mathcal{D}_N^{\beta}(t, \beta) \left\{ \{N(t) - N(s) \in \{k\}\} \right\} = \frac{(t-s)^k}{k!} e^{-t+s} \tag{42}$$

**Proof** : Direct application of Remark 2.17 and equation (14). ■

The following assertions characterize the basic properties of the Poisson process.

**Theorem 2.29** : For an homogeneous Poisson process with intensity $\nu$, the interarrival times $t_1, t_2, \ldots, t_n$ are independent and identically distributed random variables with the common distribution being exponential with parameter $\nu$. ■

**Proof** : For the proof we refer to SNYDER, D.L., (Random Point Processes). ■

**Remark 2.30** : It is however, worthy adding that interarrival times for an inhomogeneous Poisson process are not independent. ■

The fundamental properties of the realizations of a Poisson s.p. are summarized in the following

**Theorem 2.31** : Almost all sample functions of the Poisson process are non-decreasing with integer increments; they increase by jumps equal to one and the number of jumps in every finite interval is finite. ■

**Proof** : For the proof we refer to SNYDER, D.L., (Random Point Processes). ■
2.1.3.d. Martingales

An important class of processes is formed by martingales. We first recall the definition of a current of $\sigma-$algebras of a probability space. Let $I$ denotes the set of real numbers corresponding to the times at which the stochastic experiments were carried out.

**Definition 2.32**: Let the probability space $\left(\Omega, \mathcal{B}(\Omega), \mathbb{P}\right)$. A monotonically decreasing family of $\sigma-$algebras $\mathcal{B}_t(\Omega), t \in I$ with

a) $\mathcal{B}_t(\Omega) \subset \mathcal{B}(\Omega)$

b) $\mathcal{B}_{t_1}(\Omega) \subset \mathcal{B}_{t_2}(\Omega)$ if $t_1 < t_2$

is called a current of $\sigma-$algebras. ■

**Definition 2.33**: Let the s.p. $\left\{X(t;\beta), t \in I\right\}$, defined on the probability space $\left(\Omega, \mathcal{B}(\Omega), \mathbb{P}\right)$, which is equipped with a current of $\sigma-$algebras $\mathcal{B}_t(\Omega), t \in I$. The s.p. $X(t;\beta), t \in I$ is adapted to the current of $\sigma-$algebras $\mathcal{B}_t(\Omega), t \in I$ iff the random variable $X(t;\beta)$ is $\mathcal{B}_t(\Omega)$-measurable for every $t \in I$. ■

**Remark 2.34**: It follows from the definition that $\mathcal{B}_t(\Omega)$ always contains the $\sigma-$algebra generated by the random variables $\{X(s;\beta), s < t\}$. ■

**Remark 2.35**: For every real valued s.p. $X(t;\beta), t \in I$, one current of $\sigma-$algebras $\mathcal{B}_t(\Omega)$ is the following,

$$\mathcal{B}^X_t(\Omega) = \sigma\left(\left\{X^{-1}(s;A) : s \in [0,t], A \in \mathcal{B}(\mathbb{R})\right\}\right) \quad (43)$$

where $\sigma(A)$ denotes the generated $\sigma-$algebra by the set $A$. The inverse sign is for the random argument $A$. As we can understand $\mathcal{B}^X_t(\Omega)$ is representing a $\sigma-$algebra of the physical events, as they have been evolved, until time $t$, i.e. $\mathcal{B}^X_t(\Omega)$ characterizes a collection of events representing the known information at time $t$. The above remark can be extended for the case of a vector or complex valued s.p., with appropriate change to the Borel sets $\mathcal{B}(\mathbb{R})$. ■

**Definition 2.36**: Let the s.p. $X(t;\beta), t \in I$. We assume that $X(t;\beta)$ is $\mathcal{B}_t(\Omega)$-measurable. The above s.p. is called a martingale iff

$$E^\beta\left(X(t;\beta) | \mathcal{B}_s(\Omega)\right) = E^\beta\left(X(s;\beta) | \mathcal{B}_s(\Omega)\right) = X(s) \quad (44)$$

for all $s,t \in T$, and $s \leq t$. ■

**Remark 2.37**: Often the term martingale is referred to processes for which relation (44) holds for the family of $\sigma-$algebras generated by the processes $X(t;\beta)$ itself (Remark 2.34.), what means that the history of the process $X(t;\beta)$ prior to the instant $t$ is chosen as condition. ■

**Example 2.38**: A Wiener s.p. $W(t;\beta), t \in I$, is a martingale, since
2.2. CLASSIFICATION OF STOCHASTIC PROCESSES

\[ E^\beta \left( W(t; \beta) \mid \mathcal{B}_s(\Omega) \right) = E^\beta \left( W(s; \beta) + [W(t; \beta) - W(s; \beta)] \mid \mathcal{B}_s(\Omega) \right) = \]
\[ = W(s) + E^\beta \left( [W(t; \beta) - W(s; \beta)] \mid \mathcal{B}_s(\Omega) \right) = \]
\[ = W(s) + E^\beta \left( [W(t; \beta) - W(s; \beta)] \right) = W(s) \]
since the random variable \( W(t; \beta) - W(s; \beta) \) is independent of all events of \( \sigma \)-algebra \( \mathcal{B}_s(\Omega) \). ■

**Preposition 2.39**: The stochastic process \( X(t; \beta), t \in I \) is a square integrable martingale iff is an orthogonal increment process. Thus, every martingale can be described by a characteristic and satisfies the Levy property.

**Proof**: If \( X(t; \beta), t \in I \) is a square integrable martingale then we clearly have for every \( t, s \in I \)
\[ E^\beta \left( [X(t; \beta) - X(s; \beta)] \right) = 0 \]
Moreover for any disjoint sets \([t_1, t_2] \subset I \) and \([s_1, s_2] \subset I \) we have
\[ E^\beta \left( [X(t_1; \beta) - X(t_2; \beta)] \cdot [X(s_1; \beta) - X(s_2; \beta)] \right) = 0. \]
Hence \( X(t; \beta), t \in I \) is an orthogonal increment process.

If \( X(t; \beta), t \in I \) is an orthogonal increment process then the independence of increments implies that for \( t, s \in I \) with \( s < t \), we have the independence of \( \mathcal{B}_s(\Omega) \) and \( X(t; \beta) - X(s; \beta) \). (See Theorem 13.10 of YEH, J., Martingales and Stochastic Analysis). Thus
\[ E^\beta \left( [X(t; \beta) - X(s; \beta)] \mid \mathcal{B}_s(\Omega) \right) = E^\beta \left( [X(t; \beta) - X(s; \beta)] \right) = 0 \]
Then
\[ E^\beta \left( X(t; \beta) \right) = E^\beta \left( [X(t; \beta) - X(s; \beta)] \mid \mathcal{B}_s(\Omega) \right) + E^\beta \left( X(s; \beta) \mid \mathcal{B}_s(\Omega) \right) \]
\[ = X(s; \beta) \]
Therefore \( X(t; \beta) \) is a martingale. For more details on the proof of the above theorem we refer to PROHOROV, YU.V. & ROZANOV, YU.A., (Probability Theory. p.137) and YEH, J. (Martingales and Stochastic Analysis. p.86). ■

**Preposition 2.40**: Let \( X(t; \beta), t \in I \) be a martingale s.p, \( \mathcal{B}_s(\Omega) \)- measurable. Then some basic properties, it possesses, are as follows:

a) Sample functions of separable martingales have no discontinuities of the second kind; so, they have at worst jumps.

b) \( E^\beta \left( X(t; \beta) \right), t \in I \) is constant. So, for every \( t_1, t_2 \in I \) we have
\[ E^\beta \left( X(t_2; \beta) - X(t_1; \beta) \right) = 0. \]
c) If \( Y(t;\beta) \) is another martingale with respect to the same family \( \mathcal{B}_t(\Omega) \), then 
\[ a_1X(t;\beta) + a_2Y(t;\beta) \] (where \( a_1, a_2 \in \mathbb{R} \)) is also a martingale, \( \mathcal{B}_t(\Omega) \)-measurable.

d) Moreover if \( \Psi(\beta) \) is a \( \mathcal{B}_t(\Omega) \)-measurable random variable, then for every \( t_1, t_2 \in I \), greater than \( t \) we have
\[ E^\beta \left( \Psi(\beta) \cdot [X(t_2;\beta) - X(t_1;\beta)] \right) = 0 \]

**Proof:** The first assertions follows from Chentsov Theorem that it will be formulated in Theorem 2.9 of Chapter 4. The next two assertions follow directly from the definition of martingales. The proof of assertion d) can be found at PROHOROV, YU.V. & ROZANOV, YU.A., (Probability Theory. p.137).
2.2.2. Classification Based Upon Regularity

In terms of statistical regularities, stochastic processes can be grouped into two classes: *stationary stochastic processes* and *nonstationary stochastic processes.*

The $n$th probability distribution function, $F_{x_1,x_2,...,x_n}(x_1,x_2,...,x_n)$ of a s.p. $X(t;\beta), t \in I$, is in general a function of $x_1,x_2,...,x_n$ and $t_1,t_2,...,t_n$. A nonstationary s.p. is one whose distribution functions depend upon the values of time parameters explicitly. The statistical behaviour of a nonstationary s.p. thus depends upon the absolute origin of time. Clearly, most stochastic processes modelling physical phenomena are nonstationary. In particular, all physical processes having a transient period or certain damping characteristics are of the nonstationary type.

On the other hand, many stochastic processes occurring in nature have the property that their statistical behaviour does not vary significantly with respect to their parameters. The surface of the sea in spatial and time coordinates, noise in time in electric circuits under steady state operation have the appearance that their fluctuations as functions of time or spatial positions stay roughly the same. Because of powerful mathematical tools which exist for treating stationary s.p.’s, this class of s.p.’s is of great practical importance, and we consider it in some detail in the following sections. More exhaustive accounts of stationary stochastic processes can be found in the work of PUGACHEV, V.S. & SINITSYN, Stochastic Differential Systems.

2.2.2.a. Stationary and Wide-Sense Stationary Processes

**Definition 2.41.** A s.p. $X(t;\beta), t \in I$, is said to be *stationary* or *strictly stationary* if its collection of probability distributions stay invariant under an arbitrary translation of the time parameter, that is, for each $n$ and for an arbitrary $\tau$,

$$F_{x_1,x_2,...,x_n}(x_1,x_2,...,x_n) = F_{x_{\tau+1},x_{\tau+2},...,x_{\tau+n}}(x_{\tau+1},x_{\tau+2},...,x_{\tau+n}) \quad t_j + \tau \in I, \quad j = 1,2,...,n \quad (45)$$

Let $\tau = -t_1$, in the above. We see that the probability distributions depend upon the time parameters only through their differences. In other words, the statistical properties of a stationary s.p. are independent of the absolute time origin.

We can easily see that a stationary s.p. $X(t;\beta), t \in I$, possesses the following important properties for its moments, if they exist. Since the first distribution function $F_t(x)$ is not a function of $t$, we have

$$E^\delta \{X(t;\beta)\} = m_x = \text{const.}$$

Since

$$F_{t_1,t_2}(x_1,x_2) = F_{0,t_2-t_1}(x_1,x_2)$$

we have

$$E^\delta \{X(t;\beta)X(t+\tau;\beta)\} = C_{XX}(\tau)$$

In view of the general relation (18), it follows that
The correlation function of a stationary stochastic process is thus an even function of \( \tau \). Properties of higher moments can also be easily derived.

Given a physical problem, it is often quite difficult to ascertain whether the stationary property holds since Eq. (45) must hold for all \( n \). For practical purposes, we are often interested in a wider class of stationary stochastic processes.

**Definition 2.42.** A s.p. \( X(t; \beta), \ t \in I, \) is said to be wide-sense stationary if

\[
E^3(X(t; \beta)) = m_x = \text{const.} < \infty
\]

and

\[
E^3\left(\left(X(t; \beta)\right)^2\right) < \infty, \quad E^3\left(X(t_1; \beta) \cdot X(t_2; \beta)\right) = C_{xx}(t_2 - t_1)
\]

A wide-sense stationary s.p. is sometimes called second-order stationary, weakly stationary, or covariance stationary. It is clear that a strictly stationary process whose second moment is finite is also wide-sense stationary, but the converse is not true in general. An important exception is the Gaussian s.p. As we shall see, a Gaussian stochastic process is completely specified by its means and covariance functions. Hence, a wide-sense stationary Gaussian stochastic process is also strictly stationary.

**Example 2.43.** Consider a s.p. \( X(t; \beta), \ t \in I, \) defined by

\[
X(t; \beta) = a \cos(\omega t + \Phi)
\]

where \( a \) and \( \omega \) are real constants and \( \Phi \) is a r.v. uniformly distributed in the interval \([0, 2\pi]\). It is a simplified version of the Pierson Longuet-Higgins model discussed in Example 1.2 of Chapter 1. We easily obtain

\[
E^3(X(t; \beta)) = \frac{a}{2\pi} \int_0^{2\pi} \cos(\omega t + \phi) d\phi = 0
\]

\[
C_{xx}(t_1, t_2) = a^2 E^4(\cos(\omega t_1 + \phi) \cos(\omega t_2 + \phi)) = \frac{1}{2} a^2 \cos(\omega(t_1 - t_2))
\]

Equation (49) is thus wide-sense stationary.

### 2.2.2.b. Spectral Densities and Correlation Functions

We restrict ourselves here to the study of wide-sense stationary stochastic processes and develop further useful concepts in the analysis of this class of processes. It has been noted that the covariance or correlation function of a stochastic process plays a central role. Let us first develop a number of additional properties associated with the correlation function

\[
C_{xx}(\tau) = E^3(X(t + \tau; \beta) \cdot X(t; \beta))
\]

of a wide-sense stationary s.p. \( X(t; \beta), \ t \in I, \)
1) \( C_{XX}(\tau) \) is an even function of \( \tau \), that is
\[ C_{XX}(\tau) = C_{XX}(-\tau) \]
This has been derived in Section 5.2.1 Eq. (46).

2) \( C_{XX}(\tau) \) always exists and is finite. Furthermore
\[
|C_{XX}(\tau)| \leq C_{XX}(0) = E^\beta \left( \left[ X(t; \beta) \right]^2 \right) \tag{50}
\]
This result is a direct consequence of Eq. (3.19).

3) If \( C_{XX}(\tau) \) is continuous at the origin, then it is uniformly continuous in \( \tau \). In order to show this, let us consider the difference
\[
C_{XX}(\tau + \epsilon) - C_{XX}(\tau) = E^\beta \left( X(t + \tau; \beta) - X(t - \epsilon; \beta) - X(t; \beta) \right)
\]
Using the Schwarz inequality we obtain the inequality
\[
|C_{XX}(\tau + \epsilon) - C_{XX}(\tau)|^2 \leq 2C_{XX}(0)[C_{XX}(0) - C_{XX}(\epsilon)]
\]
Since \( C_{XX}(\tau) \) is continuous at the origin, \( C_{XX}(0) - C_{XX}(\epsilon) \to 0 \) as \( \epsilon \to 0 \). Hence, the difference \( C_{XX}(\tau + \epsilon) - C_{XX}(\tau) \) tends to zero with \( \epsilon \) uniformly in \( \tau \).

4) As indicated by Eq. (3.20), \( C_{XX}(\tau) \) is nonnegative definite in \( \tau \).

5) Let \( C_{XX}(\tau) \) be continuous at the origin. It can be represented in the form
\[
C_{XX}(\tau) = \frac{1}{2} \int_{-\infty}^{\infty} e^{i\omega \tau} d\Phi_{XX}(\omega) \tag{51}
\]
If \( \Phi(\omega) \) is absolutely continuous, we have
\[
S_{XX}(\omega) = d\Phi_{XX}(\omega) / d\omega
\]
and
\[
C_{XX}(\tau) = \frac{1}{2} \int_{-\infty}^{\infty} e^{i\omega \tau} S_{XX}(\omega) d\omega \tag{52}
\]
where \( S(\omega) \) is real and nonnegative.

By virtue of Properties 2), 3), and 4) possessed by \( C_{XX}(\tau) \) this result is a direct consequence of the following important theorem due to Bochner.

**Theorem 2.44:** A function \( C(\tau) \) is continuous, real, and nonnegative definite if and only if, it has the representation (51), where \( \Phi_{XX}(\omega) \) is real, non-decreasing and bounded.

**Proof:** For the proof we refer to DOOB, J.L., (Stochastic Processes) ■

We see from Eq. (52) that the functions \( C_{XX}(\tau) \) and \( S_{XX}(\omega) \) form a Fourier transform pair. The relations
\[
C_{XX}(\tau) = \frac{1}{2} \int_{-\infty}^{\infty} e^{i\omega \tau} S_{XX}(\omega) d\omega \quad S_{XX}(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} e^{i\omega \tau} C_{XX}(\tau) d\tau \tag{53}
\]
are called the **Wiener-Khintchine formulas**.

Equations (53) lead to the definition of the power spectral density function of a wide-sense stationary process whose correlation function is continuous.
Definition 2.45: The power spectral density function of a wider-sense stationary s.p. \( X(t; \beta) \), \( t \in I \) is defined by the first equation of (53), that is,

\[
S_{XX}(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} e^{i\omega \tau} C_{XX}(\tau) d\tau
\]

We have already noted that the power spectral density function \( S_{XX}(\omega) \) is real and nonnegative. Furthermore, since \( C_{XX}(\tau) \) is an even function of \( \tau \), the Wiener-Khintchine formulas may be written as

\[
S_{XX}(\omega) = \frac{2}{\pi} \int_{0}^{\infty} C_{XX}(\tau) \cos(\omega \tau) d\tau \quad C_{XX}(\tau) = \int_{0}^{\infty} S(\omega) \cos(\omega \tau) d\omega
\]  

(54)

Hence, \( S_{XX}(\omega) \) is also an even function of \( \omega \). When \( \tau = 0 \), the second of Eqs. (54) reduces to

\[
\int_{0}^{\infty} S(\omega) d\omega = C_{XX}(0)
\]

For a detailed discussion for stochastic processes we refer to ATHANASSOULIS, G.A., (Stochastic Modeling and Forecasting of Ship Systems) and PUGACHEV, V.S. & SINITSYN, (Stochastic Differential Systems).
2.2.3. Classification Based Upon Ergodic Properties

Ergodicity deals with the specific question of relating statistical or ensemble averages of a stationary stochastic process to time averages of its individual sample functions. The interchangeability of ensemble and time averages has considerable appeal in practice because, when statistical averages of a stochastic process need to be computed, what is generally available is not a representative collection of sample functions but rather certain pieces of sample functions or a long single observation of one sample function. One naturally asks if certain statistical averages of a stochastic process can be determined from appropriate time averages associated with a single sample function.

Let \( x(t) \) be a sample function of a stationary s.p. \( X(t;\beta), t \in I = [0,T] \). The time average of a given function of \( x(t) \), \( g(x(t)) \), denoted by \( g(x(t)) \), is defined by

\[
\langle g(x(t)) \rangle_I = \frac{1}{T} \int_{-T/2}^{T/2} g(x(t)) \, dt
\]

if the limit exists. In order to define the ergodic property of a stationary stochastic process, the sample function \( x(t) \) is allowed to range over all possible sample functions and, for a fixed \( T \), over all possible time translations; thus defines a random variable.

**Definition 2.46:** A s.p. \( X(t;\beta), t \in I \) is said to be ergodic relative to \( G \) if, for every \( g \in G \),

\[
\langle g[X(t;\beta)] \rangle_{\infty} = \langle E^\beta \{g[X(t;\beta)]\} \rangle_{\infty} \quad \blacksquare
\]

It is, of course, entirely possible that, given a s.p. \( X(t;\beta) \), the ergodicity condition stated above is satisfied for certain functions of \( X(t;\beta) \) and is not satisfied for others. In physical applications we are primarily interested in the following:

1. **Ergodicity in the Mean:** \( g[X(t;\beta)] = X(t;\beta) \) The ergodicity condition becomes

\[
\langle X(t;\beta) \rangle_{\infty} = \langle E^\beta \{X(t;\beta)\} \rangle_{\infty} \quad (57)
\]

2. **Ergodicity in the Correlation Function:** \( g[X(t;\beta)] = X(t;\beta) \cdot X(t + \tau;\beta) \). The ergodicity condition is

\[
\langle X(t;\beta) \cdot X(t + \tau;\beta) \rangle_{\infty} = \langle E^\beta \{X(t;\beta) \cdot X(t + \tau;\beta)\} \rangle_{\infty} = \langle R_{xx}(t,t + \tau) \rangle_{\infty} \quad (58)
\]

For these ergodicity criteria, mathematical conditions can be derived for the purpose of verifying the ergodicity conditions. These conditions are direct consequence of the following

**Theorem 2.47 [Ergodicity in the Mean]:** Let \( X(t;\beta), t \in I \) be an integrable, second order, stochastic process and \( m_x(T;\beta) = \langle X(t;\beta) \rangle_I \). We also assume that the correlation function \( C_{xx}(t_1,t_2) \) verify the condition
\[
\lim_{T \to \infty} \frac{1}{T^2} \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} C_{XX}(t_1, t_2) \, dt_1 \, dt_2 = 0 \tag{59}
\]

Then the ergodicity condition in the mean holds, i.e.

\[
\langle X(t; \beta) \rangle_T' = \left\langle E^\beta \left(X(t; \beta)\right) \right\rangle_T'
\]

**Proof:** For the proof we refer to ATHANASSOULIS, G.A., (Stochastic Modeling and Forecasting of Ship Systems. Lecture Notes NTUA).

**Theorem 2.48 [Statistical Estimation of mean value for a stationary s.p.]** Let \( X(t; \beta), \ t \in I \) be a second order, m.s. continuous and stationary stochastic process. Then, for every \( \varepsilon > 0 \) and for every \( T > 0 \), the following inequality holds,

\[
\mathcal{D}_X \left( \left\langle X(t) \right\rangle_T' - m_X \right) > \varepsilon \right) \leq \frac{2}{T} \int_0^T \left( 1 - \frac{t}{T} \right) C_{XX}(\tau) \, d\tau \tag{60}
\]

Concerning the Ergodicity criterion for the correlation function of a stochastic process the following theorem holds,

**Theorem 2.48 [Ergodicity in the Correlation Function]** Let \( X(t; \beta), \ t \in I \) be an integrable, fourth order, stochastic process and \( \hat{R}_{XX}(\tau, T; \beta) = \left\langle X(t + \tau; \beta) \cdot X(t; \beta) \right\rangle_T' \). Let also the correlation function \( C_{YY}(t_1, t_2; \tau) \) of the stochastic process \( Y(t, \tau; \beta) = X(t + \tau; \beta) \cdot X(t; \beta) \), verify the condition

\[
\lim_{T \to \infty} \frac{1}{T^2} \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} C_{YY}(t_1, t_2; \tau) \, dt_1 \, dt_2 = 0 \tag{61}
\]

Then the ergodicity condition in the Correlation Function holds, i.e.

\[
\left\langle X(t; \beta) \cdot X(t + \tau; \beta) \right\rangle_T' = \left\langle E^\beta \left(X(t; \beta) \cdot X(t + \tau; \beta)\right) \right\rangle_T'
\]

**Proof:** Direct application of Theorem 2.47 for the s.p. \( Y(t, \tau; \beta) = X(t + \tau; \beta) \cdot X(t; \beta) \)

**Example 2.49:** Consider the stochastic process

\[
X(t; \beta) = A(\beta) \cos t + B(\beta) \sin t, \ t \in I
\]

where \( A(\beta) \) and \( B(\beta) \) are independent Gaussian random variables with means zero and variances \( \sigma^2 \). The correlation function of \( X(t; \beta) \) is

\[
C_{XX}(t_1, t_2) = \sigma^2 \cos(t_1 - t_2)
\]

It is easily shown that Eq. (57) is satisfied and thus the stochastic process is ergodic in the mean.

For a detailed discussion and more conditions for ergodicity criteria we refer to ASH, R.B., (Probability and Measure Theory) and ATHANASSOULIS, G.A., (Stochastic Modeling and Forecasting of Ship Systems. Lecture Notes NTUA).
2.2.4. Gaussian Stochastic Process

A discussion of stochastic processes is incomplete without giving considerations to Gaussian stochastic processes. Classification of stochastic processes based upon regularity and memory do not make specific references to the detailed probability distributions associated with a stochastic process. Gaussian stochastic processes are characterized by a distinct probabilistic law and they are extremely useful and most often encountered in practice.

**Definition 2.50**: A stochastic process \( X(t; \beta) \), \( t \in I \) is called a Gaussian stochastic process if for every finite set \( t_1, t_2, \ldots, t_n \in I \) the random vector defined by the projection

\[
\Pi_{(t_1, t_2, \ldots, t_n)}\{x(*)\} = \{x(t_1), \ldots, x(t_n)\} \quad , \quad x(*) \in D(I)
\]

is a Gaussian random vector. ■

Thus, the probability density function of the random vector \( (x(t_1), \ldots, x(t_n)) \), i.e. the \( n \)-joint probability density of the s.p. \( X(t; \beta) \) is given by

\[
f_{t_1, t_2, \ldots, t_n}(x_1, x_2, \ldots, x_n) = \left(2\pi\right)^{-n/2}|\tilde{C}|^{-1/2}\exp\left[-\frac{1}{2}(x-m)^T \tilde{C}^{-1}(x-m)\right]
\]

where

\[
m^T = [m_t \ m_t \ \ldots \ m_t] = [E^\beta(X(t; \beta)) \ E^\beta(X_b(\beta)) \ \ldots \ E^\beta(X_n(\beta))]
\]

(63)

and \( \tilde{C} = [\mu_{ij}] \) is the \( n \times n \) covariance matrix of \( X(t; \beta) \) with

\[
\mu_{ij} = E^\beta\left(\left[X(t_i; \beta) - m_{t_i}\right]\left[X(t_j; \beta) - m_{t_j}\right]\right)
\]

(64)

**Remark 2.51**: From the above equations it is clear that the probability structure of a Gaussian stochastic process is described by the moments of first and second order. ■

This is a typical characteristic of the Gaussian stochastic processes that makes them particularly useful for applications. Direct consequence of the above remark is that the higher order moments of every Gaussian characteristic function are expressed through the moments of first and second order. The theorem presented below, due to Isserlis 1918, give the necessary formula for the calculation of higher order moments.

**Theorem 2.52**: Let the Gaussian stochastic process \( X(t; \beta) \), \( t \in I \). Then

I. The central moments of odd order, \( j = 2k+1 \), are identically zero,

II. The central moments of even order, \( j = 2k \), are given by the formula,

\[
C_{X_{XX}X}^{\beta; \beta; \beta}(t_1, t_2, \ldots, t_n) = E^\beta\left((X(t_1; \beta) - m_{t_1})^h (X(t_2; \beta) - m_{t_2})^{l_2} \ldots (X(t_n; \beta) - m_{t_n})^{l_n}\right) = \\
= \sum R_{XX}(t_{k_1}, t_{k_2})R_{XX}(t_{k_3}, t_{k_4}) \ldots R_{XX}(t_{k_{2k-1}}, t_{k_{2k}})
\]
where the summation is over all the possible partitions of the $2k$ indices $\underbrace{1,1,\ldots,1}_{h}; \underbrace{2,2,\ldots,2}_{h}; \ldots; \underbrace{n,n,\ldots,n}_{h}$ at pairs $(i_1, i_2), (i_3, i_4), \ldots, (i_{2k-1}, i_{2k})$. The number of terms is

$$\frac{(2k)!}{(2^k k!)} = 1 \cdot 3 \cdot 5 \cdots (2k - 1).$$

For the Gaussian stochastic processes the following two important theorems hold,

**Theorem 2.53 :** Let the functions $\mu(\cdot), I \rightarrow \mathbb{C}$ and $R(\cdot,\cdot), I \times I \rightarrow \mathbb{C}$ with the last one, satisfies the conditions

1) $R(s,t) = \overline{R(t,s)}$, for any $t, s \in I$

2) For any $t_1, t_2, \ldots, t_n \in I$ the matrix $[R(t_m, t_n)]_{m,n}$ is non-negative definite.

Then, there always exists a Gaussian stochastic process $X(t; \beta), t \in I$, with mean value function $m(t) = \mu(t), t \in I$ and covariance function $R(\cdot,\cdot), I \times I \rightarrow \mathbb{C}$, i.e. it satisfies the relations

$$E^\beta(X(t; \beta)) = \mu(t), \quad t \in I \quad (65)$$

$$E^\beta(X(t; \beta)X(s; \beta)) = R(t,s), \quad t, s \in I \quad (66)$$

If the functions $\mu(\cdot)$ and $R(\cdot,\cdot)$ are real then there is always a real stochastic process $X(t; \beta), t \in I$ that satisfies the above relations. In this case equation (66) becomes

$$E^\beta(X(t; \beta)X(s; \beta)) = R(t,s), \quad t, s \in I \quad (67)$$

**Theorem 2.54 :** Let the stochastic process $X(t; \beta), t \in I$ with finite variance

$$E^\beta(X(t; \beta)^2) < \infty, \quad t \in I \quad (68)$$

Then there is always exists a Gaussian stochastic process $Y(t; \beta), t \in I$ such that

$$m_Y(t) = m_X(t), \quad t \in I \quad (69)$$

$$R_{YY}(t,s) = R_{XX}(t,s), \quad t, s \in I \quad (70)$$

If the the stochastic process $X(t; \beta), t \in I$ is real then there is always a real stochastic process $Y(t; \beta), t \in I$ that satisfies the above relations. 

The above theorems are formulated and proved at the monograph of DOOB, J.L., **(Stochastic Processes).**
2.3. References

Chapter 3

Probability Measures in Hilbert Spaces

As we have discussed at Section 1 of Chapter 2 the definition of the stochastic process as a measurable function which maps the space of elementary events $\Omega$ into the metric space $\mathcal{M}$ lead us to the concept of a probability measure over an infinite dimensional space. In this chapter we will consider the case of probability measures defined on a Hilbert space. Every result presented above can be generalized to the case of a Banach space by replacing the inner product with the duality pair and making suitable changes and restrictions.

In the first section we will give the definition of a stochastic process using probability measures. Moreover we will give some special examples of probability measures to get a clearer picture of the issue. Based on these examples we will introduce the finite dimensional relative measures, i.e. induced probability measures for finite dimensional subspaces of the Hilbert space. The next step will be to give the necessary and sufficient conditions for a family of probability measures, defined on every subspace of a Hilbert space, to define a probability measure.

Section 2 deals with the special case of cylinder functions and connects their integral over an infinite dimensional Hilbert space, using induced probability measures for finite dimensional subspaces of the Hilbert space.

In Section 3 we define the notion of the characteristic functional for probability measures over Hilbert spaces. Again, using finite dimensional relatives we make exact calculations of the characteristic functional. Finally, using nuclear operators we prove Minlos-Sazonov Theorem which is the generalization of Bochner’s Theorem for the infinite dimensional case and gives the necessary and sufficient conditions for a functional to be characteristic functional of some probability measure.

In Section 4 we generalize the concepts of mean value and covariance of a random variable for probability measures defined over infinite dimensional spaces. As before, we can calculate this quantities using induced probability measures over finite subspaces.

Section 5 and 6 present exact formulas for the calculation of classical moments and characteristic functions. The interesting part of this section is that we can calculate, using generalized functions the above quantities for the (time) derivatives of the stochastic process.

Section 7 deals with special case of Gaussian measures and their properties. We prove that for this particular case of measure an exact formula for the characteristic functional. We also calculate the mean value and covariance operators and study their properties.

In Section 8 a special case of characteristic functional is studied named the Tatarskii characteristic functional. These functionals have the important property to represent a wide class of stochastic processes.

Finally in section 9 we study the reduction of the Tatarskii characteristic functional to some very important characteristic functionals, such as the characteristic functionals due to Abel, Cauchy, and Feynman.
3.1. Probability Measures in Polish Spaces and their finite-dimensional relatives

3.1.1. Stochastic Processes as Random variables defined on Polish Spaces

In what follows we will consider the case of probability measures defined on a Polish space \( \mathcal{H} \), i.e. a complete metric space that has a countable dense subset. We also denote with \( \mathcal{H}^* \) the dual space of linear continuous functionals on \( \mathcal{H} \). We recall the definition of a stochastic process with values in \( \mathcal{H} \).

**Definition 1.1**: Let \( (\Omega, \mathcal{B}(\Omega), \mathcal{P}) \) be a probability space, \( X : \Omega \to \mathcal{H} \) be a point function from the sample space \( \Omega \) to the Polish space \( \mathcal{H} \), and \( \mathcal{B}(\mathcal{H}) \) be a Borel field over \( \mathcal{H} \). The mapping \( X \) is said to be a stochastic process, iff is \( (\Omega, \mathcal{B}(\Omega), \mathcal{B}(\mathcal{H})) \)-measurable.

Any stochastic process induces a probability measure on the space it takes values.

**Definition 1.2**: Let \( (\Omega, \mathcal{B}(\Omega), \mathcal{P}) \) be a probability space, \( (\mathcal{H}, \mathcal{B}(\mathcal{H})) \) a measurable space, and \( X : \Omega \to \mathcal{H} \) a stochastic process. On the Borel field \( \mathcal{B}(\mathcal{H}) \) of the measurable space \( (\mathcal{H}, \mathcal{B}(\mathcal{H})) \) we define the induced (by \( X : \Omega \to \mathcal{H} \)) probability measure on \( \mathcal{B}(\mathcal{H}) \) is defined by

\[
P_y (B) = \mathcal{P} \left( X^{-1}(B) \right), \quad \text{for every } B \in \mathcal{B}(\mathcal{H})
\]

(1)

The triplet \( (\mathcal{H}, \mathcal{B}(\mathcal{H}), \mathcal{P}_y) \) is called the induced probability space.

We shall now present some simple, specific examples of probability measures on Polish spaces. These examples are given by means of specific constructions permitting appropriate extensions of a finite-dimensional probability measure on the whole Polish space.

**Example 1.3**: [The 1-D induced measure]: Let \( \mathcal{P}^R \) be a probability measure in \( \mathbb{R} \), and \( \hat{e} \) be a unit vector in \( \mathcal{H} \). Let, also, \( \zeta: [\hat{e}] \to \mathbb{R} \) be a bijection between the span of \( \hat{e} \), (denoted by \( [\hat{e}] \)) and the real axis \( \mathbb{R} \), given by \( \zeta(a \hat{e}) = a \). Then, for any \( E \in \mathcal{B}(\mathcal{H}) \) we define

\[
P_{\hat{e}} (E) = \mathcal{P}^R \left( \zeta(E \cap [\hat{e}]) \right)
\]

(2)

\( P_{\hat{e}} \) is a probability measure in \( \mathcal{H} \) since,

a) \( P_{\hat{e}} (E) = \mathcal{P}^R \left( \zeta(E \cap [\hat{e}]) \right) \geq 0, \quad E \in \mathcal{B}(\mathcal{H}) \)

b) \( P_{\hat{e}} (\mathcal{H}) = \mathcal{P}^R \left( \zeta(\mathcal{H} \cap [\hat{e}]) \right) = \mathcal{P}^R \left( \zeta([\hat{e}]) \right) = \mathcal{P}^R (\mathbb{R}) = 1 \)

c) for any countable collection of mutually disjoint sets \( E_1, E_2, \ldots, \) in \( \mathcal{B}(\mathcal{H}) \)

\[
P_{\hat{e}} \left( \bigcup_{n} E_n \right) = \mathcal{P}^R \left( \zeta \left( \bigcup_{n} E_n \cap [\hat{e}] \right) \right) = \mathcal{P}^R \left( \bigcup_{n} \zeta \left( E_n \cap [\hat{e}] \right) \right) = \sum_{n} P_{\hat{e}} (E_n) \]
Example 1.4: **[The N-D induced measure]**: Let $\mathcal{D}_{\mathbb{R}^N}$ be a probability measure in $\mathbb{R}^N$, and $L_N$ be a linear subspace of $\mathcal{H}$, of dimension $N$. Let $\{e_1, e_2, ..., e_N\}$ be a basis of $L_N$, and $\zeta : L_N = [e_1, e_2, ..., e_N] \rightarrow \mathbb{R}^N$ be the bijection $\zeta(a_1e_1 + a_2e_2 + ... + a_ne_N) = (a_1, a_2, ..., a_N)$. Then, for any $E \in \mathcal{B}(\mathcal{H})$ we define

$$\mathcal{D}_{\mathbb{R}^N}(e_{e_1, e_2, ..., e_N})(E) = \mathcal{D}_{\mathbb{R}^N}\left(\zeta(E \cap L_N)\right).$$

(3)

The set function $\mathcal{D}_{\mathbb{R}^N}(e_{e_1, e_2, ..., e_N})$ is a legitimate probability measure in $\mathcal{H}$ obtained as an extension of the $N$-D probability measure $\mathcal{D}_{\mathbb{R}^N}$ over the whole Polish space. In what follows we will examine the existence of a probability measure which is not an induced measure from a finite dimensional measure, but it has its own infinite dimensional structure. Before we proceed we recall some basic results from probability theory.

3.1.2. Cylindric Sets

To begin our discussion we first need to give some essential definitions, concerning a special class of sets. Thus we have,

**Definition 1.5**: A cylindric set on $\mathcal{H}$ is defined to be a set $Q_{\{l\}_{i=1}^n}(B) \subset \mathcal{H}$ of the form

$$Q_{\{l\}_{i=1}^n}(B) = Q_l(B) = \{x \in \mathcal{H} : \langle l_1, x \rangle, \langle l_2, x \rangle, ..., \langle l_n, x \rangle \in B \subset \mathbb{R}^n \}, \quad B \in \mathcal{B}(\mathbb{R}^n)$$

for every $n = 1, 2, ...$ and $l = (l_1, l_2, ..., l_n)$ with $l_i \in \mathcal{H}^*$, $i = 1, ..., n$.

The class of all cylindric sets for given $l_i \in \mathcal{H}^*$, $i = 1, ..., n$ is obviously a $\sigma$–algebra which we will write as $\mathcal{B}_{\{l\}_{i=1}^n}$. The union of all $\sigma$–algebras is an algebra. Indeed, if $B_1$ and $B_2$ belong to $\mathcal{B}_1$ and $\mathcal{B}_{l_2}$, then choosing $\tilde{l} = (l_1, l_2)$ we get,

a) $B_i \in \mathcal{B}_{l_i}, \quad i = 1, 2$
We denote this algebra of sets by $\mathcal{B}_0$. It is called the algebra of cylinder sets. Sets from $\mathcal{B}_0$ are also considered as elementary for the construction of the measure on $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$ as we shall see at the sequel.

### 3.1.3. Restriction of a probability measure over a Polish space, to finite-dimensional subspaces

We shall now study the restriction of a probability measure over a Polish space (assuming it exists), to finite-dimensional subspaces. Let $(\Omega, \mathcal{B}(\Omega), \mathbb{P})$ be a probability space, $X : \Omega \to \mathcal{H}$ a stochastic process with values in a separable, real, Polish space $\mathcal{H}$, and $\mathcal{B}(\mathcal{H})$ a Borel field over $\mathcal{H}$.

**Definition 1.6**: Let $(\mathcal{H}, \mathcal{B}(\mathcal{H}), \mathbb{P}_{\mathcal{H}})$ be the induced probability space from the s.p. $X : \Omega \to \mathcal{H}$. Let $\Pi_L$ be a projection of $\mathcal{H}$ into the finite-dimensional subspace $L$. The probability measure defined on $(L, \mathcal{B}(L))$ as

$$\mathbb{P}_{\Pi_L}(B) = \mathbb{P}_{\mathcal{H}}(\Pi_L^{-1}[B]) = \mathbb{P}_{\mathcal{H}}(\{x \in \mathcal{H} : \Pi_L[x] \in B\})$$

is called the projection of the measure $\mathbb{P}_{\mathcal{H}}$ on the subspace $L$.

**Remark 1.7**: The fact that $\mathbb{P}_{\Pi_L}(B)$ is a measure follows from the fact that for a sequence of non-overlapping sets $B_n \in \mathcal{B}(L)$ the sets

$$\Pi_L^{-1}\left[\bigcup_n B_n\right] = \bigcup_n \Pi_L^{-1}[B_n]$$

are also non-overlapping.

**Remark 1.8**: Using a bijective mapping $\zeta : L \to \mathbb{R}^n$ where $n = \dim[L]$ we can define the above probability measure of the finite-dimensional subspace $L$ on the borel field of $\mathbb{R}^n$, as follows,

$$\mathbb{P}_{\mathbb{R}^n, \Pi_L}(B) = \mathbb{P}_{\Pi_L}(\zeta^{-1}(B))$$

Hence, using the projection $\Pi_L$ we define a r.v. into the probability space $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ since the mapping defined is $(\mathcal{B}(\mathcal{H}), \mathcal{B}(\mathbb{R}^n))$-measurable. We denote this r.v. as $\Xi_{\Pi_L}(\omega)$.
The measures $\mathcal{P}_{\Pi_n}$, being projections of the same measure, are compatible in a certain sense, for different $n$. This compatibility condition is the generalization of the consistency condition of Kolmogorov Theorem (Chapter 2/ Theorem 1.7).

**Compatibility Condition 1.9**: Let the probability measure $\mathcal{P}_\mu$ defined on the measurable space $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$ and the family of measures $\{\mathcal{P}_{\Pi_L}\}$ defined for all finite-dimensional subspaces $L$ of $\mathcal{H}$. Then for all $L_1 \subset L_2$ and $B \in \mathcal{B}(L_1)$ we have,

$$\mathcal{P}_{\Pi_{L_1}}(B) = \mathcal{P}_{\Pi_{L_2}}\left(\Pi_{L_1}^{-1}[B] \cap L_2\right)$$  \hspace{1cm} (5)

**Proof**: Let $L_1 \subset L_2$ and $B \in \mathcal{B}(L_1)$. Then the set $\Pi_{L_1}^{-1}[B]$ can also be written as $\Pi_{L_2}^{-1}[B]$, where $B_2 \in \mathcal{B}(L_2)$ is defined by the equality

$$B_2 = \{x \in L_2 : \Pi_{L_1}[x] \in B\}$$

Since $\Pi_{L_1}^{-1}[B] = \Pi_{L_2}^{-1}[B_2]$, we have

$$\mathcal{P}_{\Pi_{L_1}}(B) = \mathcal{P}_\mu\left(\Pi_{L_1}^{-1}[B]\right) = \mathcal{P}_\mu\left(\Pi_{L_2}^{-1}[B_2]\right) = \mathcal{P}_{\Pi_{L_2}}(B_2)$$

Since $B_2 = \{x \in L_2 : \Pi_{L_1}[x] \in B\}$, we have prove the condition (5).  \hspace{1cm} ■

### 3.1.4. Conditions for the existence of probability measures over a Polish space

We will now study the inverse problem, that is to investigate the conditions under those, a family of probability measures $\mathcal{P}_{\Pi_{L_n}}$, for a sequence $L_n$ of linear subspaces of $\mathcal{H}$, for which $L_n \subset L_{n+1}$ and $\bigcup_n L_n$ is dense in $\mathcal{H}$, can define a probability measure on $\mathcal{H}$. This family of probability measures is redundancy, so we cannot associate it with a unique probability measure of $\mathcal{H}$, unless it satisfies the Compatibility Condition 1.9. Using the
arguments of Example 1.4, one can determine $P_{\Pi L}$ on $\mathcal{B}(L)$, by knowing $P_{\Pi L}$. Hence knowing $P_{\Pi L}$, for a sequence $L_n$ of linear subspaces of $\mathcal{H}$, we can define $P_{\mathcal{H}}$ on $\bigcup_n \mathcal{B}(L_n)$ and since the $\sigma$-closure of this algebra coincides with $\mathcal{B}(\mathcal{H})$ we can define $P_{\mathcal{H}}$ by the same token. Unfortunately, there are extra requirements for a family of probability measures, satisfying Compatibility Condition 1.9, to correspond to a probability measure on $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$. The next theorem gives a necessary and sufficient condition for a weak distribution to correspond to some measure.

**Theorem 1.10**: Let $S_{\rho}$ be a sphere of radius $\rho : S_{\rho} = \{x : |x| \leq \rho\}$. A family of probability measures $\{P_{\Pi L}\}$, satisfying Compatibility Condition 1.9, will be generated by some measure $P_{\mathcal{H}}$ onto the measurable space $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$ iff for every $\varepsilon > 0$ there exists an $\eta > 0$ such that for all $L$

$$P_{\Pi L}(S_{\rho} \cap L) \geq 1 - \varepsilon \quad \text{when} \quad \rho > \eta \quad (6)$$

**Proof**: Necessity. If $\{P_{\Pi L}\}$ is generated by the measure $P_{\mathcal{H}}$, then choosing $\eta > 0$ such that $P_{\mathcal{H}}(S_{\eta}) > 1 - \varepsilon$ (this is possible since $\lim_{\eta \to +\infty} P_{\mathcal{H}}(S_{\eta}) = P_{\mathcal{H}}(\mathcal{H}) = 1$), we obtain

$$P_{\Pi L}(S_{\rho} \cap L) = P_{\mathcal{H}}(\Pi_L^{-1}(S_{\rho} \cap L)) \geq P_{\mathcal{H}}(S_{\rho}) \geq P_{\mathcal{H}}(S_{\eta}) > 1 - \varepsilon.$$

The proof of sufficiency is more difficult. We define on the algebra $\mathcal{B}_0 = \bigcup_n \mathcal{B}(L_n)$ a finitely additive function $P_{\mathcal{H}}$ by means of

$$P_{\mathcal{H}}(A) = P_{\Pi L}(A), \quad A \subset \mathcal{B}_0.$$

To convince ourselves that $P_{\mathcal{H}}$ can be extended to a measure defined on $(\mathcal{H}, \mathcal{B}_{\mathcal{H}})$, it is sufficient to show that $P_{\mathcal{H}}$ is continuous on $\mathcal{B}_0$, i.e., that for an arbitrary sequence of sets $A_n \subset \mathcal{B}_0$ for which $A_n \supset A_{n+1}$ and $\bigcap_n A_n = \emptyset$,

$$\lim_{n \to \infty} P_{\mathcal{H}}(A_n) = 0 \quad (7)$$

Let $A_n$ be a cylinder set with base in $L_n$ and $L_n \subset L_{n+1}$. Let $B_n \subset L_n$ be the base of $A_n$. We remark that it is sufficient to prove (7) merely for sets with closed bases. Indeed, choosing closed sets $C_n \subset B_n$ such that $P_{\Pi L}(B_n - C_n) = P_{\mathcal{H}}(B_n - C_n) < \varepsilon_n$ and then taking

$$D_n = \bigcap_{m=1}^{n} \{x : \Pi_{L_m} x \in C_m\} \bigcap L_n,$$

we obtain closed sets for which

$$P_{L_n}(B_n - C_n) \leq \sum_{m=1}^{n} P_{L_m}(B_m - C_m) \leq \sum_{m=1}^{n} \varepsilon_m.$$
Hence, if $A'_n = \Pi_n^{-1}(D_n)$, then

$$A_{n+1}' \subset A'_n, \bigcap_n A'_n = \bigcap_n A_n \text{ and } \mathcal{P}^\circ\left(A'_n\right) \leq \mathcal{P}^\circ\left(A'_n\right) + \sum_{m=1}^{n} E_m$$

If for cylinder sets with closed bases $\lim_{n \to \infty} \mathcal{P}^\circ\left(A'_n\right) = 0$, then since $\sum_{n=1}^{\infty} E_n$ can be taken as suitably small, (7) is also fulfilled. Hence, it will be assumed that the $B_n$ are closed sets. Then the $A_n$ will be weakly closed sets (if $x_k \xrightarrow{w} x$ and $x_k \in A_n$, then $x \in A_n$). For all $\rho$ the sets $S_{\rho}$ is also weakly closed and weakly compact. Since,

$$\bigcup_{n=1}^{\infty} S_{\rho} \cap A_n = \emptyset$$

and since the sets $S_{\rho} \cap A_n$ are weakly closed and weakly compact and $S_{\rho} \cap A_n \supset S_{\rho} \cap A_{n+1}$, for some $n$, $S_{\rho} \cap A_n = \emptyset$. This means that,

$$\mathcal{P}^\circ\left(A_n\right) = \mathcal{P}^\circ_{L_n}\left(A_n\right) \leq \mathcal{P}^\circ_{L_n}\left(L_n\right) - \mathcal{P}^\circ_{L_n}\left(L_n \cap S_{\rho}\right) \leq \varepsilon$$

provided that $\rho > \eta$. From the arbitrariness of $\varepsilon > 0$ there follows (7). ■

To understand more precisely the role of condition (6) let us note that, for an arbitrary family of measures, is equivalent with the property of “uniformly tight” described by the following

**Definition 1.11**: A family $\{\mathcal{P}\}$ of probability measures on $\mathcal{H}$ is uniformly tight if, for every $\varepsilon > 0$ there exists a compact subset $K$ of $\mathcal{H}$ such that $\mathcal{P}(K') < \varepsilon$ for all $\mathcal{P} \in \{\mathcal{P}\}$. ■

Concerning the property of uniformly tightness the following important theorem holds,

**Theorem 1.12 [Prohorov]**: A family $\{\mathcal{P}\}$ of probability measures on a Polish space $\mathcal{H}$ is relatively sequentially compact if and only if is uniformly tight. ■

Hence condition (6) assures that for every sequence $\mathcal{P}_n$ ($n=1,2,\ldots$) of members of $\{\mathcal{P}\}$, there exists a convergent subsequence. Now consider a family of probability measures $\{\mathcal{P}_{\Pi_{L_n}}\}$, for a sequence $L_n$ of linear subspaces of $\mathcal{H}$, for which $L_n \subset L_{n+1}$ and $\bigcup_{n} L_n$ is dense in $\mathcal{H}$. Then, to assure the uniqueness of the limit of every convergent subsequence we need Compatibility Condition 1.9. Indeed, if compatibility condition 1.9 holds and for two subsequences there are two different limits, it can be easily seen that we are leading to an absurd.

Thus, the property of “uniformly tight” for a family of probability measures, assures the existence of a convergent subsequence, while, the compatibility condition assures the uniqueness of the limit for every convergent subsequence. Finally the existence of a probability measure is assured by the following
**Preposition 1.13**: Let \( \mathcal{P}_n \ (n=1,2,\ldots) \) be a relatively sequentially compact sequence of probability measures on a Polish space such that every convergent subsequence has the same limiting probability measure \( \mathcal{P} \). Then \( \mathcal{P}_n \to \mathcal{P} \) as \( n \to \infty \). ■

**Example 1.14**: Let \( \mathcal{H} \) be the \( C([a,b]) \) space of continuous functions on the interval \([a,b]\) and \( \mathcal{B}(C) \) the \( \sigma \)-field generated by the open subsets of \( C([a,b]) \). Let \( \Pi_{(t_1,t_2,\ldots,t_n)} : C([a,b]) \to \mathbb{R}^n \) be the projection defined by

\[
\Pi_{(t_1,t_2,\ldots,t_n)}[x(\bullet)] = \{x(t_1),\ldots,x(t_n)\} \quad , \quad x(\bullet) \in C([a,b])
\]

For the moment we assume that \( \mathcal{P} \) exists and it’s a probability measure over \( C([a,b]) \). We thus can define the sequence of finite probability measures

\[
\mathcal{P}_{(t_1,t_2,\ldots,t_n)}(E) = \mathcal{P}\left(E \cap \Pi_{(t_1,t_2,\ldots,t_n)}[C([a,b])]\right) =
\int_{A_{t_1}} \cdots \int_{A_{t_n}} f_{t_1,t_2,\ldots,t_n}(x_1,x_2,\ldots,x_n) \, dx_1 \cdots dx_n \quad , \quad E \in \mathcal{B}(C)
\]

Now, \( \mathcal{P} \) may be thought of as the limit of the function

\[
\int_{A_{t_1}} \cdots \int_{A_{t_n}} f_{t_1,t_2,\ldots,t_n}(x_1,x_2,\ldots,x_n) \, dx_1 \cdots dx_n ,
\]

which has discrete arguments. The above limit exists and it is a probability measure if and only if the family of distributions \( f_{t_1,t_2,\ldots,t_n}(x_1,x_2,\ldots,x_n) \) fulfils the conditions of Kolmogorov theorem (Chapter 2/ Theorem 1.7), which are special case of the conditions presented above, for general measures. ■

**Remark 1.15**: We note that for the definition of the probability measure \( \mathcal{P}_y \) on the measurable space \( (\mathcal{H}, \mathcal{B}(\mathcal{H})) \) it is sufficient to know the family of measures \( \{\mathcal{P}_L\} \) defined for all one-dimensional subspaces \( L \) of \( \mathcal{H} \) spanned by a vector \( z \in \mathcal{H} \). Thus, the measure \( \mathcal{P}_y \) is defined by its one-dimensional distributions (i.e. by its projections onto one-dimensional subspaces). The proof of the above basically relies to the characteristic functional approach that will be studied in the next section. Using this remark we could define characteristic functionals by knowing their one-dimensional distributions. However this cannot be done in general, since the resulted measure must satisfy the definition of a probability measure, a condition which is not, in general, fulfilled.
3.2. Cylinder functions and their Integral

It is interesting to note that some classes of functions can also be integrated with respect to some projection of the probability measure, defined on the Hilbert space. To these functions, for example, are related the cylinder functions defined below.

Definition 2.1 : A function $G : \mathcal{H} \to V$, where $V^1$ is a linear space, is called a cylinder function if there is a finite dimensional subspace $L \subset \mathcal{H}$ and a finite dimensional $B(L)$-measurable function $g_L : L \to V$, such that

$$G(x) = g_L(\Pi_L [x]) , \quad x \in \mathcal{H}$$

where $\Pi_L$ is the projection of $\mathcal{H}$ on $L$.

In a Hilbert space, a finite dimensional projection $\Pi_L[\ast]$ can be represented as $\{\langle l_1, y \rangle, \ldots, \langle l_n, y \rangle \}$ where $l_i$ are $n$ linearly independent elements of the Hilbert space. With this form of projection it’s possible to integrate cylinder functions of the form $g(\langle l_1, y \rangle, \ldots, \langle l_n, y \rangle)$ which will play an important role for the sequel.

Let $(\Omega, B(\Omega), \mathcal{P})$ be a probability space, $X : \Omega \to \mathcal{H}$ a stochastic process from the space of elementary events $\Omega$ to the separable, real, Hilbert space $\mathcal{H}$ and $B(\mathcal{H})$ a Borel field over $\mathcal{H}$. Let also, $\mathcal{P}_\mathcal{H}$ be the induced probability measure. For every $n$, and for every vector element $I = \{l_1, l_2, \ldots, l_n\} \in \mathcal{H}^n$ we consider the projection $\Pi_I : \mathcal{H} \to L$,

$$\Pi_I [x] = \langle l_1, x \rangle \frac{l_1}{\|l_1\|} + \langle l_2, x \rangle \frac{l_2}{\|l_2\|} + \ldots + \langle l_n, x \rangle \frac{l_n}{\|l_n\|} , \quad x \in \mathcal{H}$$

and the injection $\zeta : L \to \mathbb{R}^n$ defined by,

$$\zeta \left( \sum_{i=1}^{n} a_i l_i \right) = \{a_1, a_2, \ldots, a_n\}$$

Based on the above discussion we have the restriction of the probability measure $\mathcal{P}_\mathcal{H}$ onto $\mathbb{R}^n$,

$$\mathcal{P}_I (E) = \mathcal{P}_\mathcal{H} \left( x \in \mathcal{H} : \{\langle l_1, y \rangle, \ldots, \langle l_n, y \rangle \} \in E \right) , \quad E \in B(\mathbb{R}^n)$$

or

$$F_I (a) = F_I (a_1, \ldots, a_n) = \mathcal{P}_\mathcal{H} \left( x \in \mathcal{H} : \langle l_i, x \rangle \leq a_i , i = 1, \ldots, n \right) , \quad a \in \mathbb{R}^n$$

We then have the following

Definition 2.2 : Let $g : \mathbb{R}^n \to \mathbb{C}$ be an arbitrary integrable, measurable function, defined in $\mathbb{R}^n$, with values in $\mathbb{C}$. For every $n$, and for every vector $I = \{l_1, l_2, \ldots, l_n\} \in \mathcal{H}^n$ with $l_i \neq 0$ for

$^1$ In our applications $V$ is usually $\mathbb{R}$ or $\mathbb{C}$. 


every $i = 1, \ldots, n$, we define the infinite dimensional integral of the function $g: \mathbb{R}^n \to \mathbb{C}$ by the following relation

$$\int_{\mathcal{H}} g(\langle l_1, x \rangle, \ldots, \langle l_n, x \rangle) d\mathcal{P}(dx) = \int_{\mathbb{R}^n} g(t_1, \ldots, t_n) d\mathcal{P}_1(dt_1, \ldots, dt_n)$$

where, $d\mathcal{P}_1$ is defined by eq. (4).

**Remark 2.3**: The above definition physically connects the integral over an infinite dimensional space with a finite dimensional subspace. To understand more precisely the above connection suppose that in some way we could define the infinite dimensional integral in the l.h.s. of equation (6). Then we would have $\Pi_l[x] = \{\langle l_1, x \rangle, \ldots, \langle l_n, x \rangle\}$. Making the set valued transformation

$$x = \Pi_l^{-1}[t]: \mathbb{R}^n \to \mathcal{H}$$

we have,

$$\int_{\mathcal{H}} g(P_l[x]) d\mathcal{P}(dx) = \int_{\mathbb{R}^n} g(t_1, t_2, \ldots, t_n) d\mathcal{P}_1(P_l^{-1}[dt_1, dt_2, \ldots, dt_n])$$

Using the definition of eq. (4) we see that

$$d\mathcal{P}(\Pi_l^{-1}[E]) = d\mathcal{P}_1\left(x \in \mathcal{H}: \{\langle l_1, x \rangle, \ldots, \langle l_n, x \rangle\} \in E\right) = d\mathcal{P}_1(E) \quad E \in \mathcal{B}(\mathbb{R}^n).$$

Thus, the reduction of the infinite dimensional integral is due to the fact that the integrable function $g: \mathbb{R}^n \to \mathbb{C}$ depends from a finite dimensional subspace of $\mathcal{H}$. ■

**Remark 2.4**: As was illustrated above, one of the most important applications of the projection of a measure is the calculation of integrals of cylinder functions. These finite-dimensional measures are adapted to the form of the integrable cylinder function, using a suitable projection. The above current of ‘information’ can be summarized to the diagram below,

\[ \begin{array}{c}
(\Omega, \mathcal{B}(\Omega), \mathcal{P}) \\
\uparrow \quad \uparrow \quad \uparrow \\
(\mathcal{H}, \mathcal{B}(\mathcal{H}), \mathcal{P}_\mathcal{H}) \quad \mathcal{P}(\cdot) \quad \mathcal{P}_\mathcal{H}(\cdot) \\
\downarrow \quad \downarrow \\
(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mathcal{P}_1) \\
\end{array} \]

**Remark 2.5**: For $n = 1$ the above definition reduce to the equation

$$\int_{\mathcal{H}} g(\langle l, x \rangle) d\mathcal{P}(dx) = \int_{\mathbb{R}} g(t) d\mathcal{P}_1(dt) \quad l \in \mathcal{H}$$

where

$$d\mathcal{P}_1(a) = d\mathcal{P}\left\{x \in \mathcal{H}: \langle l, x \rangle \leq a\right\} \quad \text{for every} \quad l \in \mathcal{H}, \quad a \in \mathbb{R} \quad (8)$$

The above equation will be used extensively in what follows. ■
3.3. Characteristic Functional of the Probability Measure

In this section we will generalize the concept of the characteristic function for probability measures on real, separable, Hilbert spaces. First we recall the definition of a positive definite functional.

**Definition 3.1** : A function $\phi$ from $\mathcal{H}$ into $\mathbb{C}$ (or $\mathbb{R}$) is called a **functional of positive type** if for any $x_1, x_2, ..., x_n \in \mathcal{H}$, $n = 1, 2, ...$ and any numbers $c_1, c_2, ..., c_n \in \mathbb{C}$ we have:

$$\sum_{j, k=1}^{n} \phi(x_j - x_k) \cdot c_j \cdot c_k \geq 0$$  \hspace{1cm} (1)

We are, now, going to define the characteristic functional for a probability measure. Let $\mathcal{H}$ be a real, separable Hilbert space and $\mathcal{B}(\mathcal{H})$ the $\sigma$-field generated by the open subsets of $\mathcal{H}$. Thus, we have the following

**Definition 3.2** : The characteristic functional $\mathcal{F}: \mathcal{H} \rightarrow \mathbb{C}$ of a probability measure $P_x$ on $\mathcal{H}$, is defined by

$$\mathcal{F}(x) = E \left( e^{i\langle x, \cdot \rangle} \right) = \int_{\mathcal{H}} e^{i\langle x, y \rangle} P_{\cdot y} \, dy \quad x \in \mathcal{H}$$  \hspace{1cm} (2)

**Remark 3.3** : Since $\mathcal{P}(\mathcal{H}) < \infty$, $\mathcal{F}$ always exists. Moreover,

$$|\mathcal{F}(x)| \leq \mathcal{P}(\mathcal{H}) = 1 \quad \forall x \in \mathcal{H}$$  \hspace{1cm} (3)

**Remark 3.4** : Using the distribution of the measurable function $\langle x, \cdot \rangle$ and the remark 2.5. we have that

$$\mathcal{F}(x) = \int_{\mathcal{H}} e^{i\langle x, y \rangle} P_{\cdot y} \, dy = \int_{-\infty}^{\infty} e^{it} P_{\cdot t} \, dt \quad x \in \mathcal{H}$$  \hspace{1cm} (4)

Hence, this is how the probability measure can be defined, by using the family of measures $\left\{P_L \right\}$ defined for all one-dimensional subspaces $L$ of $\mathcal{H}$ spanned by a vector $x \in \mathcal{H}$. (See Remark 1.12).

Concerning the smoothness of $\mathcal{F}$ we have the following theorem.

**Theorem 3.5** : Let $P^x$ be a probability measure in $\mathcal{H}$ and let $\mathcal{F}$ be the characteristic functional of $P^x$. Then $\mathcal{F}$ is uniformly continuous.

**Proof** : Let $\varepsilon > 0$ be given. Then there exists $r > 0$ such that $\mathcal{P}(S_r) > 1 - \frac{\varepsilon}{4}$, where $S_r = \left\{ x : |x| \leq r \right\}$. Then

$$\left| \mathcal{F}(x) - \mathcal{F}(y) \right| \leq \int_{\mathcal{H}} \left| e^{i\langle x, z \rangle} - e^{i\langle y, z \rangle} \right| P^x (dz) =$$
\[
\frac{1}{r_1} e^{i(x, y)} - e^{i(x, y)} dz + \int_{S_r} e^{i(x, y)} |P^+(dz) | \leq \\
\leq \int_{S_r} |(x, z) - (y, z) | |P^+(dz) | + 2 \cdot \frac{\varepsilon}{4} \leq \\
\leq r |x - y| + \frac{\varepsilon}{2}.
\]

Choose \( \delta > 0 \) such that \( \delta < \frac{\varepsilon}{2r} \). Then, \( |x - y| < \delta \) implies \( |F(x) - F(y)| < \varepsilon \). Hence, is uniformly continuous. ■

If \( \mathcal{H} \) is an infinite dimensional Hilbert space we can show something more for \( F \). For the sake of convenience we introduce the following

**Definition 3.6**: An operator is called a nuclear operator of \( \mathcal{H} \) if

1. it has finite trace, i.e., for some orthonormal basis \( \{e_i\} \), \( \sum_{i=1}^{\infty} \langle S_{e_i}, e_i \rangle < \infty \),
2. it is self-adjoint ,
3. it is positive-definite.

\( \mathcal{S} \) denotes the nuclear operators of \( \mathcal{H} \). ■

This theorem gives necessary and sufficient conditions for a functional \( F \) be the characteristic functional of some measure on \( (\mathcal{H}, \mathcal{B}(\mathcal{H})) \). It is thus a generalization of the Bochner’s theorem on positive definite functions in Hilbert space.

**Theorem. 3.7 [Minlos-Sazonov]**: A functional \( F \) in \( \mathcal{H} \) is the characteristic functional of some probability measure in \( \mathcal{B}(\mathcal{H}) \) if and only if

a) \( F(0) = 1 \) and \( F \) is a functional of positive type,

b) For every \( \varepsilon > 0 \), exists \( C_\varepsilon \in \mathcal{S} \) such that

\[
1 - \text{Re} F(x) \leq \langle C_\varepsilon x, x \rangle + \varepsilon \quad \text{for all} \quad x \in \mathcal{H}.
\]

**Remark. 3.8**: b) implies, in particular, that when \( \langle C_\varepsilon x, x \rangle \) is small, then \( 1 - \text{Re} F(x) \) is small. Note that \( \langle C_\varepsilon x, x \rangle \) can be small even if \( |x| \) is big. Note also that \( \{\langle C_\varepsilon x, x \rangle < a\} \) is an ellipsoid with semi-axes \( (a/\lambda_1)^{\frac{1}{2}}, (a/\lambda_2)^{\frac{1}{2}}, \ldots \), where \( \lambda_n \)'s are the eigenvalues of \( C_\varepsilon \).

**Proof**: Necessity. Let \( F(x) = \int_{\mathcal{H}} e^{i(x, y)} |P^+(dy) | \) and \( \varepsilon > 0 \) be given. Choose \( 0 < r < \infty \) such that \( |P^+(S_r)| > 1 - \frac{\varepsilon}{2} \) where \( S_r \) is the ball \( \{|x| \leq r\} \). Then

\[
F(x) = \int_{S_r} e^{i(x, y)} |P^+(dy) | + \int_{S_r^c} e^{i(x, y)} |P^+(dy) |
\]
Note that
\[ \left| \int_{S'} e^{i \langle x, y \rangle} d\mathcal{P} (dy) \right| \leq \mathcal{P} (S') < \frac{\varepsilon}{2} \]
Hence it is sufficient to show that
\[ 1 - \text{Re} \int_{S} e^{i \langle x, y \rangle} d\mathcal{P} (dy) \leq \langle C_e x, x \rangle + \frac{\varepsilon}{2} \]
for some operator $C_e \in \mathcal{S}$. But
\[ 1 - \text{Re} \int_{S} e^{i \langle x, y \rangle} d\mathcal{P} (dy) = \int_{S} [1 - \cos \langle x, y \rangle] d\mathcal{P} (dy) + \mathcal{P} (S') \leq \int_{S} [1 - \cos \langle x, y \rangle] d\mathcal{P} (dy) + \frac{\varepsilon}{2} \]
Recall that $1 - \cos \theta \leq \frac{1}{2} \theta^2$ for all real $\theta$. Hence
\[ \int_{S} [1 - \cos \langle x, y \rangle] d\mathcal{P} (dy) \leq \frac{1}{2} \int_{S} \langle x, y \rangle^2 d\mathcal{P} (dy) \]
The same argument in the proof of Theorem 4.10 below shows that there exists $C_e \in \mathcal{S}$ such that
\[ \langle C_e x, x \rangle = \frac{1}{2} \int_{S} \langle x, z \rangle \langle y, z \rangle d\mathcal{P} (dz) \]
The desired conclusions follow immediately.

**Remark**: Observe that if the covariance operator $C_{\rho'}$ of $d\mathcal{P}$ exists, then we have
\[ 1 - \text{Re} \mathcal{F} (x) \leq \frac{1}{2} \langle C_{\rho'} x, x \rangle \quad \text{for all } x \in \mathcal{H} \]
But $C_{\rho'}$, in general, is not an $\mathcal{S}$-operator. Of course, if $C_{\rho'} \in \mathcal{S}$, then b) is trivially satisfied by taking $C_e = \frac{1}{2} C_{\rho'}$ for all $\varepsilon > 0$.

**Remark**: Observe also that $C_e$ is the covariance operator of the following Borel measure $\overline{\mathcal{P}}$ in $\mathcal{H}$,
\[ \overline{\mathcal{P}} (E) = \frac{1}{2} \mathcal{P} (E \cap S_e), \quad E \in \mathcal{B} (\mathcal{H}) \]
Clearly $\overline{\mathcal{P}}$ satisfies the hypothesis of Theorem 4.10 and thus its covariance operator $C_e$ is an $\mathcal{S}$-operator. This is another proof for $C_e \in \mathcal{S}$.

**Sufficiency.** (We will need to use Bochner’s Theorem for the finite dimensional space $\mathbb{R}^n$)

**Step 1.** We first derived some properties implied by a).
a1) \( |\mathcal{F}(x)| \leq 1 \), and \( \mathcal{F}(x) = \overline{\mathcal{F}(-x)} \) for all \( x \in \mathcal{H} \),
a2) \( |\mathcal{F}(x) - \mathcal{F}(y)| \leq 2 \sqrt{|1 - \mathcal{F}(x - y)|} \) for all \( x, y \in \mathcal{H} \),
a3) \( |1 - \mathcal{F}(x)| \leq \sqrt{2} \sqrt{|1 - \Re \mathcal{F}(x)|} \) for all \( x \in \mathcal{H} \).

a1) : Take \( n = 2 \), \( x_1 = 0 \) and \( x_2 = x \). Part a) implies that the following matrix is positive definite in the sense of linear algebra,

\[
\begin{pmatrix}
1 & \mathcal{F}(x) \\
\mathcal{F}(-x) & 1
\end{pmatrix}
\]

Hence \( \mathcal{F}(x) = \mathcal{F}(-x) \). Moreover the matrix determinant is non-negative, i.e. \( 1 - \mathcal{F}(x) \mathcal{F}(x) \geq 0 \). Hence \( |\mathcal{F}(x)| \leq 1 \).

a2) : Take \( n = 3 \), \( x_1 = 0 \), \( x_2 = x \) and \( x_3 = y \). Part a) implies that the following matrix

\[
\begin{pmatrix}
1 & \mathcal{F}(x) & \mathcal{F}(y) \\
\mathcal{F}(-x) & 1 & \mathcal{F}(y - x) \\
\mathcal{F}(-y) & \mathcal{F}(x - y) & 1
\end{pmatrix} = \begin{pmatrix}
1 & \mathcal{F}(x) & \mathcal{F}(y) \\
\mathcal{F}(y) & 1 & \mathcal{F}(x - y) \\
\mathcal{F}(y) & \mathcal{F}(x - y) & 1
\end{pmatrix}
\]

is positive definite. Hence the matrix determinant \( D \geq 0 \). But

\[
D = 1 + \mathcal{F}(x) \overline{\mathcal{F}(y)} \mathcal{F}(x - y) + \mathcal{F}(x) \overline{\mathcal{F}(y)} \mathcal{F}(x - y) - \mathcal{F}(y)^2 -
- \mathcal{F}(x)^2 - \overline{\mathcal{F}(x - y)^2} =
= 1 + 2 \Re \left[ \mathcal{F}(x) \overline{\mathcal{F}(y)} \mathcal{F}(x - y) \right] - \left| \mathcal{F}(x) \right|^2 + \left| \mathcal{F}(y) \right|^2 - \left| \mathcal{F}(x - y) \right|^2 =
= 1 + 2 \Re \left[ \mathcal{F}(x) \overline{\mathcal{F}(y)} \mathcal{F}(x - y) \right] - \left| \mathcal{F}(x) - \mathcal{F}(y) \right|^2 - \left| \mathcal{F}(x - y) \right|^2 -
- 2 \Re \left[ \mathcal{F}(x) \overline{\mathcal{F}(y)} \right] =
= 1 + 2 \Re \left[ \mathcal{F}(x) \overline{\mathcal{F}(y)} \right] \left( \mathcal{F}(x - y) - 1 \right) - \left| \mathcal{F}(x) - \mathcal{F}(y) \right|^2 - \left| \mathcal{F}(x - y) \right|^2
\]

Note that

\[
1 - \left| \mathcal{F}(x - y) \right|^2 = (1 - \left| \mathcal{F}(x - y) \right|) \left( 1 + \left| \mathcal{F}(x - y) \right| \right) \leq 2 \left| 1 - \mathcal{F}(x - y) \right|
\]
and

\[
2 \Re \left[ \mathcal{F}(x) \overline{\mathcal{F}(y)} \left( \mathcal{F}(x - y) - 1 \right) \right] \leq 2 \left| \mathcal{F}(x) \right| \left| \mathcal{F}(y) \right| \left| \mathcal{F}(x - y) - 1 \right| \leq
\leq 2 \left| 1 - \mathcal{F}(x - y) \right|
\]

i.e. \( \left| \mathcal{F}(x) - \mathcal{F}(y) \right| \leq 2 \sqrt{|1 - \mathcal{F}(x - y)|} \)

a3) : Note that if \( |z| \leq 1 \) then
From a1) we know that $|\mathcal{F}(x)| \leq 1$, therefore,

$$
|1 - \mathcal{F}(x)|^2 \leq 2(1 - \text{Re}\mathcal{F}(x))
$$

Hence,

$$
|1 - \mathcal{F}(x)| \leq \sqrt{2}\sqrt{1 - \text{Re}\mathcal{F}(x)}.
$$

Remark: Property a2) says that if $\mathcal{F}(x)$ is continuous at the origin (w.r.t. whatever topology) then $\mathcal{F}(x)$ is continuous in the whole $\mathcal{H}$ (w.r.t. the same topology). Property a3) says that the continuity of $\text{Re}\mathcal{F}(x)$ impies the continuity of $\mathcal{F}(x)$. Compare b) of this Theorem.

Step 2. Let $\{e_n\}$ be a fixed orthonormal basis of $\mathcal{H}$. For each $n \geq 1$, define $\mathcal{F}_{e_1,\ldots,e_n}(x)$ in $\mathbb{R}^n$ by

$$
\mathcal{F}_{e_1,\ldots,e_n}(a_1,\ldots,a_n) = \mathcal{F}(a_1e_1 + \ldots + a_ne_n)
$$

Note that $\text{Re}\mathcal{F}$ is continuous at the origin by b). Hence $\mathcal{F}$ is continuous in $\mathcal{H}$ by a2) and a3). Hence $\mathcal{F}_{e_1,\ldots,e_n}$ is positive definite and $\mathcal{F}_{e_1,\ldots,e_n}(0,\ldots,0) = 1$. Bochner’s Theorem for $\mathbb{R}^n$ gives us a family of probability measures $\{\mathcal{P}_n\}$ such that

$$
\mathcal{F}_{e_1,\ldots,e_n}(a_1,\ldots,a_n) = \int_{\mathbb{R}^n} e^{i\langle y, a \rangle} \mathcal{P}_n(dy), \quad a = (a_1,\ldots,a_n)
$$

It is easy to see that $\{\mathcal{P}_n\}$ is a consistent family. Hence Kolmogorov’s Theorem implies the existence of a probability space $(\Omega, \mathcal{P}_\Omega)$ and a sequence of random variables $\{X_n\}$ such that

$$
\mathcal{P}_n = \mathcal{P}_\Omega^{X_n}(X_1,X_2,\ldots,X_n)^{-1}, \quad n = 1, 2, \ldots
$$

Therefore,

$$
\mathcal{F}(a_1e_1 + a_2e_2 + \ldots + a_ne_n) = \int_{\mathbb{R}^n} e^{i\langle y, a \rangle} \mathcal{P}_n(dy) = \int_{\Omega} e^{i\langle a, X \rangle} \mathcal{P}(d\mathcal{P})
$$

Step 3. Suppose we can show that $\sum_{n=1}^{\infty} X_n^2 < \infty$ almost surely (this will be shown in Step 4) then we are done. To see this define,

$$
X(\omega) = \sum_{n=1}^{\infty} X_n(\omega) e_n, \quad \omega \in \Omega
$$

Then $X$ is measurable from $\Omega$ into $\mathcal{H}$. Define $\mathcal{P} = \mathcal{P}_\Omega^{X^{-1}}$. $\mathcal{P}$ is a probability Borel measure of $\mathcal{H}$. Let $\Pi_n$ be the orthogonal projection of $\mathcal{H}$ into the span of $e_1,e_2,\ldots,e_n$; i.e.

$$
\Pi_n = \langle x, e_1 \rangle e_1 + \langle x, e_2 \rangle e_2 + \ldots + \langle x, e_n \rangle e_n, \quad x \in \mathcal{H}.
$$
Then
\[ \Pi_n X = \sum_{k=1}^{n} X_k e_k \]

By Step 2 we have,
\[ \mathcal{F}(\Pi_n x) = \int_{\Omega} e^{i\langle x, \Pi_n x \rangle} d\mathbb{P} \cdot \]

Now letting \( n \to \infty \) and observing \( \Pi_n x \to x \) in \( \mathcal{H} \) as \( n \to \infty \), we have \( \mathcal{F}(\Pi_n x) \) by the continuity of \( \mathcal{F} \). Apply Lebesgue’s Dominated Convergence Theorem,
\[ \int_{\Omega} e^{i\langle x, \Pi_n x \rangle} d\mathbb{P} \to \int_{\Omega} e^{i\langle x, x \rangle} d\mathbb{P} \quad \text{as} \quad n \to \infty \]

Hence we have,
\[ \mathcal{F}(x) = \int_{\Omega} e^{i\langle x, x \rangle} d\mathbb{P} = \int_{\mathcal{H}} e^{i\langle x, y \rangle} \mathbb{P}(dy), \quad x \in \mathcal{H}. \]

**Step 4.** To show that \( \sum_{n=1}^{\infty} X_n^2 < \infty \) almost surely.

First note that
\[ \int_{\mathbb{R}^n} e^{i(a_1 x_1 + a_2 x_2 + \ldots + a_n x_n)} \left( \frac{1}{\sqrt{2\pi}} \right)^n e^{-\frac{1}{2}(y_1^2 + y_2^2 + \ldots + y_n^2)} \, dy_1 dy_2 \ldots dy_n = e^{-\frac{1}{2}(a_1^2 + a_2^2 + \ldots + a_n^2)} \]

Therefore
\[ \int_{\Omega} e^{-\frac{1}{2}(X_{k_1}^2 + X_{k_2}^2 + \ldots + X_{k_n}^2)} d\mathbb{P} = \]
\[ = \int_{\mathbb{R}^n} \left[ \frac{1}{\sqrt{2\pi}} \right]^n \exp \left\{ \sum_{j=1}^{n} X_{k+j} Y_j \right\} \exp \left\{ -\frac{1}{2} \sum_{j=1}^{n} Y_j^2 \right\} dy_1 dy_2 \ldots dy_n \, d\mathbb{P} = \]
\[ = \int_{\mathbb{R}^n} \exp \left\{ \sum_{j=1}^{n} X_{k+j} Y_j \right\} \left[ \frac{1}{\sqrt{2\pi}} \right]^n \exp \left\{ -\frac{1}{2} \sum_{j=1}^{n} Y_j^2 \right\} dy_1 dy_2 \ldots dy_n = \]
\[ = \int_{\mathbb{R}^n} \mathcal{F}(y_1 e_{k+1} + y_2 e_{k+2} + \ldots + y_n e_{k+n}) \left[ \frac{1}{\sqrt{2\pi}} \right]^n \exp \left\{ -\frac{1}{2} \sum_{j=1}^{n} y_j^2 \right\} dy_1 dy_2 \ldots dy_n = \]
\[ = \int_{\mathbb{R}^n} \Re \left[ \mathcal{F}(y_1 e_{k+1} + y_2 e_{k+2} + \ldots + y_n e_{k+n}) \right] \left[ \frac{1}{\sqrt{2\pi}} \right]^n \exp \left\{ -\frac{1}{2} \sum_{j=1}^{n} y_j^2 \right\} dy_1 dy_2 \ldots dy_n, \]

since the left hand side is real.

Now let \( \varepsilon > 0 \) be given, we use assumption b) to get \( C_\varepsilon \in \mathcal{S} \) such that
\[ 1 - \Re \mathcal{F}(x) \leq \langle C_\varepsilon x, x \rangle + \varepsilon \quad \text{for all} \quad x \in \mathcal{H}. \]

Therefore,
3.3. CHARACTERISTIC FUNCTIONAL OF THE PROBABILITY MEASURE

\[
1 - \int_{\Omega} e^{-\frac{1}{2}(X_{k+1}^2 + X_{k+2}^2 + \ldots + X_{k+n}^2)} \, d\mathcal{P} \leq \\
\leq \int_{\mathbb{R}^n} \left(1 - \text{Re} \left[ F (y_1 e_{k+1} + y_2 e_{k+2} + \ldots + y_n e_{k+n}) \right] \right) p_1(dy),
\]

where
\[
p_i(dy) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2}(y_1^2 + y_2^2 + \ldots + y_n^2)\right) dy_1 dy_2 \ldots dy_n
\]
is a probability measure in \(\mathbb{R}^n\). Then we have,
\[
\int_{\mathbb{R}^n} \left(1 - \text{Re} \left[ F (y_1 e_{k+1} + y_2 e_{k+2} + \ldots + y_n e_{k+n}) \right] \right) p_1(dy) \leq \\
\leq \int_{\mathbb{R}^n} \left[ \{C_\varepsilon (y_1 e_{k+1} + y_2 e_{k+2} + \ldots + y_n e_{k+n}) \}, y_1 e_{k+1} + y_2 e_{k+2} + \ldots + y_n e_{k+n} + \varepsilon \} p_1(dy) = \\
= \varepsilon + \int_{\mathbb{R}^n} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i y_j \left\langle C_\varepsilon e_{k+i}, e_{k+j} \right\rangle p_i(dy) = \\
= \varepsilon + \sum_{i=1}^{n} \sum_{j=1}^{n} \left\langle C_\varepsilon e_{k+i}, e_{k+j} \right\rangle \int_{\mathbb{R}^n} y_i y_j p_i(dy) = \\
= \varepsilon + \sum_{j=1}^{n} \left\langle C_\varepsilon e_{k+j}, e_{k+j} \right\rangle.
\]

Let \(n \to \infty\), apply Lebesgue’s Dominated Convergence Theorem to the left hand side and remember that \(C_\varepsilon \in S\). Hence,
\[
1 - \int_{\Omega} e^{-\sum_{j=1}^{k} X_{k+j}^2} \, d\mathcal{P} \leq \varepsilon + \sum_{m=k+1}^{\infty} \left\langle C_\varepsilon e_m, e_m \right\rangle \leq \\
\leq 2\varepsilon \quad \text{whenever} \ k \geq k_0 \ \text{say}.
\]

Hence,
\[
\int_{\Omega} e^{-\sum_{j=1}^{k} X_{k+j}^2} \, d\mathcal{P} \geq 1 - 2\varepsilon, \quad k \geq k_0.
\]

Finally,
\[
\mathcal{P} \left\{ \sum_{n=1}^{\infty} X_n^2 < \infty \right\} \geq \int_{\Omega} e^{-\sum_{j=1}^{\infty} X_{k+j}^2} \, d\mathcal{P} \geq 1 - 2\varepsilon
\]

Thus,
\[
\mathcal{P} \left\{ X_1^2 + \ldots + X_n^2 + \ldots < \infty \right\} \geq 1 - 2\varepsilon, \quad \text{for any} \ \varepsilon > 0.
\]

Of course we must have
\[
\mathcal{P} \left\{ X_1^2 + \ldots + X_n^2 + \ldots < \infty \right\} = 1,
\]

which is what we wanted to prove. \(\blacksquare\)
Finally we present some basic properties concerning operations with characteristic functionals

**Theorem 3.9**:  a) If $\mathcal{F}_1(x)$ and $\mathcal{F}_2(x)$ are characteristic functionals and $p_1 + p_2 = 1$, where $p_1, p_2$ are non-negative numbers, then $p_1 \mathcal{F}_1(x) + p_2 \mathcal{F}_2(x)$ is also a characteristic functional.

b) If $\mathcal{F}_1(x)$ and $\mathcal{F}_2(x)$ are characteristic functionals, then $\mathcal{F}_1(x) \cdot \mathcal{F}_2(x)$ is also a characteristic functional.

c) If $\mathcal{F}_s(x)$ is a family of characteristic functionals depending upon a random parameter $s$, and $F(s)$ is the distribution of $s$, then

$$\int_s \mathcal{F}_s(x) dF(s)$$

is also a characteristic functional.

d) Let $\mathcal{F}_X(x)$ be a characteristic functional of the process $X(t; \beta)$ and $\mathcal{F}_Y(x)$ a characteristic functional of another process $Y(t; \beta)$ being a linear transformation of $X(t; \beta)$.

1) If $Y(t; \beta) = X(t; \beta) + a(t)$, where $a(t)$ a non-random function, then

$$\mathcal{F}_Y(x) = \mathcal{F}_X(x) \cdot \exp \left[ i \int \overline{a(t)} x(t) dt \right]$$

2) If $A_L$ is an integral linear operator of the form

$$Y(t; \beta) = A_L X(t; \beta) = \int_\gamma G(s, t) X(s; \beta) ds$$

where $G(s, t)$ is a suitable non-random function, then

$$\mathcal{F}_Y(x) = \mathcal{F}_X(A_L^* x) = \mathcal{F}_X \left( \int_\gamma G(t, s) x(s) ds \right)$$

where $A_L^*$ is a conjugate to $A_L$.

**Proof**: Properties a) and c) follows from the corresponding properties of the probability measures and the continuity properties of the infinite dimensional Fourier transform. Property b) is a direct consequence of the linearity property of the Stochastic Processes space, since $\mathcal{F}_1(x) \cdot \mathcal{F}_2(x)$ is the characteristic functional of the sum $X_1(t; \beta) + X_2(t; \beta)$ with characteristic functional $\mathcal{F}_1(x), \mathcal{F}_2(x)$ respectively.

Finally, properties d) can be proven immediately using the definition of the characteristic functional.

**Remark 3.10**: The last group of properties are very essential tools for the combination of different characteristic functionals, to create new ones, more efficient for solving Functional Differential Equations. As we shall see at Chapter 5 the simplest application of those properties is the convex superposition of Gaussian Functionals to make a representation with local characteristics (kernel characteristic functionals).
3.4. Mean and Covariance Operator for Probability Measures

Now we are going to generalize the concept of the covariance functions for a probability measure defined on a real, separable, Hilbert space \( \mathcal{H} \).

**Definition 4.1**: Let \( \mathcal{P} \) be a probability measure in \( \mathcal{H} \). The mean value of \( \mathcal{P} \) is an element \( m_\mathcal{P} \in \mathcal{H} \) such that

\[
\langle m_\mathcal{P}, x \rangle = \int_{\mathcal{H}} \langle z, x \rangle \mathcal{P}(dz) \quad x \in \mathcal{H} \quad (1)
\]

**Remark 4.2**: In general \( m_\mathcal{P} \) does not exist. However, if \( \int_{\mathcal{H}} |x| \mathcal{P}(dx) \) then \( m_\mathcal{P} \) exists and

\[
|m_\mathcal{P}| \leq \int_{\mathcal{H}} |x| \mathcal{P}(dx) \quad (2)
\]

**Remark 4.3**: Using the distribution of the measurable function \( \langle x, \cdot \rangle \)

\[
\mathcal{P}_x(a) = \mathcal{P}\left\{ y \in \mathcal{H}; \langle x, y \rangle \leq a \right\}
\]

We have that

\[
\langle m_\mathcal{P}, x \rangle = \int_{\mathcal{H}} \langle z, x \rangle \mathcal{P}(dz) = \int_{-\infty}^{\infty} t \mathcal{P}_x(dt) \quad x \in \mathcal{H} \quad (3)
\]

Analogously we define the covariance operator.

**Definition 4.4**: Let \( \mathcal{P} \) be a probability measure in \( \mathcal{H} \). The covariance operator \( C_\mathcal{P} \) of \( \mathcal{P} \) is defined by

\[
\langle C_\mathcal{P}, x, y \rangle = \int_{\mathcal{H}} \langle x, z \rangle \langle y, z \rangle \mathcal{P}(dz) \quad x, y \in \mathcal{H} \quad (4)
\]

**Remark 4.5**: The operator \( C_\mathcal{P} \) may not exist. If \( C_\mathcal{P} \) exists, it is positive definite and self–adjoint.

**Remark 4.6**: Using the distribution of the measurable function \( \langle x, \cdot \rangle \)

\[
\mathcal{P}_x(a) = \mathcal{P}\left\{ y \in \mathcal{H}; \langle x, y \rangle \leq a \right\}
\]

We have that

\[
\langle C_\mathcal{P}, x, x \rangle = \int_{\mathcal{H}} \langle x, z \rangle^2 \mathcal{P}(dz) = \int_{-\infty}^{\infty} t^2 \mathcal{P}_x(dt) \quad x, y \in \mathcal{H} \quad (5)
\]

**Example 4.7**: The covariance operator of \( \delta_{x_0} \) is \( S_{\delta_{x_0}} x = \langle x, x_0 \rangle x_0 \) since

\[
\langle C_\delta_{x_0}, x, y \rangle = \langle x, x_0 \rangle \langle y, x_0 \rangle = \langle \langle x, x_0 \rangle x_0, y \rangle \quad (6)
\]

**Example 4.8**: The covariance operator of \( \mathcal{P}_e \) (example 1.4.) is
\[ C_{\mathcal{P}} x = \left( \int_{\mathbb{R}} t^2 \mathcal{P}(dt) \right) \langle x, \hat{\epsilon} \rangle \hat{\epsilon} \]  

(8)

since,

\[ \mathcal{C}_{\mathcal{P}}, x, y \rangle = \int_{\mathbb{R}} \langle x, z \rangle \langle y, z \rangle \mathcal{P}(dz) = \int_{\mathbb{R}} \langle x, z \rangle \langle y, z \rangle \mathcal{P}(dz) = \langle x, \hat{\epsilon} \rangle \langle y, \hat{\epsilon} \rangle \int_{\mathbb{R}} t^2 \mathcal{P}(dt) \]

\[ \mathcal{C}_{\mathcal{P}}, x, y \rangle = \int_{\mathbb{R}} \langle x, z \rangle \langle y, z \rangle \mathcal{P}(dz) = \int_{\mathbb{R}} \langle x, z \rangle \langle y, z \rangle \mathcal{P}(dz) = \langle x, \hat{\epsilon} \rangle \langle y, \hat{\epsilon} \rangle \int_{\mathbb{R}} t^2 \mathcal{P}(dt) \]

**Remark 4.9** : Suppose that \( \int_{\mathbb{R}} t^2 \mathcal{P}(dt) = 1 \). Then \( C_{\mathcal{P}}, x, y \rangle = C_{\delta_0} \) even though \( \mathcal{P}, e \) and \( \delta_0 \) are two different measures. Hence \( \mathcal{P}, e \) is not uniquely determined by its covariance operator. However, we will see later on that a Gaussian measure of mean zero is completely determined by its covariant operator.

The following theorem gives us all the necessary condition for an operator to be the covariance operator of a probability measure.

**Theorem 4.10** : Let \( \mathcal{P} \) be a probability measure in \( \mathcal{H} \) and \( C_{\mathcal{P}} \) the covariance operator of \( \mathcal{P} \). Then \( \int_{\mathcal{H}} |x|^2 \mathcal{P}(dx) < \infty \) if and only if \( C_{\mathcal{P}} \in \mathcal{S} \). In fact, \( \text{trace}[C_{\mathcal{P}}] = \int_{\mathcal{H}} |x|^2 \mathcal{P}(dx) \).

**Proof** : **Sufficiency.** Let \( \{e_n\} \) be an orthogonal basis of \( \mathcal{H} \). By Monotone Convergence Theorem we have,

\[ \int_{\mathcal{H}} |x|^2 \mathcal{P}(dx) = \lim_{n \to \infty} \int_{\mathcal{H}} \left[ \langle x, e_1 \rangle^2 + \langle x, e_2 \rangle^2 + \cdots + \langle x, e_n \rangle^2 \right] \mathcal{P}(dx) \]

But

\[ \int_{\mathcal{H}} \langle x, e_1 \rangle^2 \mathcal{P}(dx) = \langle C_{\mathcal{P}}, e_1, e_1 \rangle \]

Hence,

\[ \int_{\mathcal{H}} |x|^2 \mathcal{P}(dx) = \lim_{n \to \infty} \sum_{j=1}^{n} \langle C_{\mathcal{P}}, e_j, e_j \rangle = \sum_{j=1}^{\infty} \langle C_{\mathcal{P}}, e_j, e_j \rangle = \text{trace}[C_{\mathcal{P}}] < \infty \]

**Necessity.** First of all, we have to show the existence of the covariance operator. Clearly,

\[ |\langle x, z \rangle| \leq |x| \cdot |y| \cdot |z|^2 \]

Hence,
Therefore the bilinear form \( \int\langle x,z\rangle \cdot \langle y,z\rangle \mathcal{P}^x (dz) \) is continuous. Hence there exists a linear operator \( C_{\mathcal{P}^x} \) in \( \mathcal{H} \) such that,

\[
\langle C_{\mathcal{P}^x} x, y \rangle = \int \langle x,z\rangle \cdot \langle y,z\rangle \mathcal{P}^x (dz)
\]

Obviously, \( C_{\mathcal{P}^x} \) is self-adjoint and positive definite. To show that \( C_{\mathcal{P}^x} \in \mathcal{S} \), it is sufficient to show that if \( \{e_n\} \) is an orthonormal basis of \( \mathcal{H} \) then the series \( \sum_{j=1}^{\infty} \langle C_{\mathcal{P}^x} e_j, e_j \rangle \) is convergent.

But

\[
\sum_{j=1}^{\infty} \langle C_{\mathcal{P}^x} e_j, e_j \rangle = \sum_{j=1}^{\infty} \int \langle x,e_j \rangle^2 \mathcal{P}^x (dx) = \\
= \int \sum_{j=1}^{\infty} \langle x,e_j \rangle^2 \mathcal{P}^x (dx) = (\text{Monotone Convergence Theorem}) \\
= \int |x|^2 \mathcal{P}^x (dx).
\]

Hence we have not only shown that \( C_{\mathcal{P}^x} \in \mathcal{S} \), but also proven that

\[
\text{trace}[C_{\mathcal{P}^x}] = \int |x|^2 \mathcal{P}^x (dx)
\]

It’s possible for a probability measure to have a well defined correlation operator, while the mean value can not be defined. The next example illustrates the above,

**Example 4.11** : Let \( \mathcal{P}^x \) be a probability measure in \( \mathbb{R} \) such that \( \int_{-\infty}^{\infty} t^2 \mathcal{P}^x (dt) < \infty \) and \( \int_{-\infty}^{\infty} t \mathcal{P}^x (dt) = \infty \). Then \( C_{\mathcal{P}^x} \) exists, but \( m_{\mathcal{P}^x} \) doesn’t exist.
3.5. **Moments from the Characteristic Functional**

In the theory of random variables we have seen that the characteristic function includes all the information of the stochastic process. As a result we can have all the moments of the random variable from the characteristic function. In this section we will see how we can take analogous results for the characteristic functional.

Let $F$ be the characteristic functional of the probability measure $\mathcal{P}$ defined on $\mathcal{H}$. Let, also $F \in C^2$ (i.e. the space of twice differentiable functionals). Calculating the first Frechet derivative we have

$$
\delta F(x)[z] = \left. \frac{d F(x+tz)}{dt} \right|_{t=0} = \left. \frac{d}{dt} \left( \int_{\mathcal{H}} e^{i(x+tz,y)} \mathcal{P}(dy) \right) \right|_{t=0} = i \cdot \int_{\mathcal{H}} \langle z, y \rangle e^{i(x,y)} \mathcal{P}(dy)
$$

Thus, we have,

$$
\delta F(0)[z] = i \cdot \int_{\mathcal{H}} \langle z, y \rangle \mathcal{P}(dy) = i \cdot \langle m_{\alpha}, z \rangle \Leftrightarrow 
\langle m_{\alpha}, z \rangle = -i \cdot \delta F(0)[z] \quad z \in \mathcal{H}
$$

Getting the second Frechet derivative we have,

$$
\delta^2 F(x)[z,w] = \left. \frac{d \delta F(x+tw)}{dt} \right|_{t=0} = \left. \frac{d}{dt} \left( i \int_{\mathcal{H}} \langle z, y \rangle e^{i(x+tw,y)} \mathcal{P}(dy) \right) \right|_{t=0} = -\int_{\mathcal{H}} \langle w, y \rangle \langle z, y \rangle e^{i(x,y)} \mathcal{P}(dy)
$$

Thus, we have,

$$
\delta^2 F(0)[z,w] = -\int_{\mathcal{H}} \langle w, y \rangle \langle z, y \rangle \mathcal{P}(dy) = -\langle C_{\alpha} z, w \rangle 
\langle C_{\alpha} z, w \rangle = -\delta^2 F(0)[z,w] \quad z, w \in \mathcal{H}
$$

Similarly, we can get greater order quantities, depending on the smoothness of the functional.

Another very useful property of the characteristic functional is the ability to get moments, in the usual sense, for specific values of the time or spatial variable of the stochastic process. Let $\mathcal{H}$ be the space of $C^1(\mathbb{R}^N, \mathbb{R})$ and $\langle \cdot, \cdot \rangle$ the corresponding inner product. Then, following the notation of Section 2.1.2/Chapter 2 we have

$$
E^\beta [X(t;\beta)] = m_\chi(t) = \int_{-\infty}^{\infty} x \cdot P_{X(t)}(dx)
$$

But from equation (2) we have

$$
\int_{\mathcal{H}} \langle z, y \rangle \mathcal{P}(dy) = \langle m_{\alpha}, z \rangle = -i \cdot \delta F(0)[z] \quad z \in \mathcal{H}
$$

Using Remark 2.5 we have
3.5. MOMENTS FROM THE CHARACTERISTIC FUNCTIONAL

\[ \int_{-\infty}^{\infty} t \mathcal{P}_z(dt) = -i \cdot \delta \mathcal{F}(0)[z] \quad z \in \mathcal{H} \]  
(5)

where

\[ \mathcal{P}_z(a) = \mathcal{P}\{y \in \mathcal{H}; \langle z, y \rangle \leq a\} \]

Setting \( z = \delta(t-t_0) \) we have

\[ \mathcal{P}_{\delta(t-t_0)}(a) = \mathcal{P}\{y \in \mathcal{H}; y(t_0) \leq a\} = P_{x(t)}(a) \]  
(6)

Thus,

\[ E^\delta[X(t_0; \beta)] = -i \cdot \delta \mathcal{F}(0)[\delta(t-t_0)] = -i \cdot \frac{\delta \mathcal{F}}{\delta x(t_0)}(0) \]  
(7)

where \( \frac{\delta \mathcal{F}}{\delta x(t)}(\bullet) \) denotes the Volterra derivative of the functional \( \mathcal{F}(\bullet) \) at the point \( t \).

**Remark 5.1** : Hence, we see that the mean value element is of the same kind as the Hilbert space. For a space of functions, the mean value element is also a function and the classical moments are integrals with respect to measures defined by the probability of the values at an instant time.

**Example 5.2** : Consider, for example the space of continuous functions defined on the time interval \( [0, T] \), bounded above and below from the functions \( c_U(t) \) and \( c_D(t) \), i.e. the space of functions

\[ \mathcal{U} = \{u \in C([0, T]): c_D(t) \leq u(t) \leq c_U(t), \forall t \in [0, T]\} \]

Then the mean value element \( m(t) \) will be a function into the space \( \mathcal{U} \), as illustrated to the figure below. The mean value at time instants \( t_i \) will be the value of the mean element, \( m(t_i) \), or equivalently the mean value of the distribution connected with the values along the line \( \varepsilon_i \).

![Diagram](image)

For the general case we have

\[ E^\delta[X(t_1; \beta)^n \cdot X(t_2; \beta)^{n_2} \cdot \ldots \cdot X(t_M; \beta)^{n_M}] = \int_{-\infty}^{\infty} x_1^{n_1} \cdot x_2^{n_2} \cdot \ldots \cdot x_M^{n_M} \cdot P_{x(t_1), \ldots, x(t_M)}(dx) \]
where \( P_{X(t_1), \ldots, X(t_M)} \) is the joint probability distribution for the time instants \( t_1, t_2, \ldots, t_M \).

Using Definition 2.2 we will have
\[
P_{\delta(t-t_1), \ldots, \delta(t-t_M)}(a) = \left\{ y \in \mathcal{H} : y(t_1) \leq a_1, y(t_2) \leq a_2, \ldots, y(t_M) \leq a_M \right\} = P_{X(t_1), \ldots, X(t_M)}(a), \quad a = (a_1, a_2, \ldots, a_M) \in \mathbb{R}^M
\]

Using same arguments, as before we can show that
\[
E^\beta \left[ X(t_1; \beta)^{a_1} \cdot X(t_2; \beta)^{a_2} \cdots X(t_M; \beta)^{a_M} \right] = \frac{1}{i^{a_1 + \ldots + a_M}} \delta^{a_1 + \ldots + a_M} \mathcal{F}^{-1}(0) \left[ \frac{\delta(t-t_1), \ldots, \delta(t-t_1), \ldots, \delta(t-t_M), \ldots, \delta(t-t_M)}{a_1 \ldots a_M} \right] = \frac{1}{i^{a_1 + \ldots + a_M}} \delta^{a_1 + \ldots + a_M} \mathcal{F}^{-1}(0) \delta(t-t_1)^{a_1} \delta(t-t_2)^{a_2} \cdots \delta(t-t_M)^{a_M} \tag{8}
\]

Again we have two consistent interpretations of the moments \( E^\beta \left[ X(t_1; \beta)^{a_1} \cdots X(t_M; \beta)^{a_M} \right] \) as in Example 5.2. More specifically, the above moment can be interpreted as the value of the corresponding operator at \( t_1, t_2, \ldots, t_M \) or the moment of the joint distribution connected with the values along the lines \( \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_M \).

Another important property of the characterization of a stochastic process through the characteristic functional or the probability measure is the possibility to get moments for the time derivatives of the stochastic process. From equation (5) we have
\[
\int_{-\infty}^{t} t \mathcal{D}^2_z dt = -i \cdot \delta \mathcal{F}^{-1}(0)[z] \quad z \in \mathcal{H}
\]
where
\[
\mathcal{D}^2_z(a) = \mathcal{D}^2 \left\{ y \in \mathcal{H} : \langle z, y \rangle \leq a \right\}
\]

Setting \( z = \delta'(t-t_0) \) we have
\[
\mathcal{D}^2_{\delta(t-t_0)}(a) = \mathcal{D}^2 \left\{ y \in \mathcal{H} : y'(t_0) \leq a \right\} = P_{X(t)}(a) \tag{9}
\]

Thus
\[
E^\beta \left[ X'(t_0; \beta) \right] = -i \cdot \delta \mathcal{F}^{-1}(0)[\delta'(t-t_0)] \tag{10}
\]

Similarly we can get general formulas
\[
i^{a_1 + \ldots + a_M} \cdot E^\beta \left[ X(t_1; \beta)^{a_1} \cdot X(t_1; \beta)^{a_2} \cdots X(t_1; \beta)^{a_M} \right] = \delta^{a_1 + \ldots + a_M} \mathcal{F}^{-1}(0) \left[ \frac{\delta^{a_1}(t-t_1), \ldots, \delta^{a_2}(t-t_1), \ldots, \delta^{a_M}(t-t_M)}{a_1 \ldots a_M} \right] \tag{11}
\]
3.6. Characteristic Functions from the Characteristic Functional

Another essential property of the characteristic functional is the ability to get characteristic functions, in the usual sense, for specific values of the time or spatial variable of the stochastic process. However more important is the possibility of calculating characteristic functions for the derivatives of the responses.

Using the definition of Section 2.1.3/Chapter 2. for the characteristic function, we have,

\[ \phi_{X(t)}(v) \equiv E^\beta(e^{ivx}) = \int_{-\infty}^{+\infty} e^{ivx} P_{X(t)}(dx) \]  

(1)

For the characteristic functional of the probability measure, we have, by definition,

\[ \mathcal{F}(x) = \int_{\mathcal{H}} e^{ix} \mathcal{P}(dy) \quad x \in \mathcal{H} \]

Setting \( x = \nu z \), where \( \nu \in \mathbb{R} \), and using Remark 2.5. we have

\[ \mathcal{F}(\nu z) = \int_{\mathcal{H}} e^{i\nu(z,y)} \mathcal{P}(dy) = \int_{-\infty}^{\infty} e^{i\nu t} \mathcal{P}_z(dt) \quad z \in \mathcal{H} \]  

(2)

where

\[ \mathcal{P}_z(a) = \mathcal{P}\{y \in \mathcal{H}; \langle z, y \rangle \leq a\}, \quad a \in \mathbb{R} \]

Setting \( z = \delta(t-t_0) \) we have

\[ \mathcal{P}_{\delta(t-t_0)}(a) = \mathcal{P}\{y \in \mathcal{H}; y(t_0) \leq a\} = P_{X(t_0)}(a), \quad a \in \mathbb{R} \]  

(3)

Thus,

\[ \phi_{X(t)}(v) = \mathcal{F}(v \delta(t-t_0)), \quad v \in \mathbb{R} \]  

(4)

For higher dimensions we have analogous formulas. More specifically, by setting to the characteristic functional \( x = v_1 z_1 + \ldots + v_M z_M \), where \( v = (v_1, v_2, \ldots, v_M) \in \mathbb{R}^M \) and \( z = (z_1, z_2, \ldots, z_M) \in \mathcal{H}^M \) we have

\[ \mathcal{F}(v_1 z_1 + \ldots + v_M z_M) = \int_{\mathcal{H}^M} e^{i[v_1(z_1, y_1) + v_2(z_2, y_2) + \ldots + v_M(z_M, y_M)]} \mathcal{P}(dy) = \]

\[ \int_{-\infty}^{\infty} e^{i[v_1 t_1 + v_2 t_2 + \ldots + v_M t_M]} \mathcal{P}_{z_1, z_2, \ldots, z_M}(dt) \]  

(5)

where

\[ \mathcal{P}_{z_1, z_2, \ldots, z_M}(a) = \mathcal{P}\{y \in \mathcal{H}; \langle z_1, y \rangle \leq a_1, \langle z_2, y \rangle \leq a_2, \ldots, \langle z_M, y \rangle \leq a_M\} \]  

(6)

Setting \( z_i = \delta(t-t_i), \quad i = 1, 2, \ldots, M \) and using Definition 2.2 we have

\[ \mathcal{P}_{\delta(t-t_1) \ldots \delta(t-t_M)}(a) = \mathcal{P}\{y \in \mathcal{H}; y(t_1) \leq a_1, y(t_2) \leq a_2, \ldots, y(t_M) \leq a_M\} = \]
\[
P_{X(t_1),\ldots,X(t_M)}(a) = \Phi_X(t_1, t_M)(a), \quad a = (a_1, a_2, \ldots, a_M) \in \mathbb{R}^M
\]

Hence,
\[
\phi_{X(t_1),\ldots,X(t_M)}(v) = \mathcal{F} \left( v_1 \delta(t-t_1) + \ldots + v_M \delta(t-t_M) \right), \quad v = (v_1, v_2, \ldots, v_M) \in \mathbb{R}^M
\]

(7)

Additionally we can calculate characteristic functions for the time derivatives of the stochastic process. Setting to equations (5), (6) \( z_i = \delta'(t-t_i) \), \( i = 1, 2, \ldots, M \) and using Definition 2.2 we have
\[
\mathcal{D}^{(t-t_i),\ldots,(t-t_M)}(a) = \mathcal{D}^{(t-t_i)} \left\{ y \in \mathcal{Y}; y'(t_1) \leq a_1, y'(t_2) \leq a_2, \ldots, y'(t_M) \leq a_M \right\} = \\
= P_{X'(t_1),\ldots,X'(t_M)}(a), \quad a = (a_1, a_2, \ldots, a_M) \in \mathbb{R}^M
\]

Hence,
\[
\phi_{X'(t_1),\ldots,X'(t_M)}(v) = \mathcal{F} \left( v_1 \delta'(t-t_1) + \ldots + v_M \delta'(t-t_M) \right), \quad v = (v_1, v_2, \ldots, v_M) \in \mathbb{R}^M
\]

(8)

Finally, we can prove the general formulae, in the same way as before. We have from Definition 2.2 and equation (5) for \( z_i = \delta^{(n_i)}(t-t_i) \),
\[
\mathcal{D}^{(t-t_i),\ldots,(t-t_M)}(a) = \mathcal{D}^{(t-t_i)} \left\{ y \in \mathcal{Y}; y^{(n_1)}(t_1) \leq a_1, y^{(n_2)}(t_2) \leq a_2, \ldots, y^{(n_M)}(t_M) \leq a_M \right\} = \\
= P_{X'(t_1),\ldots,X'(t_M)}(a), \quad a = (a_1, a_2, \ldots, a_M) \in \mathbb{R}^M
\]

Hence we get the equality,
\[
\phi_{X^{(n_1)}(t_1),\ldots,X^{(n_M)}(t_M)}(v) = \mathcal{F} \left( v_1 \delta^{(n_1)}(t-t_1) + \ldots + v_M \delta^{(n_M)}(t-t_M) \right), \quad v \in \mathbb{R}^M
\]

(9)

Thus we have general expressions for the calculation of characteristic functions from the characteristic functionals for the values of the stochastic process as well as for the derivatives.

**Remark 6.1**: It is very useful for applications to work in the level of characteristic functional. Then all the above results gives us a method to interpret or visualize the resulting characteristic functional by taking the characteristic functions for finite dimensional subspaces and then calculate the Fourier transform to get the corresponding probability density functions. ■
3.7. Gaussian Measure

Gaussian measures play a very important role in the theory of probability measures in infinite dimensional spaces. Additionally, as we will see at Chapter 5, it is an essential tool, for the representation of the responses of stochastic dynamical systems. Their practical importance basically relies to the fact that we have analytical expressions for the infinite dimensional integrals of cylinder functions with respect to Gaussian measures. Another important feature is the localization of the probability mass into the infinite dimensional space of functions. In other words Gaussian measures have the property of kernel characteristic functionals, that it will be discussed extensively at Chapter 5/Section 5.4. The last property in combination with the property of analytical computation of integrals gives us a very promising tool for the numerical solution of functional differential equations, discussed in Chapter 5. Hence its very essential to study extensively Gaussian measures in $\mathcal{H}$. We start with the following

**Definition 7.1** : A Gaussian measure $\mathcal{P}_G$ in $\mathcal{H}$ is a probability measure in $\mathcal{H}$ such that for each $x \in \mathcal{H}$, the measurable function $\langle x, \cdot \rangle$ is normally distributed, i.e. there exists real numbers $m_x$ and $\sigma_x$ such that

$$\mathcal{P}_G(a) = \mathcal{P}_G \{ y \in \mathcal{H} : \langle x, y \rangle \leq a \} = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi\sigma_x}} e^{-\frac{(t-m_x)^2}{2\sigma_x}} dt \quad (1)$$

We will now compute the mean and covariant operator for the Gaussian measure. We have,

$$\langle m_{x'}, x \rangle = \int_{\mathcal{H}} \langle z, x \rangle \mathcal{P}_G(\,dz) \quad x \in \mathcal{H}$$

Using Remark 2.5 the above integral turns out, to be

$$\langle m_{x'}, x \rangle = \int_{-\infty}^{\infty} t \mathcal{P}_x(\,dt) = \int_{-\infty}^{\infty} t \cdot \frac{1}{\sqrt{2\pi\sigma_x}} e^{-\frac{(t-m_x)^2}{2\sigma_x}} dt = m_x \quad (2)$$

For the covariance operator we will have

$$\langle C_{x'}, x, x \rangle = \int_{\mathcal{H}} \langle x, z \rangle^2 \mathcal{P}_G(\,dz) \quad x \in \mathcal{H}$$

Using Remark 2.5, we will have

$$\langle C_{x'}, x, x \rangle = \int_{-\infty}^{\infty} t^2 \mathcal{P}_x(\,dt) = \int_{-\infty}^{\infty} t^2 \cdot \frac{1}{\sqrt{2\pi\sigma_x}} e^{-\frac{(t-m_x)^2}{2\sigma_x}} dt = \sigma_x \quad (3)$$

Concerning the characteristic functional of a Gaussian measure $\mathcal{P}_G$ we have the following

**Lemma 7.2** : Let $\mathcal{P}_G$ be a Gaussian measure in $\mathcal{H}$. Then its characteristic functional is given by

$$\mathcal{F}_G(x) = e^{i\langle m_{x'}, x \rangle - \frac{\langle C_{x'}, x, x \rangle}{2}} \quad (4)$$
where $m_{\varphi}$ is the mean of $\mathcal{P}^\varphi$ and $C_{\varphi}$ the covariance operator of $\mathcal{P}^\varphi$.

**Proof:** Using Remark 2.5 we have

$$\mathcal{F}^\varphi_G(x) = \int_{\mathcal{Q}} e^{i(x,y)} \mathcal{P}^\varphi_G(dy) = \int_{\mathcal{Q}} e^{i \pi} \mathcal{P}^\varphi_G(dt)$$

where $\mathcal{P}^\varphi_G(a) = \mathcal{P}^\varphi_G \{ y \in \mathcal{H}; \langle x,y \rangle \leq a \}$ is the distribution of $\langle x,x \rangle$. Hence,

$$\mathcal{F}^\varphi_G(x) = \int_{\mathcal{Q}} e^{i \pi} \frac{1}{\sqrt{2\pi\sigma_x}} e^{-\frac{(t-m_y)^2}{2\sigma}} dt$$

Make a change of variables first and use contour integration to conclude

$$\mathcal{F}^\varphi_G(x) = e^{i m_y - \frac{\sigma}{2}}.$$

But from (2) - (3) we have

$$m_x = \langle m_{\varphi}, x \rangle \quad \text{and} \quad \sigma_x = \langle C_{\varphi} x, x \rangle$$

Hence

$$\mathcal{F}^\varphi_G(x) = e^{i \langle m_{\varphi}, x \rangle - \frac{1}{2} \langle C_{\varphi} x, x \rangle} \quad \blacksquare$$

**Remark 7.3:** Since a probability measure is uniquely defined by its characteristic functional, then, in view of the above lemma, a Gaussian measure is uniquely defined by its mean and covariance operator. \(\blacksquare\)

For the characterization of a Gaussian measure we have the following

**Theorem 7.4 [Prohorov]:** Let $x_0 \in \mathcal{H}$ and $C_{\varphi} \in \mathcal{S}$. Then $\mathcal{F}(x) = e^{i \langle x, y \rangle - \frac{1}{2} \langle C_{\varphi} x, x \rangle}$ is the characteristic functional of a Gaussian measure in $\mathcal{H}$.

**Proof:** Clearly $F(0) = 1$ and $\mathcal{F}$ is positive definite. Consider first the case $x_0 = 0$. Then

$$1 - \text{Re} \mathcal{F}(x) \leq 1 - \frac{1}{2} \langle C_{\varphi}, x \rangle \leq \frac{1}{2} \langle C_{\varphi}, x \rangle$$

because $1 - e^{-y} \leq y$ for all $y \geq 0$.

Since $\frac{1}{2} C_{\varphi} \in \mathcal{S}$ part (b) of Minlos-Sazonov Theorem. (Theorem 3.7) is trivally satisfied.

Therefore by the conclusion of the Theorem, there exists a Borel measure $\mathcal{P}$ such that

$$\mathcal{F}(x) = \int_{\mathcal{Q}} e^{i(x,y)} \mathcal{P}(dy) \quad x \in \mathcal{H}$$

That is,

$$\int_{\mathcal{Q}} e^{i(x,y)} \mathcal{P}(dy) = e^{\frac{1}{2} \langle C_{\varphi}, x \rangle} \quad x \in \mathcal{H}$$
Let $P_x(a) = \{ y \in \mathcal{H} : \langle x, y \rangle \leq a \}$ be the distribution of $\langle x, \cdot \rangle$. Then using Remark 2.4 we have

$$\int_{\mathcal{H}} e^{it} \, dP_x(t) = e^{\langle C_{0,x}, t \rangle}$$

Hence $P_x$ is a normal distribution with mean 0 and variance $\langle C_{0,x}, \cdot \rangle$. That is $P_x$ is a Gaussian measure in $\mathcal{H}$. Consider now the case $x_0 \neq 0$. Let

$$\mathcal{F}(x) = e^{\frac{\langle C_{0,x}, x \rangle}{2}}$$

Then

$$\mathcal{F}(x) = e^{\langle t, x \rangle} \mathcal{F}(x)$$

By the case already proved, there exists a Gaussian measure $\mathcal{F}$ in $\mathcal{H}$ such that

$$\mathcal{F}(x) = \int e^{\langle t, x \rangle} \, d\mathcal{P}(t) \quad x \in \mathcal{H}$$

Define a Borel measure $\mathcal{P}$ in $\mathcal{H}$ as follows

$$\mathcal{P}(E) = \mathcal{F}(E - x_0), \quad E \in B(\mathcal{H})$$

It is easy to see that $\mathcal{P}$ is a Gaussian measure and

$$\int e^{\langle t, x \rangle} \, d\mathcal{P}(dy) = e^{\langle t, x \rangle} \mathcal{F}(x) = \mathcal{F}(x)$$

Finally, as we mention at the beginning, analytical integration of cylinder functions over Gaussian measures can be carried out. Hence for Gaussians measures we have

$$\int g(\langle x_1, y \rangle, \ldots, \langle x_n, y \rangle) P_{\mathcal{H}}(dy) = \int g(t_1, \ldots, t_n) P_x(dt_1, \ldots, dt_n) =$$

$$(2\pi)^{-n/2} (\det \tilde{C})^{-1/2} \int g(u) \exp \left\{ -\frac{1}{2} \left[ (u - m)^T \tilde{C}^{-1} (u - m)^T \right] \right\} du$$

Where, $\tilde{C}^{-1}$ is the inverse for matrix $[\tilde{C}_{ij}] = \langle C_{x_i}, x_j \rangle$, $i, j = 1, \ldots, n$, $m = (\langle m, x_1 \rangle, \ldots, \langle m, x_n \rangle)$. $\langle m, \cdot \rangle$ is the mean value of the measure and $\langle C_{\cdot, \cdot} \rangle$ is the correlation functional of the measure.

In Chapter 5/Section 5.1 we will give the precise formulation as well as the proof of the above result.
3.8. Tatarskii Characteristic Functional

As we saw at the previous section the characteristic functional for the Gaussian measure can be represented by explicit formulae. There exist few other cases for which it is possible to present such formulas. But the problem of extending the class of random functions for which it is possible to write down the characteristic functional explicitly is important for many applications. If we possess an explicit formula for the characteristic functional that contains several arbitrary functions, it is possible to use this model for an approximate solution of different statistical problems.

In this section we will present a characteristic functional representation, first introduced by V.I. Tatarskii (1995). This general expression creates a wide class of characteristic functionals. The above representation is based on the construction of a random function, which we will describe in the sequel.

Let the random function depending on vector argument $r \in \mathbb{R}^n$ of the form

$$X(r; \beta) = \begin{cases} 0 & \text{with probability } p_0 \\ \xi_k g(r, r_k) & \text{with probability } p_k \\ \vdots & \vdots \\ \sum_{k=0}^{N} \xi_k g(r, r_k) & \text{with probability } p_N \\ \vdots & \vdots \end{cases}$$

(1)

where the numbers $p_N$ obey the conditions

$$p_N \geq 0 \quad \sum_{N=0}^{\infty} p_N = 1$$

(2)

We assume the characteristic function

$$\chi(u) = \sum_{N=0}^{\infty} p_N \exp(iuN)$$

(3)

is known. The same information is available from the function

$$\Psi(A) \equiv \langle A^N \rangle = \sum_{N=0}^{\infty} p_N A^N = \chi(-i \log A)$$

(4)

The function $\Psi(A)$, the generating function of factorial moments, is related to the moments of random $N$ by the formula

$$\frac{d^n \Psi(A)}{dA^n} = \langle (N-1)(N-2)\ldots(N-n+1) A^{N-n} \rangle$$

(5)

Thus,

$$\left. \frac{d^n \Psi(A)}{dA^n} \right|_{A=1} = \langle (N-1)(N-2)\ldots(N-n+1) \rangle \equiv F_n$$

(6)
The numbers $F_n$ are called the factorial moments of the random variable $N$. The amplitudes $\xi_k$ are assumed to be statistically independent and have the same probability distribution and characteristic function independent of $k$:
\[
\varphi(i\nu) = \langle \exp(i\nu \xi_k) \rangle
\] (7)

We assume that all $r_k$ are statistically independent and have the same probability density function $W(r_k)$. Let us consider the characteristic functional for $X(r;\beta)$ given by (2). Substituting (1) into (2) we obtain
\[
\mathcal{F}(x) = \exp \left[ i \sum_{k=0}^{N} \xi_k \int g(r,r_k)x(r)dr \right]
\] (8)

We are able to average over $\xi$, and using (7) we obtain
\[
\mathcal{F}(x) = \prod_{k=0}^{N} \varphi \left( \int g(r,r_k)x(r)dr \right)
\] (9)

Here, $N$ and $r_k$ remain random. Because all $r_k$ are statistically independent and have the same probability density function $W(r_k)$, averaging over $r_k$ results in
\[
\mathcal{F}(x) = \prod_{k=0}^{N} \int dr_k W(r_k) \varphi \left( \int g(r,r_k)x(r)dr \right)
\] (10)

All factors in this product differ only by the notation for the variable of integration and are equal to each other. Thus,
\[
\mathcal{F}(x) = \left( \int dr' W(r') \varphi \left( \int g(r',r')x(r)dr \right) \right)^N
\] (11)

To fulfil the averaging with respect to $N$, we can use (4) with the value $A$ determined by the expression
\[
A = \int dr' W(r') \varphi \left( \int g(r',r')x(r)dr \right)
\] (12)

Substituting (12) into (11) we obtain
\[
\mathcal{F}(x) = \Psi \left( \int dr' W(r') \varphi \left( \int g(r',r')x(r)dr \right) \right)
\] (13)

This lead us to the following

**Theorem 8.1**: Let
a) $W : \mathbb{R}^n \rightarrow \mathbb{R}_+$, be an arbitrary non-negative function, that is normalized to 1
b) $\varphi(i\nu)$ be an arbitrary characteristic function
c) $g(\bullet,\bullet) : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ a completely arbitrary function and
d) $\Psi(A)$ be an arbitrary generating function of factorial moments, i.e. be of the form
\[
\Psi(A) = \chi(-i \log A),
\] where $\chi$ is an arbitrary characteristic function.

Then the functional
\[ \mathcal{F}(x) = \Psi \left( \int dr' W(r') \varphi \left( \int g(r, r') x(r) dr \right) \right) \]

is a characteristic functional. ■

**Remark 8.2:** Although the proof of the above theorem was carried out for the case of \( r \in \mathbb{R}^n \) being a spatial variable the same results hold for the case when we substitute \( r \) with \( t \) being time variable. ■

For the calculations that follow we will assume that the functions are defined onto the time interval \( I = [0, T] \).

### 3.8.1. Mean and Correlation Operators

As we compute at Section 4 of the present chapter the mean value operator of a probability measure is

\[ \langle m_{\varphi}, z \rangle = -i \cdot \delta \mathcal{F}(0)[z] \quad z \in \mathcal{H} \]

By direct calculation we can find that the mean operator will have the form

\[ \langle m_{\varphi}, z \rangle = -i \cdot \Psi'(1) \cdot \varphi'(0) \cdot \left( \int \int W(t') g(t', t) z(t) dt' dt \right) \quad z \in \mathcal{H} \quad (14) \]

Using the above result and equation (7) we will have that

\[ E^\beta \left[ X(t_0, \beta) \right] = -i \cdot \Psi'(1) \cdot \varphi'(0) \cdot \left( \int W(t) g(t_0, t) dt \right) \quad (15) \]

Concerning the covariance operator, we have proved from Section 4 that

\[ \langle C_{\varphi}, z, w \rangle = -\delta^2 \mathcal{F}(0)[z, w] \quad z, w \in \mathcal{H} \]

So, by calculating the second Frechet derivative, we will wave

\[ \delta^2 \mathcal{F}(0)[z_1, z_2] = \Psi''(1) \cdot \varphi'(0)^2 \cdot \left( \int \int W(t') \left( \sum_{k=1}^2 \int g(t', t) z_k(t) dt' \right) \right) + \Psi'(1) \cdot \varphi''(0) \cdot \left( \int \int W(t') \left( \sum_{k=1}^2 \int g(t', t) z_k(t) dt' \right) \right) \quad (16) \]

### 3.8.2. Characteristic functions

As we compute at Section 6 the joint characteristic function at time instants \( t_1, t_2, \ldots, t_M \) of a probability measure is

\[ \phi_{X(t_1), \ldots, X(t_M)}(v) = \mathcal{F} \left( v_1 \delta(t-t_1) + \ldots + v_M \delta(t-t_M) \right), \quad v = (v_1, v_2, \ldots, v_M) \in \mathbb{R}^M \]

By direct calculation we can find that the joint characteristic function will have the form

\[ \phi_{X(t_1), \ldots, X(t_M)}(v) = \mathcal{F} \left( v_1 \delta(t-t_1) + \ldots + v_M \delta(t-t_M) \right) = \Psi \left( \int W(t) \varphi \left( v_1 g(t_1, t) + \ldots + v_M g(t_M, t) \right) dt \right) \quad v \in \mathbb{R}^M \quad (17) \]
3.8. THE TATARKII REPRESENTATION OF A CHARACTERISTIC FUNCTIONAL

For characteristic functions concerning time derivatives a of the stochastic process we have using equation (9)

\[
\phi_{\chi^{(a)}(x_{(t)})}(v) = \mathcal{F} \left( v_1 \delta^{(a_1)}(t-t_1) + \cdots + v_M \delta^{(a_M)}(t-t_M) \right) = \Psi \left( \int W(t) \phi \left( v_1 \partial^n_{t_1} g(t_1, t) + \cdots + v_M \partial^n_{t_M} g(t_M, t) \right) dt \right), \quad v \in \mathbb{R}^M
\]  

(18)

3.8.3. The 2D Tatarskii Representation of a Characteristic Functional

For applications of characteristic functionals connected with stochastic differential equations, it is usually needed to have representations of vector valued stochastic processes. In this section we generalize the Tatarskii representation for the two dimensional case. Based on the same arguments we can easily extend the generalization presented above to the \( n \) – dimensional case.

We consider the random function depending on the \( n \)-dimensional vector argument \( r \) of the form

\[
X(r; \beta) = \begin{pmatrix}
(0, 0) \\
\cdots \\
\xi_k g\left(r, r'_k\right), \zeta_k h\left(r, r''_k\right) \\
\cdots \\
\sum_{k=0}^{N} \xi_k g\left(r, r'_k\right) + \sum_{k=0}^{N} \zeta_k h\left(r, r''_k\right)
\end{pmatrix}
\]

(19)

\( (\text{there is a chance to generalize further by assuming different } N \text{ and } M \text{ for every component, but the things gets very complicated. For our purpose the above generalization is the most suitable.)} \)

The numbers \( p_N \) obey the conditions

\[
p_N \geq 0, \quad \sum_{N,M=0}^{\infty} p_N = 1
\]  

(20)

We assume the characteristic function

\[
\chi(u) = \sum_{N,M=0}^{\infty} p_N \exp\left( iuN \right)
\]  

(21)

is known. The same information is available from the function

\[
\Psi(A) \equiv \left\langle A^N \right\rangle = \sum_{N,M=0}^{\infty} p_N A^N
\]  

(22)

The r.v.’s \( (\xi_k, \zeta_k) \) are assumed to be statistically independent and have probability distribution and characteristic function independent of \( k \):

\[
\varphi(u, \nu) = \left\langle \exp\left[ i[u\xi_k + \nu\zeta_k] \right] \right\rangle
\]  

(23)
We assume that all $\{r_k', r_k''\}$ are statistically independent and have the same probability density function $W(r_k', r_k'')$. Let us consider the characteristic functional for $X(r; \beta)$ given by (2). Substituting (19) into (2) we obtain

$$
\mathcal{F}(x, y) = \exp \left[ i \sum_{k=0}^{N} \xi_k \int g(r, r_k') x(r) \, dr + \sum_{k=0}^{N} \zeta_k \int h(r, r_k'') y(r) \, dr \right] 
$$

(24)

We are able to average over $\xi, \zeta$, and using (23) we obtain

$$
\mathcal{F}(x, y) = \prod_{k=0}^{N} \varphi \left( \int g(r, r_k') x(r) \, dr, \int h(r, r_k'') y(r) \, dr \right) 
$$

(25)

Here, $N, M$ and $r_k', r_k''$ remain random. Because all $r_k', r_k''$ are statistically independent and have the same probability density function $W(r_k', r_k'')$, averaging over $r_k'$ and $r_k''$ results in

$$
\mathcal{F}(x, y) = \prod_{k=0}^{N} \varphi \left( \int g(r, r_k') x(r) \, dr, \int h(r, r_k'') y(r) \, dr \right) 
$$

(26)

All factors in this product differ only by the notation for the variable of integration and are equal to each other. Thus,

$$
\mathcal{F}(x, y) = \left( \prod_{k=0}^{N} \int W(r_k', r_k'') \varphi \left( \int g(r, r_k') x(r) \, dr, \int h(r, r_k'') y(r) \, dr \right) \, dr' \, dr'' \right)^N 
$$

(27)

To fulfill the averaging with respect to $N$, we can use (22) with the value $A$ determined by the expression

$$
A = \int_{r'} \int_{r''} \varphi \left( \int g(r, r') x(r) \, dr, \int h(r, r'') y(r) \, dr \right) \, dr' \, dr'' 
$$

Substituting we obtain

$$
\mathcal{F}(x, y) = \Psi \left( \int_{r'} \int_{r''} \varphi \left( \int g(r, r') x(r) \, dr, \int h(r, r'') y(r) \, dr \right) \, dr' \, dr'' \right) 
$$

(28)

This formula represents the characteristic functional for a random function of the type (19). Analogous generalizations for the $n-$dimensional case can be carried out leading to results of the form

$$
\mathcal{F}(x) = \Psi \left[ \int_{r_1} \cdots \int_{r_n} \varphi \left( \int g_1(r, r_1) x_1(r) \, dr, \cdots, \int g_n(r, r_n) x_n(r) \, dr \right) \, dr_1 \cdots dr_n \right] 
$$

with $x \in H \times \cdots \times H$.
3.9. Characteristic Functionals derived from Tatarskii Functional

As we saw in the previous section the Tatarskii characteristic functional represents a wide class of stochastic processes. In what follows we will study how the general representation can be reduced to characteristic functions corresponding to useful measures for applications. These measures will be used at Chapter 5/Section 4 for the representation of the responses of the stochastic dynamical systems. These specific representations are very useful for this aim since Gaussian measures are very restricted, while, Tatarskii characteristic functional is very impractical for calculations, since it involves many unknown functions. The analysis presented below will follow EGOROV, A.D. & SOBOLEVSKY, P.I. & YANOVICH, L.A. (Functional Integrals: Approximate Evaluation and Applications).

Consider the special form of Tatarksii characteristic functional with
\[
\Psi(A) = \exp \{ A \} \tag{1}
\]
and the function \( g(r, r') \) depends only from the first argument \( r \), i.e.
\[
g(r, r') = l(r) \tag{3}
\]
Based on the above assumptions the characteristic functional takes the form
\[
\mathcal{F}(x) = \exp \left[ \sum_{n=1}^{\infty} \frac{i^n \sigma_n}{n!} K_n \left( \langle l, x \rangle, \ldots, \langle l, x \rangle \right) \right] \tag{4}
\]
where \( K_n \left( \langle l_1, x \rangle, \ldots, \langle l_n, x \rangle \right) \) is a real symmetric continuous \( n \)-linear form on \( \mathcal{H} \times \ldots \times \mathcal{H} \) \( n \) times and \( \sigma_n \ (n = 1, 2, \ldots) \) are real and complex parameters. Some \( \sigma_n \) may be equal to zero here, and the numbers of summands may be finite. Let us now consider some more restricted classes of functionals of form (4).

Let \( \mathcal{H} \) be a set with a finite measure \( \mathcal{Q} \) defined on it, \( a \in \mathcal{H} \), \( \rho(u) \) is a mapping from \( \mathcal{H} \) into \( \mathcal{H} \) and parameters \( \sigma_n \) are real. Then we may define
\[
\mathcal{F}(x) = \exp \left[ \sum_{n=1}^{\infty} \frac{i^n \sigma_n}{n!} \int \langle \rho(u), x \rangle^n \mathcal{Q}(du) \right] \tag{5}
\]
A possible example of such functional is
\[
\mathcal{F}(x) = \exp \left[ \sum_{n=1}^{\infty} \frac{i^n \sigma_n}{n!} \int g \left( \langle \rho(u), x \rangle \right) \mathcal{Q}(du) \right] \tag{6}
\]
where \( g(z) \) is a function which is analytic in the vicinity of zero ((6) is reduced to (5), if we set \( \sigma_n = \frac{d^n}{i^n dz^n} g(z) \mid_{z=0}, \sigma_0 = \sigma_1 = 0 \)). The following two parameter family of functionals
\[ \mathcal{F}(x) = \exp \left\{ i \langle a, x \rangle + \frac{1}{\alpha - \beta} \int_{\mathcal{H}} \frac{1}{\alpha} \left( e^{i\alpha \langle \rho(u), x \rangle} - 1 \right) - \frac{1}{\beta} \left( e^{i\beta \langle \rho(u), x \rangle} - 1 \right) \mathcal{Q}(du) \right\} \] 

(7)

where \( \alpha, \beta \) are real parameters, \( \alpha \geq \beta \), is the special case of (5). If \( \beta \to 0 \), we obtain the functional

\[ \mathcal{F}(x) = \exp \left\{ i \langle a, x \rangle + \frac{1}{\alpha^2} \int_{\mathcal{H}} \left( e^{i\alpha \langle \rho(u), x \rangle} - 1 - i\alpha \langle \rho(u), x \rangle \right) \mathcal{Q}(du) \right\} \]

(8)

which corresponds to Poisson measure. If \( \alpha \to 0 \) and \( \beta \to 0 \), then we obtain the functional

\[ \mathcal{F}(x) = \exp \left\{ i \langle a, x \rangle - \frac{1}{2} \int_{\mathcal{H}} \langle \rho(u), x \rangle^2 \mathcal{Q}(du) \right\} \]

(9)

which corresponds to Gaussian measure. As we see at previous sections, the characteristic functional of a Gaussian measure has the form

\[ \mathcal{F}(x) = \exp \left\{ i \langle m, x \rangle - \frac{1}{2} K(x, x) \right\} \]

(10)

where \( K(x, x) \) is an positive definite bilinear one. One more example of functionals of type (6)

\[ \mathcal{F}(x) = \exp \left\{ i \langle a, x \rangle - \int_{\mathcal{H}} \left[ \ln \left( 1 - i\sigma \langle \rho(u), x \rangle \right) + i\sigma \langle \rho(u), x \rangle \right] \mathcal{Q}(du) \right\} \]

(11)

We would like to mention separately the class when \( \mathcal{H} \) is some set of functions \( x(t) \) on segment \( [0, T] \subset \mathbb{R} \). Then

\[ \langle a, x \rangle = \int_0^T a(t) x(t) dt \ , \quad \langle a, \rho(u) \rangle = \int_0^T \rho(u) a(t) dt \]

(12)

where the functions \( a(t), \rho(u) \) may also be the generalized ones. In this case a more special form of the Poisson measure is for

\[ \alpha = 1, \]

\[ \rho(u) = \begin{cases} 1, & \text{if } t \in [u, T] \\ 0, & \text{otherwise} \end{cases} \]

and

\[ \mathcal{H} = [0, T] \text{ with } \mathcal{Q}(du) = \lambda \cdot du \ , \quad \lambda \in \mathbb{C} \]

Then

\[ \int_{\mathcal{H}} \langle \rho(u), x \rangle \mathcal{Q}(du) = \int_0^T \int_u^T x(t) dt du = \langle a^*, x \rangle \]

Setting \( a = a^* \) we will get

\[ \mathcal{F}(x) = \exp \left\{ \lambda \int_T^x e^{i\alpha \langle \rho(u), x \rangle} - 1 \right\} \]

(13)
which is a simpler version of the general Poisson characteristic functional of equation (8).

Another simplification, for the case when \( \mathcal{H} \) is some set of functions \( x(t) \) on segment \([0,T] \subset \mathbb{R}\) can be made to equation (11) by choosing again as

\[
\rho_{\iota}(u) = \begin{cases} 
1, & \text{if } t \in [u,T] \\
0, & \text{otherwise}
\end{cases}
\]

and

\[ \mathcal{H} = [0,T] \text{ with } \mathcal{Q}(du) = du \, . \]

Choosing \( a \) in a suitable way we can get the Gamma characteristic functional for the infinite dimensional Gamma distribution

\[
\mathcal{F}(x) = \exp \left[ \int_{\mathcal{H}} \ln \left( 1 - i \sigma \langle \rho_{\iota}(u), x \rangle \right) \right] du
\]

Finally, for the case of \( \mathcal{H} \) being set of functions \( x(t) \) on segment \([0,T] \subset \mathbb{R}\) functional (6) may be considered as the characteristic functional of a random process \( X(t) \) which may represented as follows,

\[
X(t) = \int_{\mathcal{H}} \rho_{\iota}(u) \mathcal{H}(du) + a(t) \, , \quad t \in [0,T]
\]

where \( \mathcal{H} \) is a stochastic orthogonal measure,

\[
E^{\mathcal{H}} \left[ \exp[i\lambda \mathcal{H}(A)] \right] = \exp \left[ g(\lambda) \mathcal{Q}(A) \right] \, , \quad A \subset \mathcal{H}
\]

\[
E^{\mathcal{H}} \left[ \mathcal{H}(A)^2 \right] = \sigma_{\mathcal{H}}^2 \mathcal{Q}(A) \, , \quad A \subset \mathcal{H}
\]

Another interesting class of functionals is of form (6) with

\[
g(z) = iaz - \frac{b}{2}z^2 + \int_{\mathbb{R}} \left( e^{i\lambda z} - 1 - \frac{iz\lambda}{1 + \lambda^2} \right) \pi(d\lambda)
\]

where \( a \in \mathbb{R} \, , \, b \geq 0 \, , \, \pi(d\lambda) = dM(\lambda) \) for \( \lambda < 0 \) and \( \pi(d\lambda) = dN(\lambda) \) for \( \lambda > 0 \); \( M(\lambda), N(\lambda) \) satisfy the following conditions

1. \( M(\lambda) \) and \( N(\lambda) \) are nondecreasing functions on \((-\infty,0)\) and \((0,\infty)\), respectively,
2. \( M(-\infty) = N(\infty) = 0 \, , \)
3. \( \int_{-\varepsilon}^{0} \lambda^2 dM(\lambda) < \infty \, , \, \int_{0}^{\varepsilon} \lambda^2 dN(\lambda) < \infty \) for any \( \varepsilon > 0 \).

If \( \mathcal{H} = [0,T] \, , \, \mathcal{Q}(du) = du \, , \) then bearing in mind (12), we obtain
\[ F(x) = \exp \left\{ \int_0^b \int_0^b \rho_s(u) x(s) \, ds \, du - \frac{b}{2} \int_0^b \int_0^b B(t,s) x(t) x(s) \, ds \, dt + \right. \]
\[ + \int_0^b \int_0^b \exp \left\{ i \lambda \int_0^b \rho_s(u) x(s) \, ds \right\} - 1 - \frac{i \lambda}{1 + \lambda^2} \int_0^b \rho_s(u) x(s) \, ds \, d\pi(d\lambda) \right\} \tag{17} \]

where \( B(t,s) = \sigma_2 \int_0^b \rho_s(u) \rho_t(u) \, du \), \( \sigma_2 = b + \int_\mathbb{R} \lambda^2 \pi(d\lambda) \). This functional is the characteristic one of the random process
\[ X(t) = \int_0^b \rho_s(s) d\xi_s \tag{18} \]

where \( \xi_s \) is a homogeneous process with independent increments which satisfies the condition \( \xi_0 = 0 \).

Note that (17) for
\[ M(\lambda) = N(\lambda) = 0 \]
gives rise to functionals of type (9); it gives rise to functionals of type (8) for
\[ b = 0, \quad M(\lambda) = 0, \quad N(\lambda) = \begin{cases} -a & \text{if } \lambda \leq 1 \\ 0 & \text{otherwise} \end{cases} \]

and to functionals of type (11) for
\[ b = 0, \quad M(\lambda) = 0, \quad N(\lambda) = -\int_{-\lambda}^\infty \frac{1}{v} \exp \left\{ -\frac{v}{\alpha} \right\} dv, \quad \alpha > 0. \]

We shall mention still another characteristic functional which corresponds to the measure which we shall call the \textit{Abel measure},
\[ F(x) = \exp \left\{ 1 + \frac{1}{2} K(x,x) \right\}^{-1} \tag{19} \]

The characteristic functional of the \textit{Cauchy measure},
\[ F(x) = \exp \left\{ -\sum_{j=1}^{\infty} a_j \langle x, e_j \rangle \right\}, \tag{20} \]

where \( a_j \) are positive numbers, \( \sum_{j=1}^{\infty} a_j < \infty \), \( \{e_j\}, j = 1,2,\ldots \), is a basis in \( \mathcal{H} \), is an example of the characteristic functional which cannot be represented in form (4).

Characteristic functional of form (4) give examples of measures, when the number of summands in the exponent is finite, the last summand has the index \( n = 2p \), \( p > 1 \) and \( \sigma_{2p} = (-1)^{p+1} \sigma^{2p} \). In particular let us mention the case, when
\[ \mathcal{F}(x) = \exp \left\{ -\frac{\sigma^2}{(2\pi)!} \int_{\mathcal{H}} \left\{ \rho(u), x \right\}^{2p} \mathcal{Q}(du) \right\}, \]  

(21)

where the notation is the same as for (5). For the space of functions on \([0, T]\), (21) will assume the form,

\[ \mathcal{F}(x) = \exp \left\{ -\frac{\sigma^2}{(2\pi)!} \int_{\mathcal{H}} \left( \int_0^T \rho_t(u) x(t) dt \right)^{2p} \mathcal{Q}(du) \right\}, \]  

(22)

Finally, when there is a finite series in the exponent of (4) and the last summand enters with the imaginary term then we obtain pseudomeasures (see EGOROV, A.D. & SOBOLEVSKY, P.I. & YANOVIČ, L.A., Functional Integrals: Approximate Evaluation and Applications [page 5] for details), namely, the Feynman measures. For example,

\[ \mathcal{F}(x) = \exp \left\{ -\frac{i}{2} \int_{\mathcal{H}} \left\{ \rho(u), x \right\}^2 \mathcal{Q}(du) \right\}, \]  

(23)

in particular,

\[ \mathcal{F}(x) = \exp \left\{ -\frac{i}{2} \int_{\mathcal{H}} \left( \int_0^T \rho_t(u) x(t) dt \right)^2 \mathcal{Q}(du) \right\}. \]  

(24)
3.10. References

PFEFFER, W.F., 1977, Integrals and Measures. Marcel Dekker, INC.
PROHOROV, YU.V. & SAZONOV, V.V., 1961, Some Results Associated with Bochner’s Theorem. Theory of Probability and Its Applications, VI, 82-87.
Chapter 4

Stochastic Calculus – Principles and Results

In various problems of mathematical physics, many physical quantities are represented using stochastic processes. In the context of mathematical modelling we come naturally to concepts such as, continuity, derivatives and integrals of the functions that represent these physical quantities. It is therefore essential to define the above concepts for stochastic processes. In this direction three kinds of convergence have been developed, namely

- Mean Square Convergence
- Convergence in Probability
- Almost Sure Convergence

The definitions of the above are immediate generalizations of the convergence of sequences of random variables, given at Chapter 1/Section 1.3.5.

In this chapter we will present the basic principles and results of mean square calculus. Using this calculus we reduce the problem of differentiation or integration of a stochastic process to the differentiation or integration of the covariance function. In Section 4.1.1-4.1.4 we present some basic definitions and results concerning the classical notions of m.s.-convergence, continuity and differentiation of second-order stochastic processes. Sections 4.1.5 deals with the issue of m.s.-integration of a stochastic process with respect to another (independent) process and Section 4.1.6 for the case of integration over martingales with non-anticipative dependence.

Section 4.2 includes some fundamental results associated with analytical properties of sample functions, such as continuity, differentiation and integration. In what follows we assume the existence of a probability space \( (\Omega, \mathcal{B}(\Omega), \mathcal{P}) \). The stochastic process \( X: \Omega \to C(I) \) is defined on it, and the induced probability space is \( (C(I), \mathcal{B}(C(I)), \mathcal{P}_{C(I)}) \).
4.1. Process of 2nd order – Mean-Square Calculus

4.1.1. Preliminaries

**Definition 1.1:** A stochastic process \( X(t;\beta), \ t \in I \) is called a second-order or Hilbert stochastic process, iff

\[
E^\beta \left( X(t;\beta)^2 \right) < \infty
\]

We shall denote this space of stochastic processes as \( \mathcal{Y}(I) \). ■

(1)

The above definition can easily be extended for stochastic processes taking complex or vector values.

**Remark 1.2:** The Schwarz inequality (Chapter 1/Section 1.3.2) implies that a second order stochastic process always has a covariance function. ■

**Theorem 1.3 [Completeness property]:** For a sequence \( \{X_n\} \) of stochastic processes in \( \mathcal{Y}(I) \) to converge in \( \mathcal{Y}(I) \) to some limit, it is necessary and sufficient that the sequence \( \{X_n\} \) is fundamental. In other words the space \( \mathcal{Y}(I) \) is complete.

**Proof:** The necessity of the above theorem is general for metric spaces as a consequence of the triangle inequality.

Let us prove the sufficiency. Since

\[
\left| E^\beta \left( X_n^2 \right) - E^\beta \left( X_n^2 \right) \right| \leq E^\beta \left( [X_n - X_n]^2 \right)
\]

if a sequence is a fundamental sequence, the sequence \( \{X_n\} \) is bounded; thus \( E^\beta \left( X_n^2 \right) < \infty \).

Furthermore the fact that the sequence is fundamental in \( \mathcal{Y}(I) \) implies that it is fundamental in measure. Consequently (See GIKHMAN, I.I. & SKOROKHOD, A.V., 1996, Theory of Random Processes) the sequence \( \{X_n\} \) contains a subsequence \( \{X_{n_k}\} \) that converges almost uniformly in \( \mathcal{Y}(I) \). Using Fatou’s lemma, we obtain

\[
c^2 \geq \lim E^\beta \left( X_n^2 \right) \geq E^\beta \left( \lim X_n^2 \right) = E^\beta \left( X^2 \right)
\]

from which it follows that \( X \in \mathcal{Y}(I) \). Let \( \varepsilon \) denote an arbitrary positive number. Since the sequence \( \{X_n\} \) is fundamental, it follows that for all \( n \geq n_0 \),

\[
\varepsilon \geq \lim_{k \to \infty} E^\beta \left( [X_{n_k} - X_n]^2 \right) \geq E^\beta \left( \lim_{k \to \infty} [X_{n_k} - X_n]^2 \right) = E^\beta \left( [X - X_n]^2 \right)
\]

which proves the convergence of the sequence \( \{X_n\} \) to \( X \in \mathcal{Y}(I) \). ■

4.1.2 Mean-Square Convergence

**Definition 1.4:** Let \( X(t;\beta), \ t \in I \) be a second-order stochastic process. A process \( X(t;\beta) \) converges in the mean-square (m.s.) to a random variable \( A(\beta) \) as \( t \to t_0 \) iff
4.1. PROCESS OF 2ND ORDER – MEAN-SQUARE CALCULUS

\[
\lim_{t \to t_0} E^\beta \left( \left[ X(t; \beta) - A(\beta) \right]^2 \right) = 0
\]  

(2)

For the sake of convenience, the m.s. convergence will be denoted as

\[
\lim_{t \to t_0} E^\beta \left( \left[ X(t; \beta) - A(\beta) \right]^2 \right) = 0 \iff \text{l.i.m.} X(t; \beta) = A(\beta)
\]

(3)

**Remark 1.5:** Necessary and sufficient condition for the m.s.-convergence of the s.p. \( X(t; \beta), t \in T \), when \( t \to t_0 \), is the Cauchy criterion,

\[
\lim_{t_1 \to t_0} \lim_{t_2 \to t_0} E^\beta \left( \left[ X(t_1; \beta) - X(t_2; \beta) \right]^2 \right) = 0
\]

(4)

since the space of second order stochastic processes is a Hilbert space (Theorem 1.3).

**Theorem 1.6:** Let \( X(t; \beta), t \in I \) be a second-order stochastic process. A process \( X(t; \beta) \) converges in the mean-square to a random variable \( A(\beta) \) as \( t \to t_0 \) iff the covariance function \( R_{XX}(t_1, t_2) \) convergence to a finite limit as \( t_1 \to t_0, t_2 \to t_0 \).

**Proof:** Let \( X(t; \beta) \to^{m.s.} A(\beta) \), as \( t \to t_0 \). Then,

\[
E^\beta \left( X(t_1; \beta) X(t_2; \beta) - A(\beta)^2 \right) = E^\beta \left( \left[ X(t_1; \beta) - A(\beta) \right] \left[ X(t_2; \beta) - A(\beta) \right] \right)
\]

\[
+ E^\beta \left( \left[ X(t_1; \beta) - A(\beta) \right] A(\beta) + E^\beta \left( \left[ X(t_2; \beta) - A(\beta) \right] A(\beta) \right) \right)
\]

Using Schwartz inequality one obtains

\[
R_{XX}(t_1, t_2) = E^\beta \left( X(t_1; \beta) X(t_2; \beta) \right) \to E^\beta \left( A(\beta)^2 \right) \text{ as } t_1 \to t_0, t_2 \to t_0.
\]

Conversely, let \( R_{XX}(t_1, t_2) \to c \) as \( t_1 \to t_0, t_2 \to t_0 \), where \( c < \infty \). Then,

\[
E^\beta \left( \left[ X(t_1; \beta) - X(t_2; \beta) \right]^2 \right) = E^\beta \left( X(t_1; \beta)^2 \right) + E^\beta \left( X(t_2; \beta)^2 \right) - 2 E^\beta \left( X(t_1; \beta) X(t_2; \beta) \right) =
\]

\[
= R_{XX}(t_1, t_1) + R_{XX}(t_2, t_2) - 2 R_{XX}(t_1, t_2) \to c + c - 2c = 0
\]

as \( t_1 \to t_0, t_2 \to t_0 \). This relationship and the Cauchy m.s. criterion imply that \( X(t; \beta) \to^{m.s.} A(\beta) \), as \( t \to t_0 \).

**Proposition 1.7:** The most important properties of m.s.-convergence are

a) For deterministic functions the m.s.-convergence is equivalent with the simple convergence.

b) The operator l.i.m. is linear, i.e. if \( a \) and \( b \) are deterministic constants, \( X(t; \beta), Y(t; \beta), t \in T \) two second-order stochastic processes and the limits \( \text{l.i.m.} X(t; \beta) = X_0(\beta) \) and \( \text{l.i.m.} Y(t; \beta) = Y_0(\beta) \) exists, then

\[
\text{l.i.m.} \left\{ a X(t; \beta) + b Y(t; \beta) \right\} = a X_0(\beta) + b Y_0(\beta)
\]

(5)
c) If \( \lim_{t \to t_0} X(t; \beta) = X_0(\beta) \) exists, then the limit \( \lim_{t \to t_0} E^\beta \left( X(t; \beta) \right) \) exists and the following relation holds,

\[
E^\beta \left( \lim_{t \to t_0} X(t; \beta) \right) = \lim_{t \to t_0} E^\beta \left( X(t; \beta) \right)
\]

i.e. the operators \( E^\beta \) and \( \lim \) are permutable.

d) If \( X(t; \beta), Y(t; \beta), t \in T \) are two second-order stochastic processes and the limits \( \lim_{t \to t_0} X(t; \beta) = X_0(\beta) \) and \( \lim_{t \to t_0} Y(t; \beta) = Y_0(\beta) \) exists, then the limit

\[
\lim_{t \to t_0} E^\beta \left( X(t; \beta)Y(t; \beta) \right)
\]

exists and the following relation holds,

\[
E^\beta \left( \lim_{t \to t_0} X(t; \beta)Y(t; \beta) \right) = \lim_{t \to t_0} E^\beta \left( X(t; \beta)Y(t; \beta) \right)
\]

**Proof:** The above assertions are direct sequences of the definition of m.s.-convergence. For details we refer to ATHANASSOULIS, G.A., (Stochastic Modeling and Forecasting of Ship Systems. Lecture Notes NTUA).

### 4.1.3. Mean-Square Continuity

**Definition 1.8:** A second-order stochastic process \( X(t; \beta), t \in I, \) is continuous in a mean square sense (m.s.-continuous) at \( t \in I, \) iff

\[
\lim_{t \to t_0} E^\beta \left[ \left( X(t; \beta) - X(t_0; \beta) \right)^2 \right] = 0
\]

A s.p. \( X(t; \beta) \) is m.s.-continuous on \( I, \) if it is m.s.-continuous at every point \( t \in I. \)

**Theorem 1.9:** (m.s.-continuity criterion) A second-order s.p. \( X(t; \beta), t \in I, \) is m.s.-continuous at \( t \in I, \) iff the covariance function \( R_{XX}(t_1, t_2) \) is continuous at \( t_1 = t_2. \)

**Proof:** Let \( R_{XX}(t_1, t_2) \) be continuous at \( t_1 = t_2. \) Then

\[
E^\beta \left[ \left( X(t+h; \beta) - X(t; \beta) \right)^2 \right] = E^\beta \left[ \left( X(t+h; \beta) - X(t; \beta) \right) \cdot \left( X(t+h; \beta) - X(t; \beta) \right) \right] = R_{XX}(t+h, t+h) - R_{XX}(t+h, t) - R_{XX}(t, t+h) + R_{XX}(t, t) \to 0
\]

therefore \( X(t; \beta) \) is m.s. continuous at \( t. \)

Let us assume that \( X(t; \beta) \) is m.s.-continuous at \( t. \) Since

\[
R_{XX}(t+h, t+k) - R_{XX}(t, t) = E^\beta \left[ \left( X(t+h; \beta) - X(t; \beta) \right) \cdot \left( X(t+k; \beta) - X(t; \beta) \right) \right] = E^\beta \left[ X(t+h; \beta) \cdot X(t; \beta) - X(t+k; \beta) \cdot X(t; \beta) - X(t; \beta) \cdot X(t+k; \beta) + X(t; \beta) \cdot X(t+k; \beta) \right]
\]

use of the Schwartz inequality implies that \( R_{XX}(t_1, t_2) \) is continuous at \( t \in T. \)
**Theorem 1.10** : If the covariance function $R_{xx}(t_1,t_2)$ is continuous at every diagonal point $t_1 = t_2$ then it is continuous everywhere on $I \times I$.

**Proof** :

\[
\begin{align*}
|R_{xx}(t_1 + h, t_2 + k) - R_{xx}(t_1, t_2)| &= \left| E^\beta \left( X(t_1 + h; \beta)X(t_2 + k; \beta) - E^\beta(X(t_1; \beta)X(t_2; \beta)) \right) \right| \\
&\leq E^\beta \left( \left| X(t_1 + h; \beta) - X(t_1; \beta) \right| \cdot \left| X(t_2 + k; \beta) \right| \right) \\
&+ \left| E^\beta \left( \left[ X(t_2 + k; \beta) - X(t_2; \beta) \right] \cdot X(t_1; \beta) \right) \right|
\end{align*}
\]

Using the Schwartz inequality it is easily seen that the above expression tends to zero as $h \to 0$ and $k \to 0$.

**Remark 1.11** : It should be emphasized that m.s.-continuity of a s.p. $X(t; \beta), t \in I$ implies that it is continuous with respect to probability. Indeed the Tchebycheff inequality gives

\[
\mathbb{P} \left( \left| X(t_0 + h; \beta) - X(t_0; \beta) \right| > \varepsilon \right) \leq \frac{E^\beta \left( \left| X(t_0 + h; \beta) - X(t_0; \beta) \right|^2 \right)}{\varepsilon^2}
\]

The m.s.-continuity does not, however, imply continuity of the realizations of the process. As an example one can take the Poisson process. Its covariance function $R_{xx}(t_1, t_2) = \nu \min(t_1, t_2)$ is continuous for all $t_1, t_2 > 0$, but almost all realizations of this process have discontinuities over a finite interval of time (see Chapter 2/Section 2.2.1.c).

### 4.1.4. Mean-Square Differentiation

**Definition 1.12** : A second-order s.p. $X(t; \beta), t \in I$, has mean-square derivative $X'(t; \beta)$ at $t_0 \in I$, if there is a r.v. $X'(t_0; \beta)$ such that

\[
\lim_{h \to 0} E^\beta \left( \left[ \frac{X(t_0 + h; \beta) - X(t_0; \beta)}{h} \right]^2 \right) = X'(t_0; \beta)
\]

A s.p. $X(t; \beta)$ is m.s.-differentiable on $I$, if it is m.s.-continuous at every point $t \in I$. Higher order derivatives can be defined analogously.

**Theorem 1.13 [m.s.-Differentiation criterion]** : A second-order s.p. $X(t; \beta), t \in I$ is m.s.-differentiable at $t_0 \in I$, iff there exist the second-order derivative of the covariance function

\[
\frac{\partial^2}{\partial t_1 \partial t_2} R_{xx}(t_1, t_2).
\]

**Proof** : This theorem follows directly from Theorem 1.6, which asserts that condition (10) is satisfied iff
\[ E^\beta \left( Y_h \cdot Y_k \right) \xrightarrow{k \to 0 \atop h \to 0} \text{finite limit,} \]

where,

\[ Y_k = \frac{1}{h} \left[ X(t + h; \beta) - X(t; \beta) \right] \tag{11} \]

But

\[ E^\beta (Y_h \cdot Y_k) = E^\beta \left( \frac{1}{kh} \left[ X(t + h; \beta) - X(t; \beta) \right] \cdot \left[ X(t + k; \beta) - X(t; \beta) \right] \right) = \frac{1}{kh} \left( R_{xx}(t + h, t + k) - R_{xx}(t, t + k) - R_{xx}(t + h, t) + R_{xx}(t, t) \right) \]

Therefore, in order to ensure the existence if the m.s.-derivative \( X'(t; \beta) \) of the s.p. \( X(t; \beta) \) is necessary and sufficient that the above expression has finite limit, i.e. that the second-order derivative defined as

\[ \frac{\partial^2}{\partial t_1 \partial t_2} R_{xx}(t_1, t_2) = \lim_{k \to 0} \frac{1}{kh} \left( R_{xx}(t + h, t + k) - R_{xx}(t, t + k) - R_{xx}(t + h, t) + R_{xx}(t, t) \right) \]

exists and is finite at the point \((t, t)\). ■

**Preposition 1.14:** The m.s.-derivatives have the following properties:

a) m.s.-differentiability of \( X(t; \beta) \) at \( t_0 \in I \), implies m.s.-continuity of \( X(t; \beta) \) at \( t_0 \in I \).

b) If \( X(t; \beta) \), \( Y(t; \beta) \), \( t \in I \) are two second-order s.p. m.s.-differentiable, then the m.s.-derivative of \( aX(t; \beta) + bY(t; \beta) \) exists and

\[ \frac{d}{dt} \left( aX(t; \beta) + bY(t; \beta) \right) = a \frac{dX(t; \beta)}{dt} + b \frac{dY(t; \beta)}{dt} \tag{12} \]

where \( a \) and \( b \) are constants.

c) If \( X(t; \beta) \), \( t \in I \) is \( n \)-times m.s.-differentiable on \( I \), the means of these m.s.-derivatives of \( X(t; \beta) \) exist on \( I \) and moreover the following equation holds,

\[ E^\beta \left( \frac{d^n X(t; \beta)}{dt^n} \right) = \frac{d^n}{dt^n} E^\beta \left( X(t; \beta) \right) \tag{13} \]

**Proof:** a) is direct consequence of Theorems 1.9 and 1.13. b) comes as result of Preposition 1.7b and the definition of m.s.-derivative. c) can be proved using Preposition 1.7c and the definition of m.s.-derivative. ■

**Theorem 1.15:** If \( X(t; \beta), t \in I \) has m.s.-derivative for all \( t \in I \), then the partial derivatives

\[ \frac{\partial R_{xx}(t_1, t_2)}{\partial t_1}, \frac{\partial R_{xx}(t_1, t_2)}{\partial t_2}, \frac{\partial^2 R_{xx}(t_1, t_2)}{\partial t_1 \partial t_2} \]

exist and

\[ R_{xx}(t_1, t_2) = E^\beta \left( X'(t_1; \beta) X(t_2; \beta) \right) = \frac{\partial R_{xx}(t_1, t_2)}{\partial t_1} \tag{14} \]
4.1. PROCESS OF 2\textsuperscript{nd} ORDER – MEAN-SQUARE CALCULUS

\[ R_{XX}(t_1, t_2) = E^\beta (X(t_1 ; \beta) X'(t_2 ; \beta)) = \frac{\partial R_{XX}(t_1, t_2)}{\partial t_2} \] (15)

\[ R_{XX'}(t_1, t_2) = E^\beta (X'(t_1 ; \beta) X'(t_2 ; \beta)) = \frac{\partial R_{XX}(t_1, t_2)}{\partial t_1 \partial t_2} \] (16)

**Proof:** If \( X(t; \beta), t \in I \) has m.s.-derivative \( X'(t; \beta) \) at \( t \in I \), then

\[
\frac{1}{h}(X(t + h; \beta) - X(t; \beta)) \xrightarrow{m.s.} X'(t; \beta) \quad \text{as} \quad h \to 0
\]

Therefore,

\[
\lim_{h \to 0} \frac{R_{XX}(t_1 + h, t_2) - R_{XX}(t_1, t_2)}{h} = \lim_{h \to 0} E^\beta \left( \frac{X(t_1 + h; \beta) - X(t_1; \beta)}{h} \right) X(t_2; \beta) = E^\beta \left( X'(t_1; \beta) X(t_2; \beta) \right)
\]

The above implies that the derivative \( \frac{\partial R_{XX}(t_1, t_2)}{\partial t_1} \) exists and eq.(14) holds.

Similarly, it can be shown that the derivative \( \frac{\partial R_{XX}(t_1, t_2)}{\partial t_2} \) exists and the formula (15) holds.

Finally,

\[
\lim_{k \to 0} \frac{1}{k} \left( \frac{\partial R_{XX}(t_1, t_2 + k)}{\partial t_1} - \frac{\partial R_{XX}(t_1, t_2)}{\partial t_1} \right) = \lim_{k \to 0} E^\beta \left( X'(t_1; \beta) \right) \frac{X(t_2 + h; \beta) - X(t_2; \beta)}{k} = E^\beta \left( X'(t_1; \beta) X'(t_2; \beta) \right)
\]

shows that formula (16) is valid. ■

**Remark 1.16:** Analogous results for the derivatives of higher order can be easily derived. The covariance function of the derivatives of order \( k \) and \( l \) of a given process is expressed by the general formulas

\[ R_{X^{(k)}X^{(l)}}(t_1, t_2) = E^\beta (X^{(k)}(t_1 ; \beta) X^{(l)}(t_2 ; \beta)) = \frac{\partial^{k+l}}{\partial t_1^k \partial t_2^l} R_{XX}(t_1, t_2) \] (17)

if appropriate m.s.-derivatives exist. ■

Let us assume that \( R_{XX}(t_1, t_2) \) has an infinite number of derivatives on \( I \times I \). Let

\[ X_n(t; \beta) = X(0; \beta) + \frac{t}{1!} X'(0; \beta) + \ldots + \frac{t^n}{n!} X^{(n)}(0; \beta) \].

Simple transformation of the expression

\[ E^\beta \left( \left[ X(t; \beta) - X_n(t; \beta) \right]^2 \right) \]

leads to the following

**Theorem 1.17:** A second-order s.p. \( X(t; \beta), t \in I \) is m.s.-analytic iff its covariance function \( R_{XX}(t_1, t_2) \) is analytic at each diagonal point \( (t, t) \in I \times I \). ■

**Remark 1.18:** As in the case of m.s.-continuity it should be emphasized that m.s.-differentiation of a s.p. \( X(t; \beta), t \in I \) does not implies differentiation of the realizations of the
stochastic process. For the conditions required for differentiation of the realizations see Section 2.

**Remark 1.19:** It follows directly from Theorem 1.13 that, a weakly-stationary s.p. \( X(t; \beta), t \in I \) is differentiable at \( t_0 \in I \) iff its covariance function \( R_{XX}(\tau), \tau = t_1 - t_2 \), has the derivative of a second order at \( \tau = 0 \). In this case formula (17) reduce to the following one

\[
R_{X(t; \beta)}^{(k)}(t_1, t_2) = E^\beta \left( X^{(k)}(t_1; \beta)X^{(l)}(t_2; \beta) \right) = (-1)^k \frac{d^{k+l}}{dt^{k+l}} R_{XX}(\tau)
\]  

(18)

4.1.5. M.S. Integration I – Integration over an Independent Stochastic Process

Let the second-order stochastic processes \( X(t; \beta), t \in I \) and \( \Psi(t; \beta), t \in I \). In what follows we will consider, for various cases of the above s.p.’s, the meaning of the integral

\[
I(\beta) = \int_I X(t; \beta) d\Psi(t; \beta)
\]  

(19)

One very important case, that gives us many of the known integrals, of practical importance is when the s.p.’s \( X(t; \beta), t \in I \) and \( \Psi(t; \beta), t \in I \), are mutually independent, i.e.

\[
E^\beta \left( X(t; \beta) \cdot \Psi(t; \beta) \right) = E^\beta \left( X(t; \beta) \right) \cdot E^\beta \left( \Psi(t; \beta) \right)
\]

For this case we have the following

**Theorem 1.20:** Let \( X(t; \beta), t \in I \) and \( \Psi(t; \beta), t \in I \) be orthogonal second-order stochastic processes. Then the s.p. \( X(t; \beta) \) is integrable with respect to \( \Psi(t; \beta) \) iff

1. \( R_{XX}(t_1, t_2) \) is integrable on \( I \times I \)
2. \( R_{\Psi\Psi}(t_1, t_2) \) has bounded variation on \( I \times I \)

In this case we have,

\[
E^\beta \left( \left[ \int_I X(t; \beta) d\Psi(t; \beta) \right]^2 \right) = \int_I \int_I R_{XX}(t_1, t_2) d^2R_{\Psi\Psi}(t_1, t_2)
\]

(20)

**Proof:** Let \([a, b] \subset I\) be a finite interval and let the points \( t_0, t_1, \ldots, t_n \) define a partition of \([a, b]\) such that

\[
a = t_0 < t_1 < \ldots < t_n = b
\]

Let \( \xi_i \in [t_i, t_{i+1}] \). We form the random variable

\[
I_n(\beta) = \sum_{i=1}^{n-1} X(\xi_i; \beta)[\Psi(t_{i+1}; \beta) - \Psi(t_i; \beta)](21)
\]

By virtue of Theorem 1.6 \( I_n(\beta) \overset{m.s.}{\rightarrow} I(\beta) \) iff \( E^\beta(\left[I_n(\beta)I_m(\beta)\right]) \overset{m.s.}{\rightarrow} E^\beta\left(\left[I(\beta)\right]^2\right) \). Thus we have from eq. (21),
4.1. PROCESS OF 2nd ORDER – MEAN-SQUARE CALCULUS

\[
E^{\beta} \left( I_n (\beta) I_m (\beta) \right) = \\
= \sum_{j=1}^{m-1} \sum_{i=1}^{n-1} E^{\beta} \left( X(\xi^n_j;\beta) X(\xi^m_i;\beta) \left[ \Psi (t^n_{j+1};\beta) - \Psi (t^n_j;\beta) \right] \left[ \Psi (t^m_{i+1};\beta) - \Psi (t^m_i;\beta) \right] \right)
\]

(22)

Since \( X(t;\beta) \) and \( \Psi(t;\beta) \) are orthogonal we will have

\[
E^{\beta} \left( I_n (\beta) I_m (\beta) \right) = \\
= \sum_{j=1}^{m-1} \sum_{i=1}^{n-1} R_{XX} (\xi^n_j, \xi^m_i) \left[ R_{\psi\psi} (t^n_j, t^n_{j+1}) - R_{\psi\psi} (t^n_j, t^n_i) - R_{\psi\psi} (t^m_{i+1}, t^m_i) + R_{\psi\psi} (t^m_i, t^n_i) \right]
\]

(23)

It can be shown that the above sum converges iff the conditions mentioned above holds.

Some basic properties of the integral of a s.p. over an independent s.p. are as follows,

**Theorem 1.21**: Let \( X_i(t;\beta), t \in I \) and \( \Psi_i(t;\beta), t \in I \) be orthogonal second-order stochastic processes with the assumptions of Theorem 1.20. Then

a) \[
\int_I \left( a_1 X_1(t;\beta) + a_2 X_2(t;\beta) \right) d\Psi(t;\beta) = a_1 \int_I X_1(t;\beta) d\Psi(t;\beta) + a_2 \int_I X_2(t;\beta) d\Psi(t;\beta)
\]

b) \[
\int_I X(t;\beta) d \left( a_1 \Psi_1(t;\beta) + a_2 \Psi_2(t;\beta) \right) = a_1 \int_I X(t;\beta) d\Psi_1(t;\beta) + a_2 \int_I X(t;\beta) d\Psi_2(t;\beta)
\]

c) \[
E^{\beta} \left( \int_I X(t;\beta) d\Psi(t;\beta) \right) = \int_I m_X(t) dm_{\Psi}(t)
\]

d) Moreover if \( X(t;\beta) \) is differentiable with derivative \( X'(t;\beta) \), the following integration by parts formula holds

\[
\int_I X(t;\beta) d\Psi(t;\beta) = X(t;\beta) \Psi(t;\beta) \bigg|_I - \int_I X'(t;\beta) \Psi(t;\beta) dt
\]

**Proof**: Assertions a), b) and c) follow directly, by considering the sequence of eq. (21). For the proof of d) we refer to SOBCZYK, K., (Stochastic Differential Equations. p.111) ■

We will now study specific integrals that come from different cases of orthogonal s.p.’s \( X(t;\beta), t \in I \) and \( \Psi(t;\beta), t \in I \).

**I. Mean Square Integral of the s.p.** \( X(t;\beta) \) with respect to the function \( \psi \in C^0(I, \mathbb{R}) \)

By setting \( \Psi(t;\beta) \equiv \psi(t), t \in I \) where \( \psi \in C^0(I, \mathbb{R}) \) the integral (19) becomes

\[
I(\beta) = \int_I X(t;\beta) d\psi(t)
\]

(24)

By virtue of Theorem 1.20 the above integral exists iff

- \( R_{XX}(t_1, t_2) \) is integrable on \( I \times I \)
- \( \psi(t) \) has bounded variation on \( I \)

In this case eq. (20) takes the form,
\[ E^\beta \left( \left[ \int_t X(t;\beta) d\psi(t) \right]^2 \right) = \int_{t_1}^{t_2} R_{XX}(t_1, t_2) d\psi(t_1) d\psi(t_2) \]  \hspace{1cm} (25)

II. Riemann-Stieltjes Integral of the function \( f \in C^0(I, \mathbb{R}) \) with respect to the s.p. \( \Psi(t;\beta) \)

By setting \( X(t;\beta) \equiv f(t), \ t \in I \) where \( f \in C^0(I, \mathbb{R}) \) the integral (19) becomes

\[ I(\beta) = \int_I f(t) d\Psi(t;\beta) \]  \hspace{1cm} (26)

By virtue of Theorem 1.20 the above integral exists iff

- \( f(t) \) is integrable on \( I \)
- \( R_{\psi\psi}(t_1, t_2) \) has bounded variation on \( I \times I \)

In this case eq. (20) takes the form,

\[ E^\beta \left( \left[ \int_t f(t) d\Psi(t;\beta) \right]^2 \right) = \int_{t_1}^{t_2} f(t_1) f(t_2) d^2 R_{\psi\psi}(t_1, t_2) \]  \hspace{1cm} (27)

III. Riemann-Stieltjes Integral of the s.p. \( X(t;\beta) \) with respect to the martingale \( \Psi(t;\beta) \) (Orthogonal Increment Process).

We will study the form of the quantity \( \Delta^2_{h,w} R_{\psi\psi}(t_1, t_2) \) as \( h, w \to 0 \). We have

\[ \Delta^2_{h,w} R_{\psi\psi}(t_1, t_2) = E^\beta \left( \left[ \Psi(t_1 + h;\beta) - \Psi(t_1;\beta) \right] \cdot \left[ \Psi(t_2 + w;\beta) - \Psi(t_2;\beta) \right] \right) \]

For every \( t_1 \neq t_2 \) we can find appropriately small \( h, w \), such that

\[ E^\beta \left( \left[ \Psi(t_1 + h;\beta) - \Psi(t_1;\beta) \right] \cdot \left[ \Psi(t_2 + w;\beta) - \Psi(t_2;\beta) \right] \right) = 0 \]

On the other hand for \( t_1 = t_2 \) we will have

\[ \Delta^2_{h,w} R_{\psi\psi}(t, t) = E^\beta \left( \left[ \Psi(t + h;\beta) - \Psi(t;\beta) \right] \cdot \left[ \Psi(t + w;\beta) - \Psi(t;\beta) \right] \right) = \]

\[ = E^\beta \left( \left[ \Psi(t + \min(h,w);\beta) - \Psi(t;\beta) \right]^2 \right) = \]

\[ = F(t + \min(h,w)) - F(t) = \Delta_{\min(h,w)} F(t) \]

where the function is defined at Chapter 2/Section 2.2.1.c. So, the conditions for the existence of the integral are fulfilled iff

- \( R_{xx}(t,t) \) is integrable on \( I \)
- \( F(t) \) has bounded variation on \( I \)

In this case eq. (20) takes the form,
4.1. PROCESS OF $2^{nd}$ ORDER – MEAN-SQUARE CALCULUS

$$E^\beta \left( \int I X(t;\beta)d\Psi(t;\beta) \right)^2 = \int I R_{XX}(t,t)dF(t)$$  \hspace{1cm} (28)

4.1.6. M.S. Integration II–Integration over Martingales with non-anticipative dependence

Another important case of stochastic integral is when the stochastic process is being integrated over martingales. One special case with various applications in Mechanics is the Ito Integral which, as we shall see, is the integration of a s.p. over a Wiener Measure. The essential feature of the construction of this integral is that the dependence of the stochastic process on the martingale should be non-anticipative; the stochastic process can depend at most, on the present and past values of the martingale process.

Let $\Psi(t;\beta)$ be a second order, martingale, $\mathcal{B}_t(\Omega)$-measurable. For every $t,s \in I$ we define the stochastic process $\eta(t,s;\beta) = \Psi(t;\beta) - \Psi(s;\beta)$. Using Prepositions 2.39-2.40/Chapter 2, we clearly have

- a) $E^\beta(\eta(t,s;\beta)) = 0$, for every $t,s \in I$
- b) $E^\beta(\eta(t,s;\beta)^2) = F(t) - F(s) < \infty$, for every $s < t$ with $t,s \in I$
- c) $E^\beta(\eta(t,s;\beta) | \mathcal{B}_t(\Omega)) = 0$, for every $t,s, \tau \in I$ with $\tau < \min\{t,s\}$.

Let as denote by $H_2(I)$ the class of stochastic processes $X(t;\beta), t \in I$ satisfying the following conditions,

- a) $X(t;\beta)$ is $\mathcal{B}_t(\Omega)$-measurable,
- b) the integral $\int I [X(t;\beta)]^2 dt$ is finite with probability one.

Remark 1.22: The last two conditions and the Preposition 2.40d/Chapter 2, implies that for every $\tau \in I$ the process $X(\tau;\beta)$ is independent of the increments $\eta(t,s;\beta)$ for all $t,s \in I$ such that $\tau < \min\{t,s\}$.

Theorem 1.23: Let $X(\bullet;\bullet) \in H_2(I)$. Then there is a sequence of simple random functions $\tilde{X}_n(\bullet;\bullet) \in H_2(I)$ such that

$$\lim_{n \to \infty} \int I [\tilde{X}_n(t;\beta) - X(t;\beta)]^2 dt = 0$$ \hspace{1cm} (29)

Proof: We will now introduce the stochastic integral over a martingale process in two stages. First we define it for a simple function. Let $\tilde{X}(\bullet;\bullet) \in H_2(I)$ be a simple function, i.e. $\tilde{X}(t;\beta) = \tilde{X}(t_i;\beta)$ for $t \in (t_i,t_{i+1}]$ where $\{t_i\}_{i=1}^n$ is a partition of $I$. In this case the above integral is defined as follows:

$$\int I \tilde{X}(t;\beta)d\Psi(t;\beta) = \sum_{i=1}^{n-1} \tilde{X}(t_i;\beta)[\Psi(t_{i+1};\beta) - \Psi(t_i;\beta)]$$ \hspace{1cm} (30)
Since for every function \( X(\, \cdot \, \cdot) \in H_2(I) \) there is a convergent sequence of simple functions \( \tilde{X}_n(\, \cdot \, \cdot) \in H_2(I) \) we form the sequence

\[
I_n = \int_I \tilde{X}_n(t; \beta) d\Psi(t; \beta) = \sum_{i=1}^{n-1} \tilde{X}_n(t_i; \beta) \left[ \Psi(t_{i+1}; \beta) - \Psi(t_i; \beta) \right]
\]  

(31)

**Definition 1.24** : Let the stochastic process \( X(\, \cdot \, \cdot) \in H_2(I) \) and \( \tilde{X}_n(\, \cdot \, \cdot) \in H_2(I) \) a sequence of simple functions. Let also \( I_n \) the sequence of random variables of eq. (31). If

\[
\lim_{n \to \infty} I_n = I_M
\]

exists, the random variable \( I_M \) is called the *integral of the stochastic process \( X(\, \cdot \, \cdot) \in H_2(I) \) over the martingale \( \Psi(t; \beta) \). ■

**Remark 1.25** : We note that, for the above integral we use the values of the stochastic process at \( t_i \). The value of the m.s. Integral of a stochastic process over a martingale process depends on the choice of the intermediate points \( \xi_i \). To illustrate this dependence let consider the two integrals on the interval \( I = [a, b] \)

\[
I_{M,0} = \lim_{n \to \infty} \sum_{i=1}^{n-1} \Psi(t_i; \beta) \left[ \Psi(t_{i+1}; \beta) - \Psi(t_i; \beta) \right]
\]

\[
I_{M,1} = \lim_{n \to \infty} \sum_{i=1}^{n-1} \Psi(t_{i+1}; \beta) \left[ \Psi(t_{i+1}; \beta) - \Psi(t_i; \beta) \right]
\]

Getting the difference of the two above integrals, we have

\[
I_{M,1} - I_{M,0} = \lim_{n \to \infty} \sum_{i=1}^{n-1} \left[ \Psi(t_{i+1}; \beta) - \Psi(t_i; \beta) \right]^2 \xrightarrow{\text{Levy Property}} F(b) - F(a)
\]  ■

**Remark 1.26** : An alternative approach is also been used by some authors based on the convergence of the above limits with respect to probability instead of mean-square calculus. (SOIZE, C., The Fokker-Planck Equation and its Explicit Steady State Solutions). ■

**Theorem 1.27** : Let \( X(\, \cdot \, \cdot) \in H_2(I), \ I = [a, b] \) and \( \Psi(t; \beta) \) a second-order, \( \mathcal{B}(\Omega) \)-measurable, martingale with continuous characteristic \( F(t) \). Then the integral \( I_M \) as defined above exists and is unique, in the mean-square sense. Moreover

\[
E^{\beta} \left[ \left( \int_I X(t; \beta) d\Psi(t; \beta) \right)^2 \right] = \int_I R(\chi, t) dF(t)
\]

**Proof** : By virtue of Theorem 1.6 \( I_n(\beta) \xrightarrow{m.s.} I(\beta) \) iff

\[
E^{\beta} \left( I_n(\beta) I_m(\beta) \right) \xrightarrow{m.s.} E^{\beta} \left[ I(\beta)^2 \right].
\]

Thus we have from eq.(31),

\[
E^{\beta} \left( I_n(\beta) I_m(\beta) \right) = \sum_{j=1}^{m-1} \sum_{i=1}^{n-1} E^{\beta} \left[ X(t_i; \beta) X(t_j; \beta) \left[ \Psi(t_{i+1}; \beta) - \Psi(t_i; \beta) \right] \left[ \Psi(t_{j+1}; \beta) - \Psi(t_j; \beta) \right] \right]
\]

(33)
It is to be evaluated over the region \([a, b] \times [a, b]\) as indicated in the figure below. Each term in the sum is to be evaluated over a rectangle such as \(R_1\) or \(R_2\). In evaluating these terms, we distinguish two types of rectangles: those containing a segment of the diagonal line \(t^n = t^m\), such as \(R_1\) and those containing at most one point of the diagonal line, such as \(R_2\).

For the second type, a typical situation is one where \(t^m < t^n_{i+1} \leq t^n_i < t^n_{i+1}\). This implies that 
\[
[\Psi(t^n_i; \beta) - \Psi(t^n_{i+1}; \beta)]
\]
is independent of \(X(\xi^n_i; \beta), X(\xi^n_{i+1}; \beta)\) and 
\[
[\Psi(t^n_{i+1}; \beta) - \Psi(t^n_i; \beta)].
\]
Hence
\[
E^\beta\left(X(t^n_i; \beta)X(t^n_j; \beta)[\Psi(t^n_{i+1}; \beta) - \Psi(t^n_i; \beta)][\Psi(t^n_{i+1}; \beta) - \Psi(t^n_i; \beta)]\right) = 0
\]
Hence, the contributions of all the terms in eq. (33) over rectangles of the second type are zero.

Turning now to rectangles of the first type, consider a typical situation where \(t^m_i < t^n_i < t^n_{i+1} < t^n_{i+1}\). It is clear that we have
\[
E^\beta\left(X(t^n_i; \beta)X(t^n_j; \beta)[\Psi(t^n_{i+1}; \beta) - \Psi(t^n_i; \beta)][\Psi(t^n_{i+1}; \beta) - \Psi(t^n_i; \beta)]\right) =
\[
eq E^\beta\left(X(t^n_i; \beta)X(t^n_j; \beta)[\Psi(t^n_{i+1}; \beta) - \Psi(t^n_i; \beta)]^2\right) =
\[
eq E^\beta\left(X(t^n_i; \beta)X(t^n_j; \beta)\right)[F(t^n_{i+1}) - F(t^n_i)]
\]
Thus,
\[
E^\beta(I_n(\beta)I_m(\beta)) = \sum_{i=1}^{n-1} \sum_{j=1}^{m-1} E^\beta\left(X(t^n_i; \beta)X(t^n_j; \beta)\right)[F(t^n_{i+1}) - F(t^n_i)]
\]
So, we have the same summation as in case III of the above paragraph. Since \(F(t)\) is continuous, it has finite variation on \(I\). Moreover \(X(\cdot; \cdot) \in H_2(I)\), so \(R_{xx}(t, t)\) is integrable on \(I\).

Some basic properties of the integral of a s.p. over a martingale are as follows,

**Theorem 1.28**: Let \(X(\cdot; \cdot) \in H_2(I)\) and \(\Psi(t; \beta), t \in I, I = [a, b],\) is a second-order, \(\mathcal{B}_i(\Omega)\)-measurable, martingale with continuous characteristic \(F(t)\). Then
a)  \[ \int X(t;\beta)d\Psi(t;\beta) = 0 \]

b)  \[ \int (a_1 X_1(t;\beta) + a_2 X_2(t;\beta))d\Psi(t;\beta) = a_1 \int X_1(t;\beta)d\Psi(t;\beta) + a_2 \int X_2(t;\beta)d\Psi(t;\beta) \]

c)  \[ \int X(t;\beta)d\left(a_1 \Psi_1(t;\beta) + a_2 \Psi_2(t;\beta)\right) = a_1 \int X(t;\beta)d\Psi_1(t;\beta) + a_2 \int X(t;\beta)d\Psi_2(t;\beta) \]

d)  \[ \int X(t;\beta)d\Psi(t;\beta) = X(t;\beta)\Psi(t;\beta)|_0^\tau - \int_0^\tau X'(t;\beta)\Psi(t;\beta)dt \quad \text{if} \ X(\cdot;\cdot) \text{ is m.s. differentiable.} \]

e)  \[ \int_{\alpha} X_{[a,b]}(t)d\Psi(t;\beta) = \Psi(\tau;\beta) - \Psi(\alpha;\beta) \]

f) The integral \( I_M \) of \( X_i(t;\beta) \) over \( \Psi(t;\beta) \) is also a martingale with characteristic
\[ G(\tau) = \int_\alpha^\tau R_{XX}(t,t)dF(t) \]

g) The sample functions of the s.p. \( I_M(\tau;\beta) = \int_\alpha^\tau X(t;\beta)d\Psi(t;\beta) \) are continuous with probability one.

**Proof**: Assertions a), b) and c) follow directly, by considering eq. (31). The proof of d) and e) can be found at SOBCZYK, K., (Stochastic Differential Equations. p.111 and 108 respectively). Proof of f) and g) can be found at GIKHMAN, I.I. & SKOROKHOD, A.V., (The Theory of Stochastic Processes III. p. 54 and 57 respectively.). An alternative proof of the last assertion could be based on direct application of the Kolmogorov Theorem with the use of mean value theorem for integrals.

**Remark 1.29**: As we have mention at the beginning a very important case of stochastic integral, for applications, is the Ito Integral. For this case, the martingale is a standard Wiener process \( \mathcal{B}_t(\Omega) \)-measurable. The Ito Integral can be defined for every process \( X(\cdot;\cdot) \in H_2(I) \), since the characteristic of the standard Wiener process \( F(t) = t, \ t \in I \), has bounded variation on \( I \). Of course all the properties derived above for the general case of a martingale process, holds for the Ito Integral.

**4.1.6.a. A formula for changing variables (Ito Formula)**

In this section analogs and corollaries of the chain rule of differentiation of composite functions are studied in the case when differentiation is interpreted as the inverse operation of stochastic integration. The results obtained are of importance in what follows.

**Theorem 1.30 [Ito Formula]**: Let

1. \( \psi_0(\beta) \) be a second-order random variable,
2. \( a(t;\beta), \ t \in I \) be a second-order stochastic process that its correlation function \( R_{aa}(t_1,t_2) \) has bounded variation on \( I \times I \),
3. \( \mu(t; \beta), t \in I \) be a second-order, martingale, \( \mathcal{B}_I(\Omega) \)-measurable with continuous characteristic \( F(t), t \in I \)

4. \( \Psi(t; \beta) = \psi_0(\beta) + a(t; \beta) + \mu(t; \beta), t \in I \)

We also assume that \( f(x), x \in \mathbb{R} \) is a twice continuously differentiable function. Then

\[
    f(\Psi(t; \beta)) = f(\psi_0(\beta)) + \int_{I} \frac{df}{dx}(\Psi(s; \beta))d\Psi(s; \beta) + \frac{1}{2} \int_{I} \frac{d^2f}{dx^2}(\Psi(s; \beta))dF(s)
\]  

(34)


**Corollary 1.31**: If \( f(x, t), x \in \mathbb{R}, t \in I \), is continuously differentiable with respect to \( t \in I \) and twice continuously differentiable with respect to \( x \in \mathbb{R} \), then

\[
    f(\Psi(t; \beta), t) - f(\psi_0(\beta), t) = \tilde{a}(t; \beta) + \tilde{b}(t; \beta)
\]  

(35)

where

\[
    \tilde{a}(t; \beta) = \int_{I} \frac{\partial f}{\partial t}(\Psi(s; \beta), t)ds + \int_{I} \frac{\partial f}{\partial x}(\Psi(s; \beta), s)da(s; \beta) + \frac{1}{2} \int_{I} \frac{\partial^2 f}{\partial x^2}(\Psi(s; \beta), s)dF(s)
\]

and

\[
    \tilde{b}(t; \beta) = \int_{I} \frac{\partial f}{\partial x}(\Psi(s; \beta), s)d\mu(s; \beta)
\]


Concerning the special case of Ito Integral, i.e. integration over a standard Wiener process we have the following result

**Corollary 1.32**: Let \( W(t; \beta), t \in I \) be a standard Wiener process, \( \mathcal{B}_I(\Omega) \)-measurable and \( f(x), x \in \mathbb{R} \) is a twice continuously differentiable function. Then

\[
    f(W(t; \beta)) = f(0) + \int_{I} \frac{df}{dx}(W(s; \beta))dW(s; \beta) + \frac{1}{2} \int_{I} \frac{d^2f}{dx^2}(W(s; \beta))ds
\]  

(36)

**Proof**: Direct application of Theorem 1.30.

4.1.6.b Stochastic Differentials

**Definition 1.33**: Let

1. \( a(t; \beta), t \in I = [a, b] \) be a second-order stochastic process that its correlation function \( R_{aa}(t_1, t_2) \) has bounded variation on \( I \times I \),

2. \( \mu(t; \beta), t \in I \) be a second-order, martingale, \( \mathcal{B}_I(\Omega) \)-measurable with continuous characteristic \( F(t), t \in I \).
We say that a stochastic process \( X(t;\beta), \ t \in I \) possesses a stochastic differential (of a continuous type)

\[
d X(t;\beta) = (\varphi, da) + (\psi, d\mu) = \varphi(t;\beta)da(t;\beta) + \psi(t;\beta)d\mu(t;\beta), \ t \in I
\]

where, \( \psi(\bullet, \bullet) \in H_2(I) \) and the process \( \varphi(t;\beta) \) is progressively measurable with respect to \( \mathcal{B}_s(\Omega) \) whose realizations are (with probability one) bounded functions iff

\[
X(t;\beta) = X(0;\beta) + \int_a^t \varphi(s;\beta)da(s;\beta) + \int_a^t \psi(s;\beta)d\mu(s;\beta)
\]

for each and every \( t \in I \).

**Remark 1.34**: We recall that a stochastic process \( \varphi(t;\beta), \ t \in [a,b] \) is progressively measurable iff the restriction of the function \( \varphi(t;\beta) \) to the set \( [a,s] \times \Omega \) is \( \mathcal{B}_s(I) \times \mathcal{B}_s(\Omega) \), i.e. if

\[
\{(t,\beta) : \varphi(t;\beta) \in B, t \in [a,s]\} \in \mathcal{B}_s(I) \times \mathcal{B}_s(\Omega)
\]

for every \( B \in \mathcal{B}_s(D(T)) \) and \( s \in T \).

Clearly a process possessing a stochastic differential of continuous type has continuous modification. In what follows we shall consider just this kind of modification of the process \( \Psi(t;\beta), \ t \in I \). Theorem 1.30 can be formulated using the concept of a stochastic differential in the following manner.

If process \( \Psi(t;\beta), \ t \in I \) possesses stochastic differential \( d\Psi(t;\beta) = da(t;\beta) + d\mu(t;\beta) \) and the function \( f(x,t), x \in \mathbb{R}, t \in I \), is continuously differentiable with respect to \( t \in I \) and twice continuously differentiable with respect to \( x \) then the process \( X(t;\beta) = f(\Psi(t;\beta),t) \), \( t \in I \) also posses a stochastic differential and, moreover, for every \( t \in I \),

\[
d X(t;\beta) = \frac{\partial f}{\partial t}(\Psi(t;\beta),t)dt + \frac{\partial f}{\partial x}(\Psi(t;\beta),t)da(t;\beta) + \\
+ \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(\Psi(t;\beta),t)dF(t) + \frac{\partial f}{\partial x}(\Psi(t;\beta),t)d\mu(t;\beta)
\]

This implies

**Theorem 1.35**: If the stochastic processes \( \Psi_i(t;\beta), \ t \in I \) (we write \( \Psi_i \) for simplicity) possess stochastic differentials \( d\Psi_i(t;\beta) = da_i(t;\beta) + d\mu_i(t;\beta) \) then

\[
d(\Psi_1 + \Psi_2) = d\Psi_1 + d\Psi_2
\]

\[
d(\Psi_1, \Psi_2) = \Psi_1d\Psi_2 + \Psi_2d\Psi_1 + dF_{12}(t)
\]

\[
d\left(\frac{\Psi_1}{\Psi_2}\right) = \frac{\Psi_2d\Psi_1 - \Psi_1d\Psi_2}{\Psi_2^2} + \frac{\Psi_1dF_2(t) - \Psi_2dF_1(t)}{\Psi_2^3}
\]
\[ d(e^{\varphi u}) = e^{\varphi u} u d\Psi + \frac{1}{2}e^{\varphi u} \mu^2 dF(t) \]  

(44)

where \( F_{12}(t) \) is the joint characteristic of \( \mu_1(t;\beta) \) and \( \mu_2(t;\beta) \), i.e.

\[ F_{12}(t) = \frac{1}{2} [F_{S_1,2}(t) - F_1(t) - F_2(t)] \]

with \( F_{S_1,2}(t) \) denoting the characteristic of the sum \( \mu_1(t;\beta) + \mu_2(t;\beta) \). The formula for \( d\left(\frac{\Psi_1}{\Psi_2}\right) \) is applicable provided \( \Psi_2(t;\beta) \geq \delta > 0 \).

**Proof:** The formulas follow directly from Theorem 1.30 when applied successively to functions \( x_1 + x_2, x_1 \cdot x_2, x_1/x_2 \) and \( e^{au} \). ■
4.2. Analytical Properties of Sample Functions

Second order properties of a stochastic process formulated by use of m.s.-convergence in terms of correlation functions do not provide, in general, sufficient information about the behavior of the realizations of a process. This fact can be illustrated by a Poisson process which – as we have seen in Section 2.2.1.c/Chapter 2 – is m.s.-continuous, but its sample functions are discontinuous. It is therefore of interest to throw some light on the analysis and the results associated with sample function properties. As we mention at the beginning of the chapter we assume the existence of a probability space \((\Omega, \mathcal{B}(\Omega), \mathbb{P})\) and the stochastic process \(X: \Omega \rightarrow D(T)\) defined on it, with induced probability space \(\left(D(T), \mathcal{B}(D(T)), \mathbb{P}_{D(T)}\right)\).

4.2.1. Sample Function Integration

In the theory and applications we often deal with ordinary Lebesgue Integral of realization of stochastic process. We, now, recall the definition of a measurable stochastic process.

**Definition 2.1**: A stochastic process \(X(t; \beta), t \in I\) is a measurable stochastic process iff is measurable with respect to the product measure \(\mathcal{P}_{D(I)} \times \mathcal{P}_{I}\), where \(\mathcal{P}_{I}\) is the Lebesgue measure on \(I\).

The mathematical base for sample function integration gives a theorem following directly from the known Fubini Theorem.

**Theorem 2.2**: Let \(X(t; \beta), t \in I\) be a measurable stochastic process. Then almost all realizations of \(X(t; \beta)\) are measurable with respect to the Lebesgue measure \(\mathcal{P}_{I}\) on \(I\). If \(E^\beta(X(t; \beta))\) exists for \(t \in I\), then this average defines a \(\mathcal{P}_{I}\)-measurable function of \(t \in I\). If \(A\) is a measurable set on \(I\) and

\[
\int_A E^\beta(X(t; \beta))\,dt < \infty
\]

then almost all realizations of \(X(t; \beta)\) are Lebesgue integrable on \(A\) and

\[
E^\beta\left(\int_A X(t; \beta)\,dt\right) = \int_A E^\beta(X(t; \beta))\,dt
\]

**Proof**: According to the assumption \(X(t; \beta), t \in I\) is a measurable stochastic process. So, the Fubini Theorem implies that the ‘intersection’ \(X(t; \beta)\) is, for almost all \(\beta\), a measurable function of \(t \in I\). The Fubini Theorem also implies that if \(E^\beta(X(t; \beta))\) exists, then it is a measurable function of \(t\). Since the average of \(X(t; \beta)\) is the integral

\[
E^\beta(X(t; \beta)) = \int_\Omega X(t; \beta)\,d\mathbb{P}(d\beta)
\]
the assumption (1) asserts that the double integral (first – with respect to $\beta$, and then with respect to $t$) of $X(t;\beta)$ is finite. This double integral taken in the opposite order is also finite; therefore the integral

$$\int X(t;\beta) dt$$

is finite for almost all $\beta \in \Omega$. This means that almost all realizations of $X(t;\beta)$ are Lebesgue integrable on $I$. The Fubini theorem implies that the value of a double integral is independent of the order of the integration; so we obtain (2).

Let us note that the basics assumption of the above theorem requires the process $X(t;\beta)$ to be measurable. The following theorem gives a condition for an arbitrary stochastic process $X(t;\beta)$ to be equivalent to measurable process $\tilde{X}(t;\beta)$.

**Theorem 2.3**: If for almost all $t$ (with respect to the measure $\mathcal{P}(\mathcal{R})$) a stochastic process $X(t;\beta)$ is continuous in probability then a measurable and separable stochastic process $\tilde{X}(t;\beta)$ exists which is stochastically equivalent to $X(t;\beta)$.


**Remark 2.4**: The condition of this theorem (continuity of $X(t;\beta)$ in probability) is not very restrictive. It is satisfied, for instance, by all processes whose correlation function is continuous.

**Remark 2.5**: If $f(t)$ is a non-random Lebesgue measurable function defined on $I$ with values in $\mathbb{R}$ and $X(t;\beta)$ is a measurable stochastic process, then the sample function integral

$$\int_I f(t) X(t;\beta) dt$$

exists and by virtue of Fubini theorem

$$E^{\beta}\left(\int_I f(t) X(t;\beta) dt \int_I f(s) X(s;\beta) ds\right) = E^{\beta}\left(\int_I \int_I f(t)f(s) X(s;\beta) X(t;\beta) dt ds\right) = \int_I \int_I f(t)f(s) R_{xx}(t,s) dt ds$$

4.2.2. Sample Function Continuity

We shall now study the conditions that a s.p. should impose, so that its realizations are continuous. The following theorem gives a sufficient condition for a s.p. to have continuous realizations.

**Theorem 2.6**: Let $X(t;\beta)$, $t \in I$ be a separable stochastic process. If for all $t, t+h$ in the interval $I$

$$\mathcal{P}\left(\left|X(t+h;\beta) - X(t;\beta)\right| \geq g(h)\right) \leq q(h)$$

(5)
where \( g(h) \) and \( q(h) \) are even functions of \( h \), non-increasing as \( h \to 0 \) and such that
\[
\sum_{n=1}^{\infty} g(2^{-n}) < \infty \quad \text{and} \quad \sum_{n=1}^{\infty} 2^n q(2^{-n}) < \infty \quad (*)
\]
then almost all realizations of a process \( X(t; \beta) \) are continuous on \( I \).

**Proof:** See LOÈVE, M., (Probability Theory II).

The above theorem and the Tchebycheff Inequality:
\[
\mathbb{P}\left( |X(\beta)| \geq k \right) \leq \frac{E\left[ |X(\beta)|^n \right]}{k^n}
\]
implication the following corollary which constitutes a useful sufficient condition for the validity of the theorem.

**Corollary 2.7 [Kolmogorov Theorem]:** Let \( X(t; \beta), t \in I \) be a separable stochastic process such that for certain number \( p > 0 \) and for all \( t, t+h \) in the interval \( I \)
\[
E^\beta\left( |X(t+h; \beta) - X(t; \beta)|^p \right) \leq \rho(h)
\]
where
\[a) \quad \rho(h) = c|h|^{1+r}, \quad c > 0, \quad r > 0
\]
\[b) \quad \rho(h) = \frac{c|h|}{\log |h|^{1+r}}, \quad r > p
\]
then almost all realizations of a process \( X(t; \beta) \) are continuous on \( I \).

**Proof:** After application of the Tchebycheff Inequality, condition (5) will be satisfied if
\[
q(h) = \frac{\rho(h)}{[g(h)]^p}
\]
Taking in case a) \( g(h) = h^a, \quad 0 < a < \frac{r}{p} \) one obtains \( q(h) = c|h|^{1+r-a} \). Taking the case b)
\[
g(h) = \log |h|^{-\beta}, \quad 1 < \beta < \frac{r}{p}
\]
one gets \( q(h) = \frac{h}{\log |h|^{1+r-\beta p}} \). These expressions for \( q(h) \) agree with the requirements of the theorem.

**Example 2.8:** Simple calculations show that for the Wiener process
\[
\mathbb{P}_{D(I)}\left( \{W(\cdot) \in D(I) : W(t) - W(s) < x\} \right) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi|t-s|}} \exp \left( -\frac{1}{2} \frac{\xi^2}{|t-s|} \right) d\xi
\]
and
\[
E^\beta\left( |X(t+h; \beta) - X(t; \beta)|^4 \right) = \int_{-\infty}^{\infty} \frac{\xi^4}{\sqrt{2\pi|t-s|}} \exp \left( -\frac{1}{2} \frac{\xi^2}{|t-s|} \right) d\xi = 3(t-s)^2
\]
which means that,
4.2. ANALYTICAL PROPERTIES OF SAMPLE FUNCTIONS

\[ E^\delta \left( \left| X(t + h; \beta) - X(t; \beta) \right|^{4} \right) = 3h^2 \]

Therefore Kolmogorov condition a) is satisfied with constants \( c = 3, \ p = 4, \ r = 1 \). It proves that almost all realizations of the Wiener process are continuous on each finite interval. ■

If sample functions of a s.p. are discontinuous it is of interest to know whether the discontinuities are of the first kind only (finite jumps). The following theorem holds.

Theorem 2.9 [Chentsov] : Let \( X(t; \beta), \ t \in I \) be a separable stochastic process. If for \( t_1 < t_2 < t_3 \) with \( t_1, t_2, t_3 \in I \) and \( t_3 - t_1 = h \)

\[ E^\delta \left( \left| X(t_1; \beta) - X(t_2; \beta) \right| \left| X(t_2; \beta) - X(t_3; \beta) \right|^p \right) \leq c |h|^{p+r} \]

where \( p, r, c \) are positive constants, then almost all realizations of \( X(t; \beta) \) do not have discontinuities the second kind.


Remark 2.10 : The conditions of the above theorem are, in particular, satisfied for a process with independent increments and such that for any \( t, t + h \in I \),

\[ E^\delta \left( \left| X(t + h; \beta) - X(t; \beta) \right|^2 \right) \leq Ah \]

where \( A \) is a constant. It can be verified that Poisson process satisfies the condition.

4.2.3. Sample Function Differentiation

We shall now state here a general theorem giving a sufficient condition in order that almost all realizations of a process are continuously differentiable.

Theorem 2.11 : Let \( X(t; \beta), \ t \in I \) be a separable stochastic process. If for all \( t, t + h \) in the interval \( I \) the conditions of the Theorem 2.6 are satisfied and for all \( t - h, t, t + h \) in the interval \( I \)

\[ \mathbb{P} \left( \left| X(t + h; \beta) + X(t - h; \beta) - 2X(t; \beta) \right| \geq g_1(h) \right) \leq q_1(h) \]

where \( g_1(h) \) and \( q_1(h) \) are even functions of \( h \), non-increasing as \( h \to 0 \) and such that

\[ \sum_{n=1}^{\infty} 2^n g_1 \left( 2^{-n} \right) < \infty \quad \text{and} \quad \sum_{n=1}^{\infty} 2^n q_1 \left( 2^{-n} \right) < \infty \]

then almost all realizations of a process \( X(t; \beta) \) have continuous derivatives in \( I \).

Proof : See LOÈVE, M., (Probability Theory II). ■
4.2.4. Relation to Second-Order properties

For the sequel we assume that \( X(t; \beta), t \in I \) is a second-order stochastic process. We shall now study the mutual connection between analytical properties of the sample functions of a process and its second-order properties.

Let us consider first integration of the process \( X(t; \beta), t \in I \). In general, may happens, the realizations of a m.s. integrable s.p. to be non-integrable. But, even if the realizations are integrable nevertheless one should differentiate the integrals,

\[
I = (\text{m.s.}) \int I \ X(t; \beta) \ dt \tag{12}
\]

\[
I^* = \int I \ X(t; \beta) \ dt \tag{13}
\]

where \( I \) is a m.s. Riemann integral of \( X(t; \beta) \), that is it is a random variable defined as a m.s. limit of appropriate approximating integral sums and \( I^* \) is a function making a correspondence between the elementary event \( \beta \in \Omega \) and a Lebesque integral of the corresponding realization. In general, without assumption on measurability of a process, the integral \( I^* \) needs not to be a random variable. The following theorem shows the relation between the two integrals.

**Theorem 2.12** : Let \( X(t; \beta), t \in I \) be a measurable and m.s.-continuous stochastic process. Then the integrals \( I \) and \( I^* \) exist and they are equal with probability one.

**Proof** : The existence of integrals \( I \) and \( I^* \) follows directly from the assumptions. To show the equality with probability one it is sufficient to show that

\[
E^\beta \left( \left| I_n - I^* \right|^2 \right) \xrightarrow{n \to \infty} 0 \tag{14}
\]

where \( I_n \) denotes an approximating sum of \( I \). Of course,

\[
E^\beta \left( \left| I_n - I^* \right|^2 \right) = E^\beta \left( \left( I_n \right)^2 \right) + E^\beta \left( \left( I^* \right)^2 \right) - 2E^\beta \left( I^* I_n \right)
\]

By virtue of eq. (1.25)

\[
E^\beta \left( \left( I_n \right)^2 \right) \to \iint_{t_1, t_2} R_{XX}(t_1, t_2) \ dt_1 \ dt_2
\]

Making use of (4) one has

\[
E^\beta \left( \left( I^* \right)^2 \right) = \iint_{t_1, t_2} R_{XX}(t_1, t_2) \ dt_1 \ dt_2
\]

It can be easily shown that \( E^\beta \left( I^* I_n \right) \) approach the same quantity equal to \( E^\beta \left( \left| I^* \right|^2 \right) \) an \( n \to \infty \).

**Remark 2.13** : It can only be assumed that a process \( X(t; \beta), t \in I \) is m.s.-continuous. Hence by virtue of Theorem 2.3 process \( X(t; \beta), t \in I \) has stochastically equivalent measurable representation.
Let us consider now the questions associated with the continuity of a stochastic process \( X(t; \beta), t \in I \).

**Theorem 2.14**: Let \( X(t; \beta), t \in I \) be a separable second-order stochastic process, and there exists constants \( C > 0 \) and \( r > 0 \) such that, for every \( t_1, t_2 \in I \)
\[
\left| R_{xx}(t_2, t_2) + R_{xx}(t_1, t_1) - 2R_{xx}(t_1, t_2) \right| \leq C|t_2 - t_1|^{2r}
\]  \( (15) \)
then almost all realizations of \( X(t; \beta), t \in I \) are continuous on \( I \).

**Proof**: The proof is direct consequence of Kolmogorov Theorem with \( p = 2 \). \( \blacksquare \)

**Remark 2.15**: Actually it is sufficient for the correlation function to have a second order derivative \( \frac{\partial^2}{\partial t_1 \partial t_2} R_{xx}(t_1, t_2) \). The above conditions can be easily proved, using mean value theorem, to be equivalent with condition (15).

For the specific case of weakly stationary stochastic processes, there is the following

**Theorem 2.16**: Let \( X(t; \beta), t \in I \) be a weakly stationary process, m.s.-differentiable. Then its stochastically equivalent representation with continuous realizations exists.

**Proof**: Since \( X(t; \beta), t \in I \) is a weakly stationary, m.s.-differentiable, stochastic process, we have
\[
\left( R_{xx}(0) - R_{xx}(\tau) \right) \leq C|\tau|
\]
Thus, the Kolmogorov condition a) takes the form \( (p = 2) \)
\[
E^\beta \left[ \left| X(t + \tau; \beta) - X(t; \beta) \right|^2 \right] \leq \tilde{C}|\tau|
\]
So, we have proved the assertion. \( \blacksquare \)

We finish this section by presenting analogous results for differentiation of realizations of a stochastic process \( X(t; \beta), t \in I \).

**Theorem 2.13**: Let \( X(t; \beta), t \in I \) be a separable second-order stochastic process, and there exists a finite derivative \( \frac{\partial^{2n+2}}{\partial t_1^{n+1} \partial t_2^n} R_{xx}(t_1, t_2) \). Then almost all realizations of \( X(t; \beta) \) are \( n \)-times differentiable.

**Proof**: See SOBCZYK, K., (Stochastic Differential Equations). \( \blacksquare \)
4.3. References

Chapter 5

Probabilistic Analysis of the responses of Dynamical Systems under Stochastic Excitation

As we saw in Chapter 3 the characteristic functional gives a complete description of a stochastic process. In this chapter we shall see, how we can derive functional differential equations for the characteristic functional of a continuous time stochastic process whose sample functions satisfy a given Ordinary Differential Equation. This fact is of great importance, since we can formulate the full statistical problem into one equation. We must also note that, having this equation, all the statistical information of the problem can be recovered. Hence, form the functional differential equation we can, for example, derive the moments equations associated with the problem.

The above formulation, although it contains the whole statistical information, it’s not yet of practical importance. The difficulty comes from the functional differential equation itself, since up to the present time very little has been done to solve this kind of equations. On the other hand, very few characteristic functions, can be generalized in infinite dimensional Hilbert spaces, in order to give representations of the characteristic functional.

In Section 5.1 we introduce the notion of a stochastic differential equation and prove some general theorem concerning the existence and uniqueness of solutions for equations of this kind. Both existence and uniqueness is proved in the sense of probability measure. Moreover we study the necessary conditions for bounded of all moments of the probability measure that describes the system (ODE) response. Additionally we investigate the dependence of the solution of the system from a parameter inhered into the ODE. The analysis follows the steps of GIKHMAN, I.I. & SKOROKHOD, A.V., (The Theory of Stochastic Processes III).

Section 5.2 deals with the formulation of functional differential equations for the joint characteristic functional that describes the joint probability structure of the excitation and the response. The formulation concerns stochastic systems described by nonlinear ODE’s with polynomial nonlinearities. We distinguish between the case of a m.s. continuous stochastic excitation and an orthogonal increment excitation. Before the general results we present two specific systems, the Duffing Oscillator and the Van der Poll Oscillator.

In the next section 5.3 we prove how the general functional differential equation describing the probability structure of a system can be reduced to partial differential equations for the characteristic functions of various orders. Special cases of the above reduction are the general form of the Fokker-Planck equation as well as the Liouville equation. Of great importance are the partial differential equations proved at Section 5.3.4 that describes the probability structure of the system for the case of m.s. continuous excitation.

In the last Section 5.5 we introduce the notion of the kernel characteristic functional that generalizes the corresponding idea of kernel density functions in infinite dimensional spaces. We prove that the Gaussian characteristic functional has the kernel property, based on existent
theorems for extreme values of stochastic processes. In Section 5.4.4 we present how the kernel characteristic functionals can be used efficiently for the solution of functional differential equations. The basic idea is the localization in the physical phase space domain (probability measure) expressed in terms of the characteristic functional. We must note that the Gaussian measures are of great importance for these constructions since every analytical calculation concerning infinite dimensional integrals holds for Gaussian measures. In this way we derive a set of partial differential equations that governs the functional parameters of the kernel characteristic functionals. Finally in Section 5.4.5 we make special assumptions to derive a more simplified set of equations. Moreover a simple nonlinear ODE is studied and numerical results are presented.
5.1. Stochastic Differential Equations – General Theory

In the present section we introduce the notion of a stochastic differential equation and prove some general theorems concerning the existence and uniqueness of solutions of these equations. For this purpose it is necessary to generalize the notion of a stochastic integral introduced above. The following analysis is based on the approach of Gikhman, I. I. & Skorokhod, A. V., (The Theory of Stochastic Processes III).


Assume that we are dealing with a motion of a system \( S \) in the phase space \( \mathbb{R}^N \) and let \( x(t) = (x_1(t), x_2(t), \ldots, x_N(t)) \) denote the location of this system in \( \mathbb{R}^N \) at time \( t \). Assume also that a displacement of system \( \Sigma \) located at point \( x \) during the time interval \((t, t+\Delta t)\) can be represented in the form,

\[
x(t + \Delta t) - x(t) = A(x, t + \Delta t) - A(x, t) + \delta
\]

Here, \( A(x, t) \) is in general a random function; the difference \( A(x, t + \Delta t) - A(x, t) \) characterizes the action of an “external field of forces” at point \( x \) on \( \Sigma \) during the time period \((t, t+\Delta t)\) and \( \delta \) is a quantity which is of higher order of smallness in a certain sense than the difference \( A(x, t + \Delta t) - A(x, t) \). If \( A(x, t) \) as a function of \( t \) is absolutely continuous, then relation (1) can be replaced by the ordinary differential equation

\[
\frac{d x(t)}{dt} = A'_i(x(t), t)
\]

Equation (2) defines the motion of \( \Sigma \) in \( \mathbb{R}^N \) for \( t > t_0 \) under the initial condition \( x(t_0) = x_0 \), while \( A'_i(x, t) \) determines the “velocity field” in the phase space at time \( t \).

It is obvious that equation (2) cannot describe motions such as Brownian, i.e. motions which do not possess a finite velocity in the phase space or motions which possess discontinuities in the phase space. To obtain an equation which will describe a motion of systems of this kind it is expedient to replace relation (1) with an equation of an integral type. For this purpose we visualize that the time interval \([t_0, t]\) is subdivided into subintervals by the subdividing points \( t_1, t_2, \ldots, t_n = t \). It then follows from (1) that

\[
x(t) - x(t_0) = \sum_{i=0}^{n-1} A(x(t_i), t_{i+1}) - A(x(t_i), t_i) + \sum_{i=0}^{n-1} \delta_i
\]

Since \( \delta_i \) are of a small order, it is natural to assume that, \( \sum_{i=0}^{n-1} \delta_i \to 0 \) as \( n \to \infty \). In this case the last equality formally becomes,

\[
x(t) - x(t_0) = \int_{t_0}^{t} A(x(s), s) \, ds
\]
and the expression
\[ \int_{t_0}^t A(x(s), s) \, ds \]
can be called a stochastic integral in the random field \( A(x, t) \) along the random curve \( x(s), s \in [t_0, t] \); the integral should be interpreted as the limit, in a certain sense, to be defined more precisely, of sums of the form
\[ \sum_{j=0}^{n-1} A(x(t_j), t_{j+1}) - A(x(t_j), t_j) \]
Relation (3) is called a Stochastic Differential Equation and is written in the form
\[ d_x(x(t)) = A(x(t), t) \, dt, \quad x(t_0) = x_0, \quad t \geq t_0 \]
Under sufficiently general assumptions, for example, if \( A(x, t) \) is a martingale for each \( x \in \mathbb{R}^N \), one can assume that
\[ A(x, t) = a(x, t) + \beta(x, t) \] (4)
where \( \beta(x, t) \) as a function of \( t \) is a martingale and process \( a(x, t) \) is representable as the difference of two monotonically nondecreasing natural processes. In this connection it makes sense to suppose that the right-hand side of equation (3) can be represented according to formula (4) and introduce further restrictions on functions \( a(x, t) \) and \( \beta(x, t) \) in various ways. For example, we may assume that the function \( a(x, t) \) appearing in expression (4) is an absolute continuous function of \( t \) while \( \beta(x, t) \) as a function of \( t \) is a local square integrable martingale. (Some more general assumptions concerning \( \beta(x, t) \) are considered below).
In what follows, equation (3) will be written in the form
\[ x(t) = x(t_0) + \int_{t_0}^t a(x(s), s) \, ds + \int_{t_0}^t \beta(x(s), ds) \] (5)
or,
\[ d_x(x(t)) = a(x(t), t) \, dt + \beta(x(t), dt) \, , \quad x(t_0) = x_0 \]
In the case when \( \beta(x(t), dt) = 0 \), equation (5) is called an ordinary differential equation (with a random right-hand side).
Often fields \( \beta(x, t) = (\beta_1(x, t), \beta_2(x, t), \ldots, \beta_N(x, t)) \) of the form
\[ \beta_n(x, t) = \int \sum_{m=1}^M \gamma_{nm}(x, s) \, d\mu_m(s) \, , \quad n = 1, \ldots, N \] (6)
are considered. Here \( \mu_m(s), m = 1, \ldots, M \) are local mutually orthogonal square integrable martingales and \( \gamma_{nm}(x, s) \) are random functions satisfying conditions which assure the
existence of corresponding integrals. In this case, the second integral in equation (5) may be
defined as the vector-valued integral with components
\[ \int_{t_0}^{t} \beta_n(x(s), ds) = \int_{t_0}^{t} \sum_{m=1}^{M} \gamma_{nm}(x(s), s) d\mu_m(s), \quad n = 1, \ldots, N \]
and the theory of stochastic integrals described in Section 1 of Chapter 4 can be utilized.
However, if we confine ourselves to functions \( \beta(x, t) \) of type (6) a substantial amount of
generality is lost. This can be seen from the fact that the joint characteristic of processes
\( \beta_n(x, t) \) and \( \beta_n(y, t) \) defined by formula (6) is of the form
\[
F_{\beta_n(x, \cdot) \beta_n(y, \cdot)}(t) = E^\beta \left( \beta_n(x, t) \cdot \beta_n(y, t) \right) = \int_{t_0}^{t} \sum_{m=1}^{M} \gamma_{nm}(x, s) \gamma_{nm}(y, s) E^\beta (d\mu_m(s) \cdot d\mu_m(s)) = \\
= \int_{0}^{t} \sum_{m=1}^{M} \gamma_{nm}(x, s) \gamma_{nm}(y, s) dF_{\mu_m(x), \mu_m(y)}(s)
\]
while in the general case it is given by a function \( \Gamma^n(x, y, t) \) which for fixed \( t \) is an arbitrary
nonnegative-defined kernel of arguments \( x \) and \( y \),
\[
\sum_{i,j=1}^{t} \Gamma^n(x_i, y_j, t) z_i z_j \geq 0 \quad \text{for all } z_i \in \mathbb{R}, \ n = 1, 2, \ldots, N, \ I = 1, 2, \ldots
\]
For simplicity we consider here the one-dimensional case) let functions, \( \gamma_{nm}(x, t) = c_{nm}(x, t), \ n = 1, \ldots, N, \ m = 1, \ldots, M \) be nonrandom and \( \mu_m(t) = w_m(t) \) be
independent Wiener process. In this case the correlation function \( R_n(x, y, t) \) of the field
\[
\beta_n(x, t) = \int_{t_0}^{t} \sum_{m=1}^{M} c_{nm}(x, s) dw_m(s), \quad n = 1, \ldots, N
\]
equals
\[
R_n(x, y, t) = E^\beta \left( \beta_n(x, t) \cdot \beta_n(y, t) \right) = \int_{t_0}^{t} \sum_{m=1}^{M} c_{nm}(x, s) c_{nm}(y, s) ds, \quad n = 1, \ldots, N
\]
On the other hand, if we set \( \beta_n(x, t) = w_n(x, t) \), where \( w_n(x, t) \) is an arbitrary Gaussian field
with independent increments in \( t \), its correlation function \( R_n(x, y, t) = E^\beta \left( w_n(x, t) \cdot w_n(y, t) \right) \) is
then an arbitrary nonnegative-definite kernel (for a fixed \( t \)).
Thus, the restrictions, when considering stochastic integrals along a process \( x(t) \) imposed by
fields of the type (6) lead to a substantial narrowing of the class of problems under
consideration. Therefore it is expedient to introduce a direct definition and investigate
properties of the stochastic integral
\[
\int_{t_0}^{t} \beta(x(s), ds)
\]
by interpreting it in the simplest case as the limit in probability of the sums
\[
\sigma(x) = \sum_{k=1}^{n} \beta(x(s_{k-1}), s_k) - \beta(x(s_{k-1}), s_{k-1})
\]
The sums \( \sigma \) are called integral sums.
It is appropriate to observe that remarks about the insufficient generality of random fields as
given by relation (6) are not fully justified in the case when the stochastic processes are
represented by equation (5), with \( a(x,t) = \alpha(x,t) \) being a nonrandom function and \( \beta(x,t) \)
being a function with independent increments in \( t \). Indeed, the increment \( \Delta x(t) \) of a solution of
equation (5) at each time \( t \) depends on \( x(t) \) and on the value of the field \( \beta(x,t) \) at the point
\( y = x(t) \) (provided the probabilistic characteristics of the field \( \beta(x,t) \) as a function
of \( x \) are sufficiently smooth). Therefore one may expect that solutions of equations (5) will be
stochastically equivalent for any two fields \( \beta(x,t) = \beta_1(x,t) \) and \( \beta(x,t) = \beta_2(x,t) \) under the
condition that the joint distributions of the sequence of vectors
\[
(\beta_1(x,t_1), \beta_1(x,t_2), \ldots, \beta_1(x,t_n)) \quad \forall x \in \mathbb{R}^N, \quad \forall n \in \mathbb{Z}_+
\]
coincide for \( i = 1 \) and \( i = 2 \) and that the fields, \( \beta_i(x,t) \) possess independent increments in \( t \).
For example, let \( w(x,t) \) be an arbitrary Gaussian field possessing independent increments in \( t \),
\( B(x,t) = \{ E^n \left( w_i(x,t) \cdot w_j(x,t) \right) \} = \{ B_{ij}(x,t) \} \) and the functions \( B_{ij}(x,t) \) be differentiable
with respect to \( t \), \( B_{ij}(x,t) = \frac{d}{dt} B_{ij}(x,t) \). Denote by \( \sigma(x,t) \) a symmetric matrix such that
\( \sigma^2(x,t) = b(x,t) \) and introduce independent Wiener process \( w_j(t) \), \( j = 1, \ldots, N \). Set
\[
\beta_i(x,t) = \int_0^t \sigma(x,s) dw(s), \quad w(t) = (w_1(t), w_2(t), \ldots, w_N(t))
\]
Then
\[
E^3 \left( \beta_1(x,t) \otimes \beta_1(x,t) \right) = \int_0^t \sigma(x,s) \sigma(x,s) ds = \int_0^t b(x,s) ds
\]
and one can expect that the solutions of the differential equations,
\[
\begin{align*}
dx &= a(x,t) dt + w(x,dt) \\
dx &= a(x,t) dt + \sigma(x,t) dw(t)
\end{align*}
\]
will be stochastically equivalent, although the fields \( w(x,t) \) and \( \beta_i(x,t) \) in general are not.
Analogous observations can be made also in the case when \( \beta_i(x,t) \) is an arbitrary field with
independent increments in \( t \) with finite moments of the second order. Assume that the
increment \( \beta(x,t + \Delta t) - \beta(x,t) \) possesses the characteristic function
\[
E^3 \left( \exp \left( iz \cdot \beta(x,t + \Delta t) - iz \cdot \beta(x,t) \right) \right) = \exp \left( \frac{1}{2} \int_0^{t+\Delta t} z^T b(x,s) z ds + \int_0^{t+\Delta t} ds \int_{\mathbb{R}^n} \left[ e^{i z \cdot e(x,s,u)} - 1 - iz \cdot e(x,s,u) \right] \Pi(s,du) \right)
\]
(if \( \beta(x,t) \) possess finite moments of the second order and is absolutely continuous in \( t \) one
can reduce an arbitrary characteristic function of a process with independent increments to this
In this case, one can expect for sufficiently smooth functions \( a(x,t) \), \( b(x,t) \) and \( c(x,t,u) \) that solutions of the stochastic equations
\[
d_x x = a(x,t) \, dt + b(x, dt)
\]
and
\[
d_x x = a(x,t) \, dt + \sigma(x,t) \, dt + \int_{\mathbb{R}^n} c(x,t,u) \, \nu(dt, du)
\]
will be stochastically equivalent. Here \( \sigma(x,t) \) is a symmetric matrix, \( \sigma^2(x,t) = b(x,t) \), \( \nu(t, A) \) is a centered Poisson measure with \( \text{Var}[\nu(t, A)] = \int_0^t \Pi(s, A) \, ds \).

It is also clear that if the increments in \( t \) of the field \( \beta(x,t) \) are dependent then the remarks above concerning the possibility of replacing the field \( \beta(x,t) \) in equation (5) by a simpler field without restricting the class of obtained solution are no longer valid.

The preceding outline of a definition of a stochastic differential equation is expedient to extend in yet another direction. At present, "feedback" systems play an important part in a number of scientific-engineering problems. For such systems the "exterior field of forces" acting on the system at a given time depends not only on the instantaneous location of the system in the phase space but also on its phase trajectory in "the past":
\[
(\beta_{\ast} x_{\ast} s, t)\}
\]
Next let \( \alpha(x,t) = \alpha(x,t, \omega) \) be a random function defined on \( \mathcal{D}^N \times [0,T] \times \Omega \). Relation (7) can be written as follows,
\[
x(t + \Delta t) - x(t) = \alpha(t, t + \Delta t) - \alpha(t, t) + \delta_t
\]
and equation (5) can be represented by the equation
\[
x(t) = x_0 + \int_{t_0}^t \alpha(\theta, x, s) \, ds + \int_{t_0}^t \beta(\theta, x, ds) \quad , \quad t > t_0
\]
Here it becomes necessary to define the process \( x(t) \) over the whole "past", i.e., up to the time \( t_0 \). In this connection one should adjoin to equation (8) relation
\[
x(t) = x_{\text{past}}(t) \quad , \quad t \leq t_0
\]
which will be called from now on the initial condition for the stochastic differential equation (8).

5.1.2. The Existence and Uniqueness Theorems of Stochastic Differential Equations

Let the probability space \((\Omega, \mathcal{B}(\Omega), \mathbb{P})\) which is equipped with a current of \(\sigma\)–algebras \(\mathcal{B}_i(\Omega), \ t \in [0,T]\). Let also the random functions \(a(\phi,t)\) and \(\beta(\phi,t)\), with values in \(\mathbb{R}^N\), adopted to \(\mathcal{B}_i(\Omega)\) with \(\phi \in D^X, \ t \in [0,T]\) be given. For what follows we will use the concept of random time, so we have the following

**Definition 1.1**: Let \(\mathcal{B}_i(\Omega), \ t \in [0,T]\) be a current of \(\sigma\)–algebras. A function \(\tau(\omega) = t\), \(\tau \in \Omega \subset \Omega\), with values in \([0,T]\) is called a random time on \(\mathcal{B}_i(\Omega), \ t \in [0,T]\) or an \(\mathcal{B}_i\)–random time, if \(\{\tau \leq t\} \in \mathcal{B}_i\). ■

Hence we can now give the definition of a solution of a stochastic differential equation.

**Definition 1.2**: A random time \(\tau, \ 0 < \tau \leq T\), on \(\mathcal{B}_i(\Omega), \ t \in [0,T]\) and a random process \(x(t), \ t \in [0,T]\), \(\mathcal{B}_i \times \mathcal{B}([0,T])\)–measurable and satisfying with probability 1 relations

\[
x(t) = x_{past}(t), \ t \leq t_0 \\
x(t) = x_0 + \int_0^t \alpha(\theta, x, s)ds + \int_0^t \beta(\theta, x, ds), \ t > t_0
\]

for each \(t \leq \tau\), is called a solution of the stochastic differential equation

\[
dx = \alpha(\theta, x, t)dt + \beta(\theta, x, dt), \ t > t_0
\]

satisfying the initial condition

\[
x(t) = x_{past}(t), \ t \leq t_0
\]

It is assumed that integrals in the r.h.s. of (11) are well defined, the first as a Lebesgue integral and the second as a stochastic integral. The random variable \(\tau\) is called a lifetime of the process \(x(t)\) (i.e., of a solution of a stochastic differential equation). Equation (12) is called regular on \([0,T]\) provided it possess a unique solution on the whole time interval \([0,T]\) (i.e., if a unique solution of equation (12) with \(\tau = T\) exists).

We now introduce some general assumptions on functions \(a(\phi,t)\) and \(\beta(\phi,t)\) under which the r.h.s. of equality (11) will be well defined for a sufficiently wide class of processes \(x(t)\). Observe that there is no point to consider the most general classes of processes \(x(t)\) since the r.h.s. of equation (11) represents a process which possesses a continuous modification or a modification with sample functions in \(D_T^X\) and hence the process \(x(t)\) must also possess these properties.

We shall first discuss the function \(a(\phi,t)\). We introduce two sets of assumptions.
5.1. STOCHASTIC DIFFERENTIAL EQUATIONS – GENERAL THEORY

\[\alpha \]

\textbf{a) The function } \(a(\phi, s) = a(\phi, s, \omega)\) \text{ is defined on } \mathcal{D}^N \times [0, T] \times \Omega \text{ and is measurable for } s \in [0, T].

\[\alpha \]

\textbf{b) The function } a(\phi, s) \text{ is defined on } \mathcal{D}^N \times [0, T] \times \Omega \text{ and is measurable for } s \in [0, T].

\[\alpha \]

\textbf{c) For a fixed } \omega \text{ the family of functions } \{a(\bullet,t), t \in [0,T]\} \text{ of arguments } \phi \text{ is uniformly continuous on } \mathcal{D}^N \text{ relative to metric } d_{\mathcal{D}}.

\[\alpha \]

\textbf{2) The function } a(\phi, s) \text{ satisfies assumptions a1 provided } \mathcal{D}^N, \mathcal{B}(\mathcal{D}^N) \text{ and } \mathcal{D}[0, T] \text{ are replaced by } \mathcal{C}^N, \mathcal{B}(\mathcal{C}^N) \text{ and } \mathcal{C}[0, T] \text{ respectively.}

We say that the function \(a(\phi, t)\) is linearly bounded (relative to a uniform norm or a seminorm if in the succeeding inequalities the norm \(\|\|\) can be replaced by the seminorm \(\|\|\)) provided a continuous monotonically nondecreasing process \(\lambda_\nu(t)\) exists adopted to \(\mathcal{B}_i(\Omega), t \in [0, T]\) such that \(\lambda_\nu(t) < \infty\) with probability 1 and

\[
\int_a^b a(\phi, t) dt \leq (1 + \|\phi\|) \int_a^b \lambda_\nu(t) dt 
\]

We say that the function \(a(\phi, t)\) satisfies the Local Lipschitz condition (in a uniform norm or seminorm) if for any \(\nu > 0\) there exists a monotonically \(\lambda_\nu(t)\) exists adopted to \(\mathcal{B}_i(\Omega), t \in [0, T]\) such that

\[
\int_a^b a(\phi, t) - a(\psi, t) dt \leq (1 + \|\phi - \psi\|) \int_a^b \lambda_\nu(t) dt
\]

for all \(\phi\) and \(\psi\) satisfying \(\|\phi\| \leq N\) and \(\|\psi\| \leq N\). If we choose the process which is independent of \(\nu\) in place of \(\lambda_\nu(t)\) then we shall refer to the process \(a(\phi, t)\) as one satisfying the uniform Lipschitz condition.

The class of processes \(a(\phi, t)\) which satisfy condition a2, (14) and (15) will be denoted by \(S^c_\nu (\lambda_\nu, \lambda_\nu)\). We denote by \(S_\nu (\lambda_\nu, \lambda_\nu)\) the class of processes which satisfy a1 and the condition obtained from (14) and (15) when the uniform norm \(\|\|\) is replaced by \(\|\|\) in the corresponding inequalities. In the case when we shall be dealing with random functions \(a(\phi, t)\) satisfying only one of the inequalities (14) or (15), for instance (14), we shall write \(a(\phi, t) \in S^c_\nu (\lambda_\nu, \bullet)\) and analogously in other cases.

Note that \(\phi_t = \theta_t \psi (\psi \in \mathcal{D}^N, t \in [0,T])\), with values in \(\mathbb{R}^N\) is a Borel function in argument \(t\).

Indeed if \(B\) is a cylinder in \(\mathcal{D}^N\) with the basis \(B = \prod_{j=1}^n B_j\) in coordinates \((s_1, s_2, \ldots, s_n)\), \(s_i \leq 0\), then

\[
\{t: \phi(t) \in B\} = \bigcap_{j=1}^n \{t: \phi(t) + s_j \in B_j\}
\]
Since the sets \( \{ z : \phi(z) \in B_i \} = Z_i \) are Borel sets provided \( B_i \in \mathbb{R}^N \), are such, the set \( \{ t : \phi_i \in B \} = \bigcap_{i=1}^n \{ Z_i - s \} \), where \( Z - s \) denotes the set \( \{ z : z + s \in Z \} \), will also be a Borel set. Thus, if \( g(\phi,t) \) is a \( \mathcal{B}(\mathcal{D}^N) \times \mathcal{B}([0,T]) \) – measurable function in the arguments \((\phi,t)\) where \( \mathcal{B}(\mathcal{D}^N) \) is the minimal \( \sigma \) – algebra generated by the cylinders in \( \mathcal{D}^N \), then \( g(\theta,\phi,t) \) will be a Borel function in argument \( t \).

Consequently if the field \( a(\phi,t) \) satisfies condition a1 or a2 then the integral

\[
\int_a^b a(\theta,\psi,t) \, dt
\]

exists with probability 1.

Next, if \( a(\phi,t) \in \mathcal{S}_a(\lambda_{\psi\star}) \), then we have for \( 0 \leq a < b \leq T \),

\[
\left| \int_a^b a(\theta,\psi,t) \, dt \right| \leq \int_a^b (1 + \|\theta,\psi\|) \lambda_{\psi}(t) \, dt
\]

(16)

The proof follows easily from the fact that \( \theta,\psi \), \( t \in [0,T] \) is a continuous function with values in \( \mathcal{D}^N \) with respect to a metric \( d_{\psi\nu} \) in \( \mathcal{D}^N \), and \( a(\phi,t) \) is a continuous function of argument \( \phi \) (uniformly in \( t \)).

Analogously if \( a(\phi,t) \in \mathcal{S}_a^c(\lambda_{\psi\star}) \), then

\[
\left| \int_a^b a(\theta,\psi,t) \, dt \right| \leq \int_a^b (1 + \|\theta,\psi\|) \lambda_\psi(t) \, dt
\]

(17)

If, however \( a(\phi,t) \in \mathcal{S}_a(\star,\lambda_{\psi}) \) and \( \|\psi_1\| \vee \|\psi_2\| < N \), then

\[
\left| \int_a^b a(\theta,\psi_1,t) \, dt - \int_a^b a(\theta,\psi_2,t) \, dt \right| \leq \int_a^b \|\psi_1 - \psi_2\| \lambda_\psi(t) \, dt
\]

(18)

and an analogous inequality is valid for the case \( a(\phi,t) \in \mathcal{S}_a^c(\star,\lambda_\nu) \).

As far as the integral \( \int_a^b \beta(\theta,x,ds) \) is concerned, conditions for its existence and its properties were discussed in the preceding chapter.

We introduce the notation for the classes of fields \( \beta(\phi,t) \) analogous to the notation for the classes of functions \( a(\phi,t) \) introduced above. Namely we write \( \beta(\phi,t) \in \mathcal{S}_\beta(\lambda_\nu) \) (or \( \beta(\phi,t) \in \mathcal{S}_\beta^c(\lambda_\nu) \) provided the field \( \beta(\phi,t) \) satisfies conditions \( \beta_1 \) and \( \beta_2 \), is linearly bounded, satisfies the local Lipschitz condition relative to a seminorm (norm) and the dominating processes \( \Lambda_\psi(t) \) and \( \Lambda_\nu(t) \) are absolutely continuous with \( \lambda_\psi(t) = \Lambda_\psi'(t) \) and \( \lambda_\nu(t) = \Lambda_\nu'(t) \).
Set
\[
A(\phi,t) = \int_{t_0}^{t} a(\phi,s) ds + \beta(\phi,t)
\]
and write \( A(\phi,t) \in S \left( \lambda_0, \lambda_v \right) \) (or \( A(\phi,t) \in S^c \left( \lambda_0, \lambda_v \right) \)) provided
\[
a(\phi,t) \in S_{\nu} \left( \lambda_0, \lambda_v \right) \quad \text{and} \quad \beta(\phi,t) \in S_{\beta} \left( \lambda_0, \lambda_v \right)
\]
\[
a(\phi,t) \in S_{\nu}^c \left( \lambda_0, \lambda_v \right) \quad \text{and} \quad \beta(\phi,t) \in S_{\beta}^c \left( \lambda_0, \lambda_v \right)
\]
Let the process \( x(t), t \in [0,T] \), be adopted to a current of \( \sigma \)-algebras \( \mathcal{B}_i(\Omega), t \in [0,T] \) with sample functions belonging with probability 1 to \( \mathcal{D}^V [0,T] \). We complete the definition of \( x(t) \) for \( t \leq 0 \) by setting \( x(t) = \phi(t) \) for \( t \leq t_0 \) where \( \phi(t) \) is a given function in \( \mathcal{D}^V \).

Assume that \( a(\phi,t) \) satisfies condition \( \alpha1 \) a) and that \( \beta(\phi,t) \in S_{\beta} \left( \bullet, \lambda_v \right) \). In what follows these condition will always be assumed to be valid.

We now formulate the first basic theorems for the existence and uniqueness of a solution in accordance with the operator \( A(\phi,t) \).

**Theorem 1.3 :** Assume that \( A(\phi,t) \in S \left( \lambda_0, \lambda_v \right) \) (or that \( A(\phi,t) \in S^c \left( \lambda_0, \lambda_v \right) \)). The stochastic differential equation (12) possesses in \( H^* \left( \mathcal{D}^V \right) \) a unique solution which is defined for all \( t \in [0,T] \). ■

**Proof :** For the proof see GIKHMAN, I.I. & SKOROKHOD, A.V., (The Theory of Stochastic Processes III. p.141) ■

**Theorem 1.4 :** If \( A(\phi,t) \in S \left( \bullet, \lambda_v \right) \) (or \( A(\phi,t) \in S^c \left( \bullet, \lambda_v \right) \)) then a random time \( \tau_\infty \) exists on \( \mathcal{B}_i(\Omega), t \in [0,T] \) and a random process \( x(t) \) adopted to \( \mathcal{B}_i(\Omega), t \in [0,T] \) defined for \( t < \tau_\infty \) such that \( x(t) \) satisfies (12) for all \( t < \tau_\infty \), and the sample functions of \( x(t) \) possess with probability 1 left – hand limits and are continuous from the right for all \( t < \tau_\infty \). Moreover \( \mathcal{D}^V (\tau_\infty > 0) = 1 \).

**Proof :** For the proof see GIKHMAN, I.I. & SKOROKHOD, A.V., (The Theory of Stochastic Processes III. p.142) ■

**Theorem 1.5 :** If under the condition of Theorem 1.3 one can set \( \lambda_0 = C \), then the solution of equation (12) belongs to \( H^*_2 \left( \mathcal{D}^V \right) \) and
\[
E^3 \left( \sup_{t \in [0,T]} \left| x(t) \right|^2 \right) \leq B \left( 1 + \left\| \phi \right\|^2 \right)
\]
(19)
\[
E^3 \left( \sup_{s \in [t,t+h]} \left| x(s) - x(t) \right|^2 \right) \leq B \left( 1 + \left\| \phi \right\|^2 \right) h
\]
(20)

where \( B \) is a constant which depends on \( C, K \) and \( T \) only.

For the proof see GIKHMAN, I.I. & SKOROKHOD, A.V., (The Theory of Stochastic Processes III. p.142) ■
5.1.3. **Bounds on the Moments of Solutions of Stochastic Differential Equations**

A very important issue for a solution of a stochastic differential equation is whether it possesses or not finite order moments. Based on theorems for the existence of moments, we can solve moment equations and have numerical results for the stochastic differential equation. The theorem presented below guarantees the existence of all moments for a very wide class of systems including Ito equations as well as equations with continuous excitation process.

Consider equation (12) - (13)

\[
dx = \alpha(\theta, x, t)dt + \beta(\theta, x, dt), \quad t > t_0
\]

\[
x(t) = x_{past}(t), \quad t \leq t_0
\]

satisfying the conditions of Theorem 1.3. Assume first that for a fixed \(\phi\), \(\beta(\phi, t)\) is a martingale process with continuous sample functions with probability 1. (GIKHMAN, I. I. & SKOROKHOD, A. V., The Theory of Stochastic Processes III. p.37, 144). Furthermore let the characteristic \(F^k(t)\), of the process \(\beta^k(\phi, t)\) be absolutely continuous with respect to the Lebesgue measure

\[
F^k(t) = \int_0^t \beta_k(\phi, s)ds
\]

and

\[
\left| \beta_k(\phi, s) \right| \leq \lambda^2(s)(1 + \|\phi\|)
\]

For the existence of moments of the solution of an SDE we have the following

**Theorem 1.6 :** A process \(x(t)\) satisfying equation (12) - (13) and conditions (21) - (22) for \(t \leq \tau_N = \inf \{t : \lambda(t) > N\}\) possess moments of all orders.

**Proof :** For the proof see GIKHMAN, I. I. & SKOROKHOD, A. V., (The Theory of Stochastic Processes III. p.144)

**Corollary 1.7 :** If

\[
|\beta^k(\phi, t)| \leq C(1 + \|\phi\|)
\]

\[
E\left(\beta^k(\phi, \Delta t) | \mathcal{B}_t(\Omega)\right) \leq C \left(1 + \|\phi\|^2\right) \Delta t
\]

where \(C\) is an absolute constant, then the solution of equation (12) - (13) possess finite moments of all orders for \(t \in [t_0, T]\).

More extended results concerning the existence of all or of finite number moments and the corresponding conditions can be found at GIKHMAN, I. I. & SKOROKHOD, A. V., (The Theory of Stochastic Processes III. p.145). Moreover we must emphasize the issue of moment stabilities although we will not study it in the present work. For a detailed introductory discussion we refer to SOOOG, T. T., (Random Differential Equations, p.256). Some interesting results for special cases of stochastic differential equations as well as stochastic partial differential equation can be found at the paper of KOTULSKI, Z. & SOBCZYK, K., (On the Moment Stability of Vibratory Systems with Random Parametric Impulsive Excitation). Another approach based
on the Fourier transform of the moment equations with respect to time have been studied. The transformed moments have been studied extensively in homogeneous turbulence theory, and we refer to Batchelor, G.K., (Theory of Homogeneous Turbulence) for a thorough discussion of the single-time equation.

5.1.4. Continuous dependence on a parameter of solutions of stochastic equations

Consider the equation of the form

\[ \eta_u(t) = x_u(t) + \int_{t_0}^{t} A_u(\theta_s, \eta_u(s), s) \, ds, \quad t \geq t_0 \tag{25} \]

\[ \eta_u(t) = x_{\text{past}}(t), \quad t < t_0 \tag{26} \]

where \( u \) is a scalar parameter, \( u \in [0, u_0] \), the field

\[ A_u(\psi, t) = \int_{t_0}^{t} a_u(\psi, s) \, ds + \beta_u(\psi, t) \]

and the function \( x_u(t) \) both depend on the parameter \( u \) and the initial condition \( (x_{\text{past}}(t), t < t_0) \) does not depend on \( u \).

**Theorem 1.8**: Assume that \( A_u(x, t) \in S(C, C) \) and, moreover, let

a) \( \sup_{t_0 \leq t \leq T} E \left( \left| x_u(t) \right|^2 \right) \leq C \)

b) \( \lim_{u 
arrow 0} \sup_{t_0 \leq t \leq T} E \left( \left| x_u(t) - x_0(t) \right|^2 \right) = 0 \quad \forall t \in [t_0, T] \)

c) \( E \left( \left| \Delta \beta_u(\psi, t) - \Delta \beta_0(\psi, t) \right|^2 \mid \Omega \right) \leq E \left( \int_{t}^{t+\Delta t} \gamma_u(\psi, s) \, ds \mid \Omega \right) \)

and, for arbitrary \( N > 0 \) and \( t \in [t_0, T] \), let

\[ \lim_{u 
arrow 0} \sup_{\|\psi\| \leq N} E \left( \left| a_u(\psi, t) - a_0(\psi, t) \right|^2 \right) = 0 \]

Then

\[ \lim_{u \narrow 0} \sup_{t_0 \leq t \leq T} E \left( \left| \eta_u(t) - \eta_0(t) \right|^2 \right) = 0 \]

**Proof**: The variables \( \eta_u(t) \) possess finite moments of the second order since equations (25) - (26) satisfy the conditions of Theorem 1.6. We represent the difference \( \eta_u(t) - \eta_0(t) \) in the form

\[ \eta_u(t) - \eta_0(t) = \sigma_u(t) + \int_{t_0}^{t} A_u(\theta_s, \eta_u(s), s) \, ds - \int_{t_0}^{t} A_u(\theta_s, \eta_0(s), s) \, ds, \]

where,

\[ \sigma_u(t) = x_u(t) - x_0(t) + \int_{t_0}^{t} A_u(\theta_s, \eta_u(s), s) \, ds - \int_{t_0}^{t} A_0(\theta_s, \eta_0(s), s) \, ds \]
It is easy to verify that
\[ E\left(\left|\eta_u(t) - \eta_0(t)\right|^2\right) \leq 3E\left(\left|\sigma_u(t)\right|^2\right) + 3C^2(T + 1)E\int_{t_0}^t \|\theta_u(\eta_u - \eta_0)\|_u^2 \, ds = \]
\[ = 3E\left(\left|\sigma_u(t)\right|^2\right) + C'\int_{t_0}^t \int_{s-t}^s E\left(\left|\eta_u(s + u) + \eta_0(s + u)\right|^2\right) K(du) \, ds = \]

Set \( \nu_u(t) = \sup_{t_0 \leq s \leq t} E\left(\left|\eta_u(s) - \eta_0(s)\right|^2\right) \). It follows from the last inequality that
\[ \nu_u(t) \leq 3E\left(\left|\sigma_u(t)\right|^2\right) + \int_{t_0}^t \nu_u(s) \, ds \]

Then it can be easily proved that
\[ \nu_u(t) \leq C'' \sup_{t_0 \leq s \leq t} E\left(\left|\sigma_u(s)\right|^2\right) \]
where \( C'' \) does not depend on \( u \). Furthermore
\[ \sup_{t_0 \leq s \leq t} E\left(\left|\sigma_u(s)\right|^2\right) \leq 3\left(\sup_{t_0 \leq s \leq t} E\left(\left|x_u(s) - x_0(s)\right|^2\right) + \right. \]
\[ + E\left(\left|\int_{t_0}^s a_u(\theta, \eta_0, s) - a_0(\theta, \eta_0, s) \, ds\right|^2\right) + \]
\[ + \sup_{t_0 \leq s \leq t} E\left(\left|\int_{t_0}^s \beta_u(\theta, \eta_0, ds) - \int_{t_0}^s \beta_0(\theta, \eta_0, ds) \right|^2\right) \]
\[ = 3(I_1 + I_2 + I_3) \]

By the condition, the quantity \( I_1 \to 0 \). Next
\[ I_2 \leq TE\left(\int_{t_0}^t \left|a_u(\theta, \eta_0, s) - a_0(\theta, \eta_0, s)\right|^2 \, ds\right) \]
where the integrand is dominated by the quantity
\[ 4C^2 \left(1 + \int_{-\infty}^{t_0} \eta_0(s + u) K(du)\right) \]
which is independent of \( u \) and is integrable with respect to the measure \( d\Theta \times ds \). On the other hand,
\[ \left|a_u(\theta, \eta_0, s) - a_0(\theta, \eta_0, s)\right| \to 0 \]
in probability for each \( s \) and hence in measure \( d\Theta \times ds \). Therefore \( I_2 \to 0 \) as \( u \to 0 \). Finally
\[ I_3 \leq 4E\left(\left|\int_{t_0}^t \beta_u(\theta, \eta_0, ds) - \int_{t_0}^t \beta_0(\theta, \eta_0, ds) \right|^2\right) = 4E\left(\int_{t_0}^t \gamma_u(\theta, \eta_0, s) \, ds\right) \]
and as in the case of quantity \( I_2 \) it is easy to verify \( I_3 \to 0 \) as \( u \to 0 \). \( \square \)

**Remark 1.9:** We shall strengthen the assumption of Theorem 1.8 and assume that in addition to the conditions stipulated in the theorem
\begin{align*}
\lim_{u \to 0} E \left( \sup_{t \leq T} \left| x_u(t) - x_0(t) \right|^2 \right) &= 0
\end{align*}

In this case
\begin{align*}
\lim_{u \to 0} E \left( \sup_{t \leq T} \left| \eta_u(t) - \eta_0(t) \right|^2 \right) &= 0
\end{align*}

The proof of this assertion is analogous to the proof of Theorem 1.8. ■

**Theorem 1.10:** Consider the stochastic equation.
\[ dx_u = A_u(\theta_x x_u, dt) , \quad x_u(t) = x_{past}(t), \ t \leq t_0 \quad (u \in [0,u_0]) \]

satisfying the conditions of Theorem 1.3 and for all \( N > 0 \) let
\[ \lim_{p \to \infty} \sup_u \left[ \mathcal{P} \left( \sup_{t \leq T} \lambda_u^u(t) \geq p \right) + \mathcal{P} \left( \sup_{t \leq T} \lambda_u^u(t) \geq p \right) \right] = 0 \]

and also let condition c) of Theorem 1.8 be satisfied. Then, for any \( \varepsilon > 0 \)
\[ \mathcal{P} \left( \sup_{t \leq T} \left| x_u(t) - x_0(t) \right| > \varepsilon \right) \to 0 \quad \text{as} \quad u \to 0. \]

**Proof:** Let
\[ \tau_p = \inf \left\{ t : \lambda_u^u(t) \geq p, \lambda_u^u(t) \geq p, \left| x_u(t) \right| \geq N \right\} \quad (\inf \emptyset = T) \]
\[ a^p_u(\phi, t) = a_u(\phi, t) \quad \text{for} \quad t < \tau_p \]
\[ a^p_u(\phi, t) = 0 \quad \text{for} \quad t \geq \tau_p \quad , \quad \beta^p_u(\phi, t) = \beta_u(\phi, t \land \tau_p) \]
\[ A^p_u(\phi, t) = \int_{t_0}^t a^p_u(\phi, s) ds + \beta^p_u(\phi, t) \]

Theorem 1.8 (or the corresponding remarks of this theorem) is applicable to equation
\[ dx_u^p = A^p_u(\theta_x x_u, dt) , \quad x_u^p(t) = x_{past}(t), \ t \leq t_0 \]

On the other hand, in view of Theorem 1.3 \( x_u^p(t) = x_u(t) \) with probability 1 for all \( t < \tau_p \).

Therefore for any \( \varepsilon > 0 \)
\[ \mathcal{P} \left( \sup_{t \leq T} \left| x_u(t) - x_u^p(t) \right| > \varepsilon \right) \leq \mathcal{P} \left\{ \tau_p < T \right\} \]

Furthermore,
\[ \mathcal{P} \left( \sup_{t \leq T} \left| x_u(t) - x_u^p(t) \right| > \varepsilon \right) \leq \mathcal{P} \left( \sup_{t \leq T} \left| x_u(t) - x_u^p(t) \right| > \frac{\varepsilon}{3} \right) + 
\]
\[ + \mathcal{P} \left( \sup_{t \leq T} \left| x_u^p(t) - x_u^p(t) \right| > \frac{\varepsilon}{3} \right) + \mathcal{P} \left( \sup_{t \leq T} \left| x_u^p(t) - x(t) \right| > \frac{\varepsilon}{3} \right) \leq 
\]
\[ \leq \mathcal{P} \left( \sup_{t \leq T} \left| x_u^p(t) - x_u^p(t) \right| > \frac{\varepsilon}{3} \right) + 2\mathcal{P} \left\{ \tau_p < T \right\}. \]
The uniform stochastic boundness of the processes $\lambda^u(t)$ and $\lambda(t)$ and Theorem 1.8 imply that one can first choose a sufficiently large value of $p$ and $N$ such that $\mathbb{P}\{\tau_p < T\} < \varepsilon/3$ and then choose a $\delta > 0$ such that $\mathbb{P}\left\{\sup_{t \in [0, T]} |x^u(t) - x(t)| > \frac{\varepsilon}{3}\right\} < \varepsilon$ for $u \in [0, \delta]$.
5.2. Hopf-type Functional Differential Equation (FDE) for the Characteristic Functional of Dynamical Systems

In this section we will derive the Functional Differential Equation for a wide class of autonomous stochastic dynamical systems with external forcing. To begin our discussion let us consider the special cases of a Duffing Oscillator and a van der Poll oscillator. Let \((\Omega, \mathcal{B}(\Omega), \mathcal{P})\) be a probability space and \(C_c^\infty(I)\) be the separable, real, inner product \(L^2\) (inner product) space of infinitely differentiable functions with values in \(U \subset \mathbb{R}\). \(\mathcal{B}\left(C_c^\infty(I)\right)\) will denote the \(\sigma\)–field generated by the open subsets of \(C_c^\infty(I)\).

5.2.1. Duffing oscillator

The Duffing oscillator has the following equation of motion

\[
m\ddot{x}(t; \beta) + c\dot{x}(t; \beta) + kx(t; \beta) + ax^3(t; \beta) = y(t; \beta)
\]

(1)

\[
[x(t_0; \beta), \dot{x}(t_0; \beta)] = [x_0(\beta), \dot{x}_0(\beta)] \quad \text{a.s.}
\]

(2)

We assume that \(y(t; \beta)\) is a stochastic process defined onto the probability space \((\Omega, \mathcal{B}(\Omega), \mathcal{P})\) with the associated probability space \(\left(C_c^\infty(I), \mathcal{B}\left(C_c^\infty(I)\right), \mathcal{P}_y\right)\), whose characteristic functional is \(\mathcal{F}_y(v)\). Using the results of the previous section we can conclude that there is a unique solution for the SDE (1)-(2). We desire to calculate the joint characteristic functional \(\mathcal{F}_{xy}(u, v)\). We first note that

\[
\mathcal{F}_{xy}(u, v) = \int \int_{C_c^\infty \times C_c^\infty} \exp \left\{ i \langle x, u \rangle + i \langle y, v \rangle \right\} \mathcal{P}_{xy}(dx, dy)
\]

(3)

where the integrals are over the infinite dimensional space \(C_c^\infty\). The brackets \(\langle \cdot, \cdot \rangle\) denotes the inner product of the Hilbert space \(L^2\), i.e. \(\langle x, y \rangle = \int x(t) y(t) dt\). The Volterra derivative of \(\mathcal{F}_{xy}(u, v)\) with respect to \(u(\xi)\) will be denoted as \(\frac{\delta \mathcal{F}_{xy}(u, v)}{\delta u(\xi)}\) or simply \(\frac{\delta \mathcal{F}_{xy}(u, v)}{\delta u_\xi}\). Direct calculations shows

\[
\frac{\delta \mathcal{F}_{xy}(u, v)}{\delta u_\xi} = \int_{C_c^\infty} \int_{C_c^\infty} i x(\xi) \exp \left\{ i \langle x, u \rangle + i \langle y, v \rangle \right\} \mathcal{P}_{xy}(dx, dy)
\]

(4)

Taking the \(n\)th ordinary derivative of \(\frac{\delta \mathcal{F}_{xy}(u, v)}{\delta u_\xi}\) with respect to \(\xi\) we will have

\[
\frac{d^n}{d\xi^n} \frac{\delta \mathcal{F}_{xy}(u, v)}{\delta u_\xi} = \int_{C_c^\infty} \int_{C_c^\infty} i \frac{d^n}{d\xi^n} x(\xi) \exp \left\{ i \langle x, u \rangle + i \langle y, v \rangle \right\} \mathcal{P}_{xy}(dx, dy)
\]

(5)

Additionally we note that the third Volterra derivative of \(\mathcal{F}_{xy}(u, v)\) with respect to \(u(\xi_1), u(\xi_2), u(\xi_3)\) is
\[
\delta \mathcal{F}_{xy}(u,v) \over \delta u_\xi \delta u_\eta \delta u_\zeta = \int_{\mathbb{C}^n} \int_{\mathbb{C}^n} i \xi(x_1)i\xi(x_2)i\xi(x_3) \exp \{i\langle x,u \rangle + i\langle y,v \rangle \} \mathcal{P}_{xy}(dx,dy)
\]

(6)

Setting \( \xi = \xi_2 = \xi_3 = \xi \) we will have

\[
\delta^3 \mathcal{F}_{xy}(u,v) \over \delta u_\xi^3 = \int_{\mathbb{C}^n} \int_{\mathbb{C}^n} [i\xi(x)]^3 \exp \{i\langle x,u \rangle + i\langle y,v \rangle \} \mathcal{P}_{xy}(dx,dy)
\]

(7)

Thus, we have

\[
m \over i \over d\xi^2 \left( \delta \mathcal{F}_{xy}(u,v) \over \delta u_\xi + c \over i \over d\xi \delta \mathcal{F}_{xy}(u,v) \over \delta u_\xi + k \delta \mathcal{F}_{xy}(u,v) \over \delta u_\xi + a \delta^3 \mathcal{F}_{xy}(u,v) \over \delta u_\xi^3 \right) = \int_{\mathbb{C}^n} \int_{\mathbb{C}^n} \left\{ m \over d\xi^2 \left( \delta \mathcal{F}_{xy}(u,v) \over \delta u_\xi + c \over d\xi \delta \mathcal{F}_{xy}(u,v) \over \delta u_\xi + k \delta \mathcal{F}_{xy}(u,v) \over \delta u_\xi + a \delta^3 \mathcal{F}_{xy}(u,v) \over \delta u_\xi^3 \right) \right\} \exp \{i\langle x,u \rangle + i\langle y,v \rangle \} \mathcal{P}_{xy}(dx,dy)
\]

(8)

Using the equation (1) governing the system we will have

\[
m \over i \over d\xi^2 \left( \delta \mathcal{F}_{xy}(u,v) \over \delta u_\xi + c \over i \over d\xi \delta \mathcal{F}_{xy}(u,v) \over \delta u_\xi + k \delta \mathcal{F}_{xy}(u,v) \over \delta u_\xi + a \delta^3 \mathcal{F}_{xy}(u,v) \over \delta u_\xi^3 \right) = \int_{\mathbb{C}^n} \int_{\mathbb{C}^n} y(x) \exp \{i\langle x,u \rangle + i\langle y,v \rangle \} \mathcal{P}_{xy}(dx,dy)
\]

(9)

Hence, the Functional Differential equation for the Duffing Oscillator will have the form

\[
m \over i \over d\xi^2 \delta \mathcal{F}_{xy}(u,v) \over \delta u_\xi + c \over i \over d\xi \delta \mathcal{F}_{xy}(u,v) \over \delta u_\xi + k \delta \mathcal{F}_{xy}(u,v) \over \delta u_\xi + a \delta^3 \mathcal{F}_{xy}(u,v) \over \delta u_\xi^3 + \frac{a}{i^3} \left[ \delta \mathcal{F}_{xy}(u,v) \over \delta u_\xi \right]^3 = \frac{1}{i} \delta \mathcal{F}_{xy}(u,v) \over \delta v_\xi
\]

(10)

Concerning the initial conditions of the problem we note that

\[
\mathcal{F}_{xy}(v_1_\delta_0(\bullet) + v_2_\delta_0'(\bullet),0) = \int_{\mathbb{C}^n} \int_{\mathbb{C}^n} \exp \{i\langle x,v_1_\delta_0(\bullet) + v_2_\delta_0'(\bullet) + i\langle y,0 \rangle \} \mathcal{P}_{xy}(dx,dy) = \int_{\mathbb{C}^n} \exp i\langle x,t_0 \rangle + v_2 \dot{x}(t_0) \rangle \mathcal{P}_{x}(dx) = \mathcal{F}_0(v_1,v_2)
\]

(11)
5.2.2. Van der Poll Oscillator

The van der Pol oscillator has the following equation of motion
\[
m\ddot{x}(t;\beta)+c\left(x(t;\beta)^2-1\right)\dot{x}(t;\beta)+kx(t;\beta) = y(t;\beta) \quad (12)
\]
x(t₀;\beta) = x₀(\beta) \quad \text{and} \quad \dot{x}(t₀;\beta) = \dot{x₀}(\beta) \quad \text{a.s.} \quad (13)

Assuming that \( y(t;\beta) \) is a stochastic process defined onto the probability space \((\Omega, \mathcal{B}(\Omega), \mathbb{P})\) with the associated probability space \((C_\infty^\omega, \mathcal{B}(C_\infty^\omega), \mathbb{P})\), there is a unique joint Characteristic Functional
\[
\mathcal{F}_{xy}(u,v) = \int \int \exp \left\{ i \langle x, u \rangle + i \langle y, v \rangle \right\} \mathcal{F}_{xy} \left( dx, dy \right)
\]

Similar calculations on the characteristic functional, as before, shows that
\[
\frac{m \cdot d^2}{i \cdot d\xi^2} \frac{\delta \mathcal{F}_{xy}(u,v)}{\delta u(\xi)} + \frac{c}{i^3} \cdot \frac{d}{dt} \frac{\delta^3 \mathcal{F}_{xy}(u,v)}{\delta u(t) \delta u(\xi)^3} \bigg|_{t=\xi} - \frac{c}{i \cdot d\xi} \frac{\delta \mathcal{F}_{xy}(u,v)}{\delta u(\xi)} + k \cdot \frac{\delta \mathcal{F}_{xy}(u,v)}{\delta u(\xi)} = \int_{C_\infty^\omega \times C_\infty^\omega} \left\{ m \frac{d^2 x(\xi)}{d\xi^2} + c \left( x(\xi)^2 - 1 \right) \frac{dx(\xi)}{d\xi} + kx(\xi) \right\} e^{i\langle x, u \rangle + i\langle y, v \rangle} \mathcal{F}_{xy} \left( dx, dy \right) \quad (14)
\]

By using the equation (12) we will have
\[
\frac{m \cdot d^2}{i \cdot d\xi^2} \frac{\delta \mathcal{F}_{xy}(u,v)}{\delta u(\xi)} + \frac{c}{i^3} \cdot \frac{d}{dt} \frac{\delta^3 \mathcal{F}_{xy}(u,v)}{\delta u(t) \delta u(\xi)^3} \bigg|_{t=\xi} - \frac{c}{i \cdot d\xi} \frac{\delta \mathcal{F}_{xy}(u,v)}{\delta u(\xi)} + k \cdot \frac{\delta \mathcal{F}_{xy}(u,v)}{\delta u(\xi)} = \int_{C_\infty^\omega \times C_\infty^\omega} y(\xi) e^{i\langle x, u \rangle + i\langle y, v \rangle} \mathcal{F}_{xy} \left( dx, dy \right) \quad (15)
\]

Hence, the Functional Differential equation for the Duffing Oscillator will have the form
\[
\frac{m \cdot d^2}{i \cdot d\xi^2} \frac{\delta \mathcal{F}_{xy}(u,v)}{\delta u(\xi)} + \frac{c}{i^3} \cdot \frac{d}{dt} \frac{\delta^3 \mathcal{F}_{xy}(u,v)}{\delta u(t) \delta u(\xi)^3} \bigg|_{t=\xi} - \frac{c}{i \cdot d\xi} \frac{\delta \mathcal{F}_{xy}(u,v)}{\delta u(\xi)} + k \cdot \frac{\delta \mathcal{F}_{xy}(u,v)}{\delta u(\xi)} = \frac{1}{i} \frac{\delta \mathcal{F}_{xy}(u,v)}{\delta \nu_\xi} \quad (16)
\]

Concerning the initial conditions of the problem we note that
\[
\mathcal{F}_{xy} \left( v_1 \delta_\nu(\bullet) + v_2 \delta_\nu^\prime(\bullet), 0 \right) = \int \int \exp \left\{ i \langle x, v_1 \delta_\nu(\bullet) + v_2 \delta_\nu^\prime(\bullet) \rangle + i \langle y, 0 \rangle \right\} \mathcal{F}_{xy} \left( dx, dy \right) = \mathcal{F}_{0,v_1,v_2} = \mathcal{F}_{x}(dx) = \mathcal{F}_0(v_1, v_2) \quad (17)
\]
5.2.3. Non-Autonomous Dynamical Systems under Stochastic Excitation I – Case of m.s. Continuous Excitation

Let \( (\Omega, \mathcal{B}(\Omega), \mathcal{F}) \) be a probability space and \( C_c^{\infty}(I) \), \( I = [t_0, \infty) \) be the separable, real, inner product space of infinitely differentiable functions with values in \( U \subset \mathbb{R}^n \). \( \mathcal{B}(C_c^{\infty}) \) will denote the \( \sigma \) – field generated by the open subsets of \( C_c^{\infty}(I) \). Let the dynamical system

\[
\dot{x}(t; \beta) = F(t, x(t; \beta), y(t; \beta)) \quad \text{a.s.} \tag{18}
\]

(\( \Sigma_1 \))

\[
x(t_0; \beta) = x_0(\beta), \quad \text{a.s.} \tag{19}
\]

where,

a) \( y(\cdot, \cdot): I \times \Omega \rightarrow \mathbb{R}^M \) is a mean square continuous stochastic process defined onto the probability space \( (\Omega, \mathcal{B}(\Omega), \mathcal{F}) \) with the associated probability space \( (C_c^{\infty}, \mathcal{B}(C_c^{\infty}), \mathcal{F}_y) \), and with characteristic functional \( \mathcal{F}_y \).

b) \( x_0(\cdot): \Omega \rightarrow \mathbb{R}^N \) is a random variable with the associated probability space \( (\mathbb{R}^M, \mathcal{B}(\mathbb{R}^M), \mathcal{F}_{x_0}) \) with characteristic function \( \mathcal{F}_{x_0} \).

c) \( F(\cdot, \cdot, \cdot): I \times \mathbb{R}^N \times \mathbb{R}^M \rightarrow \mathbb{R}^N \) is a measurable, deterministic function that can be expanded into a series of the form

\[
F(t, x, y) = \sum_{k=0}^{K} \sum_{l=0}^{L} \left( \nabla_y \nabla_x^l F(t, x_0, y_0) \right) \circ \left( y - y_0 \right) \left( x - x_0 \right) \tag{20}
\]

where \( \nabla_m \nabla_n^m \) is the tensor product \( \nabla_y \otimes \cdots \otimes \nabla_y \otimes \nabla_x \otimes \cdots \otimes \nabla_x \) and \( \circ \) is the tensor product \( (y - y_0) \otimes \cdots \otimes (y - y_0) \otimes (x - x_0) \otimes \cdots \otimes (x - x_0) \) and the symbol \( \circ \) denotes the inner product between the two tensors. For the sequel we assume, for simplicity that \( x_0 = 0 \) and \( y_0 = 0 \).

**Theorem 2.1** : The problem \( \Sigma_1 \) is equivalent to the functional differential equation

\[
\frac{1}{i} D_t \mathcal{F}_{xy}(u, v) = \sum_{k=0}^{K} \sum_{l=0}^{L} \left( \nabla_y \nabla_x^l F(t, x_0, y_0) \right) \circ \left( y - y_0 \right) \left( x - x_0 \right) \mathcal{F}_{xy}(u, v) \tag{21}
\]

\[
\mathcal{F}_{xy}(v, \delta_u, 0) = \mathcal{F}_0(v), \quad v \in \mathbb{R}^N \tag{22}
\]

for every \( u \in C_c^{\infty}(I), v \in C_c^{\infty}(I) \) and \( t \in I, \quad I = [t_0, \infty) \). Where \( \mathcal{F}_{xy} \) is the joint characteristic functional of the s.p.’s \( x(t; \beta), y(t; \beta), t \in I \)

\[
\nabla_{u(t)} = \left[ \frac{\delta \mathcal{F}_{xy}}{\delta u_1(\xi)}, \frac{\delta \mathcal{F}_{xy}}{\delta u_2(\xi)}, \ldots, \frac{\delta \mathcal{F}_{xy}}{\delta u_N(\xi)} \right] \quad \text{and} \quad \nabla_{v(t)} = \left[ \frac{\delta \mathcal{F}_{xy}}{\delta v_1(\xi)}, \frac{\delta \mathcal{F}_{xy}}{\delta v_2(\xi)}, \ldots, \frac{\delta \mathcal{F}_{xy}}{\delta v_M(\xi)} \right]
\]
are the Volterra gradients of the functional $\mathcal{F}_{xy}$ with respect to every argument $u, v$ and
\[ \nabla u(i) \nabla v(k) \] is the operator defined by the $(l, k)$-tensor product of the operators $\nabla u(\xi)$ and $\nabla v(\xi)$, i.e.
\[ \nabla u(i) \nabla v(k) = \nabla u(i) \otimes \cdots \otimes \nabla u(i) \otimes \nabla v(i) \cdots \otimes \nabla v(i) \]
With $\mathbf{v} \cdot \delta_0(\bullet) = (v_1, v_2, \ldots, v_N) \cdot \delta (\bullet - t_0)$ we denote the vector valued Dirac delta function centred at zero.

**Proof:** Let $\mathcal{P}$ be the joint probability measure. Multiplying every equation
\[ \dot{x}_n(t; \beta) = F_n(t, x(t; \beta), y(t; \beta)) \quad n = 1, 2, \ldots, N \]
with $\exp \left\{ i \langle x, u \rangle + i \langle y, v \rangle \right\}$ and integrating over $C_{c}^{\infty, N} \times C_{c}^{\infty, M}$ with respect to the probability measure $\mathcal{P}$ we will have for every $n = 1, 2, \ldots, N$,
\[
\int_{C_{c}^{\infty, N} \times C_{c}^{\infty, M}} \dot{x}_n(t; \beta) \exp \left\{ i \langle x, u \rangle + i \langle y, v \rangle \right\} \mathcal{P}(dx, dy) = \int_{C_{c}^{\infty, N} \times C_{c}^{\infty, M}} F_n(t, x, y) \exp \left\{ i \langle x, u \rangle + i \langle y, v \rangle \right\} \mathcal{P}(dx, dy) \quad (23)
\]
For the l.h.s. of the above equation we can easily notice that,
\[
\int_{C_{c}^{\infty, N} \times C_{c}^{\infty, M}} \dot{x}_n(t; \beta) \exp \left\{ i \langle x, u \rangle + i \langle y, v \rangle \right\} \mathcal{P}(dx, dy) = \frac{1}{i} \frac{d}{dt} \frac{\delta \mathcal{F}_{xy}(u, v)}{\delta u_n(t)} \quad (24)
\]
For the r.h.s. we have using (20)
\[
\int_{C_{c}^{\infty, N} \times C_{c}^{\infty, M}} F_n(t, x, y) \exp \left\{ i \langle x, u \rangle + i \langle y, v \rangle \right\} \mathcal{P}(dx, dy) = \sum_{k=0}^{K} \sum_{l=0}^{L} \frac{1}{l!} \frac{\delta^{k+l} F_n(t, 0, 0)}{\delta x_l \delta y_j} \frac{\delta^{k+l} \mathcal{F}_{xy}(u, v)}{\delta u_n(t) \delta v_j(t)} \quad (25)
\]
\[
\int_{C_{c}^{\infty, N} \times C_{c}^{\infty, M}} F_n(t, x, y) \exp \left\{ i \langle x, u \rangle + i \langle y, v \rangle \right\} \mathcal{P}(dx, dy) = \sum_{k=0}^{K} \sum_{l=0}^{L} \frac{1}{l!} \frac{\delta^{k+l} F_n(t, 0, 0)}{\delta x_l \delta y_j} \frac{\delta^{k+l} \mathcal{F}_{xy}(u, v)}{\delta u_n(t) \delta v_j(t)} = \sum_{k=0}^{K} \frac{1}{l!} \frac{\delta^{k+l} \mathcal{F}_{xy}(u, v)}{\delta u_n(t) \delta v_j(t)}
\]
Where we have used the notation,
\[
\partial x_i = \partial x_i^1 \cdot \partial x_i^2 \cdots \cdot \partial x_i^N \quad \mathbf{x}^l = x_i^1 \cdot x_i^2 \cdots x_i^N
\]
Hence, we have proved the equality,

\[
\frac{1}{i} \frac{d}{dt} \delta \mathcal{F}_{xy}(u, v) \bigg|_{u(t)} = \sum_{k=0}^{K} \sum_{l=0}^{L} \left\{ \nabla_{y}^{i} \nabla_{x}^{j} F_{i}^{k} \right\} \frac{1}{i^{k+l}} \frac{\partial}{\partial u} \frac{\partial}{\partial v} \mathcal{F}_{xy}(u, v) \tag{25}
\]

or equivalently,

\[
\frac{1}{i} \frac{d}{dt} \left[ \nabla_{u(i)} \mathcal{F}_{xy}(u, v) \right] = \sum_{k=0}^{K} \sum_{l=0}^{L} \left\{ \nabla_{y}^{i} \nabla_{x}^{j} F_{i} \right\} \frac{1}{i^{k+l}} \nabla_{u(i)} \nabla_{v(i)} \mathcal{F}_{xy}(u, v) \tag{26}
\]

Concerning the initial conditions, we have by direct calculation of the l.h.s.of equation (22)

\[
\mathcal{F}_{xy}(v \cdot \delta_{u}(\star), 0) = \int_{C_{x}^{N} \times C_{t}^{M}} \exp \left\{ i \left( x, v \cdot \delta_{u} \right) \right\} \mathbb{P}(dxdy) = \int_{C_{x}^{N} \times C_{t}^{M}} \exp \left\{ i \sum_{n=1}^{N} x_{n} v_{n} \right\} \mathbb{P}_{xy}(dxdy) = \mathcal{F}_{0}(v) \tag{27}
\]

**Example 2.2 [One - Dimensional Non-Linear Oscillator]:** To illustrate the above Theorem we will study the special case of the one-dimensional non-linear oscillator, i.e. the system

\[
m\ddot{x}(t; \beta) + f(x(t; \beta)) = y(t; \beta) \quad \text{a.s.} \quad (\Sigma 1a) \tag{28}
\]

\[
[x(t_0; \beta), \dot{x}(t_0; \beta)] = [x_0(\beta), \dot{x}_0(\beta)] \quad \text{a.s.} \tag{29}
\]

where, we have made the same hypothesis as in Theorem 5.1. We also assume that \( f(\cdot, \cdot) \) can be expanded in the form

\[
f(x_1, x_2) = \sum_{i=0}^{N} \sum_{j=0}^{M} a_{ij} x_{i} x_{j} \tag{30}
\]

Transforming the above equation into a state-space form we have

\[
\dot{x}_1(t; \beta) = x_2(t; \beta)
\]

\[
m\ddot{x}_2(t; \beta) + f(x_2(t; \beta), x_1(t; \beta)) = y(t; \beta)
\]

\[
[x_1(t_0; \beta), x_2(t_0; \beta)] = [x_0(\beta), \dot{x}_0(\beta)]
\]

By direct application of the Theorem 2.1 we will have the two functional differential equations

\[
\frac{d}{dt} \delta \mathcal{F}_{xy}(u_1, u_2, v) \bigg|_{u_1(t)} = \frac{\delta \mathcal{F}_{xy}(u_1, u_2, v)}{\delta u_1(t)} \tag{21a}
\]

\[
\frac{m}{i} \frac{d}{dt} \frac{\delta \mathcal{F}_{xy}(u_1, u_2, v)}{\delta u_2(t)} + \sum_{n=0}^{N} \sum_{m=0}^{M} a_{nm} \frac{\delta \mathcal{F}_{xy}(u_1, u_2, v)}{\delta u_2(\xi_n)} \frac{\partial}{\partial u_2(\xi_m)} \mathcal{F}_{xy}(u_1, u_2, v) \bigg|_{\xi_{n+m}} = \frac{1}{i} \frac{\delta \mathcal{F}_{xy}(u_1, u_2, v)}{\delta v(t)} \tag{21b}
\]

Substituting the first Functional Equation into the second we have,
\[ \frac{m}{i} \cdot \frac{d^2}{d\xi^2} \frac{\delta \widetilde{F}_{xy}(u,v)}{\delta u(\xi)} \bigg|_{\xi=t} + \sum_{n=0}^{N} \sum_{m=0}^{M} a_{nm} \frac{d^n}{d\xi_1 \ldots d\xi_n} \frac{\delta^{n+m} \widetilde{F}_{xy}(u,v)}{\delta u(\xi_1) \ldots \delta u(\xi_{n+m})} \bigg|_{\xi_j=t}^{\xi_j=n+m} = \frac{1}{i} \frac{\delta \widetilde{F}_{xy}(u,v)}{\delta v(t)} \]

for every \( u \in C_c^\infty(I), v \in C_c^\infty(I) \) and \( t \in I \) with

\[ \widetilde{F}_{xy} \left( v_1 \delta_{\alpha}(\bullet) + v_2 \delta_{\alpha}'(\bullet), 0 \right) = \widetilde{F}_0(v_1, v_2), \quad (v_1, v_2) \in \mathbb{R}^2 \]
5.2.4. Non-Autonomous Dynamical Systems under Stochastic Excitation II – Case of Independent Increment Excitation

Let \( \left( \Omega, \mathcal{B}(\Omega), \mathcal{P} \right) \) be a probability space and \( C_c^{\infty,N}(I) \) be the separable, real, inner product space of square integrable functions with values in \( U \subset \mathbb{R}^N \). \( \mathcal{B}\left(C_c^{\infty,N}(I)\right) \) will denote the \( \sigma \) – field generated by the open subsets of \( C_c^{\infty,N}(I) \). Let the dynamical system

\[
\begin{align*}
 d_x(t;\beta) &= F\left(t,x(t;\beta)\right)dt + G\left(t,x(t;\beta)\right)d_y(t;\beta) \quad \text{a.s.} \\
x(t_0;\beta) &= x_0(\beta), \quad \text{a.s.}
\end{align*}
\]

(Σ2)

where,

a) \( y(\bullet;\bullet): I \times \Omega \rightarrow \mathbb{R}^M \) is a stochastic process with independent increments defined onto the probability space \( \left( \Omega, \mathcal{B}(\Omega), \mathcal{P} \right) \) with the associated probability space \( \left( C_c^{\infty,M}, \mathcal{B}\left(C_c^{\infty,M}\right), \mathcal{P}_y \right) \), and with characteristic functional \( \mathcal{F}_y \).

b) \( x_0(\bullet): \Omega \rightarrow \mathbb{R}^N \) is a random variable with the associated probability space \( \left( \mathbb{R}^N, \mathcal{B}\left(\mathbb{R}^N\right), \mathcal{P}_{x_0} \right) \) with characteristic function \( \mathcal{F}_{x_0} \).

c) \( F(\bullet,\bullet): T \times \mathbb{R}^N \rightarrow \mathbb{R}^N \) is a measurable, deterministic function that can be expanded into a series of the form

\[
F\left(t,x\right) = \sum_{i=0}^I \left\{ \nabla^i_x F\left(t,x_0\right) \right\} \circ \left( \otimes (x-x_0) \right)
\]

(33)

d) \( G(\bullet,\bullet): T \times \mathbb{R}^N \rightarrow \mathbb{R}^{N \times M} \) is a measurable, deterministic function that can be expanded into a series of the form

\[
G\left(t,x\right) = \sum_{k=0}^K \left\{ \nabla_x^k G\left(t,x_0\right) \right\} \circ \left( \otimes (x-x_0) \right)
\]

(34)

where \( \nabla_x^i \) is the tensor product \( \nabla_x \otimes \cdots \otimes \nabla_x \), \( \otimes (x-x_0) \) is the tensor product \( (x-x_0) \otimes \cdots \otimes (x-x_0) \) and the symbol \( \circ \) denotes the inner product between the two tensors.

For the sequel we assume, for simplicity that \( x_0 = 0 \) and \( y_0 = 0 \).

**Remark 2.3**: We must note that since \( y(\bullet;\bullet): I \times \Omega \rightarrow \mathbb{R}^m \) is a stochastic process with independent increments, it has nowhere differentiable sample functions. This explains the fact that the equation describing the dynamical system is in terms of differentials. ■
Theorem 2.4: The problem $\Sigma 2$ defined above, is equivalent with the functional differential equation
\begin{equation}
\frac{1}{i} d_t \left[ \nabla_{u(t)} \mathcal{F}_x (u, v) \right] = \sum_{l=0}^{L} \left\{ \nabla_{x}^{l} F (t, 0) \right\} \frac{1}{i} \nabla_{u(t)}^{l} \mathcal{F}_x (u, v) dt + \sum_{k=0}^{K} \left\{ \nabla_{x}^{k} G (t, 0) \right\} \frac{1}{i} \nabla_{u(t)}^{k} \mathcal{F}_x (u, v) \right] \end{equation}

for every $u \in C_{c}^{\infty, N} (I), v \in C_{c}^{\infty, M} (I)$ and $t \in I$. Where $\mathcal{F}_x$ is the joint characteristic functional of the s.p.’s $x(t; \beta), y(t; \beta), t \in I$

\[ \nabla_{u(t)} = \left[ \frac{\delta \mathcal{F}_x}{\delta u_1 (t)}, \frac{\delta \mathcal{F}_x}{\delta u_2 (t)}, \ldots, \frac{\delta \mathcal{F}_x}{\delta u_N (t)} \right] \]

are the Volterra gradients of the functional $\mathcal{F}_x$ with respect to every argument $u, v$ and $\nabla_{u(t)}^{l}$ is the operator defined by the $l$-tensor product of the operator $\nabla_{u(t)}^{l}$, i.e.

\[ \nabla_{u(t)}^{l} = \nabla_{u(t)} \otimes \ldots \otimes \nabla_{u(t)} \]

are the Dirac delta functions centred at zero.

**Proof**: Let $\mathcal{P}_{xy}$ be the joint probability measure. Multiplying every equation

\[ d_x x_n (t; \beta) = F_n (t, x(t; \beta)) dt + \sum_{m=1}^{M} G_{nm} (t, x(t; \beta)) d_y y_m (t; \beta) \quad n = 1, 2, \ldots, N \]

with $\exp \{ i \langle x, u \rangle + i \langle y, v \rangle \}$ and integrating over $C_{c}^{\infty, N} \times C_{c}^{\infty, M}$ with respect to the probability measure $\mathcal{P}_{xy}$ we will have for every $n = 1, 2, \ldots, N$,

\[ \int_{C_{c}^{\infty, N} \times C_{c}^{\infty, M}} d_x x_n (t) \exp \{ i \langle x, u \rangle + i \langle y, v \rangle \} \mathcal{P}_{xy} (dx, dy) = \]

\[ = \int_{C_{c}^{\infty, N} \times C_{c}^{\infty, M}} F_n (t, x) dt \exp \{ i \langle x, u \rangle + i \langle y, v \rangle \} \mathcal{P}_{xy} (dx, dy) + \]

\[ + \int_{C_{c}^{\infty, N} \times C_{c}^{\infty, M}} \sum_{m=1}^{M} G_{nm} (t, x) d_y y_m (t; \beta) \exp \{ i \langle x, u \rangle + i \langle y, v \rangle \} \mathcal{P}_{xy} (dx, dy) \]

For the l.h.s of the above equation we can easily notice that,

\[ \int_{C_{c}^{\infty, N} \times C_{c}^{\infty, M}} d_x x_n (t; \beta) \exp \{ i \langle x, u \rangle + i \langle y, v \rangle \} \mathcal{P}_{xy} (dx, dy) = \frac{1}{i} d_t \frac{\delta \mathcal{F}_x (u, v)}{\delta u_n (t)} \]

For the first term of the r.h.s we have using (33)

\[ \int_{C_{c}^{\infty, N} \times C_{c}^{\infty, M}} F_n (t, x) \exp \{ i \langle x, u \rangle + i \langle y, v \rangle \} \mathcal{P}_{xy} (dx, dy) = \]
\[\begin{align*}
&= \int_{C^\mathcal{L} \times C^\mathcal{L}} \sum_{l=0}^{L} \left\{ \nabla_x F_n(t,0) \right\} \circ \left( \otimes \mathbf{X} \right) \exp \left\{ i \langle \mathbf{x}, \mathbf{u} \rangle + i \langle \mathbf{y}, \mathbf{v} \rangle \right\} \mathcal{P}_{xy}(dx,dy) = \\
&= \int_{C^\mathcal{L} \times C^\mathcal{L}} \sum_{l=0}^{L} \sum_{i \in J} \frac{\partial^l F_n(t,0)}{\partial x_i} \cdot \mathbf{x}^i \exp \left\{ i \langle \mathbf{x}, \mathbf{u} \rangle + i \langle \mathbf{y}, \mathbf{v} \rangle \right\} \mathcal{P}_{xy}(dx,dy) = \\
&= \sum_{l=0}^{L} \frac{1}{i} \sum_{i \in J} \frac{\partial^l F_n(t,0)}{\partial x_i} \int_{C^\mathcal{L} \times C^\mathcal{L}} i^l \mathbf{x}^l \exp \left\{ i \langle \mathbf{x}, \mathbf{u} \rangle + i \langle \mathbf{y}, \mathbf{v} \rangle \right\} \mathcal{P}_{xy}(dx,dy) = \\
&= \sum_{l=0}^{L} \frac{1}{i} \sum_{i \in J} \frac{\partial^l F_n(t,0)}{\partial x_i} \cdot \delta^l \mathcal{F}(\mathbf{u},\mathbf{v}) = \sum_{l=0}^{L} \left\{ \nabla_x F_n(t,0) \right\} \frac{1}{i^l} \nabla_{u(t)}^l \mathcal{F}(\mathbf{u},\mathbf{v}).
\end{align*}\]

Where we have used the notation,
\[\partial x^i = \partial x^1 \cdot \partial x^2 \cdots \partial x^N \]
\[\mathbf{x}^l = x^1_l \cdot x^2_l \cdots x^N_l\]

For the second term of the r.h.s. we have using (34)
\[\begin{align*}
&= \int_{C^\mathcal{L} \times C^\mathcal{L}} \sum_{m=0}^{M} \sum_{k=0}^{K} \frac{\partial^k G_{nm}(t,0)}{\partial x_i} \cdot \mathbf{y}^k \exp \left\{ i \langle \mathbf{x}, \mathbf{u} \rangle + i \langle \mathbf{y}, \mathbf{v} \rangle \right\} \mathcal{P}_{xy}(dx,dy) = \\
&= \sum_{m=0}^{M} \sum_{k=0}^{K} \sum_{l=0}^{L} \frac{\partial^k G_{nm}(t,0)}{\partial x_i} \int_{C^\mathcal{L} \times C^\mathcal{L}} i^k \mathbf{y}^k \exp \left\{ i \langle \mathbf{x}, \mathbf{u} \rangle + i \langle \mathbf{y}, \mathbf{v} \rangle \right\} \mathcal{P}_{xy}(dx,dy) = \\
&= \sum_{m=0}^{M} \sum_{k=0}^{K} \frac{1}{i^k} \sum_{l=0}^{L} \frac{\delta^k G_{nm}(t,0)}{\partial x_i} \frac{\delta^k}{\delta u_j} \left| \frac{d}{dt} \mathcal{F}_{xy}(\mathbf{u},\mathbf{v}) \right| = \\
&= \sum_{m=0}^{M} \sum_{k=0}^{K} \left\{ \nabla_x G_{nm}(t,0) \right\} \frac{1}{i^k} \nabla_{u(t)}^k \left| \frac{d}{dt} \mathcal{F}_{xy}(\mathbf{u},\mathbf{v}) \right|.
\end{align*}\]

Hence we have proved the equality,
\[\frac{1}{i} \mathbf{d}_i \frac{\delta \mathcal{F}_{xy}(\mathbf{u},\mathbf{v})}{\delta u_j}(t) = dt \sum_{l=0}^{L} \left\{ \nabla_x F_n(t,0) \right\} \frac{1}{i^l} \nabla_{u(t)}^l \mathcal{F}_{xy}(\mathbf{u},\mathbf{v}) + \sum_{m=0}^{M} \sum_{k=0}^{K} \left\{ \nabla_x G_{nm}(t,0) \right\} \frac{1}{i^k} \nabla_{u(t)}^k \left| \frac{d}{dt} \mathcal{F}_{xy}(\mathbf{u},\mathbf{v}) \right| \]

or equivalently,
\[ \frac{1}{i} d_t \left[ \nabla_u(\mathcal{F}_{xy}(u,v)) \right] = \sum_{l=0}^{L} \left\{ \nabla_x^l F(t,0) \right\} \frac{1}{i} \nabla_u(\mathcal{F}_{xy}(u,v)) dt + \]
\[ + \sum_{k=0}^{K} \left\{ \nabla_x^k G(t,0) \right\} \nabla_u(\mathcal{F}_{xy}(u,v)) \]

Concerning the initial conditions, we have by direct calculation of the l.h.s. of equation (36)

\[ \mathcal{F}_{xy}(\nu \cdot \delta_t, 0) = \int_{\mathbb{C}^N \times \mathbb{C}^M} \exp \left\{ i \langle x, \nu \cdot \delta_t \rangle \right\} \mathcal{P}_{xy}(dx, dy) = \]
\[ = \int_{\mathbb{C}^N \times \mathbb{C}^M} \exp \left\{ i \sum_{n=1}^{N} x_n \nu_n \right\} \mathcal{P}_{xy}(dx, dy) = \mathcal{F}_0(\nu) \]  

(37)

**Example 2.5 [Linear Filtering System]**: Let us consider as a special case of the above theorem the a linear time invariant system, i.e. a system of the form,

\[ d_x(t; \beta) = F \cdot x(t; \beta) dt + G \cdot d_w(t; \beta) \text{ a.s.} \]  

(38)

\[ x(t_0; \beta) = x_0(\beta), \text{ a.s.} \]  

(39)

where \( w(\cdot; \cdot) : I \times \Omega \rightarrow \mathbb{R}^M \) is the standard Wiener process and \( F \in \mathbb{R}^{N \times N} \) and \( G \in \mathbb{R}^{N \times M} \). The above system is very essential for linear system theory and linear filter design, since we have explicit expression for the spectrum of the response in terms of the excitation spectrum (Wiener-Khinchin relations). For a detailed discussion we refer to PUGACHEV, V.S. & SINITSYN, (Stochastic Differential Systems). Now, let us derivate the associated functional differential equation for the joint characteristic functional of the response and the excitation. By direct application of Theorem 2.4 we have the functional differential equation

\[ d_t \left[ \nabla_u(\mathcal{F}_{xy}(u,v)) \right] = F \cdot \nabla_u(\mathcal{F}_{xy}(u,v)) dt + G \cdot d_t \nabla_v(\mathcal{F}_{xy}(u,v)) \]  

(40)

\[ \mathcal{F}_{xy}(\nu \cdot \delta (t-t_0), 0) = \mathcal{F}_0(\nu), \nu \in \mathbb{R}^N \]  

(41)

for every \( u \in C_c^\infty(I), v \in C_c^\infty(I) \) and \( t \in I \). Where \( \mathcal{F}_{xy} \) is the joint characteristic functional of the s.p.’s \( x(t; \beta), w(t; \beta), t \in I \).  

\[ \blacksquare \]
5.3. Reduction of the Hopf-Type Functional Differential Equation for Special Cases

In this section we will study, how the general Functional Differential Equation can be reduced to partial differential equations when the external excitation fulfills special conditions. We will start our discussion with the case of deterministic systems (absence of stochastic excitation), with random initial conditions. Using the Hopf-type FDE we will prove Liouville Equation for the characteristic function of the response.

5.3.1. Systems with Random Initial Conditions – Liouville Equation

Let \( (Ω, \mathcal{B}(Ω), \mathcal{P}) \) be a probability space and \( C^\infty_c(I) \) be the separable, real, inner product (inner product) space of infinitely differentiable functions with values in \( U \subset \mathbb{R}^N \). \( \mathcal{B}(C^\infty_c(I)) \) will denote the \( σ \)– field generated by the open subsets of \( C^\infty_c(I) \). Let the dynamical system

\[
\dot{x}(t; β) = F(t, x(t; β)) \quad \text{a.s.} \quad (1)
\]

\[
x(t_0; β) = x_0(β), \quad \text{a.s.} \quad (2)
\]

where,

a) \( x_0(•): Ω \rightarrow \mathbb{R}^N \) is a random variable with the associated probability space \( (\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \mathcal{P}_{x_0}) \) with characteristic function \( \varphi_0(υ), \ υ \in \mathbb{R}^N \).

b) \( F(•, •): I \times \mathbb{R}^N \rightarrow \mathbb{R}^N \) is a measurable, deterministic function that can be expanded into a series of the form

\[
F(t, x) = \sum_{l=0}^{K} \left\{ \nabla_x^l F(t, x_0) \right\} \circ \left\{ \otimes (x - x_0) \right\} \quad (3)
\]

For the sequel we assume, for simplicity that \( x_0 = 0 \).

For the case where an explicit solution of the above system of equations can be found, the solution would have the following form

\[
x(t; β) = g(t, x_0(β)) \quad (4)
\]

The mapping (4) characterize at arbitrary time \( t \) the state of a system which has an initial state described by the random variable (2). In other words, these functions represent at time \( t \) a transformation of the random vector (2). It is clear that the probability distribution of the response can be obtained by use of the theorem of transformations of random vectors.

Let us assume that the transformation defined by (4) is continuous with respect to \( x_0(β) \) (for each \( t \)), has continuous partial derivatives with respect to \( x_0(β) \), and defines a one-to-one mapping. Then, if the inverse transformation is written as

\[
x_0(β) = h[t, x(t; β)] \quad (5)
\]

the probability density of the response takes the form
5.3. REDUCTION OF THE HOPF-TYPE FUNCTIONAL DIFFERENTIAL EQUATIONS FOR SPECIAL CASES

For the case where no explicit solution of the (1) can be found, we return to the Hopf-type FDE associated with the stochastic initial value problem described above. Using Theorem 2.1 we have

\[
\frac{d}{dt} \left[ \nabla_{u(t)} \mathcal{F}_x (u) \right] = \sum_{i=0}^{k} \left\{ \nabla_x F_i (t, 0) \right\} \frac{1}{i!} \nabla_{u(t)} \mathcal{F}_x (u)
\]

(7)

\[
\mathcal{F}_x (v, \delta_0, 0) = \mathcal{F}_0 (v), \quad v \in \mathbb{R}^N
\]

(8)

for every \( u \in C_c^{\infty, N} (I) \) and \( t \in I \). Where \( \mathcal{F}_x \) is the characteristic functional of the s.p. \( x(t; \beta), t \in I \). For the case described above the following theorem holds

**Theorem 3.1 [Liouville equation]**: Let the dynamical system \( \Sigma_3 \) described by (1)-(2). Then the FDE (7)-(8) which describes the system can be reduced to the equation

\[
\frac{1}{i!} \nabla_x \phi_x (0, v) = \mathcal{F}_0 (v), \quad v \in \mathbb{R}^N
\]

(10)

where \( \phi(t, v) \) is the characteristic function of the response. ■

**Proof**: Let’s consider the component wise form of the FDE (9), used for the derivation at Theorem 2.1.

\[
\frac{1}{i!} \nabla_x \phi_x (t, v) + \sum_{n=1}^{N} \sum_{i=0}^{k} \frac{\partial^n F_n (t, 0)}{\partial x_i^{(n)}} \delta \mathcal{F}_x (u) = 0, \quad v \in \mathbb{R}^N, \quad t \in I
\]

(9)

Setting to the system of functional differential equations \( u = (u_1, u_2, \ldots, u_N) \cdot \delta (\bullet - t) \) and substituting \( \mathcal{F}_x (u) = \int_{C_c^{\infty}} \exp \left\{ i \int (x, u) \right\} \mathcal{P}_x (dx) \) we have

\[
\int_{C_c^{\infty,N}} \hat{x}_n (t) \exp \left\{ i \sum_{k=1}^{N} x_k (t) u_k \right\} \mathcal{P}_x (dx) = \sum_{j=1}^{n} \frac{\partial^n F_n (t, 0)}{\partial x_j^{(n)}} \int_{C_c^{\infty,N}} i \int (x(t))' \exp \left\{ i \sum_{k=1}^{N} x_k (t) u_k \right\} \mathcal{P}_x (dx) \Leftrightarrow
\]

Where we have used the notation,

\[
\partial x_j = \partial x_k^1, \partial x_k^2, \ldots, \partial x_k^N
\]

\[
x_i = x_i^1, x_i^2, \ldots, x_i^N
\]

Multiplying both sides with \( i u_n \) and summing up for all \( n = 1, 2, \ldots, N \), we will have,

\[
\sum_{n=1}^{N} \int_{C_c^{\infty,N}} i u_n \hat{x}_n (t) \exp \left\{ i \sum_{k=1}^{N} x_k (t) u_k \right\} \mathcal{P}_x (dx) =
\]
\[ = \sum_{n=1}^{N} i v_n \sum_{l=0}^{L} \int \frac{\partial^l F_n(t,0)}{\partial x_l} \cdot \int_{C_n} i^l [x(t)]^l \exp \left[ i \sum_{k=1}^{n} x_k(t) v_k \right] \mathcal{P}_x(dx) \leftrightarrow \]

But
\[ \sum_{n=1}^{N} \int \int_{C_n} i v_n \dot{x}_n(t) \exp \left[ i \sum_{k=1}^{n} x_k(t) v_k \right] \mathcal{P}_x(dx) = \frac{\partial \phi_x(t,v)}{\partial t} \]

and
\[ \sum_{n=1}^{N} \sum_{l=0}^{L} \int_{C_n} i^l [x(t)]^l \exp \left[ i \sum_{k=1}^{n} x_k(t) v_k \right] \mathcal{P}_x(dx) = \]

\[ = \sum_{n=1}^{N} \sum_{l=0}^{L} i v_n \sum_{l=0}^{L} \int_{C_n} i^l [x(t)]^l \exp \left[ i \sum_{k=1}^{n} x_k(t) v_k \right] \mathcal{P}_x(dx) = \]

Hence, we have proved the equality,
\[ \frac{\partial \phi_x(t,v)}{\partial t} = \sum_{n=1}^{N} \sum_{l=0}^{L} i v_n \sum_{l=0}^{L} \frac{\partial^l F_n(t,0)}{\partial x_l} \cdot \frac{\partial^l \phi_x(t,v)}{\partial v_l} \] (11)

or equivalently,
\[ \frac{\partial \phi_x(t,v)}{\partial t} + \sum_{n=1}^{N} \sum_{l=0}^{L} v_n \left\{ \nabla^l F_n(t,0) \right\} \cdot \left\{ \nabla^l \phi_x(t,v) \right\} = 0 \quad \blacksquare \] (12)

Remark 3.2: Taking the Fourier transform of the equation (11) or (12) we can calculate that the equation governing the probability density of the response is
\[ \frac{\partial f_x(t,x)}{\partial t} + \sum_{n=1}^{N} \frac{\partial}{\partial x_n} \left[ f_x(t,x) F_n(t,x) \right] = 0, \quad t \in I, \ x \in \mathbb{R}^N \] (13)

which is the Liouville equation from the general theory of dynamical systems. It is worth emphasizing the fact that equations of the form (13) expresses, in general conservation of some quantities in time. For example in continuum mechanics, such equation is known as the continuity equation, the equation of conservation of mass or the transport equation. Here, the Liouville equation expresses the conservation of probability during a motion of the dynamical system. \quad \blacksquare
5.3. REDUCTION OF THE HOPF-TYPE FUNCTIONAL DIFFERENTIAL EQUATIONS FOR SPECIAL CASES

5.3.2. Moment Equations from the FDE

As we mention at the beginning a very important property of the formulation based on the characteristic functional, is that we can recover all the statistical information of the problem. In this section we will show, how we can derive the set of moment equations, associated with the stochastic differential equation, from the functional differential equation. Moreover, based on this analysis we will derive an equivalent set of integrodifferential equations with the functional differential equation derived at Theorem 2.1. As we prove at the previous section the stochastic differential equation

\[\Sigma_{1}(t) = \int_{0}^{t} F(x(u(t)), x(0), 0) \, du\]

for every \(u \in C_{c}^{\infty}(I), v \in C_{c}^{\infty}(I)\) and \(t \in I\).

First we recall some important results concerning the polynomial expansion of functionals.

**Theorem 3.3 [Taylor’s formula with integral reminder]:** Let \(G: \mathcal{U} = \{C_{c}^{\infty} \times C_{c}^{\infty}\} \to \mathbb{R}\) be a functional of class \(C^{r+1}\). If the interval \([a, a+h]\) is contained in \(\mathcal{U}\) then

\[G(a+h) = G(a) + \delta G(a)[h] + \frac{1}{2} \delta^2 G(a)[h,h] + \ldots + \frac{1}{N!} \delta^r G(a)[h,\ldots,h] + \int_{0}^{1} \frac{(1-t)^{r}}{N!} \delta^{r+1} G(a+th)[h,\ldots,h] \, dt\]

**Proof:** For the proof we refer to CARTAN, H., (Differential Calculus. p.70). ■

**Theorem 3.4 [Taylor’s formula with Lagrangian reminder]:** Let the \(C^{r+1}\) functional \(G: \mathcal{U} = \{C_{c}^{\infty} \times C_{c}^{\infty}\} \to \mathbb{R}\). If

\[\|G^{r+1}(a)\| \leq M \text{ for } a \in \mathcal{U}\]

then

\[\|G(a+h) - G(a) - \delta G(a)[h] - \ldots - \frac{1}{N!} \delta^r G(a)[h,\ldots,h]\| \leq M \frac{\|h\|^{r+1}}{(r+1)!}\]

**Proof:** For the proof we refer to CARTAN, H., (Differential Calculus. p.70). ■

Now consider the functional \(G: \mathcal{U} = \{C_{c}^{\infty} \times C_{c}^{\infty}\} \to \mathbb{R}\) given by

\[G(u,v; t) = \frac{1}{i} \int_{0}^{t} d[\nabla_{u(t)} F_{xy}(u,v)] - \sum_{k=0}^{K} \sum_{l=0}^{L} \left\{ \nabla_{y} \nabla_{x} F(t,0,0) \right\} \frac{1}{i^{k+l}} \nabla_{u(t)}^{l} \nabla_{v(t)}^{k} F_{xy}(u,v)\]

for every \(t \in I\). Let us assume that \(G\) is of class \(C^{\infty}\), then, by Theorem 3.3 it can be expanded to the following form
\[ \mathcal{G}(u, v; t) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{1}{n!m!} \delta_n^m \mathcal{G}(0, 0; t)[u, \ldots, u][v, \ldots, v] \]  
(17)

for every \( t \in I \). It is clear that \( \mathcal{G}(u, v; t) = 0 \) iff

\[ \delta_n^m \mathcal{G}(0, 0; t)[u, \ldots, u][v, \ldots, v] = 0, \quad \text{for every } (n, m) \in \mathbb{Z}_+^2 \quad \text{and for every } t \in I. \]

Clearly every term \( \delta_n^m \mathcal{G}(0, 0; t)[u, \ldots, u][v, \ldots, v] \) for all \((n, m) \in \mathbb{Z}_+^2\) of the sum (17) is a homogeneous functional polynomial, i.e. there exists an \((n, m)\)-multilinear mapping

\[ \mathcal{K} : (C_{\infty,N}^c)^n \times (C_{\infty,M}^c)^m \to \mathbb{R} \]

with the symmetry property

\[ \mathcal{K}[u_1, \ldots, u_n][v_1, \ldots, v_m] = \mathcal{K}[u_1, \ldots, u_n][v_{i_1}, \ldots, v_{j_n}] \]

for every permutation \( (i_1, \ldots, i_n) \) and \( (j_1, \ldots, j_m) \), such that,

\[ \delta_n^m \mathcal{G}(0, 0; t)[u, \ldots, u][v, \ldots, v] = \mathcal{K}[u, \ldots, u][v, \ldots, v] \]

For the homogeneous functional polynomials the following important property holds,

**Preposition 3.5** : Let the homogeneous functional polynomial of degree \( n \), \( \Phi_n : C_{\infty,N}^c \to \mathbb{R} \)

described by the corresponding \( n \)-multilinear, symmetric function \( \mathcal{K} : (C_{\infty,N}^c)^n \to \mathbb{R} \). Then the polynomial \( \Phi_n : C_{\infty,N}^c \to \mathbb{R} \) is the zero polynomial iff

\[ \mathcal{K}[\delta_{t_1}, \ldots, \delta_{t_n}] = 0, \quad \text{for every } (t_1, t_2, \ldots, t_n) \text{ with } t_i \in I, \quad \text{times}\]

**Proof** : Clearly, the above assertion for a functional polynomial of degree \( n \), \( \Phi_n : C_{\infty,N}^c \to \mathbb{R} \) is equivalent with the same assertion for a functional polynomial of degree \( n \cdot N \), \( \Phi_{nN} : C_{\infty}^c \to \mathbb{R} \).

Hence, its sufficient to prove the preposition for the case when \( N = 1 \). The general form of a homogeneous functional polynomial with \( N = 1 \) will be of the form

\[ \mathcal{K}[u_1, \ldots, u_n] = \int \ldots \int K(s_1, s_2, \ldots, s_n)u_1(s_1)u_2(s_2)\ldots u_n(s_n)ds_1 \ldots ds_n \]

If

\[ \mathcal{K}[\delta_{t_1}, \delta_{t_2}, \ldots, \delta_{t_n}] = 0, \quad \text{for every } (t_1, t_2, \ldots, t_n) \text{ with } t_i \in I \]

then

\[ K(s_1, s_2, \ldots, s_n) = 0, \quad \text{for every } (s_1, s_2, \ldots, s_n) \text{ with } s_i \in I \]

Hence, \( \Phi_n \) will be the zero polynomial.

If \( \Phi_n \) is the zero polynomial then using Schwarz’s Inequality we will have

\[ \mathcal{K}[u_1, \ldots, u_n] = \left( \sum_{i=1}^{n} \int \ldots \int K(s_1, s_2, \ldots, s_n)u_i(s_i)u_2(s_2)\ldots u_n(s_n)ds_i \ldots ds_n \right) \leq \]

\[ = \prod_{i=1}^{n} \left( \int \ldots \int K(s_1, s_2, \ldots, s_n)u_i(s_i)ds_i \ldots ds_n \right) = 0 \]
Hence, we have proved the assertion.

Thus we have proved the following

**Theorem 3.6 :** Let the Functional Differential Equation

\[
\frac{1}{i} \frac{d}{dt} \left[ \nabla u(t) F_x(u, v) \right] = \sum_{k=0}^L \sum_{l=0}^L \left[ \nabla^k y \nabla^l x F(t, 0, 0) \right] \frac{1}{i^{k+l}} \nabla u(t) \nabla v(t) F_x(u, v)
\]

\[
F_x(v, \delta_0, 0) = F_0(v), \quad v \in \mathbb{R}^N
\]

for every \( u \in C_c^\infty(I), \ v \in C_c^\infty(I) \) and \( t \in I \). We assume that \( F_x \) is of class \( C^\infty \). Then the above Functional Differential Equation is equivalent with the infinite set of equations,

\[
\frac{1}{i} \frac{d}{dt} \left[ \delta^n \delta^m \nabla u(t) F_x(0, 0) \right] = \sum_{k=0}^L \sum_{l=0}^L \left[ \nabla^k y \nabla^l x F(t, 0, 0) \right] \frac{1}{i^{k+l}} \delta^n \delta^m \nabla u(t) \nabla v(t) F_x(0, 0) \left[ \delta_0, \ldots, \delta_{l}, \delta_{k}, \ldots, \delta_{n} \right]
\]

for all \((n,m) \in \mathbb{Z}^2_+\) and every \((t_1, t_2, \ldots, t_k)\) and \((s_1, s_2, \ldots, s_l)\) with \(t \in I, t_i \in I^N, s_i \in I^M\). The above set of equations will be called the corresponding moment equations of problem \( \Sigma 1 \).

**Remark 3.7 :** For the case when the first \((k,l)\) moment equations satisfied, we can get, using Theorem 3.3, useful bounds for the error of the estimation.

**Remark 3.8 :** The above Theorem presents a valuable method for solving numerically functional differential equations of the form (7)-(8).

**Example 3.9 :** Let us now apply the above Theorem to the specific case of the Duffing Oscillator. As we saw, the corresponding functional differential equation is

\[
\frac{m}{i} \frac{d^2}{dt^2} \frac{\delta F_x(u, v)}{\delta u(t)} + c \frac{d}{dt} \frac{\delta F_x(u, v)}{\delta u(t)} + k \frac{\delta F_x(u, v)}{\delta u(t)} + a \frac{1}{i} \frac{\delta^3 F_x(u, v)}{\delta u(t)^3} = \frac{1}{i} \frac{\delta F_x(u, v)}{\delta v(t)}
\]

\[
F_x(\delta_0, 0) = F_0 \quad \text{and} \quad F_x(\delta_0', 0) = \hat{F}_0
\]

for every \( u \in C_c^\infty(I), v \in C_c^\infty(I) \) and \( t \in I \). Applying Theorem 3.6 we will have the equivalent set of moment equations,

\[
\frac{m}{i} \frac{d^2}{dt^2} \frac{\delta^{n+1} F_x(0, 0)}{\delta u(t)^n} + c \frac{d}{dt} \frac{\delta^{n+1} F_x(0, 0)}{\delta u(t)^n} + k \frac{\delta^{n+1} F_x(0, 0)}{\delta u(t)^n} + a \frac{1}{i} \frac{\delta^{n+3} F_x(u, v)}{\delta u(t)^2} \frac{\delta F_x(u, v)}{\delta v(t)}
\]

\[
\frac{\delta^{n+1} F_x(0, 0)}{\delta u(t)^n} - a \frac{\delta^{n+1} F_x(0, 0)}{\delta u(t)^n} \frac{\delta^{n+1} F_x(0, 0)}{\delta u(t)^n} = \frac{\delta^{n+1} F_x(u, v)}{\delta v(t)^n}
\]

for all \((n,m) \in \mathbb{Z}^2_+\) and every \((t_1, t_2, \ldots, t_{n+m})\) with \(t \in I\), where \([D] = [\delta_0, \delta_1, \delta_2, \ldots, \delta_n] \).
5.3.3. Excitation with Independent Increments - Fokker-Planck Equation

In this section we will derive from the Functional Differential Equation the Fokker-Planck type equation for the characteristic function of the response. It is very important to note that the analysis below will be based only on the concept of Functional Differential Equations. Moreover the Fokker-Planck Equation derived can describe systems under any stochastic excitation with independent increments.

To begin the derivation of the Fokker-Planck equation let us first consider the special case of a first order equation with cubic nonlinearity

\[ d_x(t; \beta) = ax(t; \beta)^3 dt + d_y(t; \beta) \quad \text{a.s.} \] \hspace{1cm} (20)

\[ x(t_0; \beta) = x_0(\beta), \quad \text{a.s.} \] \hspace{1cm} (21)

where \( y(\cdot; \cdot) : I \times \Omega \to \mathbb{R} \) is a stochastic process with independent increments defined onto the probability space \( (\Omega, \mathcal{B}(\Omega), \mathbb{P}) \) with the induced probability space \( (\mathbb{C}^\infty, \mathcal{B}(\mathbb{C}^\infty), \mathbb{P}_y) \), and with characteristic functional \( \mathcal{F}_y \). The initial condition \( x_0(\cdot) : \Omega \to \mathbb{R} \) is described by a random variable with the induced probability space \( (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_{x_0}) \) with characteristic function \( \mathbb{F}_0 \). By Theorem 2.4 we have the equivalent Functional Differential Equation

\[ \frac{d}{dt} \left[ \frac{\delta \mathcal{F}_{xy}(u,v)}{\delta u(t)} \right] = -a \frac{\delta^3 \mathcal{F}_{xy}(u,v)}{\delta u(t)^3} dt + d \left[ \frac{\delta \mathcal{F}_{xy}(u,v)}{\delta v(t)} \right] \] \hspace{1cm} (22)

\[ \mathcal{F}_{xy}(v \delta_0, 0) = \mathcal{F}_0(v), \quad v \in \mathbb{R} \] \hspace{1cm} (23)

Setting to the above FDE \( u = v \cdot \delta(\cdot - s - t) \), \( v = 0 \) at time \( t + s \), and taking the limit \( s \to 0 \) we have for the l.h.s. term,

\[ \lim_{s \to 0} d_s \left[ \delta \mathcal{F}_{xy}(v \delta(\cdot - s - t), 0) \right] = \lim_{s \to 0} \int_{\mathbb{C}^\infty} d_s \left[ x(t+s) \right] \exp \left[ ivx(t+s) \right] \mathbb{P}_x(dx) = \lim_{s \to 0} \frac{1}{v} d_s \left[ \phi_x(v,t+s) \right] = \frac{1}{v} \frac{\partial \phi_x(v,t)}{\partial t} \] \hspace{1cm} (24)

For the first r.h.s. term we have,

\[ \lim_{s \to 0} \frac{\delta^3 \mathcal{F}_{xy}(v \delta(\cdot - s - t), 0)}{\delta u(t+s)^3} dt = \lim_{s \to 0} \int_{\mathbb{C}^\infty} \left[ ivx(t+s) \right]^3 \exp \left[ ivx(t+s) \right] \mathbb{P}_x(dx) = \frac{\partial^3 \phi_x(v,t)}{\partial v^3} \] \hspace{1cm} (25)

Finally for the second r.h.s. term we have

\[ \lim_{s \to 0} d_s \left[ \delta \mathcal{F}_{xy}(v \delta(\cdot - s - t), 0) \right] = \lim_{s \to 0} \int_{\mathbb{C}^\infty} \int_{\mathbb{C}^\infty} d_s \left[ y(t+s) \right] \exp \left[ ivx(t+s) \right] \mathbb{P}_x(1) \mathbb{P}_y(dx,dy) \] \hspace{1cm} (26)

But

\[ \exp \left[ ivx(t+s) \right] = \exp \left[ ivx(t) + ivax(t)^3 s \right] \exp \left[ iv \left[ y(t+s) - y(t) \right] \right] \] \hspace{1cm} (27)

Hence,
\[
\lim_{s \to 0} \int \int \frac{dI}{c^s} y(t+s) \exp \{ ivx(t+s) \} \mathcal{P}_{xy}(dx,dy) = \\
= \lim_{s \to 0} \int \int \frac{dI}{c^s} y(t+s) \exp \{ ivx(t) + ivax(t)^3 s \} \cdot \exp \{ iv(y(t+s) - y(t)) \} \mathcal{P}_{xy}(dx,dy) \\
= \lim_{s \to 0} \int \frac{dI}{c^s} y(t+s) \exp \{ iv(y(t+s) - y(t)) \} \mathcal{P}_{y}(dy) \cdot \lim_{s \to 0} \int \exp \{ ivx(t) \} \mathcal{P}_{x}(dx) = \\
= \lim_{v \to 0} \frac{1}{v} \int \frac{dI}{c^s} \exp \{ iv(y(t+s) - y(t)) \} \mathcal{P}_{y}(dy) \cdot \int \exp \{ ivx(t) \} \mathcal{P}_{x}(dx) = \\
= \lim_{v \to 0} \int \frac{dI}{c^s} \exp \{ iv(y(t+s) - y(t)) \} \mathcal{P}_{y}(dy) \cdot \int \exp \{ iv \xi \} \mathcal{P}_{y}(dy) = \\
= \lim_{v \to 0} \int \frac{dI}{c^s} \exp \{ iv(y(t+s) - y(t)) \} \mathcal{P}_{y}(dy) \cdot \int \exp \{ iv \xi \} \mathcal{P}_{y}(dy) = \\
= \lim_{v \to 0} \frac{\partial \phi_{\Delta v}(t,v,s)}{\partial s} \phi_x(t,v) ds \\
\] 

Thus we have proved the equality

\[
\frac{1}{v} \frac{\partial \phi_x(v,t)}{\partial t} ds = -a \frac{\partial^3 \phi_x(v,t)}{\partial v^3} ds + \lim_{v \to 0} \frac{\partial \phi_{\Delta v}(t,v,s)}{\partial s} \phi_x(t,v) ds \\
\]

or equivalently,

\[
\frac{\partial \phi_x(v,t)}{\partial t} + av \frac{\partial^3 \phi_x(v,t)}{\partial v^3} = \lim_{v \to 0} \frac{\partial \phi_{\Delta v}(t,v,s)}{\partial s} \phi_x(t,v) \\
\]

Assume now, that \( y(\cdot,\cdot) : I \times \Omega \to \mathbb{R} \) is the standard Wiener process. Then,

\[
\phi_{\Delta v}(t,v,s) = \exp \left\{ -\frac{1}{2} v^2 s \right\} \\
\]

Hence,

\[
\lim_{s \to 0} \frac{\partial \phi_{\Delta v}(t,v,s)}{\partial s} = -\frac{1}{2} v^2 \\
\]

and the Fokker-Planck equation takes the form

\[
\frac{\partial \phi_x(v,t)}{\partial t} + av \frac{\partial^3 \phi_x(v,t)}{\partial v^3} = -\frac{v^2}{2} \phi_x(t,v) \\
\]

For the general case we have the following
Theorem 3.10: Let the dynamical system $\Sigma_2$ described by (31)-(32) of Section 5.2.4. Then the FDE
\[
\frac{1}{i} \frac{d}{dt} [\mathcal{V}_{\mu(t)} \mathcal{F}_{xy} (u,v)] = \sum_{l=0}^{L} \{\nabla^l F(t,0)\} \frac{1}{i} \mathcal{V}_{\mu(t)} \mathcal{F}_{xy} (u,v) dt + \\
+ \sum_{k=0}^{K} \{\nabla^k G(t,0)\} \frac{1}{i} \mathcal{V}_{\nu(t)} \mathcal{F}_{xy} (u,v)
\]
(24)
and
\[
\mathcal{F}_0 (v, \delta_t, 0) = \mathcal{F}_0
\]
(25)
for every $u \in C_\varphi^\infty (I), v \in C_\varphi^\infty (I)$ and $t \in I$, which describes the system, can be reduced to the partial differential equation
\[
\frac{\partial \phi_x (t, v)}{\partial t} = \sum_{n=1}^{N} \sum_{l=0}^{L} \frac{v_n}{i} \left\{ \nabla^l F_n (t,0) \right\} \delta \left\{ \nabla^l \phi_x (t, v) \right\} + \\
+ \int_{C_\varphi^\infty (I)} \frac{\partial \phi_{x'} (t, v)}{\partial s} \left\{ t, \sum_{m=1}^{N} G_{nm} (t, x(t), v_{s}) \right\} \exp \left\{ i \sum_{n=1}^{N} x_n (t) v_n \right\} \mathcal{P}_x (dx)
\]
(26)
where $\phi (\bullet, \bullet) : I \times \mathbb{R}^N \rightarrow \mathbb{C}$ is the characteristic function of the system response $x(t; \beta)$ and $\phi_{x'} (\bullet, \bullet) : I \times \mathbb{R}^M \rightarrow \mathbb{C}$ is the characteristic function of the independent increments $\Delta y (t; \beta) = y (t + s) - y (t) \cdot \Delta t$. 

Proof: We consider the component wise form of the FDE (24), used for the derivation at Theorem 2.4. ($n = 1, 2, \ldots, N$)
\[
\frac{1}{i} \frac{d}{dt} \frac{\delta \mathcal{F}_{xy} (u,v)}{\delta u_n (t)} = dt \sum_{l=0}^{L} \sum_{i \in I^N} \frac{\partial^l F_n (t,0)}{\partial x_i} \frac{\delta \mathcal{F}_{xy} (u,v)}{\delta u_i (t)} + \\
+ \sum_{m=1}^{M} \sum_{l=0}^{K} \frac{\partial^l G_{nm} (t,0)}{\partial x_i} \frac{\delta \mathcal{F}_{xy} (u,v)}{\delta u_i (t)} \left[ d_l \frac{\delta \mathcal{F}_{xy} (u,v)}{\delta v_m (t)} \right]
\]
(27)
Setting to the system of functional differential equations $u = (u_1, u_2, \ldots, u_N) \cdot \delta (\bullet - t - s)$ at time $t + s$ and $t$ waking the limit $s \rightarrow 0$ we have for the l.h.s. term,
\[
\lim_{s \rightarrow 0} \frac{1}{i} \frac{d}{dt} \frac{\delta \mathcal{F}_{xy} (v \cdot \delta (\bullet - t - s), 0)}{\delta u_n (t + s)} = \lim_{s \rightarrow 0} \frac{1}{i v_n} \int_{\mathbb{R}^N} \left\{ \sum_{s=1}^{N} x_n (t + s) v_s \right\} \mathcal{P}_x (dx)
\]
For the first r.h.s term we have,
\[
\lim_{s \rightarrow 0} \sum_{l=0}^{L} \sum_{i \in I^N} \frac{\partial^l F_n (t + s,0)}{\partial x_i} \frac{\delta \mathcal{F}_{xy} (v \cdot \delta (\bullet - t - s), 0)}{\delta u_i (t + s)} = \\
= \lim_{s \rightarrow 0} \sum_{l=0}^{L} \int_{\mathbb{R}^N} \left\{ \sum_{i \in I^N} \frac{\partial^l F_n (t + s,0)}{\partial x_i} \right\} \left\{ \sum_{s=1}^{N} x_n (t + s) v_s \right\} \mathcal{P}_x (dx) =
\]
5.3. REDUCTION OF THE HOPF-TYPE FUNCTIONAL DIFFERENTIAL EQUATIONS FOR SPECIAL CASES

\[
\frac{1}{i^N} \sum_{i=0}^{K} \frac{\partial^k F_n(t,0)}{\partial x_i} \cdot \frac{\partial^j \phi_n(t,\nu)}{\partial \nu_j}.
\]

Where we have used the notation

\[
\partial x_i = \partial x_1 \cdot \partial x_2 \ldots \partial x_N
\]

\[
x^j = x_1^j \cdot x_2^j \ldots x_N^j
\]

For the second r.h.s term we have

\[
\lim_{s \to 0} \sum_{k=0}^{M} \sum_{i=0}^{K} \frac{\partial^k G_m(t+s,0)}{\partial x_i} \sum_{\nu \in \nu} \left[ \frac{\delta^m}{\delta u_i(t+s,0)} \right] \left[ \frac{\delta^m}{\delta v_m(t+s,0)} \right] =
\]

\[
\lim_{s \to 0} \sum_{k=0}^{M} \sum_{i=0}^{K} \frac{\partial^k G_m(t+s,0)}{\partial x_i} \int_{\nu \in \nu} \left[ x(t+s) \right]^j \cdot d_s y_m(t+s;\beta) \exp \left\{ \sum_{\nu \in \nu} \left[ x(t+s) \right] \nu \right\} \partial_{xy} (dx,dy).
\]

Since \( x(t) \) is m.s continuous \( \lim_{s \to 0} x(t+s) = x(t) \). Then

\[
\lim_{s \to 0} \sum_{k=0}^{M} \sum_{i=0}^{K} \frac{\partial^k G_m(t+s,0)}{\partial x_i} \int_{\nu \in \nu} \left[ x(t+s) \right]^j \cdot d_s y_m(t+s) \exp \left\{ \sum_{\nu \in \nu} \left[ x(t+s) \right] \nu \right\} \partial_{xy} (dx,dy) =
\]

\[
\lim_{s \to 0} \sum_{k=0}^{M} \sum_{i=0}^{K} \frac{\partial^k G_m(t,0)}{\partial x_i} \int_{\nu \in \nu} \left[ x(t) \right]^j y_m(t+s) \cdot \exp \left\{ \sum_{\nu \in \nu} \left[ x(t+s) \right] \nu \right\} \partial_{xy} (dx,dy) =
\]

Using the equation of the system, or the equation that results from the FDE by setting \( u = 0, \nu = 0 \) we will have for small \( s \)

\[
x_m(t+s;\beta) = \phi_m(t;\beta) + F_n(t,x(t;\beta))s + \sum_{m=1}^{M} G_m(t,x(t;\beta))[y_m(t+s;\beta) - y_m(t;\beta)]
\]

Hence the last term above will be equal with

\[
\lim_{s \to 0} \sum_{k=0}^{M} \sum_{i=0}^{K} \frac{\partial^k G_m(t,0)}{\partial x_i} \int_{\nu \in \nu} \left[ x(t) \right]^j y_m(t+s) \exp \left\{ \sum_{\nu \in \nu} \left[ x(t+s) \right] \nu \right\} \partial_{xy} (dx,dy) =
\]

\[
\exp \left\{ \sum_{\nu \in \nu} \left[ F_{\nu}(t,x(t))s + \sum_{m=1}^{M} G_{\nu m}(t,x(t))[y_m(t+s) - y_m(t)] \nu \right] \partial_{xy} (dx,dy),
\]
Thus for every $n = 1, 2, \ldots, N$ equation (27) takes the form

$$\lim_{s \to -\infty} \int_{C_{c}^{\infty \times}} i v_{\nu} d x_{\nu} (t + s) \exp \left[ i \sum_{\nu=1}^{N} x_{\nu} (t + s) v_{\nu} \right] \mathcal{P}_{\phi} (dx) = \sum_{l=0}^{K} \sum_{m=1}^{M} \int_{C_{c}^{\infty \times}} i v_{\nu} d x_{\nu} (t + s) \exp \left[ i \sum_{\nu=1}^{N} x_{\nu} (t + s) v_{\nu} \right] \mathcal{P}_{\phi} (dx) =$$

$$= \lim_{s \to -\infty} \int_{C_{c}^{\infty \times}} i v_{\nu} d x_{\nu} (t + s) \exp \left[ i \sum_{\nu=1}^{N} x_{\nu} (t + s) v_{\nu} \right] \mathcal{P}_{\phi} (dx) =$$

$$= \lim_{s \to -\infty} \int_{C_{c}^{\infty \times}} i v_{\nu} d x_{\nu} (t + s) \exp \left[ i \sum_{\nu=1}^{N} x_{\nu} (t + s) v_{\nu} \right] \mathcal{P}_{\phi} (dx) =$$

For the first r.h.s sum we get,

$$\sum_{n=1}^{N} \sum_{l=0}^{K} i v_{\nu} \sum_{m=1}^{M} \sum_{i=l}^{K} \frac{\partial^{i} F_{n} (t, 0)}{\partial x_{i}} \cdot \frac{\partial^{i} \phi_{x} (t, 0)}{\partial v_{i}} ds = \sum_{n=1}^{N} \sum_{l=0}^{K} i v_{\nu} \left\{ \nabla_{x} F_{n} (t, 0) \right\} \cdot \left\{ \nabla_{v} \phi_{x} (t, 0) \right\}$$

For the second r.h.s sum we have,

$$\lim_{s \to -\infty} \int_{C_{c}^{\infty \times \times}} i v_{\nu} d x_{\nu} (t + s) \exp \left[ i \sum_{\nu=1}^{N} x_{\nu} (t + s) v_{\nu} \right] \mathcal{P}_{\phi} (dx, dy) =$$

$$= \lim_{s \to -\infty} \int_{C_{c}^{\infty \times \times}} i v_{\nu} d x_{\nu} (t + s) \exp \left[ i \sum_{\nu=1}^{N} x_{\nu} (t + s) v_{\nu} \right] \mathcal{P}_{\phi} (dx, dy) =$$

$$= \lim_{s \to -\infty} \int_{C_{c}^{\infty \times \times}} i v_{\nu} d x_{\nu} (t + s) \exp \left[ i \sum_{\nu=1}^{N} x_{\nu} (t + s) v_{\nu} \right] \mathcal{P}_{\phi} (dx, dy) =$$
5.3. REDUCTION OF THE HOPF-TYPE FUNCTIONAL DIFFERENTIAL EQUATIONS FOR SPECIAL CASES

\[= \lim_{s \to 0} d_s \int_{C^N \times C^M \times C^N} \exp \left\{ \sum_{m=1}^{M} \sum_{s=1}^{N} G_{\nu,m}(t, x(t)) \right\} \left[ \int_{t}^{t+s} \left( y_m(t) - y_m(t) \right) \nu \right] \cdot \exp \left\{ \sum_{v=1}^{N} x_v(t) \nu \right\} \mathcal{D}_{\nu} (dx, dy)\]

Since the stochastic process \( y(t): I \times \Omega \to \mathbb{R}^M \) has independent increments, and the quantity \( \exp \left\{ \sum_{v=1}^{N} x_v(t) \nu \right\} \) depends only to prior increments, the last expression will be equal with

\[= \lim_{s \to 0} d_s \int_{C^N \times C^M \times C^N} \exp \left\{ \sum_{m=1}^{M} \sum_{s=1}^{N} G_{\nu,m}(t, x(t)) \right\} \left[ \int_{t}^{t+s} \left( y_m(t) - y_m(t) \right) \nu \right] \mathcal{D}_{\nu} (dy) \cdot \exp \left\{ \sum_{v=1}^{N} x_v(t) \nu \right\} \mathcal{D}_{\nu} (dx) =\]

\[= \lim_{s \to 0} d_s \int_{C^N \times C^M \times C^N} \phi_{\Delta \nu} \left\{ t \sum_{v=1}^{N} G_{\nu,m}(t, x(t)) \nu \right\} \exp \left\{ \sum_{v=1}^{N} x_v(t) \nu \right\} \mathcal{D}_{\nu} (dx) =\]

\[= d_s \int_{C^N \times C^M \times C^N} \frac{\partial \phi_{\Delta \nu}}{\partial s} \left\{ t \sum_{v=1}^{N} G_{\nu,m}(t, x(t)) \nu \right\} \exp \left\{ \sum_{v=1}^{N} x_v(t) \nu \right\} \mathcal{D}_{\nu} (dx)\]

As a result we have the desired equation for the characteristic function

\[\frac{\partial \phi_{\Delta \nu}(t, \nu)}{\partial t} = \sum_{m=1}^{M} \sum_{s=1}^{N} t_v \left\{ \nabla_{x} F_m(t, 0) \right\} \mathcal{O} \left\{ \nabla_{\nu} \phi_{\Delta \nu}(t, \nu) \right\} +\]

\[+ \int_{C^N \times C^M \times C^N} \frac{\partial \phi_{\Delta \nu}}{\partial s} \left\{ t \sum_{v=1}^{N} G_{\nu,m}(t, x(t)) \nu \right\} \exp \left\{ \sum_{v=1}^{N} x_v(t) \nu \right\} \mathcal{D}_{\nu} (dx)\]

**Remark 3.11:** We must note that the essential property for the proof of the Fokker-Planck equation is the independent increment property of the excitation, since we can decouple the probabilistic structures of the excitation and the response at the last term of the functional equation. Hence we can write equations for the response without taking into account the correlation structure of the two stochastic processes. This is not the case, when the stochastic excitation does not possess the independent increment property. In this case we must consider the correlation structure of the problem.

**Remark 3.12:** As we mention at the beginning the above proof of the Fokker-Plank equation is based only on the

**Remark 3.13:** The limit procedure with respect to \( s \) is carried out to take into account the discontinuity of the excitation.

**Corollary 3.14:** Let the dynamical system \( \Sigma 2 \) described by (31)-(32) of Section 5.2.4, with the excitation process being the standard Wiener process \( W(t; \beta) \) \( t \in I \) with intensity vector \( d_w(t), t \in I \), i.e., with characteristic function of increments
\[ \phi_{\Delta x}(t; y_1, \ldots, y_M) = \exp \left[ -\frac{1}{2} \left( \sum_{i=1}^{M} y_i^2 d_{W,i}(t) \right) s \right] \]

then eq. (26) takes the form

\[
\frac{\partial \phi_x(t, \nu)}{\partial t} = \sum_{n=1}^{N} \sum_{l=0}^{L} \int_{v_{j=1}}^{v_{j=1}+\Delta} \{ \nabla_x \Phi_n(t, 0) \} \cdot \{ \nabla_{\nu} \phi_x(t, \nu) \} + \sum_{i=1}^{N} v_{i=1} \sum_{m=1}^{K} \sum_{k=0}^{K} d_{W,m}(t) \frac{\partial^k G_{\nu,m}(t, 0)}{\partial x_i} \frac{\partial^k G_{\nu,m}(t, 0)}{\partial x_j} \frac{\partial^{k+\nu} \phi_x(t, \nu)}{\partial \nu_i} \]

where \( \phi_x(\cdot, \cdot): I \times \mathbb{R}^N \to \mathbb{C} \) is the characteristic function of the system response \( x(t; \beta) \).

**Proof:** By direct calculation we have

\[
\int_{C_{t,N}} \lim_{s \to 0} \frac{\partial \phi_{\Delta x}(t, N(G_{\nu,m}(t, x(t)) v_{i, j} s)}{\partial s} \exp \left[ i \sum_{i=1}^{N} x_{i}(t) v_{i} \right] \mathcal{P}_x (dx) =
\]

\[
= \int_{C_{t,N}} -\frac{1}{2} \sum_{m=1}^{M} d_{W,m}(t) \left( \sum_{i=1}^{N} G_{\nu,m}(t, x(t)) v_{i} \right)^2 \exp \left[ i \sum_{i=1}^{N} x_{i}(t) v_{i} \right] \mathcal{P}_x (dx) =
\]

\[
= -\frac{1}{2} \sum_{m=1}^{M} d_{W,m}(t) \sum_{i=1}^{N} v_{i} v_{i} \int_{C_{t,N}} G_{\nu,m}(t, x(t)) G_{\nu,m}(t, x(t)) \exp \left[ i \sum_{i=1}^{N} x_{i}(t) v_{i} \right] \mathcal{P}_x (dx) =
\]

\[
= -\frac{1}{2} \sum_{i=1}^{N} v_{i} v_{i} \int_{C_{t,N}} d_{W,m}(t) G_{\nu,m}(t, x(t)) G_{\nu,m}(t, x(t)) \exp \left[ i \sum_{i=1}^{N} x_{i}(t) v_{i} \right] \mathcal{P}_x (dx) (29)
\]

Using the component wise form the Taylor expansion of \( G_{\nu,m}(t, x(t)) \) we have

\[
\int_{C_{t,N}} d_{W,m}(t) \sum_{k=0}^{K} \sum_{j=1}^{N} \frac{\partial^k G_{\nu,m}(t, 0)}{\partial x_j} x_j \sum_{k=0}^{K} \sum_{j=1}^{N} \frac{\partial^k G_{\nu,m}(t, 0)}{\partial x_j} x_j \exp \left[ i \sum_{i=1}^{N} x_{i}(t) v_{i} \right] \mathcal{P}_x (dx) =
\]

\[
= \sum_{m=1}^{M} d_{W,m}(t) \sum_{k=0}^{K} \sum_{j=1}^{N} \frac{\partial^k G_{\nu,m}(t, 0)}{\partial v_j} x_j \sum_{k=0}^{K} \sum_{j=1}^{N} \frac{\partial^k G_{\nu,m}(t, 0)}{\partial v_j} x_j \exp \left[ i \sum_{i=1}^{N} x_{i}(t) v_{i} \right] \mathcal{P}_x (dx) =
\]

\[
= \sum_{m=1}^{K} \sum_{k=0}^{K} \frac{\partial^k G_{\nu,m}(t, 0)}{\partial x_j} \frac{\partial^k G_{\nu,m}(t, 0)}{\partial x_j} \sum_{k=0}^{K} \sum_{j=1}^{N} \frac{\partial^{k+\nu} \phi_x(t, \nu)}{\partial \nu_i} \frac{\partial^{k+\nu} \phi_x(t, \nu)}{\partial \nu_i} \]

Hence we have proved the desired equation.
Remark 3.15: Taking the Fourier transform of the equation (28) we clearly have for the l.h.s. and r.h.s. term that
\[
\mathcal{F}^{-1}\left[\frac{\partial \phi_x(t, \nu)}{\partial t}\right] = \frac{\partial f_x(t, \mathbf{x})}{\partial t}
\]
\[
\mathcal{F}^{-1}\left[\sum_{n=1}^{N} \frac{\nu_{n}}{L} \left\{ \nabla^l F_n(t, 0) \right\} \circ \left\{ \nabla^l \phi_x(t, \nu) \right\}\right] =
\]
\[
= -\sum_{n=1}^{N} \frac{\partial}{\partial x_n} \left[ f_x(t, \mathbf{x}) F_n(t, \mathbf{x}) \right], \quad t \in I, \quad \mathbf{x} \in \mathbb{R}^N
\]

For the second r.h.s. term we have from (29)
\[
\mathcal{F}^{-1}\left[ \int_{C^{\infty}_{\mathbb{R}^N}} \lim_{s \to 0} \frac{\partial \phi_{x,y}(t, s)}{\partial s} \exp \left\{ i \sum_{\nu=1}^{N} x_{\nu}(t) \nu_{\nu} \right\} \mathcal{P}_{x}(d\mathbf{x}) \right] =
\]
\[
= -\frac{1}{2} \mathcal{F}^{-1}\left[ \sum_{\nu_1=1}^{N} \nu_1 \nu_2 \int_{\mathbb{R}^N} \sum_{\nu_3=1}^{M} d_{w,m}(t) G_{w,m}(t, \mathbf{x}(t)) G_{w,m}(t, \mathbf{x}(t)) e^{i \sum_{\nu=1}^{N} x_{\nu}(t) \nu_{\nu}} \mathcal{P}_{x}(d\mathbf{x}) \right] =
\]
\[
= \frac{1}{2} \sum_{\nu_1=1}^{N} \frac{\partial}{\partial x_{\nu_1}} \frac{\partial}{\partial x_{\nu_2}} \left[ G(t, \mathbf{x}) \tilde{d}_{w,m}(t) G(t, \mathbf{x})^T f_x(t, \mathbf{x}) \right]
\]

where \( \tilde{d}_{w,m}(t) = \begin{pmatrix} d_{w,1}(t) & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & d_{w,M}(t) \end{pmatrix} \) is the intensity matrix of the Wiener process \( \mathbf{W}(t; \beta) \)

\( t \in I \). Hence we have proved the Fokker-Planck equation for the probability density function
\[
\frac{\partial f_x(t, \mathbf{x})}{\partial t} + \sum_{n=1}^{N} \frac{\partial}{\partial x_n} \left[ f_x(t, \mathbf{x}) F_n(t, \mathbf{x}) \right] - \frac{1}{2} \sum_{\nu_1=1}^{N} \frac{\partial}{\partial x_{\nu_1}} \frac{\partial}{\partial x_{\nu_2}} \left[ G(t, \mathbf{x}) d_{w,m}(t) G(t, \mathbf{x})^T f_x(t, \mathbf{x}) \right] = 0
\]

for \( (t, \mathbf{x}) \in I \times \mathbb{R}^N \). ■

Example 3.16 [One-dimensional case]: Let us now consider a specific case of a one dimensional system excited with from a Wiener stochastic process \( \mathbf{W}(t; \beta) \) \( t \in I \) with constant intensity \( d_{\nu} \) and coefficients independent of time. The Fokker-Planck equation takes the form
\[
\frac{\partial f_x(t, x)}{\partial t} + \frac{\partial}{\partial x} \left[ f_x(t, x) F(x) \right] - \frac{1}{2} \frac{\partial^2}{\partial x^2} \left[ d_{\nu} G(x)^2 f_x(t, x) \right] = 0
\]

(30)

Using the standard method of separation of variables we assume that
\[
f_x(t, x) = \Lambda(x) T(t)
\]

(31)

where \( \Lambda \) and \( T \) are functions of \( x \) and \( t \), respectively. Substitution of (31) into (30) gives the two ordinary differential equation
\[
\frac{1}{T(t)} \frac{dT}{dt} = -\lambda^2 
\]  
(32)

\[
\frac{1}{2} \frac{\partial^2}{\partial x^2} \left[ d\nu G(x)^2 \Lambda(x) \right] - \frac{\partial}{\partial x} \left[ F(x)\Lambda(x) \right] + \lambda^2 \Lambda(x) = 0 
\]  
(33)

Equation (32) is a simple first order equation with solution

\[
T(t) = \exp\{-\lambda^2 t\} 
\]  
(34)

while equation (33) is a second order equation with variable coefficients. In general, the solution \( \Lambda(x,A,B,\lambda) \) depends on two arbitrary constants \( A \) and \( B \). If the eigenvalues of equation (33) are discrete and distinct, then by virtue of linearity of the problem the solution can be represented symbolically in the form

\[
f_x(t,x) = \sum_{n=0}^{\infty} \Lambda(x,A_n,B_n,\lambda_n) \exp\{-\lambda_n^2 t\} 
\]  
(35)

where \( A_n, B_n \) and \( \lambda_n \) are determined from the initial and boundary conditions.

In general case of variable coefficients \( G(x)^2 \) and \( F(x) \) the eigenfunction-eigenvalue problem (33) is not easy to solve. Only in restricted numbers of cases (special form of \( G(x)^2 \) and \( F(x) \)) a solution can be determined in a closed form – in terms of special functions. The solution of the Fokker-Planck equation can be then expressed in terms of the eigenfunctions \( \Lambda_n(x) \).

For example, if equation (33) is supplemented by the absorbing boundary condition \( f_x(t,\pm a) = 0 \) and the initial condition is \( f_x(t_0,x) = \delta(t-t_0) \) then the solution is represented as

\[
f_x(t,x) = \sum_{n=0}^{\infty} \frac{\Lambda_n(x)\Lambda_n(x_0)}{f_{x,\text{st}}(x)} \exp\{-\lambda_n^2 (t-t_0)\} 
\]  
(36)

where \( f_{x,\text{st}}(x) \) is the stationary solution of equation (30) (which will determined in the sequel) and the functions \( \Lambda_n(x) \) are orthonormal with the weight function \( [f_{x,\text{st}}(x)]^{-1} \), i.e.,

\[
\int_{-a}^{a} \frac{1}{f_{x,\text{st}}(x)} \Lambda_n(x)\Lambda_m(x) dx = \begin{cases} 1, & n = m \\ 0, & n \neq m \end{cases} 
\]

The above form of solution is often used in applications.

Now let us determine the stationary solution probability density function for the one dimensional case described above. Note that for the existence of stationary solutions, certain conditions must be fulfilled for coefficients \( G(x)^2 \) and \( F(x) \). These conditions are described extensively for the one dimensional case, as well as, for the general case at SOBCZYK, K., (Stochastic Differential Equations. p.154). Now consider the one-dimensional case. Assuming that a stationary solution exists, then since \( f_{x,\text{st}}(x) = \lim_{t \to \infty} f_x(t,x) \), the stationary solution does not depend on time and the initial density \( f_0(x) \). The Fokker-Planck equation takes the form
5.3. REDUCTION OF THE HOPF-TYPE FUNCTIONAL DIFFERENTIAL EQUATIONS FOR SPECIAL CASES

\[
\frac{\partial}{\partial x} \left[ f_{x,m}(x) F(x) \right] - \frac{1}{2} \frac{\partial^2}{\partial x^2} \left[ d_w G(x)^2 f_{x,m}(x) \right] = 0
\]  

(37)

or, after integration with respect to \( x \)

\[
\frac{d}{dx} \left[ d_w G(x)^2 f_{x,m}(x) \right] - 2 \left[ f_{x,m}(x) F(x) \right] = -2\Phi_0
\]  

(38)

where \( \Phi_0 \) is a constant. The last equation is first order inhomogeneous equation for function

\[ h(x) = d_w G(x)^2 f_{x,m}(x), \]

\[
\frac{dh(x)}{dx} - 2 \frac{F(x)}{d_w G(x)^2} h(x) = -2\Phi_0
\]  

(39)

Hence, the general solution for \( f_{x,m}(x) \) is

\[ f_{x,m}(x) = \frac{C}{d_w G(x)^m} \exp \left\{ \int_{x'}^x \frac{2F(\xi)}{d_w G(\xi)^2} d\xi \right\} = \frac{2\Phi_0}{d_w G(x)^m} \int_{x'}^x \exp \left\{ \int_{x'}^{\xi} \frac{2F(\xi)}{d_w G(\xi)^2} d\xi \right\} dy \]  

(40)

where the constant \( C \) is determined by the normalization condition, whereas the constant \( \Phi_0 \) is found from the boundary conditions; as lower integration limit \( x' \) one can take an arbitrary point from the state space of the process.

If \( \Phi_0 = 0 \) equation (37) gives

\[
\frac{d}{dx} \left[ d_w G(x)^2 f_{x,m}(x) \right] - 2 \left[ f_{x,m}(x) F(x) \right] = -2\Phi_0
\]  

(41)

with the following general solution

\[ f_{x,m}(x) = \frac{C}{d_w G(x)^m} \exp \left\{ \int_{x'}^x \frac{2F(\xi)}{d_w G(\xi)^2} d\xi \right\} \]

(42)

The above formula is often used in applications. More general formulas describing the probability density function of Hamiltonian Dynamical Systems with Wiener excitation and/or white noise parametric uncertainty are of special interest. For results of this type, as well as the necessary conditions for their existence we refer to SOIZE, C., (The Fokker-Planck Equation and its Explicit Steady State Solutions) and to POLIDORI, D.C. & BECK, J.L. & PAPADIMITRIOU, C., (A New Stationary PDF Approximation for Non-Linear Oscillators), MOSBAH, H. & FOGLI, M., (An Original Approximate Method for Estimating the Invariant Probability Distribution of a Large Class of Multi-Dimensional Nonlinear Stochastic Oscillators). For other interesting results concerning systems with random excitation of special type we refer to DI PAOLA, M. & SOFI, A., (Approximate Solution of the Fokker-Planck-Kolmogorov Equation) and GUO-KANG ER., (Exponential Closure Method for some Randomly Excited Non-Linear Systems).
5.3.4. Partial Differential Equation for (N+M)-dimensional joint (excitation-response) characteristic functions in the case of general m.s.-Continuous Excitation

In this section we shall derive a partial differential equation for the joint characteristic function of the excitation and the response when the excitation is m.s.-continuous stochastic process. In contrast with the previous case of an independent increment process where the randomness of the excitation ‘regenerates’ every instant time, in this case the stochasticity evolves in a smoother way, as a result of the finite correlation time of the excitation. This means that, although the response \( x(t;\beta) \) on the past history of the excitation \( y(t;\beta) \) (Principle of Causality), the correlation between \( x(t;\beta) \) and \( y(t+\varepsilon;\beta) \), with \( \varepsilon > 0 \) may not be zero, since \( C_{yy}(t+\varepsilon,t) \neq 0 \).

Because of the last property of the excitation process, i.e. finite correlation time we cannot use the same arguments as for the previous case of independent increment process. Our analysis will be a generalization of Liouville equation. We shall use the same style of presentation as before, i.e., first a special case will be considered and then the general problem will be addressed. Let us consider a first order equation with cubic nonlinearity

\[
\dot{x}(t;\beta) = ax(t;\beta)^3 + y(t;\beta) \quad \text{a.s.} \quad t \geq t_0 \tag{43}
\]
\[
x(t_0;\beta) = x_0(\beta), \quad \text{a.s.} \tag{44}
\]

where \( y(\bullet;\bullet):I \times \Omega \to \mathbb{R} \) is a m.s. continuous stochastic process defined onto the probability space \( (\Omega, \mathcal{B}(\Omega), \mathcal{P}) \) with the induced probability space \( (C^\infty_c, \mathcal{B}(C^\infty_c), \mathcal{P}_y) \), and characteristic functional \( \mathcal{F}_y = \int_{C^\infty_c} \exp\{i\langle y,v \rangle\} \mathcal{P}_y(dy) \). The initial condition \( x_0(\bullet):\Omega \to \mathbb{R} \) is described by a random variable with the induced probability space \( (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{P}_{x_0}) \) and characteristic function \( \mathcal{F}_0(a), \ a \in \mathbb{R} \). By Theorem 2.1 we know that the joint characteristic functional satisfies

\[
\frac{d}{dt}\left[ \frac{\delta\mathcal{F}_{xy}(u,v)}{\delta u(t)} \right] = -a \frac{\delta^3\mathcal{F}_{xy}(u,v)}{\delta u(t)^3} + \frac{\delta\mathcal{F}_{xy}(u,v)}{\delta v(t)}, \quad u,v \in C^\infty_c, \quad t \geq t_0 \tag{45}
\]
\[
\mathcal{F}(a \cdot \delta_0,0) = \mathcal{F}_0(a), \quad a \in \mathbb{R} \tag{46}
\]

- Derivation of a P.D.E. for the joint characteristic function \( \phi_{xy}(v,t;v,t) \)

We shall now derive a partial differential equation for the characteristic function. For the l.h.s. term of (45) we clearly have

\[
\frac{d}{dt}\left[ \frac{\delta\mathcal{F}_{xy}(u,v)}{\delta u(t)} \right] = \frac{d}{dt} \int_{C^\infty_c \times C^\infty_c} ix(t) \exp\{i\langle x,u \rangle + i\langle y,v \rangle\} \mathcal{P}_{xy}(dx,dy) =
\]
\[
= \int_{C^\infty_c \times C^\infty_c} i \frac{dx(t)}{dt} \exp\{i\langle x,u \rangle + i\langle y,v \rangle\} \mathcal{P}_{xy}(dx,dy)
\]
Taking the r.h.s. of the last equality and substituting \( u(\bullet) = v \cdot \delta(\bullet - t) \) and \( v(\bullet) = v \cdot \delta(\bullet - s) \) \( v, v \in \mathbb{R} \) we obtain
\[
\int \int_{C_y \times C_x} \frac{dx(t)}{dt} \exp \{ i(x,u) + i(y,v) \} \mathcal{D}_{xy}^t (dx, dy) = \\
= \int \int_{C_y \times C_x} \frac{dx(t)}{dt} \exp \{ ivx(t) + ivy(s) \} \mathcal{D}_{xy}^t (dx, dy) = \\
= \frac{1}{v} \frac{d}{dt} \int_{C_y \times C_x} \exp \{ ivx(t) + ivy(s) \} \mathcal{D}_{xy}^t (dx, dy) = \\
= \frac{1}{v} \frac{d}{dt} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp \{ iv + iv \} f_{xy} (x,t,y,s) dx dy = \\
= \frac{1}{v} \frac{\partial \phi_{xy} (v,t,v,s)}{\partial t}
\]

For the first r.h.s. term we have,
\[
\frac{\delta^3 \mathcal{F} \left(v \cdot \delta(\bullet - t), v \cdot \delta(\bullet - s)\right)}{\delta u(t)^3} = \int \int_{C_y \times C_x} [ix(t)]^3 \exp \{ ivx(t) + ivy(s) \} \mathcal{D}_{xy}^t (dx, dy) = \\
= \frac{\partial \phi_{xy} (v,t,v,s)}{\partial v^3}
\]

Finally for the second r.h.s. term we have
\[
\frac{\delta^3 \mathcal{F} \left(v \cdot \delta(\bullet - t), v \cdot \delta(\bullet - s)\right)}{\delta v(t)} = \int \int_{C_y \times C_x} iv(t) \exp \{ ivx(t) + ivy(t) \} \mathcal{D}_{xy}^t (dx, dy)
\]

Taking the limit \( s \to t \) and since the excitation has continuous paths we will have
\[
\int \int_{C_y \times C_x} iv(t) \exp \{ ivx(t) + ivy(t) \} \mathcal{D}_{xy}^t (dx, dy) = \\
= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} iv \exp \{ iv + iv \} f_{xy} (x,t,y,t) dx dy = \frac{\partial \phi_{xy} (v,t,v,s)}{\partial v}
\]

Thus we have proved the following partial differential equation involving the joint characteristic function,
\[
\frac{\partial \phi_{xy} (v,t,v,s)}{\partial t} \bigg|_{t=s} + uv^3 \frac{\partial^3 \phi_{xy} (v,t,v,t)}{\partial v^3} = v \frac{\partial \phi_{xy} (v,t,v,t)}{\partial v}, \quad v, v \in \mathbb{R}, \quad t \in T \quad (47)
\]

with the additional condition
\[
\phi_{xy} (0,t,v,s) = \phi_y (v,s) = \mathcal{F}_y (v \cdot \delta(\bullet - s)) \quad , \quad v \in \mathbb{R} \quad (48)
\]

and the initial condition
\[
\phi_{xy} (v,0;0,s) = \phi_x (v,0) = \mathcal{F}_0 (v) \quad , \quad v \in \mathbb{R}
\]
We note that the partial differential equation for the joint characteristic function is of the Liouville’s equation form with additional condition for the marginal $\phi_y(v,s)$. In other words, we treat the excitation term $y(t;\beta)$ as a state space variable which does not participate to the dynamical part of the equation. Instead of that, the additional condition (48) is fulfilled.

- *Derivation of moment equations for eq. (43) based on eq. (47)*

It is worth noticing that equation (47) is equivalent with the infinite set of moment equations

$$\frac{1}{n+1} \frac{dE^\beta\left([x(t;\beta)]^{n+1}[y(s;\beta)]^m\right)}{dt} = aE^\beta\left([x(t;\beta)]^{n+3}[y(s;\beta)]^m\right) + E^\beta\left([x(t;\beta)]^n[y(s;\beta)]^{m+1}\right)$$

for $n = 0,1,\ldots$ and $m = 0,1,\ldots$. The proof of the above set of moment equation for eq. (43) can be found at BERAN M.J., (Statistical Continuum Mechanics). Let us now derive the above set using directly eq. (47). From Chapter I we clearly have the equation connecting the characteristic function and the moments of a stochastic process. Hence, for the examined case the suitable form of this equation will be

$$\frac{\partial^{n+m} \phi_{xy}(v,t;\nu,s)}{\partial \nu^n \partial \nu^m} \bigg|_{\nu = 0, \nu = 0} = i^n m E^\beta \left([x(t;\beta)]^n \cdot [y(s;\beta)]^m\right), \quad n,m = 0,1,\ldots$$

Moreover we have,

$$\frac{1}{i^{n+m}} \frac{\partial^{n+m+1} \phi_{xy}(v,t;\nu,s)}{\partial t \partial \nu^n \partial \nu^m} \bigg|_{\nu = 0, \nu = 0} = \frac{d}{dt} E^\beta \left([x(t;\beta)]^n \cdot [y(s;\beta)]^m\right) = n E^\beta \left(\frac{dx(t;\beta)}{dt}[x(t;\beta)]^{n-1} \cdot [y(s;\beta)]^m\right), \quad n,m = 0,1,\ldots$$

Now, taking $n$–partial derivatives with respect to $\nu$ and $m$–partial derivatives with respect to $\nu$ to eq (47) and using the two last equations (50) and (51) we obtain the set of moment equations (49).

- *Derivation of a P.D.E. for the joint density function $f_{xy}(x,t;y,t)$*

An alternative form of equation (47) can be results by applying the Fourier transform to the above equations. Before that, we recall some useful relations connecting the probability density function and the characteristic function of stochastic processes written in a suitable form for our study. Thus, based on simple properties of the Fourier transform, we have
5.3. REDUCTION OF THE HOPF-TYPE FUNCTIONAL DIFFERENTIAL EQUATIONS FOR SPECIAL CASES

\[
\frac{\partial f_{xy}(x,t;y,s)}{\partial x} = \mathcal{F}\left[u \phi_{xy}(u,t;v,s)\right]
\]

\[
\frac{\partial \phi_{xy}(u,t;v,s)}{\partial u} = \mathcal{F}\left[i f_{xy}(x,t;y,s)\right]
\]

Based on the above equations we can easily prove the more general

\[
\mathcal{F}^{-1}\left[u^m \frac{\partial^n \phi_{xy}(u,t;v,s)}{\partial u^n}\right] = \frac{1}{i^n} \frac{\partial^m}{\partial x^m} \mathcal{F}^{-1}\left[\frac{\partial^n \phi_{xy}(u,t;v,s)}{\partial u^n}\right] = \frac{i^n}{i^n} \frac{\partial^m}{\partial x^m}\left[x^n f_{xy}(x,t;y,s)\right]
\]

By direct application of the Fourier transform and using the last equation we have

\[
\frac{\partial f_{xy}(x,t;y,s)}{\partial t} + \frac{\partial}{\partial x} \left[ax^3 f_{xy}(x,t;y,t) + y f_{xy}(x,t;y,t)\right] = 0, \quad x, y \in \mathbb{R}, t \in I \tag{52}
\]

\[
\int_{-\infty}^{\infty} f_{xy}(x,t;y,s) dx = f_y(y,s) = \int_{-\infty}^{\infty} e^{ivs} \mathcal{F}_y\left(v \cdot \delta(\bullet - s)\right) dv, \quad y \in \mathbb{R} \tag{53}
\]

\[
\int_{-\infty}^{\infty} f_{xy}(x,0;y,s) dy = f_x(x,0) = \int_{-\infty}^{\infty} e^{ivs} \mathcal{F}_0(v) dv, \quad x \in \mathbb{R} \tag{54}
\]

Thus, the l.h.s. of equation (52), which is equivalent with the l.h.s. of Liouville equation express the conservation of ‘probability’ for the state space variable \(x(t;\beta)\). The r.h.s term, which is due to the excitation term, express the transfer-interaction of ‘probability’ from the stochastic process \(y(t;\beta)\) to the state space variables \(x(t;\beta)\). Of course the stochastic characteristics of the excitation \(y(t;\beta)\) are introduced with condition (53).

--- Numerical scheme for the solution of eq. (52)---

For the case of a first order equation with cubic non-linearity we have proved at the beginning of the section the following partial differential equation involving the joint characteristic function,

\[
\frac{\partial \phi_{xy}(v,t;v,t)}{\partial t} + av^3 \frac{\partial^3 \phi_{xy}(v,t;v,t)}{\partial v^3} = v \frac{\partial \phi_{xy}(v,t;v,t)}{\partial v}, \quad v, v \in \mathbb{R}, \quad t \in T
\]

with the additional condition

\[
\phi_{xy}(0,t;v,s) = \phi_y(v,s) = \mathcal{F}_y(v \cdot \delta(\bullet - s)), \quad v \in \mathbb{R}
\]

and the initial condition

\[
\phi_{xy}(v,0;0,s) = \phi_x(v,0) = \mathcal{F}_0(v), \quad v \in \mathbb{R}
\]

Taking the Fourier transform of the above equations we have the equivalent set of equations for the joint probability density function,
\[
\frac{\partial f_{xy}(x,t;y,s)}{\partial t} \bigg|_{x=t} + \frac{\partial}{\partial x} \left[ a x^3 f_{xy}(x,t;y,t) + y f_{xy}(x,t;y,t) \right] = 0, \quad x, y \in \mathbb{R}, \quad t \in T \tag{55}
\]

\[
\int_{-\infty}^{\infty} f_{xy}(x,t;y,s) \, dx = f_y(y,s) = \int_{-\infty}^{\infty} e^{iu \cdot \nu} \mathcal{F}_y(\nu \cdot \delta(\nu)) \, du , \quad y \in \mathbb{R} \tag{56}
\]

\[
\int_{-\infty}^{\infty} f_{xy}(x,0;y,s) \, dy = f_x(x,0) = \int_{-\infty}^{\infty} e^{ix \nu} \mathcal{F}_0(\nu) \, d\nu , \quad x \in \mathbb{R} \tag{57}
\]

For the numerical solution of the above problem we will use a combination of a bivariate copula function, to represent the correlation structure of the excitation and the response at time \( t \), and a kernel density function representation for the marginal \( f_x(x,t) \). We briefly, recall that a bivariate copula function is a distribution function that has given marginals. For more information we refer to JOE, H., (Multivariate Models and Dependence Concepts). For this example we will use the Plackett copula function that gives a joint density function representation at time \( t \),

\[
f_{xy}(x,t;y,t) = C(F_x(x,t), F_y(y,t); \psi(t)) \cdot f_x(x,t) \cdot f_y(y,t) \tag{58}
\]

with,

\[
C(x_1,x_2; \psi(t)) = \frac{\psi(t) \cdot (\psi(t)-1) \cdot (x_1 + x_2 - 2x_1x_2) + 1}{\left(1+(x_1 + x_2)(\psi(t)-1)^2 - 4\psi(t)(\psi(t)-1)x_1x_2\right)^{\frac{1}{2}}} , \quad \psi(t) > 0 \tag{59}
\]

Additionally we assume that the correlation structure between the response and the excitation, which is described by the function \( \psi(t) \), doesn’t change considerable for the joint density function \( f_{xy}(x,t;y,s) \) when the time difference \( t-s \) is sufficiently small. In this way we can assume that for small \( t-s \) the following representation for the joint characteristic function holds

\[
f_{xy}(x,t;y,s) = C(F_x(x,t), F_y(y,s); \psi(t)) \cdot f_x(x,t) \cdot f_y(y,s) \tag{60}
\]

Hence, substituting representation (60) into eq. (55) we will have

\[
\left[ \frac{\partial}{\partial x} f_x(x,t) + \frac{\partial}{\partial \psi} \psi(t) f_x(x,t) + C \frac{\partial f_x(x,t)}{\partial t} \right] + \frac{\partial}{\partial x} \left[ a x^3 C f_x(x,t) + y C f_x(x,t) \right] = 0, \quad x, y \in \mathbb{R}, \quad t \in I \tag{61}
\]

For the last equation the unknowns are the functions \( f_x(x,t) \) and \( \psi(t) \). Now using a kernel density representation for the marginal density function \( f_x(x,t) \) and multiplying and integrate over \( x \) with each kernel we will have a set of integrodifferential equations for the parameters of the kernel density functions and the function \( \psi(t) \). The above approach should be clearer at Section 5.4 where a kernel representation will considered at the level of characteristic functional. An alternative approach can be based on the kernel representation of the correlation structure, i.e., two dimensional kernels that include the correlation structure. For more details we refer to Section 5.4.
5.3. REDUCTION OF THE HOPF-TYPE FUNCTIONAL DIFFERENTIAL EQUATIONS FOR SPECIAL CASES

Derivation of a P.D.E. for the joint characteristic function $\phi_{x\dot{y}y}(v_1,t;v_2,t)$

For many practical applications it is often needed to know the joint probability of the response and its derivative. We shall show that for the case of the joint probability of the response and its derivative, we can easily have the desired density by a simple transformation of random variables $x(t;\beta), y(t;\beta)$ using their joint density derived before. In contrast for higher order derivatives another approach has to be considered. More precisely the joint characteristic function of the response, its derivative and the excitation can be easily derived using the equation governing the system,

$$\dot{x}(t;\beta) = ax(t;\beta)^2 + y(t;\beta) \quad \text{a.s. } t \geq t_0$$

For the joint probability of the second derivative we consider the following set of equations

$$\dot{x}(t;\beta) = ax(t;\beta)^3 + y(t;\beta) \quad \text{a.s. } t \geq t_0$$

$$\ddot{x}(t;\beta) = 3ax(t;\beta)^2 \dot{x}(t;\beta) + y(t;\beta) \quad (62)$$

where the second equations has been derived by direct differentiation of the first one. Now using same arguments as for the case of the derivation of the joint characteristic function $\phi_{x\dot{y}}(v,t;\nu,t)$ we have the equation

$$\frac{\partial \phi_{x\dot{y}y}(t,s)}{\partial t} + av_1 \frac{\partial^3 \phi_{x\dot{y}y}(t,t)}{\partial v_1^3} + 3av_2 \frac{\partial^3 \phi_{x\dot{y}y}(t,t)}{\partial v_1^3 \partial v_2} = $$

$$= v_1 \frac{\partial \phi_{x\dot{y}y}(t,t)}{\partial v_1} + v_2 \frac{\partial \phi_{x\dot{y}y}(t,t)}{\partial v_2}, \quad v_1, v_2, v_1, v_2 \in \mathbb{R}, \quad t \in T$$

where for simplicity, we write the joint characteristic function $\phi_{x\dot{y}y}(v_1,t;v_2,t)$ as $\phi_{x\dot{y}}(t,s)$.

Moreover the additional condition holds

$$\phi_{x\dot{y}y}(0,t;0,t;v_1,s;v_2,s) = \phi_{x\dot{y}}(v_1,s;v_2,s) = \mathcal{T}_y\left(v_1 \cdot \delta(\cdot - s) + v_2 \cdot \delta(\cdot - s)\right), \quad v_1, v_2 \in \mathbb{R}$$

and the initial condition

$$\phi_{x\dot{y}y}(v_1,0;v_2,0;0,0) = \phi_{x\dot{y}}(v_1,0;v_2,0) = \mathcal{T}_0(v_1,v_2), \quad v_1, v_2 \in \mathbb{R}$$

Similarly we can derive equations for joint characteristic functions of higher derivatives.

Derivation of a P.D.E. for the joint ch. function $\phi_{x\dot{y}y}(v_1,t;v_2,t+\tau;v_1,t;v_2,t+\tau)$

For many applications it is often need to know the joint probability of the response for two distinct time instants. More precisely the joint characteristic function of the response at two distinct time instants $t$ and $t + \tau$ can be derived by considering the following set of equations

$$\dot{x}(t;\beta) = ax(t;\beta)^2 + y(t;\beta) \quad \text{a.s. } t \geq t_0$$

$$\ddot{x}(t;\beta) = 3ax(t;\beta)^2 \dot{x}(t;\beta) + y(t;\beta)$$

$$\dddot{x}(t;\beta) = 3a(2ax(t;\beta)\dot{x}(t;\beta) + ay(t;\beta)) \dot{x}(t;\beta) + 3ay(t;\beta) \dot{x}(t;\beta) + y(t;\beta)$$
\[
\dot{x}(t;\beta) = ax(t;\beta)^3 + y(t;\beta) \quad \text{a.s. } t \geq t_0
\]

Now using same arguments as for the case of the derivation of the joint characteristic function \( \phi_{xy}(v,t;\nu,t) \) we have the equation

\[
\frac{\partial \phi_{xyy}(t,s)}{\partial t} = av_1 \frac{\partial^3 \phi_{xyy}(t,t)}{\partial v_1^3} + av_2 \frac{\partial^3 \phi_{xyy}(t,t)}{\partial v_2^3} = v_1 \frac{\partial \phi_{xyy}(t,t)}{\partial v_1} + v_2 \frac{\partial \phi_{xyy}(t,t)}{\partial v_2} , \quad v_1,v_2 \in \mathbb{R}, \quad t \in I
\]

where for simplicity, we write the joint characteristic function \( \phi_{xyy}(v_1,t;v_2,t+\tau;v_1,s,v_2,s+\tau) \) as \( \phi_{xyy}(t,s) \).

Moreover the additional condition holds

\[
\phi_{xyy}(0,t;0,t+\tau;v_1,s,v_2,s+\tau) = \phi_{yy}(v_1,s,v_2,s+\tau) = \mathcal{F}_s \left( v_1 \cdot \delta(s) + v_2 \cdot \delta(s) \right), \quad v_1,v_2 \in \mathbb{R}
\]

and the initial condition

\[
\phi_{xyy}(v_1,0;v_2,\tau;0,s;0,s+\tau) = \phi_{xy}(v_1,0;v_2,\tau) = \mathcal{F}_0 (v_1,v_2) , \quad v_1,v_2 \in \mathbb{R}
\]

Similarly we can derive equations for joint characteristic functions of higher derivatives.

\[\text{– Derivation of a P.D.E. for the joint characteristic function } \phi_{xy}(v,t;\nu,t) \text{ for a general system.}\]

For the general case we have the following

**Theorem 3.17.** Let the dynamical system \( \Sigma_1 \) described by (18)-(19) of Section 5.2.3. Then the FDE

\[
\frac{1}{i} \frac{d}{dt} \left[ \nabla_{u(t)} \mathcal{F}_{xy}(u,v) \right] = \sum_{k=0}^K \sum_{l=0}^L \left\{ \nabla^k_y \nabla^l_x F(t,x_0,y_0) \right\} \frac{1}{i^{k+l}} \nabla^i_{u(t)} \nabla^1_{v(t)} \mathcal{F}_{xy}(u,v)
\]

\[\mathcal{F}_{xy}(v;\delta_0,0) = \mathcal{F}_0 (v), \quad v \in \mathbb{R}^N\]

for every \( u \in C^{\infty,N}_c (I), v \in C^{\infty,M}_c (I) \) and \( t \in I \), which describes the system, can be reduced to the partial integrodifferential equation

\[
\frac{\partial \phi_{xy}(t,v;s,\nu)}{\partial t} = \sum_{n=1}^N \sum_{k=1}^L \frac{v_n}{i^{k+l-1}} \left\{ \nabla^k_y \nabla^l_x F_n(t,0,0) \right\} \circ \left\{ \nabla^i_y \nabla^j_v \phi_{xy}(t,v;s,\nu) \right\} = \frac{1}{i} \frac{d}{dt} \left[ \nabla_{u(t)} \mathcal{F}_{xy}(u,v) \right] - \sum_{n=1}^N \sum_{k=1}^L \frac{v_n}{i^{k+l-1}} \left\{ \nabla^k_y \nabla^l_x F_n(t,0,0) \right\} \circ \left\{ \nabla^i_y \nabla^j_v \phi_{xy}(t,v;s,\nu) \right\}
\]

with the additional condition,
5.3. REDUCTION OF THE HOPF-TYPE FUNCTIONAL DIFFERENTIAL EQUATIONS FOR SPECIAL CASES

where \( \phi_{xy}(\bullet, \bullet, \bullet): \mathbb{R}^N \times \mathbb{R}^M \rightarrow \mathbb{C} \) is the joint characteristic function of the system response \( x(t; \beta) \) and the excitation \( y(t; \beta) \).

**Proof:** We consider the component wise form of the FDE, used for the proof of Theorem 2.1. 

\[
\frac{1}{i} \frac{d}{dt} \frac{\delta \mathcal{F}_{xy}(u, v)}{\delta u_n(t)} = \sum_{k=0}^{K} \sum_{l=0}^{L} \frac{\partial^{k+l} F_n(t, 0, 0)}{\partial u_i(t) \partial v_j(t)} \frac{\delta^{k+l} \mathcal{F}_{xy}(u, v)}{\delta u_i(t) \delta v_j(t)}
\]

We separate the r.h.s. term as follows,

\[
\sum_{k=0}^{K} \sum_{l=0}^{L} \frac{\partial^{k+l} F_n(t, 0, 0)}{\partial u_i(t) \partial v_j(t)} \frac{\delta^{k+l} \mathcal{F}_{xy}(u, v)}{\delta u_i(t) \delta v_j(t)}
\]

Setting to the system of functional differential equations \( u = (u_1, u_2, \ldots, u_N) \cdot \delta(\bullet - t) \) and \( v = (v_1, v_2, \ldots, v_M) \cdot \delta(\bullet - s) \) we have using same arguments as for the proof of Liouville's equation,

\[
\lim_{s \to t} \frac{1}{i} \frac{d}{dt} \frac{\delta \mathcal{F}_{xy}(v, v(\bullet - t), \nu \cdot \delta(\bullet - s))}{\delta u_n(t)} =
\]

\[
= \lim_{s \to t} \int_{C^R \times \mathbb{C}^N \times \mathbb{C}^M} iv_n \hat{x}_n(t) \exp \left( i \sum_{k=1}^{N} x_k(t) v_k + i \sum_{k=1}^{M} y_k(s) v_k \right) \mathcal{F}_{xy}(dx, dy)
\]

For the first r.h.s term we have,

\[
\lim_{s \to t} \sum_{k=0}^{K} \sum_{l=0}^{L} \frac{\partial^{k+l} F_n(t, 0, 0)}{\partial x_i(t) \partial v_j(t)} \frac{\delta^{k+l} \mathcal{F}_{xy}(v, v(\bullet - t), \nu \cdot \delta(\bullet - s))}{\delta x_i(t) \delta v_j(t)} =
\]

\[
= \sum_{k=0}^{K} \sum_{l=0}^{L} \frac{\partial^{k+l} F_n(t, 0, 0)}{\partial x_i \partial v_j} \frac{\delta^{k+l} \mathcal{F}_{xy}(v, v(t), \nu \cdot \delta(\bullet - s))}{\delta x_i \delta v_j}
\]

Similarly for the second r.h.s. term we have,

\[
\lim_{s \to t} \sum_{k=0}^{K} \sum_{l=0}^{L} \frac{\partial^{k+l} F_n(t, 0, 0)}{\partial x_i \partial v_j} \frac{\delta^{k+l} \mathcal{F}_{xy}(v, v(\bullet - t), \nu \cdot \delta(\bullet - s))}{\delta x_i \delta v_j} =
\]

\[
= \sum_{k=0}^{K} \sum_{l=0}^{L} \frac{\partial^{k+l} F_n(t, 0, 0)}{\partial x_i \partial v_j} \frac{\delta^{k+l} \mathcal{F}_{xy}(v, v(t), \nu \cdot \delta(\bullet - s))}{\delta x_i \delta v_j}
\]
Where we have used the notation,
\[
\partial x_i = \partial x_i^h \cdot \partial x_i^h \cdots \partial x_N^h
\]
\[
x^i = x_i^1 \cdot x_i^2 \cdots x_N^i
\]
Multiplying all terms with \( iv_n \) and summing up for all \( n = 1, 2, \ldots, N \), we will have for the l.h.s. term,
\[
\lim_{s \to t} \sum_{n=1}^{N} \int_{C_n^+, C_n^-} iv_n \dot{x}_n(t) \exp \left\{ i \sum_{k=1}^{N} x_k(t) v_{n_k} + i \sum_{k=1}^{M} y_k(s) \nu_{n_k} \right\} \mathcal{D}_{xy}(dx, dy) =
\]
\[
= \lim_{s \to t} \frac{\partial \phi_{xy}(t, \nu; s, \nu)}{\partial t} \bigg|_{s=t}
\]
Making the summation for the r.h.s. terms we have the partial differential equation,
\[
\frac{\partial \phi_{xy}(t, \nu; s, \nu)}{\partial t} \bigg|_{s=t} = \sum_{n=1}^{N} \sum_{i=1}^{L} \sum_{j \in J} \frac{\partial F_n(t, 0, 0)}{\partial x_i} \frac{\partial \phi_{xy}(\nu, t; \nu, t)}{\partial v_i} +
\]
\[
+ \sum_{n=1}^{N} \sum_{k=1}^{K} \sum_{l=0}^{L} \sum_{(i,j) \in F_n^+, F_n^-} \frac{\partial^{k+\ell} F_n(t, 0, 0)}{\partial x_i \partial y_j} \frac{\partial \phi_{xy}(\nu, t; \nu, t)}{\partial v_i \partial v_j}
\]
It is easy to verify conditions (65) and (66). ■

Remark 3.18: The separation of the r.h.s. that we have made allows us to give the physical meaning for the above equation. Thus, the l.h.s. of equation (64), which is equivalent with the l.h.s. of Liouville equation express the conservation of ‘probability’ for the state space variables \( x(t; \beta) \). The r.h.s term, which is due to the excitation term, express the transfer-interaction of ‘probability’ from the stochastic process \( y(t; \beta) \) to the state space variables \( x(t; \beta) \). Of course the stochastic characteristics of the excitation \( y(t; \beta) \) are introduced with condition (65). ■
5.4. Kernel representations for the Characteristic Functional

We will now study methods for the efficient numerical solution of the Functional Differential Equations described at Section 5.4.2 of the present chapter. These methods, in contrast to the previous section (based on the reduction of the FDE’s to PDE’s) will be based on analytical calculations of functional integrals (EGOROV, A.D. & SOBOLEVSKY, P.I. & YANOVICH, L.A., Functional Integrals: Approximate Evaluation and Applications) presented below. More specifically, using the linearity property of the Functional Differential Equations, we will apply a convex superposition of characteristic functionals (kernel characteristic functionals). Then two different approaches will be formulated. The first one will superpose kernel characteristic functionals with given parameters, efficiently chosen, such that the corresponding measures be local in space and time, i.e. in \( \mathbb{R}^N \). We will explain this property extensively in what follows (Section 5.4.3). Thus we will get a Gallerkin type approximation at functional level. The second approach will work in two stages. In the first stage we will approximate the initial condition distribution from a superposition of kernel characteristic functionals. In the second stage we will let the characteristic functionals to evolve in time with constants amplitudes (those that we computed). After an efficiently chosen time interval we will recompute the amplitudes, by approximating the computed probabilistic stage as it was an initial condition using suitable criteria.

5.4.1. Exact Formula for Integrals of Special Functionals

a) Integration with respect to Independent Increment Processes

We will present the above theorems for the more general case of \( X \) be a linear topological space (separable, locally convex linear space). \( X' \) will denote the dual space of linear continuous functionals on \( X \). In what follows, we shall consider without any loss of generality the mean value of a Gaussian measure \( \mathcal{P} \) to be equal to zero. In may cases, we may without any loss of generality consider the correlation operator of \( \mathcal{P} \), \( C(x,x) \) to be nondegenerate, i.e., \( C(x,x) = 0 \) if and only if \( x = 0 \) (we have substitute the notation \( \langle Cx, x \rangle \) of Chapter II, for the current suction, with \( C(x,x) \) to avoid confusions with the dual product). Otherwise, we could find the subspace \( X_0 \subset X \) were the measure is concentrated, and the correlation operator will no longer be degenerate there.

The following construction is described and substained in papers (see EGOROV, A.D. & SOBOLEVSKY, P.I. & YANOVICH, L.A., Functional Integrals: Approximate Evaluation and Applications p.16) in more details. The Hilbert space which is the closure of the set of functionals of the form \( \langle \xi, x \rangle \), \( \xi \in X' \) in space \( L_2(X, \mathcal{P}) \) will be denoted by \( H \) and \( \mathcal{H} \subset X \) denotes the Hilbert subspace dual to \( H \) whose closure in \( X \) is the support of measure \( \mathcal{P} \). For almost all \( x \in X \), a functional \( \langle a, x \rangle \) \( a \in \mathcal{H}, x \in X \) is defined. It is specified by the series

\[
\langle a, x \rangle = \sum_{k=1}^{\infty} \langle \phi_k, x \rangle (a, e_k)_{\mathcal{H}}
\]  

(1)
where \( \{ \phi_k \}, \{ e_k \}, k = 1, 2, \ldots \), are orthonormal bases in \( H \) and \( \mathcal{H} \), respectively, with \( \phi_k \in X' \) for all \( k \) and \( \langle \phi_k, e_j \rangle = \delta_{kj} \). We assume that the space \( \mathcal{H} \) is sepearable for all spaces \( X \) and measures \( \mathcal{P} \) considered. In particular, this is true for separable Frechet spaces and for other spaces considered here.

Functionals of the form (1) are called \textit{measurable linear functionals}. Note that

\[
\langle e_k, x \rangle = \langle \phi_k, x \rangle, \quad \langle a, h \rangle = \sum_{k=1}^{\infty} \langle \phi_k, x \rangle (a, e_k) = (a, h)_{\mathcal{H}}
\]

for \( a, h \in \mathcal{H} \).

It is an important property of the considered spaces with Gaussian measure that they admit the following expansion

\[
x = \sum_{k=1}^{\infty} (e_k, x) e_k = \sum_{k=1}^{\infty} (\phi_k, x) e_k
\]

(2)

which converges under topology of space \( X \) for almost all \( x \in X \). Proofs of the convergence of expansion (2) may be found in papers. The definition of \( H \) implies that \( \langle \xi, \eta \rangle_H = \int \langle \xi, x \rangle \langle \eta, x \rangle \mathcal{P} (dx) = C(\xi, \eta) \). Let \( T \) define an isomorphism of space \( \mathcal{H} \) into \( H \) which assigns a basis \( \{ \phi_k \} \) to the basis \( \{ e_k \} \). Then

\[
\langle \xi, x \rangle = \sum_{k=1}^{\infty} \langle \xi, \phi_k \rangle_H \langle \phi_k, x \rangle = \sum_{k=1}^{\infty} (T^{-1} \xi, e_k)_{\mathcal{H}} \langle \phi_k, x \rangle = (T^{-1} \xi, x)
\]

(3)

\[
C(\xi, \eta) = \sum_{k=1}^{\infty} \langle \xi, \phi_k \rangle_H \langle \eta, \phi_k \rangle_H = \sum_{k=1}^{\infty} (T^{-1} \xi, e_k)_{\mathcal{H}} \langle T^{-1} \eta, e_k \rangle = (T^{-1} \xi, T^{-1} \eta)
\]

For the proofs of the theorems presented below we will need the following transformation formulae for the integrals with respect to Gaussian measure under translation and under general linear transformation may be obtained. Thus, suppose \( F(x) \) is a functional to be integrated and \( a \in \mathcal{H} \). Then the transformation formula for an integral with respect to Gaussian measure under translation is as follows

\[
\int_X F(x) \mathcal{P} (dx) = \exp \left\{ -\frac{1}{2} \| a \|_{\mathcal{H}}^2 \right\} \int_X F(x + a) \exp \left\{ -\langle a, x \rangle \right\} \mathcal{P} (dx)
\]

(4)

In order to state the next result, we need some definitions. A compact operator \( A \) in Hilbert space is called a \textit{Hilbert-Schmidt operator}, if \( \sum_{k=1}^{\infty} \lambda_k^2 < \infty \), where \( \{ \lambda_k \} (k = 1, 2, \ldots) \) is the sequence of its eigenvalues of \( A \). On the other hand, if the series \( \sum_{k=1}^{\infty} | \lambda_k | < \infty \) converges then this operator is called the operator of trace class. Clearly, any trace class operator is Hilbert-Schmidt one. The product \( D_A(\lambda) = \prod_{k=1}^{\infty} (1 - \lambda \lambda_k) \), where \( \{ \lambda_k \} (k = 1, 2, \ldots) \) is the sequence of its eigenvalues of \( A \), is called the \textit{Fredholm determinant} at point \( \lambda \) for operator \( A \). The Fredholm determinant of a trace class operator is finite for any \( \lambda \). It is an entire function of complex variable \( \lambda \) with zeroes at points \( \lambda = \lambda_k^{-1} (k = 1, 2, \ldots) \). If \( A \) is a Hilbert-Schmidt
5.4. KERNEL REPRESENTATIONS FOR THE CHARACTERISTIC FUNCTIONAL

operator then the product \( \prod_{k=1}^{\infty} (1 - \lambda \lambda_k) \) may diverge. For such operators, the determinant of the form \( \delta_{A}(\lambda) = \prod_{k=1}^{\infty} (1 - \lambda \lambda_k) e^{\lambda \lambda_k} \) is introduced and is called the Carleman determinant.

Let us proceed now to the evaluation of integrals with respect to Gaussian measure for particular functionals. First we shall evaluate integrals of functionals which are functions of measurable linear functionals

**Theorem 4.1**: Let \( a_1, \ldots, a_n \) be linearly independent elements from \( \mathcal{H} \). Then the following equality holds

\[
\int_{\mathbb{R}^n} F\left((a_1, a_2, \ldots, a_n, x)\right) \mathcal{D}^\sigma(dx) = \left(2\pi\right)^{-n/2} \left|\det A\right|^{-1/2} \int_{\mathbb{R}^n} \exp\left(-\frac{1}{2} (A^{-1} u, u)\right) F(u) du \quad (5)
\]

if any of these integrals exists, where \( A \) is a matrix with elements \( [A_{ij}] = (a_i, a_j)_{\mathbb{R}^n} \) \((i, j = 1, 2, \ldots, n)\).

**Proof**: We shall make use of formula (4) for the transformation of integrals under translation, wherein we set \( a = -\lambda \sum_{k=1}^{n} v_k a_k \) , with \( \lambda \) and \( v_k \) be real numbers, and \( F(x) = 1 \). Then we have

\[
\int_{\mathbb{R}^n} \exp\left(\frac{\lambda}{2} \left|\sum_{k=1}^{n} v_k a_k\right|^2\right) = \exp\left(\frac{\lambda^2}{2} (A v, v)\right)
\]

This equality is also valid for \( \lambda = -i \). This together with the following formula

\[
\int_{\mathbb{R}^n} \exp\left(-\frac{1}{2} (A^{-1} u, u) - i (u, v)\right) du = \left(2\pi\right)^{-n/2} \left|\det A\right|^{1/2} \exp\left(-\frac{1}{2} (Av, v)\right)
\]

implies that

\[
\int_{\mathbb{R}^n} \mathcal{D}^\sigma(dx) = \left(2\pi\right)^{-n/2} \left|\det A\right|^{-1/2} \int_{\mathbb{R}^n} \exp\left(-\frac{1}{2} (A^{-1} u, u)\right) e^{-i(u,v)} du
\]

i.e. formula (5) is valid for \( F(u) = e^{-i(u,v)} \). Let now the function be specified by its Fourier transform, i.e. \( F(u) = \int f(v) e^{-i(u,v)} dv \). Let us multiply the previous relation by \( f(v) \) and integrate it over space \( \mathbb{R}^n \) with respect to the variables \( v_1, v_2, \ldots, v_n \). After changing the order of integration we obtain (5). The validity of the formula for arbitrary functions \( F(u) \) is verified by passage to the limit. ■

**Remark 4.2**: For the case of integration with respect to a Gaussian measure with non zero mean value \( m \) we have

\[
\int_{\mathbb{R}^n} F\left((a_1, x), \ldots, (a_n, x)\right) \mathcal{D}^\sigma(dx) = \int_{\mathbb{R}^n} F\left((a_1, x) + (a_1, m), \ldots, (a_n, x) + (a_n, m)\right) \mathcal{D}^\sigma(dx) = \left(2\pi\right)^{-n/2} \left|\det A\right|^{-1/2} \int_{\mathbb{R}^n} \exp\left(-\frac{1}{2} (A^{-1} u, u)\right) F(u + (a_1, m), \ldots, u_n + (a_n, m)) du
\]
where \( \mathcal{P} \), is the Gaussian measure \( \mathcal{P} \) with zero mean value. Making the suitable translation transformation to the last integral we have
\[
\int_{x} F((a_{1}, x, \ldots, a_{n}, x)) \mathcal{P}(dx) = (2\pi)^{-n/2} \left( \det A \right)^{-1/2} \int_{\mathbb{R}^{n}} \exp \left\{ -\frac{1}{2} \left[ A^{-1} [u - (a, m)] \right] \left[ u - (a, m) \right]^\top \right\} F(u) du \quad \blacksquare \tag{6}
\]

**Remark 4.3:** For case of \( X \) be a separable real Hilbert space, matrix \( A \) with \( [A]_{ij} = (a_{j}, a_{i})_{\mathcal{H}} \) will take the simple form \( [A]_{ij} = C(a_{j}, a_{i}) \). \( \blacksquare \)

Let us mention two special cases of the obtained formula. For \( n = 1 \) we have
\[
\int_{x} F((a, x)) \mathcal{P}(dx) = (2\pi)^{-1/2} \left\| a \right\|^{-1} \int_{\mathbb{R}} \exp \left\{ -\frac{u^2}{2\left\| a \right\|^2} \right\} F(u) du
\]

Let \( a_{1}, a_{2}, \ldots, a_{n} \) be orthogonal elements of space \( \mathbb{R}^{n} \). Then
\[
\int_{x} F((a_{1}, x, \ldots, a_{n}, x)) \mathcal{P}(dx) = (2\pi)^{-n/2} \int_{\mathbb{R}^{n}} \exp \left\{ -\frac{1}{2} (u, u) \right\} F(u) du
\]

Consider now evaluation of integrals of functionals which are functions of quadratic functionals on \( X \) (for example the Tatarskii representation)
\[
A(x, x) = \sum_{k, j=1}^{\infty} a_{ij} (e_{j}, x) (e_{k}, x) \tag{7}
\]
where \( a_{ij} = (Ae_{k}, e_{j}) \); \( A \) is a self-adjoint trace class operator on \( \mathcal{H} \); \( \{e_{k}\}, k = 1, 2, \ldots \), is an orthonormal basis in \( \mathcal{H} \). The double series in (7) converges for almost all \( x \in X \). The functional \( A(x, x) \) does not depend on the choice of basis \( \{e_{k}\} \) and may be written as follows
\[
A(x, x) = \sum_{k, j=1}^{\infty} \lambda_{k} (e_{k}, x)^2
\]
where \( \{\lambda_{k}\} \) and \( \{e_{k}\} \) are the eigenvalues and the orthonormal eigenvectors of operator \( A \), respectively.

The following integrals are easy to evaluate with the help of (5)
\[
\int_{x} A(x, x) \mathcal{P}(dx) = \text{trace}[A]
\]
\[
\int_{x} A^{2}(x, x) \mathcal{P}(dx) = \text{trace}^2[A] + 2\text{trace}[A^2]
\]

Let us prove the equality
\[
\int_{x} \exp \left\{ \frac{\lambda}{2} A(x, x) \right\} \mathcal{P}(dx) = [D_{\lambda}]^{-1/2}
\]
where $D_d(\lambda)$ is the Fredholm determinant of operator $A$ at point $\lambda$; $\Re \lambda < \lambda_i^{-1}$ ($\lambda_i$ is the largest eigenvalue) and

$$\left[ D_d(\lambda) \right]^{-1/2} = \left| D_d(\lambda) \right|^{-1/2} \exp \left\{-\frac{i}{2} \text{arg} D_d(\lambda) \right\}$$

In fact,

$$\int_x \exp \left[ \frac{\lambda}{2} A(x,x) \right] \varphi^\prime (dx) = \int_x \exp \left[ \frac{\lambda}{2} \sum_{k,j=1}^\infty \lambda_k e_k(x)^2 \right] \varphi^\prime (dx) =$$

$$= \prod_{k=1}^\infty \left( \frac{2\pi}{2} \right)^{-1/2} \int_\mathbb{R} \exp \left\{ \frac{1}{2} (\lambda \lambda_k - 1) u^2 \right\} du = \prod_{k=1}^\infty (1 - \lambda \lambda_k)^{-1/2} = \left[ D_d(\lambda) \right]^{-1/2}$$

We may similarly evaluate an integral of the more general form

$$\int_x \exp [\lambda A(x,x) + b(g,g)] \varphi^\prime (dx) = \left[ D_d(2\lambda) \right]^{-1/2} \exp \left\{ \frac{b^2}{2} (B_d(2\lambda)g,g)_{\mathcal{H}} \right\}$$

where $g \in \mathcal{H}$, $B_d(2\lambda) = (1 - 2\lambda A)^{-1}$, $\Re \lambda < (2\lambda_i)^{-1}$ ($\lambda_i$ is the largest eigenvalue); $b$ is a numeric parameter; $I$ is the identity operator.

Now formula (8) will be exploited in the proof of a more general result.

**Theorem 4.4** : Let $A_1(x,x), \ldots, A_n(x,x)$ be quadratic functionals of form (7) and let $g_1, g_2, \ldots, g_m$ be linearly independent elements of space $\mathcal{H}$. Then the following equality holds

$$\int_x F(A_1(x,x), \ldots, A_n(x,x); (g_1,x), \ldots, (g_m,x)) \varphi^\prime (dx) = \int_{\mathbb{R}^{2n+m}} p(u;\nu) F(u;\nu) du d\nu$$

(9)

(If any of these integrals exists), where

$$p(u;\nu) = (2\pi)^{-n-m/2} \int_{\mathbb{R}^n} \exp \left\{ i(u,\xi) \right\} [D_d(\lambda)]^{-1/2} \exp \left\{ -\frac{1}{2} (S(\xi)^{-1} u, \nu) \right\} d\xi$$

$S(\xi)$ is a matrix with elements $[S_{ij}(\xi)] = \left[ (I + 2iA(\xi))^{-1} g_k, g_j \right]_{\mathcal{H}}$; $[D_d(\lambda)]^{-1/2}$ is the Fredholm determinant of the operator $A(\xi) = \sum_{k=1}^n \xi_k A_k$ at point $\lambda$; $u = (u_1, \ldots, u_n)$ and $\nu = (\nu_1, \ldots, \nu_m)$.

**Proof** : The proof of this theorem is similar to that of Theorem 4.1, i.e., first we prove the validity of formula (9) with the help of formula (8) for the functions $F(u;\nu)$ which are specified by their Fourier transforms; and then we prove the formula for arbitrary functions by passing to the limit. ■

**Remark 4.5** : As we notice at Remark 4.2 for the case of integration with respect to a Gaussian measure with non zero mean value $m$ we have

$$\int_x F(A_1(x,x), \ldots, A_n(x,x); (g_1,x), \ldots, (g_m,x)) \varphi^\prime (dx) =$$

$$= \int_x F(A_1(x+m,x+m), \ldots, A_n(x+m,x+m); (g_1,x+m), \ldots, (g_m,x+m)) \varphi^\prime (dx) =$$
\[ \int_{\mathbb{R}^{n+n}} p(\vec{u};\vec{v}) \tilde{F}(\vec{u};\vec{v}) d\vec{u} d\vec{v} \]

where \( \tilde{F} \), is the Gaussian measure \( \mathcal{S} \) with zero mean value and we have apply Theorem 4.4 for

\[ \tilde{F} = F(A_1(x,x) + 2A_1(m,x) + A_1(m,m),..., A_n(x,x) + 2A_n(m,x) + A_n(m,m);(g_1,x+m),..., (g_m,x+m)) \]

**Remark 4.6:** For case of \( X \) be a separable real Hilbert space, matrix \( S \) will take the simple form \( [S_{ij}(\xi)] = C\left[[I + 2iA(x)]^{-1}g_i,g_j\right] \) where \( C \) is the covariance operator of the Gaussian measure.

Consider special cases of formula (9). If \( F(\vec{u};\vec{v}) \) does not depend on \( \vec{u} \), i.e., \( F(\vec{u};\vec{v}) = F(\vec{v}) \), then we obtain formula (5). Suppose that \( F(\vec{u};\vec{v}) = F(\vec{u}) \), i.e., it does not depend on \( \vec{v} \), then

\[ \int_X F(A_1(x,x),...,A_n(x,x)) \mathcal{S}(dx) = \int_{\mathbb{R}^n} \rho(\vec{u}) F(\vec{u}) d\vec{u} \]

where \( \rho(\vec{u}) = (2\pi)^{-n} \int_{\mathbb{R}^n} \exp\{i(\vec{u},\vec{\xi})\} [D_{A(\xi)}(-2i)]^{-1/2} d\vec{\xi} \)

\( D_{A(\xi)}(-2i) \) is the Fredholm determinant of the operator \( A(\xi) = \sum_{k=1}^n \xi_1 A_k \) at point \( \lambda = -2i \).

Consider still another example

\[ \int_X \exp\left[\frac{\lambda}{2} A(x,x)\right] F((g,x)) \mathcal{S}(dx) = [2\pi D_0(\lambda) b(\lambda)]^{-1/2} \int_{\mathbb{R}} \exp\left[-\frac{\nu^2}{2b(\lambda)}\right] F(\nu) d\nu \]

where \( b(\lambda) = \left[[I - \lambda A]^{-1}g,g\right]_\nu \).

**b) Integration with respect to Independent Increment Processes**

For the completion of the Section we also report some useful formulas for the calculation of integrals with respect to measures which correspond to independent increment processes. Measures of this kind are featured by integrals of the form

\[ F(x(\bullet)) = f(\Delta x(t_1),...,\Delta x(t_n)) \]

where \( \Delta x(t_k) = x(t_k) - x(t_{k-1}) \). Thus for the measure which corresponds to Poisson process with the characteristic functional

\[ \mathcal{F}(x) = \exp\left[\lambda \int_T e^{\rho(u)\Delta x(u)} - 1\right] du \]

where

\[ \rho_i(u) = \begin{cases} 1, & \text{if } t \in [u,T] \\ 0, & \text{otherwise} \end{cases} \]
the following formula holds
\[
\int f(\Delta x(t_1), \ldots, \Delta x(t_n)) \mathcal{P}^\prime(dx) = \sum_{k_1, \ldots, k_n=0}^{\infty} f(k_1, \ldots, k_n) \prod_{i=1}^{n} \frac{(\lambda \Delta t_i)^{k_i}}{k_i!} e^{-\lambda \Delta t_i} \tag{10}
\]
and for the measures generated by the infinite-dimensional Gamma distribution with the characteristic functional
\[
\mathcal{F}(x) = \exp \left[ \int_T \left[ \ln \left( 1 - i \sigma \langle \rho_i(u), x \rangle \right) \right] du \right]
\]
where
\[
\rho_i(u) = \begin{cases} 1, & \text{if } t \in [u, T] \\ 0, & \text{otherwise} \end{cases}
\]
the following formula is valid
\[
\int f(\Delta x(t_1), \ldots, \Delta x(t_n)) \mathcal{P}^\prime(dx) = \int f(u_1, \ldots, u_n) \prod_{k=1}^{n} \frac{u_k^{2 \Delta t_i - 1}}{\Gamma(\Delta t_k) \sigma^{2 \rho_i}} e^{-\frac{u_k}{\sigma}} du \tag{11}
\]
where \( \Gamma(y) = \int_{-\infty}^{\infty} u^{y-1} e^{-u} du \) is Euler’s Gamma function, \( \mathbb{R}^+ \) is the product of the 4th positive half axes.
We shall also give an example for the case of the characteristic functional
\[
\mathcal{F}(x) = \exp \left[ -\int_0^1 \int_0^1 x(t) dt^{2^p} du \right] \tag{12}
\]
which is the special case of the characteristic functional of form (22) of Chapter 3/Section 3.9 for \( T = 1 \), \( \sigma^{2^p} \), \( \nu(du) = du \) and \( \rho_i(u) \) as we defined it before. Note that (10) may also be written in the form
\[
\mathcal{F}(x) = \exp \left[ -\int_0^1 \ldots \int_0^1 \min(t_1, \ldots, t_{2^p}) x(t_1) \ldots x(t_{2^p}) dt_1 \ldots dt_{2^p} \right] \tag{13}
\]
The following formula holds
\[
\int f(\Delta x(t_1), \ldots, \Delta x(t_n)) \mathcal{P}^\prime(dx) = \int f(u) \prod_{k=1}^{n} S(\Delta t_k, u_k - u_{k-1}) du \tag{14}
\]
where
\[
S(\tau, u) = \frac{1}{2\pi} \int_{\mathbb{R}} \exp \left\{ -\tau v^{2^p} + iuv \right\} dv
\]
is the fundamental solution of the parabolic equation
\[
\frac{\partial S}{\partial \tau} = (-1)^{p+1} \frac{\partial^{2^p} S}{\partial u^{2^p}}
\]
Using (12) we obtain the following formula
\[
\int_X f \left( \int_0^1 a_i(\tau)x(\tau)d\tau, \ldots, \int_0^1 a_k(\tau)x(\tau)d\tau \right) \mathcal{P}^{\rho}(dx) = \\
= \int_{\mathbb{R}^n} f(u) \left( (2\pi)^{-n} \int_{\mathbb{R}^n} \exp \left\{ -i(u,v) - \int_0^1 \left( \sum_{k=1}^n v_k \int_0^1 a_k(s)ds \right)^{2\rho} d\tau \right\} dv \right) du
\]

where \((u,v) = \sum_{k=1}^n u_k v_k\), which is valid under the condition of measurability of \(f(u_1,\ldots,u_n)\) and the condition

\[
\left| f(u_1,\ldots,u_n) \exp \left\{ -\varepsilon(u,u)^{\frac{\rho}{2\rho-1}} \right\} \right| \leq H(u_1,\ldots,u_n)
\]

for any \(\varepsilon > 0\), where

\[
\int_{\mathbb{R}^n} H(u_1,\ldots,u_n) du_1 \ldots du_n < \infty.
\]

5.4.2. Integrals of Variations and of Derivatives of Functionals.

Let functional \( F(x) \) be differentiable along direction \( a \in \mathcal{H} \) at any point \( x \in X \) and let 
\[
\sup_{|\lambda| < \varepsilon} \frac{d}{d\lambda} F(x + \lambda a)
\]
be summable for some \( \varepsilon > 0 \). Then the following relation holds
\[
\int_x \delta F(x)[a] \mathcal{D} \rho (dx) = \int_x F(x)(a,x) \mathcal{D} \rho (dx)
\]
The latter relation follows from the identity
\[
\int_x F(x + \lambda a) \mathcal{D} \rho (dx) = \exp \left[ -\frac{\lambda^2}{2} \|a\|_{\mathcal{H}}^2 \right] \int_x F(x) \exp \{\lambda(a,x)\} \mathcal{D} \rho (dx)
\]
(which is a modification of formula (4)), if we differentiate it with respect to \( \lambda \) and set \( \lambda = 0 \). Moreover the above condition enables differentiation under the integral sign.

The functionals Hermite polynomials may be defined by the equality
\[
H^{(n)}[x; a_1, a_2, \ldots, a_n] = \frac{\partial^n}{\partial \lambda_1 \partial \lambda_2 \ldots \partial \lambda_n} G[x; a_1, a_2, \ldots, a_n] \bigg|_{\lambda_m = \ldots = \lambda_n = 0}
\]
where
\[
G[x; a_1, a_2, \ldots, a_n] = \exp \left[ -\frac{1}{2} \sum_{i=1}^n \lambda_i \lambda_j (a_i, a_j)_{\mathcal{H}} + \sum_{i=1}^n \lambda_i (a_i, x) \right]
\]
\( a_i \in \mathcal{H} \); the functional \( G[x; a_1, a_2, \ldots, a_n] \) is defined for almost all \( x \in X \). The following recurrence relation holds for functional Hermite polynomials
\[
H^{(n)}[x; a_1, a_2, \ldots, a_n] = (a_n, x) H^{(n-1)}[x; a_1, a_2, \ldots, a_{n-1}] - \sum_{i=1}^{n-1} (a_i, a_n)_{\mathcal{H}} H^{(n-2)}[x; a_1, \ldots, a_i, \ldots, a_{n-1}]
\]
(here the “hat” over \( a_i \) means that \( a_i \) must be omitted); it is verified by an immediate calculation with the help of the above-mentioned definition of functional Hermite polynomials. The explicit form of them is as follows.
\[
H^{(n)}[x; a_1, a_2, \ldots, a_n] = \prod_{i=1}^n (a_i, x) + \sum_{s=1}^{n/2} \frac{(-1)^s}{(2s)!(n-2s)!} \times \\
\times \sum_{j_1, \ldots, j_{n-s}} \prod_{p=1}^s (a_{j_p}, a_{j_p})_{\mathcal{H}} \prod_{q=2s+1}^n (a_{j_q}, x)
\]
where \( (2s)! = 2 \cdot 4 \cdots 2s \), \( \lfloor \cdot \rfloor \) is the floor function. The first three Hermite polynomials are
\[
H^{(1)}[x; a] = (a, x),
\]
\[
H^{(2)}[x; a_1, a_2] = (a_1, x)(a_2, x) - (a_1, a_2)_{\mathcal{H}} ,
\]
\[ H^{(3)}[x; a_1, a_2, a_3] = (a_1, x) (a_2, x) (a_3, x) - (a_1, a_3)_{\mathcal{J}} (a_2, x) - (a_1, a_2)_{\mathcal{J}} (a_3, x) - (a_2, a_3)_{\mathcal{J}} (a_1, x) \]

Note also the special case of \( a_1 = a_2 = \ldots = a_n = a \)

\[ H^{(n)}[x; a] = H^{(n)}[x, a, a, \ldots, a] = \sum_{j=1}^{\left\lfloor \frac{n}{2} \right\rfloor} \frac{(-1)^j n!}{(2j)!!(n-2j)!!} ||a||_{\mathcal{J}}^{2j} (a, x)^{n-2j} \]

The following formula holds

\[
\int_{\mathcal{X}} \delta^n F \left( x \right) [a_1, a_2, \ldots, a_n] \mathcal{P} (dx) = \int_{\mathcal{X}} F (x) H^{(n)}[x; a_1, a_2, \ldots, a_n] \mathcal{P} (dx)
\]

provided that these integrals exist and

\[
\sup_{\lambda \varepsilon} \left| \frac{\partial}{\partial \lambda_1 \ldots \partial \lambda_n} F \left( x + \sum_{i=1}^{n} \lambda_i a_i \right) \right|
\]

is summable for some \( \varepsilon > 0 \). The formula follows immediately from the definition of the n-th order variation, the definition of Hermite polynomials given at the beginning of the section, and relation (4) or its modification also given at the beginning of the section.
5.4.3. Kernel Characteristic Functionals

We shall now use the theorems for analytical calculations of infinite dimensional integrals presented above. As we mention at the introduction of the section we will use a convex superposition of kernel characteristic functionals, i.e. functionals that correspond to probability measure which are local into the functional space (kernel probability measures). These analytical calculations, will allows us to get a system of differential equations for the parameter functions of the kernel characteristic functionals.

First we will give a more precise definition of kernel characteristic functionals and some simple examples. Then we will illustrate the proposed method, by applying to a first order equation with cubic nonlinearity, which has been studied at Section 5.3.4 of the present chapter. In following section, we will also present the theorems concerning the approximation of an induced probability measure of finite dimension (density) from a kernel type representation.

Thus, let us introduce a precise definition of kernel characteristic functionals

**Definition 4.7 [(ε, h)− Kernel Characteristic Functionals]**: Let $\mathcal{H}$ be a linear topological space(1) which will be mainly thought of as a space of functions defined on the time interval $I$. Let also $\mathcal{B} (\mathcal{H})$ denote the $\sigma$–field generated by the open subsets of $\mathcal{H}$ (the Borel $\sigma$–field of $\mathcal{H}$), $m \in \mathcal{H}$, $h \in \mathbb{R}_+$ and $\mathcal{U}_{m,h}$ be a neighbourhood of $m$ with concentration parameter (e.g. radius, in the case of metric spaces) $h$. [For example, $\mathcal{U}_{m,h}$ can have the form $\mathcal{U}_{m,h} = \{u \in \mathcal{H} : \|u - m\| \leq h\}$ if an appropriate norm is available.]

Given $\varepsilon > 0$, we shall say that a characteristic functional $\mathcal{F} (u;m,h)$ is an $(\varepsilon,h)$−kernel characteristic functional(2) iff the associated probability measure $\mathcal{P}$ on $\mathcal{B} (\mathcal{H})$ is concentrated around the element $m \in \mathcal{H}$ in the sense that

$$\mathcal{P} (\mathcal{U}_{m,h}) \geq 1 - \varepsilon.$$  ■

The existence of $(\varepsilon,h)$−kernel ch. functionals, for any given $\varepsilon > 0$ and $h > 0$ (relatively small), will be established in Section 5.4.3a by means of construction of a specific example.

**Remark 4.8**: To understand the above definition we study the space of real valued continuous functions with the sup norm defined on the time interval $I = [0,T]$. Then an arbitrary set of the form $\mathcal{U}_{m,h} = \{u \in \mathcal{H} : \sup_{t \in I} |u(t) - m(t)| \leq h\}$ will be a set of functions that take values within the shaded region.

---

(1) Reflexive sometimes Banach or Hilbert
(2) The terms $\varepsilon$−kernel ch. functional or -simpler- kernel ch. functional will also be used if no confusion is likely to occur.
As \( h \to 0 \) the above region will be reduced to the element \( m \). The form of the characteristic functional/probability measure is evolved from the parameter \( h \) (bandwidth) which adjust the variance of the probability mass about the element \( m \). Usually the parameter \( h \) is taken as the maximum value of the variance, \([C(t,t)]^{1/2}\) and \( m \) as the mean value element of the probability measure.

The notion of \((\varepsilon, h)\)–kernel ch. functionals will now be exploited to establish a fundamental result concerning the approximation of any probability measure (at the level of projection into finite-dimensional subspaces). More precisely, we shall prove that any finite-dimensional projection of a given probability measure \( \mathcal{P} \) on \( \mathcal{B} (\mathcal{H}) \), can be approximated by the corresponding projection of an appropriate convex superposition of kernel ch. functionals.

Let \( \mathcal{P} \) be a given probability measure defined into the measurable space \((\mathcal{H}, \mathcal{B} (\mathcal{H}))\) and let \( \Pi_L \) be a projection of \( \mathcal{H} \) into the finite-dimensional subspace \( L \). We then have the projection (See Definition 1.6/Chapter 3)

\[
\mathcal{P}_{\Pi_L} (B) = \mathcal{P}_{\Pi_L} (\Pi_L^{-1} [B]) = \mathcal{P}_{\Pi_L} \left( \{ x \in \mathcal{H} : \Pi_L [x] \in B \} \right), \quad B \in \mathcal{B} (L)
\]

which is a probability measure defined into the finite dimensional subspace \( L \). Let also the set of all finite linear combinations (convex) of the form

\[
\mathcal{P}^n (\ast) = \sum_{k=1}^{n} p_k \mathcal{P} (\ast; m_k, h_k), \quad \text{with} \quad p_n \geq 0, \quad n = 1, \ldots, N \quad \text{and} \quad \sum_n p_n = 1
\]

where every \( \mathcal{P} (\ast; m_k, h_k) \) \((k = 1, \ldots, n)\) is a kernel probability measure. Similarly we have the projection of the kernel probability measure

\[
\mathcal{P}^n_{\Pi_L} (B) = \mathcal{P}^n_{\Pi_L} (\Pi_L^{-1} [B]) = \mathcal{P}^n \left( \{ x \in \mathcal{H} : \Pi_L [x] \in B \} \right), \quad B \in \mathcal{B} (L)
\]

We then have the following

**Theorem 4.9:** For every projection \( \Pi_L \) from \( \mathcal{H} \) into the finite dimensional subspace \( L \), the set of measures \( \mathcal{P}^n_{\Pi_L} (B) \) is dense into the set of all continuous \( \mathcal{P}^n_{\Pi_L} \) with the same support.

Remark 4.10: In other words, for every measure $\mathcal{P}_{\mathcal{F}}(\bullet)$, which has continuous projection $\mathcal{P}_{\Pi_L}$ into the finite dimensional subspace $L$ and every $\varepsilon > 0$, given a kernel measure $\mathcal{P}^k$, there is a finite set of mean value elements $m_k \in \text{supp}[\mathcal{P}^k]$, and real numbers $h_k$ and $p_k$ $(k = 1, \ldots, n)$ satisfying conditions (15) such that

$$\left|\mathcal{P}_{\Pi_L}(B) - \mathcal{P}_{\Pi_L}^n(B)\right| < \varepsilon , \quad B \in \mathcal{B}(L)$$

Remark 4.11: By assuming the necessary continuity conditions we can have the corresponding results for the characteristic functional.

Remark 4.12: We must note that the kernel property of a probability measure pass along to all of its projections on finite dimensional subspaces in the sense that for every projection $\Pi_L$ the probability measure $\mathcal{P}_{\Pi_L}$ has the kernel property, i.e.,

$$\mathcal{P}_{\Pi_L}(\mathcal{U}_{m,h}) = \mathcal{P}\left(\Pi_L^{-1}[\mathcal{U}_{m,h}]\right) = \mathcal{P}\left(\{x \in \mathcal{H} : \Pi_L[\mathcal{U}_{m,h}] \in B\}\right) \geq 1 - \varepsilon .$$
5.4.3a Gaussian Kernel Characteristic Functionals

The most immediate application of the above definition is for the case of Gaussian characteristic functionals. Functionals of this type have the advantage of an explicit form for both the characteristic functional and the probability density functionals. Suppose we are working in a suitable space of functions defined on the time interval \( I = [0, T] \). Then the following theorem due to D.G. Konstant and V.I. Piterbarg (Extreme values of the cyclostationary gaussian random process) gives a set of sufficient conditions for a Gaussian probability measure to be of kernel type with the sup norm.

**Theorem 4.13**: Let \( X(t; \beta), t \in I \subseteq \mathbb{R} \), be a Gaussian random process which is differentiable in mean square, with (differentiable) mean-value function \( m(\cdot): I \rightarrow \mathbb{R} \) and set 
\[
d_i^2 = E^\beta\left[\left|X'(t; \beta) - m'(t)\right|^2\right] > 0.
\]
The covariance function \( C(\cdot, \cdot): I \times I \rightarrow \mathbb{R} \) satisfies the strict inequality 
\[
C(t, s) < \sigma(t)\sigma(s), \quad t \neq s,
\]
and the variance \( \sigma^2(t) = C(t, t) \) reaches its maximum \( \sigma^2(t_0) = \sigma_0^2 \) at a unique point \( t_0 \in I \). Moreover, we assume that there exist two constants, \( A > 0, b > 0 \), defining the local structure of \( \sigma^2(t) = C(t, t) \) at the vicinity of \( t_0 \in I \) such that 
\[
\sigma(t) = \sigma_0 - A|t - t_0|^b + o(|t - t_0|^b), \quad t \to t_0
\]
Then \( b \geq 2 \) (because of the differentiability), and
1. If \( b > 2 \), then 
\[
\mathcal{P}\left\{u \in \mathcal{U} : \max_{t \in I} |u(t) - m(t)| > h \right\} =
\]
\[
\Gamma\left(\frac{1}{b}\right) \frac{d_0}{\pi b A^2 \sigma_0^{2+\frac{2}{b}}} h^{-\frac{2}{b}} \exp\left(-\frac{h^2}{2\sigma_0^2}\right) \left[1 + o(1)\right], \quad h \to \infty,
\]
2. If \( b = 2 \), then 
\[
\mathcal{P}\left\{u \in \mathcal{U} : \max_{t \in I} |u(t) - m(t)| > h \right\} =
\]
\[
\sigma_0 \sqrt{1 + \frac{d_0^2}{2A\sigma_0^2}} h^{-1} \exp\left(-\frac{h^2}{2\sigma_0^2}\right) \left[1 + o(1)\right], \quad h \to \infty,
\]
where 
\[
d_0^2 = E^\beta\left[\left|X'(t_0; \beta) - m'(t_0)\right|^2\right] = d_{i0}^2 > 0.
\]

**Remark 4.14**: Note that the condition of differentiability is crucial. For example, if consider any stochastic process with no memory at all (see Chapter 2, Section 2.2.1.a) then the probability of the maximum will be 0 or 1 (0 or 1 Kolmogorov Law). To illustrate the above consider a white noise process with constant intensity \( D \). Then the maximum between \( n \) time instants will be
\[ \mathcal{P}\left\{ u \in \mathcal{H} : \max_{t \in [t_i, t_f]} |u(t)| > h \right\} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{n/2} D^n} \exp \left\{ -\frac{\sum_{k=1}^{n} u_k^2}{2D^2} \right\} du_1 \cdots du_k = \]
\[ = \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{n/2} D^n} \exp \left\{ -\frac{\sum_{k=1}^{n} u_k^2}{2D^2} \right\} du = \varepsilon^n \to 0 , \text{ as } m \to \infty . \]

Hence for the case of a Gaussian measure as described at Theorem 4.13 we will the description by the family of one-dimensional distributions,
\[ \mathcal{P}_x(a) = \mathcal{P}_G \left\{ y \in \mathcal{H} : \langle x, y \rangle \leq a \right\} = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi} C_1} \exp \left\{ \frac{(t-m(x))^2}{2C_1} \right\} dt , \quad x \in \mathcal{H} , \quad a \in \mathbb{R} \]

Setting to the above \( x(t) = \delta(t-t) \) we have the probability of values at a specific time \( t \),
\[ \mathcal{P}_G \left\{ y \in \mathcal{H} : y(t) \leq a \right\} = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi} C(t,t)} \exp \left\{ \frac{(t-m(t))^2}{2C(t,t)} \right\} dt , \quad a \in \mathbb{R} , \quad t \in I \]

Thus, we see that
- the probability mass is concentrated to functions that have small distance from the mean value element \( m : I \to \mathbb{R} \). By ‘small distance’ we mean that \( \sup_{t \in I} (x(t) - m(t)) \leq h^2 \).

Moreover we see that for \( h \to 0 \) the probability is concentrated to the mean value element.
- For a convex superposition the probability mass is shared between the supports of the Gaussian measures. More precisely let the convex superposition \( \mathcal{P}_G = \frac{1}{2} \mathcal{P}_{G_1} + \frac{1}{2} \mathcal{P}_{G_2} \) with the Gaussian measures \( \mathcal{P}_{G_1}, \mathcal{P}_{G_2} \) described by the parameters \( m_1, m_2 \) and \( C_1, C_2 \) respectively. This kind of description is very essential for the representation of nonlinear dynamical system’s response.
We must note that for the case described above, we have more precise information for the description of the probability mass. This is because we have the exact variation (through the variance \( C(t,t)^{1/2} \)) from the mean element \( m \) for each time instant \( t \). In other words we know that, given a covariance function with diagonal \( C(t,t) \) and the mean value element \( m(t) \), the probability mass will be concentrated into the functional set
\[
\mathcal{U} = \left\{ u \in \mathcal{H} : |u(t) - m(t)| \leq [C(t,t)]^{1/2} \right\}
\]
The figure below illustrates graphically the above assertion for a given \( C(t,t) \) and \( m(t) \).

Another interesting subject connected with the above discussion is the characterization of the derivative of the stochastic process. In Chapter II we saw that having the characteristic function of a stochastic process gives us the opportunity to get statistical information for the derivatives of the stochastic process. Now we are going to discuss how we can get qualitative results for the probability mass of the derivative of the stochastic process, by having the behaviour of the correlation operator. We must say that the same arguments hold for a general stochastic process, so we can make the desired generalizations.

Now, the correlation function of a mean-square continuous stochastic process can be expanded into a biorthogonal series of the form (Kahurent-Loeve expansion)
\[
C(t,s) = \sum_{i=1}^{\infty} \lambda_i f_i(t) f_i(s)
\]
where \( \{\lambda_i, f_i(t)\}_{i=1}^{\infty} \) are the eigenvalues – eigenfunctions of the correlation operator
\[
\langle Cx, x \rangle = \int_{I} \int_{I} C(t,s) x(t) x(s) dt ds
\]
Using the results of Chapter 3 we can have the distribution of the values for the derivative of the stochastic process.
5.4. KERNEL REPRESENTATIONS FOR THE CHARACTERISTIC FUNCTIONAL

\[ P_{x'[0]}(a) = \mathcal{D}\left\{ y \in \mathcal{H} ; y'(t_0) \leq a \right\} = \mathcal{D}_{\mathcal{H}_t(\tau_0)}(a) \]

Hence

\[ \mathcal{D}_G\left\{ y \in \mathcal{H} ; y'(t) \leq a \right\} = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{(t - m'(t))^2}{2\partial^2 C(t,s)_{t=m'}} \right\} dt, a \in \mathbb{R}, \ t \in T \]

Thus we must investigate the behaviour of the function \( \frac{\partial^2 C(t,s)}{\partial t \partial s} \) or using (17) the behaviour of the function

\[ \frac{\partial^2 C(t,s)}{\partial t \partial s} \bigg|_{t=m'} = \sum_{i=1}^{\infty} \lambda_i [f'_i(t)]^2 \]

Hence we can conclude that when the diagonal of the correlation function, \( C(t,t) \) change slowly with time, then the probability mass for the derivative of the stochastic process will concentrate very close to the derivative of the mean value element. Conversely, if the diagonal of the correlation function, \( C(t,t) \) change fast with time, even if decreases the probability mass will spread widely around the derivative of the mean value element.

The above results, although easy to derive, gives us a very important physical intuition for the distribution of the probability mass into the infinite dimensional space of functions. Moreover, based on the analysis, presented for the simple case of Gaussian measures, one can produce valuable results for general functionals (for example Gamma characteristic functional), or for the responses of dynamical systems. Hence for responses having the form of convex superposition of Gaussian functionals, which will be used in the sequel, we can immediately have the picture of the probability measure.
5.4.4. *Superposition of Kernel Characteristic Functionals for the numerical solution of FDEs*

In the present section we will apply the representation described above for the numerical solution of functional differential equation. Theorem 4.9 in combination with the linearity of the Functional Differential Equations is the basic reason for the use of such representations to the numerical solutions of stochastic differential equations. In what follows we will see how we can get equations describing efficiently the evolution of the unknown parameters.

Now let us describe in more detail the representation of the joint characteristic functional that we will use for the sequel. We assume that

a1. The random excitation is described by a stochastic process \( y(\bullet;\bullet) : I \times \Omega \rightarrow \mathbb{R}^M \) defined onto the probability space \( (\Omega, \mathcal{B}(\Omega), \mathcal{P}) \) with the associated probability space \( (C^\infty_M(I), \mathcal{B}(C^\infty_M), \mathcal{P}_y) \), and with characteristic functional \( \mathcal{F}_y \).

a2. The response is described by the stochastic process \( (x(\bullet;\bullet), y(\bullet;\bullet)) : I \times \Omega \rightarrow \mathbb{R}^N \times \mathbb{R}^M \) defined onto the probability space \( (\Omega, \mathcal{B}(\Omega), \mathcal{P}) \) with the associated probability space \( (C^\infty_N \times C^\infty_M, \mathcal{B}(C^\infty_N \times C^\infty_M), \mathcal{P}_{xy}) \), and with characteristic functional \( \mathcal{F}_{xy} \).

a3. The initial state of the system is described by a random variable \( x_0(\bullet) : \Omega \rightarrow \mathbb{R}^N \) with the associated probability space \( (\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \mathcal{P}_{x_0}) \) with characteristic function \( \mathcal{F}_0 \). (Where we have assumed that for the initial state the random variables \( x_0, y_0 \) are independent).

Consider the representation (15) of the joint probability measure using kernel probability measures with fixed real numbers \( p_k = \frac{1}{n} \), \( k = 1, \ldots, n \).

\[
\mathcal{P}_{xy}^n(x, y) = \frac{1}{n} \sum_{k=1}^{n} \mathcal{P}_{xy}(x, y; m_{x}^k, m_{y}^k, h_{x}^k, h_{y}^k), \quad x \in C^\infty_N, y \in C^\infty_M
\]  

(18)

with \( \mathcal{P}_{xy}(x, y; m_{x}^k, m_{y}^k, h_{x}^k, h_{y}^k) \) being a kernel probability measure (Definition 4.7) and \( m_{x}^k \in C^\infty_N, m_{y}^k \in C^\infty_M, h_{x}^k \in \mathbb{R}^N, h_{y}^k \in \mathbb{R}^M \) for all \( k = 1, \ldots, n \). The characteristic functional corresponding to the above probability measure is

\[
\mathcal{F}_{xy}^n(u, v) = \frac{1}{n} \sum_{k=1}^{n} \mathcal{F}_{xy}(u, v; m_{x}^k, m_{y}^k, h_{x}^k, h_{y}^k), \quad u \in C^\infty_N, v \in C^\infty_M
\]  

(19)

with \( \mathcal{F}_{xy} \) being the characteristic functional of the kernel probability measure \( \mathcal{P}_{xy} \).

**Remark 4.15 :** Note that the parameters \( m_{x}^k, m_{y}^k, h_{x}^k, h_{y}^k \) are not necessarily defining uniquely the kernel probability measures. In other words those parameters, express the kernel property of the probability measure \( \mathcal{P}_{xy} \) and they result from the specific parameters of the kernel representation. For example, in the case where \( \mathcal{P}_{xy} \) is a Gaussian kernel measure, \( m_{x}^k, m_{y}^k \), will be the mean value elements and \( h_{x}^k, h_{y}^k \) will be related with the correlation operator (more
specifically with the variance vector). Hence, for the sequel we emphasize, that, when we refer to the determination of the unknown parameters $m_y^k, m_x^k, h_x^k, h_y^k$, we will mean the determination of the unknown parameters that define uniquely the kernel probability measure (for the Gaussian measure, those will be the mean value and the correlation operator) and hence they result the above quantities.

The characteristic functional (19) must satisfy the Functional Differential Equation governing the dynamical system, as well as, conditions a) and c) described above. Hence the characteristic functional must verify the conditions

b1. Related with condition a1 for the characteristic functional of the excitation, we must have

$$\mathcal{F}_{x^n}(0, v) = \mathcal{F}_{x^n}(v) = \frac{1}{n} \sum_{k=1}^{n} \mathcal{F}_{x^n}(v; m_y^k, h_y^k) = \mathcal{F}_{x^n}(v), \quad v \in C^{\infty,M}_c$$

(20)

To verify the last condition we can solve the approximating problem defined by equation (20) and determine the unknown parameters $m_y^k \in C^{\infty,M}_c, h_y^k \in \mathbb{R}^M$ for all $k = 1, \ldots, n$. Alternatively we can set

$$\mathcal{F}_{x^n}(v; m_y^k, h_y^k) = \mathcal{F}_{x^n}(v), \quad v \in C^{\infty,M}_c, \quad k = 1, \ldots, n$$

(21)

In the latter case we must have a suitable kernel characteristic function such that (21) holds. This can be generally succeeded when the excitation process is described by a characteristic functional with explicit form (for example Gaussian or Gamma characteristic functional etc.). If this not the case we must determine the suitable parameters for the kernel characteristic functionals. In that case Theorem 4.3 guarantees that the equation (20) can be approximately satisfied.

b2. Related with condition a3 for the initial state of the system we must have

$$\mathcal{F}_{x^n}(v \cdot \delta_{t_0}, 0) = \mathcal{F}_{x^n}(v \cdot \delta_{t_0}) =$$

$$= \frac{1}{n} \sum_{k=1}^{n} \mathcal{F}_{x^n}(v \cdot \delta_{t_0}; m_y^k, h_y^k) = \mathcal{F}_0(v) \quad , v \in \mathbb{R}^N$$

(22)

As in the previous condition b1, similarly this condition will be satisfied by solving equation (22). The resulted values of the parameters will constitute the initial conditions for the set of differential equations that will be formulated using the Functional Differential Equation. Note that for the special case, where the initial states of the excitation and response are depending with each other the condition described above will take the form

$$\mathcal{F}_{x^n}(v \cdot \delta_{t_0}, \nu \cdot \delta_{t_0}) =$$

$$= \frac{1}{n} \sum_{k=1}^{n} \mathcal{F}_{x^n}(v \cdot \delta_{t_0}, \nu \cdot \delta_{t_0}; m_y^k, m_x^k, h_x^k, h_y^k) = \mathcal{F}_0(v, \nu), \quad v \in \mathbb{R}^N, \nu \in \mathbb{R}^M$$

(23)

where $\mathcal{F}_0(v, \nu)$ will be the joint characteristic functional of the random variable $(x(0;.), y(0;)) : \Omega \to \mathbb{R}^N \times \mathbb{R}^M$ describing the initial state of the system.
Equations governing the parameters of the kernel characteristic functionals

We have determined the essential conditions for the parameters of the representation (19) so that conditions a1 and a3 are verified. We shall now derive an equivalent set of differential equations with the functional differential equation for the undetermined parameters of the kernel characteristic functionals. This set of equations will be formulated using a Galerkin approach for the probability measures by direct application of kernel probability measures. In other words, if formally, we could write the linear operator equation for the probability measure of the solution as

\[ L \delta_{xy}^k(x, y) = 0, \quad \text{for every } x \in C_0^\infty, y \in C_0^\infty \]

then, using the kernel representation (18) and integrating equation (24) with respect to every probability measure we would have the equivalent set of equations

\[ \sum_{k=1}^n \int \int L \left[ \delta_{xy}^k(x, y) \right] \delta_{xy}^j(dx, dy) = 0, \quad j = 1, \ldots, n \]

Alternatively we could multiply with every kernel probability measure, instead of integrating. As a result we would the set of equations

\[ \sum_{k=1}^n \int \int L \left[ \delta_{xy}^k \right] \delta_{xy}^j(x, y) = 0, \quad j = 1, \ldots, n, \quad \text{for every } x \in C_0^\infty, y \in C_0^\infty \]

Then we can integrate with respect to every measure and get a set of equations for the parameters of the kernel probability measures. This above method seems very promising, but the main problem is that in general (an important exception is the Fokker-Planck equation), we cannot have in an explicit form the operator \( L \). Additionally it’s very difficult to operate with probability measures, since they are set functions.

Instead of this, we have for a very wide class of dynamical systems an explicit form of the operator governing the characteristic functional. Such operators were studied, extensively in the present chapter at Section 5.2. Thus, we can express the dynamics of the system with an operator equation (related to (24) if exists) of the form

\[ \tilde{L} \left[ \delta_{xy} \right] (u, v) = 0, \quad \text{for every } u \in C_0^\infty, v \in C_0^\infty \]

Moreover we have a wide class of explicit formulas for characteristic functionals as well as theorems for the analytical calculation of infinite dimensional integrals of those functionals. Hence, it would very interesting, if we could derive an equivalent set of equations with equations (25) or (26) using the characteristic functional approach.

For simplicity we will first present the subject for the case of a probability measure defined on a one-dimensional space. Such a case can be met on a one dimensional dynamical system under independent increments excitation.

Hence the operator equation (24) will take the form of a partial differential equation for the probability density function

\[ L [f](x) = 0, \quad x \in \mathbb{R} \]

For the same case we would have partial differential equation for the characteristic function \( \phi \)
\( \mathcal{L}[\phi](u) = 0 \), \( u \in \mathbb{R} \) \hspace{1cm} (29)

Of course the following relation diagram holds,

\[
\begin{align*}
\mathcal{L}[f](x) &= 0, \quad x \in \mathbb{R} \\
\xrightarrow{\mathcal{F}} &\quad \mathcal{F}^{-1}\mathcal{L}[\phi](u) = 0, \quad u \in \mathbb{R} \\
\xrightarrow{\mathcal{F}^{-1}} &\quad \phi(u), \quad u \in \mathbb{R}
\end{align*}
\]

where \( \mathcal{F} \) denotes the Fourier transform operator.

Now, consider the operator

\[
T[\mathcal{L}[\phi], \psi; \sigma](u) = \int_{-\infty}^{\infty} \mathcal{L}[\phi](u-v) \psi(v) \cdot \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{v^2}{2\sigma^2}} dv \, , \quad u \in \mathbb{R}
\]

and \( \phi, \psi \) are characteristic functions. Moreover the operator equation

\[
T[\mathcal{L}[\phi], \psi; \sigma](u) = 0 \, , \quad u \in \mathbb{R}, \quad \psi \in \mathbb{R}
\]

can be easily verified that is equivalent with (29). We will study the behaviour of the above operator for extreme values of the parameter \( \sigma \).

For \( \sigma \rightarrow 0 \) we will have

\[
\lim_{\sigma \rightarrow 0} T[\mathcal{L}[\phi], \psi; \sigma](u) = \int_{-\infty}^{\infty} \mathcal{L}[\phi](u-v) \psi(v) \cdot \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{v^2}{2\sigma^2}} dv = \mathcal{L}[\phi](u) \, , \quad u \in \mathbb{R}
\]

Hence,

\[
\lim_{\sigma \rightarrow 0} T[\mathcal{L}[\phi], \psi; \sigma](u) = 0 \, , \quad u \in \mathbb{R} \quad \Leftrightarrow \quad \mathcal{L}[\phi](u) = 0 \, , \quad u \in \mathbb{R}
\]

Thus, for \( \sigma \rightarrow 0 \) operator equation (31) interprets equation (24) in terms of the characteristic function.

In contrast for \( \sigma \rightarrow \infty \) we will have

\[
\lim_{\sigma \rightarrow \infty} \sigma \cdot T[\mathcal{L}[\phi], \psi; \sigma](u) = \sigma \int_{-\infty}^{\infty} \mathcal{L}[\phi](u-v) \psi(v) \cdot \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{v^2}{2\sigma^2}} dv =
\]

\[
= \int_{-\infty}^{\infty} \mathcal{L}[\phi](u-v) \psi(v) dv \, , \quad u \in \mathbb{R}
\]

Hence,

\[
\lim_{\sigma \rightarrow \infty} \sigma \cdot T[\mathcal{L}[\phi], \psi; \sigma](u) = 0 \, , \quad u \in \mathbb{R} \quad \Leftrightarrow \quad [\mathcal{L}[\phi] \ast \psi](u) = 0 \, , \quad u \in \mathbb{R}
\]

Thus, for \( \sigma \rightarrow 0 \) we have (when operator \( \mathcal{L} \) exists)

\[
\lim_{\sigma \rightarrow \infty} \sigma \cdot T[\mathcal{L}[\phi], \psi; \sigma](u) = 0 \, , \quad u \in \mathbb{R} \quad \Leftrightarrow \quad \mathcal{L}[f](x) \cdot g(x) = 0 \, , \quad x \in \mathbb{R}
\]

where \( g = \mathcal{F}^{-1}[\psi] \).

Now let us represent the solution of (28) using kernel representation \( \hat{f}_k \)

\[
f(x) = \frac{1}{n} \sum_{k=1}^{n} \hat{f}_k(x)
\]

(34)
Hence, the characteristic function will have the form
\[ \phi(x) = \frac{1}{n} \sum_{k=1}^{n} \hat{\phi}_k(x) \] (35)

Applying the above representation into (33) we will have the set of equations
\[ \lim_{\sigma \to \infty} \sigma \cdot T \left[ \sum_{k=1}^{n} \hat{\phi}_k \right] \left[ \phi_i ; \sigma \right] (u) = 0 , \ u \in \mathbb{R} , \ l = 1,...,n \] (36)

Because of (33) the above set will be equivalent with the set of equations
\[ \sum_{k=1}^{n} L \left[ \hat{f}_k (x) \right] \cdot \hat{f}_i (x) = 0 , \ x \in \mathbb{R} , \ l = 1,...,n \] (37)

Thus we have derived equations (36) that interpret equations (26) or (37) in terms of the characteristic function. We must note that for many kernel probability measures it’s not necessary to get \( \sigma \to \infty \) to achieve equivalence of (31) with (37). More specifically for kernels characteristic functions that have compact support (like Gaussian measures) it’s sufficient to have a value of \( \sigma \) such that \( \exp \left( -\frac{v_0^2}{2\sigma} \right) \approx 1 \) where \( v_0 \) is the radius of the support of the kernel characteristic function \( \hat{f} \) (if exists).

By inference we have achieved to derive an equivalent reformulation of the characteristic function differential equation (29) described by equation (31) which has the property to interpret (for special choice of the parameter \( \sigma \)) equation (26) or (37).

Motivated from the above discussion we generalize for the general case where no explicit formula for the operator \( L \) exists. So, let us consider the general case of a system as described at the beginning of the present section (conditions a1, a2, a3) and let the associated functional differential equation described by the operator
\[ L \left[ \mathcal{F}_{xy} \right] (u,v) = 0 \quad , \quad u \in C_c^{\infty,N} , \ v \in C_c^{\infty,M} \] (38)

Then, based on the analysis for the finite dimensional case described above we consider the operator
\[ T \left[ L \left[ \mathcal{F} \right] , \mathcal{G} ; C \right] (u,v) = \int \int_{C_c^{\infty,N} \times C_c^{\infty,M}} L \left[ \mathcal{F} \right] (u-w,v-z) \cdot \mathcal{G} (w,z) \cdot \mathcal{D}_G (dw,dz) \] (39)

for every \( u \in C_c^{\infty,N} , \ v \in C_c^{\infty,M} \), where the measure \( \mathcal{D}_G \) denotes a Gaussian measure into the measurable space \( \left( C_c^{\infty,N} \times C_c^{\infty,M} , \mathcal{B} \left( C_c^{\infty,N} \times C_c^{\infty,M} \right) \right) \) with correlation operator \( C \), centred at zero. When the norm of the correlation operator \( C \) tends to zero equation (39) will formally, interpret equation (24) and when the norm of the correlation operator \( C \) is sufficiently large equation (39) will formally tends to interpret the equation
\[ L \left[ \mathcal{D}_{xy} \right] (x,y) \cdot Q (x,y) = 0 \quad , \quad \text{for every} \quad x \in C_c^{\infty,N} , \ y \in C_c^{\infty,M} \] (40)

where \( Q \) is the measure associated with the characteristic functional \( \mathcal{G} \). Now consider the representation (19) satisfying conditions b1 and b2. Applying it at the operator equation
We emphasize that all the infinite dimensional integration described above can be carried out analytically using the results of Section 5.4.1 of the present chapter. Hence we have achieved to interpret equation (26) in terms of the characteristic functional. We must emphasize that in the above discussion we don’t prove anything about equivalence between equations (40) and and (41) or between (42) and (26). We just argue using finite dimensional relatives that an efficient way to produce equations from the characteristic functional equation is by using operator (39). In this way achieve localization in the probability measure domain.

Analogous comments concerning the correlation operator \( C \) can be made as in the case of the finite dimensional case. Thus if we choose the correlation operator such that the rate of decay of probability for the elements \( w \in C_{c}^{\infty,N} \), \( z \in C_{c}^{\infty,M} \) is less than the rate of decay of the values of the characteristic functionals \( \mathcal{F}(u-w, v-z) \cdot \mathcal{G}(w, z) \) for every \( u \in C_{c}^{\infty,N} \), \( v \in C_{c}^{\infty,M} \) then we would have achieved the desired result.

Finally since the equations (42) expresses the desired properties for the probability measures we can either integrate with respect to a Gaussian measure over the probability measures analytically (again we choose a zero centred Gaussian measure with suitable correlation operator as before), or we can take moment equations for every equation (42) using the results of Section 5.3.2 of the present chapter.

We must also notice that another approach for the derivation of simpler set of equations from the functional differential equation is based on the direct integration of the equation (38) with respect to various probability measures. Is should be emphasized that since we relate characteristic functionals with probability measures the choice of the probability measures should be very careful if we want to maintain all the necessary information including into equation (38). The above it will not be studied into the present work.

Now let us illustrate the above discussion by studying a specific system

**Example 4.16:** To describe the proposed method we will study the specific case of a first order stochastic differential equation with cubic nonlinearity, i.e. the system

\[
\begin{align*}
\dot{x}(t; \beta) &= a x(t; \beta)^3 + y(t; \beta) \quad \text{a.s.} \\
x(t_0; \beta) &= x_0(\beta), \quad \text{a.s.}
\end{align*}
\]

We have seen that the associated functional differential equation has the form

\[
\mathcal{T}[\hat{L}, \mathcal{F}_x, \mathcal{G}, \mathcal{C}](u, v) = 0 \quad \text{for every } u \in C_{c}^{\infty,N}, v \in C_{c}^{\infty,M}
\]

and setting \( \mathcal{G} = \mathcal{F}_x^k \) for \( k = 1, \ldots, n \) we will have the set of equations,

\[
\sum_{j=1}^{n} \mathcal{T}[\hat{L}, \mathcal{F}_x^j, \mathcal{G}, \mathcal{C}](u, v) = 0, \quad u \in C_{c}^{\infty,N}, v \in C_{c}^{\infty,M}, \quad k = 1, \ldots, n
\]
\[
\frac{d}{dt} \delta \mathcal{F}_{xy}(u,v) = -a \frac{\delta^3 \mathcal{F}_{xy}(u,v)}{\delta u(t)^3} + \frac{\delta \mathcal{F}_{xy}(u,v)}{\delta v(t)}
\]  
(45)

\[
\mathcal{F}_{xy}(v|0,0) = \mathcal{F}_0(v), \quad v \in \mathbb{R}
\]  
(46)

where we have assumed that the same conditions \(a_1, a_2, a_3\) mentioned before, are fulfilled.

Consider the representation of the joint characteristic functional for the response and the excitation

\[
\mathcal{F}_{xy}^n(u,v) = \frac{1}{n} \sum_{k=1}^{n} \mathcal{F}_{xy}(u,v; m_x^k, m_y^k, h_x^k, h_y^k)
\]  
(47)

where

a) \(\mathcal{F}_{xy}(u,v; m_x^k, m_y^k, h_x^k, h_y^k)\) \((k = 1, \ldots n)\) are the characteristic functionals corresponding to the kernel probability measures \(\mathcal{F}_{xy}(x,y; m_x^k, m_y^k, h_x^k, h_y^k)\) as described at Definition 4.7.

b) The characteristic functional is compatible with the characteristic functional of the excitation process, i.e., \(\mathcal{F}_{xy}^n(0,v) = \mathcal{F}_y^n(v) = \frac{1}{n} \sum_{k=1}^{n} \mathcal{F}_y(v; m_y^k, h_y^k)\). Hence the parameters \(m_y^k, h_y^k\) \((k = 1, \ldots n)\) are given.

c) The characteristic functional satisfies the initial conditions (44) or (46), i.e.,

\[
\mathcal{F}_{xy}^n(v_0(\bullet),0) = \mathcal{F}_{xy}^n(v_0(\bullet)) = \frac{1}{n} \sum_{k=1}^{n} \mathcal{F}_{xy}(v_0(\bullet); m_y^k, h_y^k)
\]  
(48)

Substituting representation (47) into the functional differential equation (45) we have

\[
\frac{d}{dt} \sum_{k=1}^{n} \frac{\delta \mathcal{F}_{xy}^k(u,v)}{\delta u(t)} = -a \sum_{k=1}^{n} \frac{\delta^3 \mathcal{F}_{xy}^k(u,v)}{\delta u(t)^3} + \sum_{k=1}^{n} \frac{\delta \mathcal{F}_{xy}^k(u,v)}{\delta v(t)}
\]  
(49)

Multiplying equation (49) with \(\mathcal{F}_{xy}^l(w-u, z-v)\) and integrating over a Gaussian measure with zero mean value and covariance operator \(C\), with respect to \(u,v\) we have the set of equations \((l = 1, \ldots n)\) for every \(u \in C_c, v \in C_c\)

\[
\frac{d}{dt} \sum_{k=1}^{n} \int_{C_c \times C_c} \frac{\delta \mathcal{F}_{xy}^k(w,z)}{\delta w(t)} \mathcal{F}_{xy}^l(u-w,v-z) \mathcal{P}_G(dw,dz) =
\]

\[
= -a \sum_{k=1}^{n} \int_{C_c \times C_c} \frac{\delta^3 \mathcal{F}_{xy}^k(w,z)}{\delta w(t)^3} \mathcal{F}_{xy}^l(u-w,v-z) \mathcal{P}_G(dw,dz) +
\]

\[
\sum_{k=1}^{n} \int_{C_c \times C_c} \frac{\delta \mathcal{F}_{xy}^k(w,z)}{\delta z(t)} \mathcal{F}_{xy}^l(u-w,v-z) \mathcal{P}_G(dw,dz)
\]  
(50)

Using the notation

\[
\int_{C_c \times C_c} \mathcal{F}_1(u-w) \mathcal{F}_2(u) \mathcal{P}_G(dw) = \mathcal{T} \left[ \mathcal{F}_1, \mathcal{F}_2, \mathcal{C} \right](u)
\]

we have the set of equation \((l = 1, \ldots n)\) for every \(u \in C_c, v \in C_c\)
\[ \frac{d}{dt} \sum_{k=1}^{n} T \left[ \frac{\delta \mathcal{F}_{xy}^{k}}{\delta u(t)} , \mathcal{F}_{xy}^{l} , C \right] (u,v) = -a \sum_{k=1}^{n} T \left[ \frac{\delta^{3} \mathcal{F}_{xy}^{k}}{\delta u(t)} , \mathcal{F}_{xy}^{l} , C \right] (u,v) + \sum_{k=1}^{n} T \left[ \frac{\delta \mathcal{F}_{xy}^{k}}{\delta v(t)} , \mathcal{F}_{xy}^{l} , C \right] (u,v) \] \quad (51)

The above set of equations formulates a set of functional differential equations for the parameters of the kernel characteristic functionals with initial conditions defined from equation (48). The difference with the prior functional differential equation (49), is that every equation of the last set, controls a specific part of the probability space and is coupled with the other parts through the coupling of the functional differential equation set (51).

Since the integrals are over Gaussian measures and the integrated functionals involve linear and quadratic forms, Theorems 4.1 and 4.4 can be applied for the analytical calculation. We will not pass into details for the analytical calculation of the above integrals.

The resulted set of functional equations can be numerically solved by either integrate them analytically with respect to a suitable Gaussian measure or by getting moment equations using the result of Section 5.3.2 of the present chapter.

In the following section we will try to simplify further the set of functional equations (42) and get numerical results without even make the analytical calculations of the functional integrals.

In the next section a special simplification will be applied to equations (50) to get numerical results without having to calculate the above integrals.
5.4.5. Simplification of the set of FDE’s governing the Kernel Characteristic Functionals.

In this section we shall simplify the set of functional differential equations (42) to obtain a set of ordinary differential equations for immediate numerical solution. Hence, based on the following simplification we will not need to evaluate analytically the functional integrals. However we must emphasize, that in this way, we just illustrate the validity of the proposed method. We don’t expect to have the best possible numerical results. For this reason the following analysis will mainly concern the first order system with cubic nonlinearity, described at the previous sections.

\[ \dot{x}(t) + kx(t) + ax(t)^3 = y(t) \]  

(52)

Analogous arguments can be used for the analysis of higher dimensional systems with more complex polynomial nonlinearities. For the numerical implementation we will use Gaussian kernel characteristic functionals.

The simplification will be based on the restriction of the concentration parameter (for Gaussian kernel characteristic functionals this parameter will be the variance) to sufficiently small values such that the interaction between every pair of Gaussian kernels could be considered negligible. However this assumption cannot hold for a great time since the equations governing every kernel leads to greater values of the concentration parameter. Thus, to avoid growth of the concentration parameter we approximate, when a variance exceed a certain value, the current probability measure by another set of kernel measures with different amplitudes and smaller variances. Then this set of kernel measures evolutes with respect to the dynamical equations until some concentration parameter exceed a certain value. The accuracy of the proposed method mainly depends on the threshold value of the concentration parameter.

Hence we use a kernel characteristic functional representation

\[ \mathcal{F}^n(\bullet) = \sum_{k=1}^{n} p_k \mathcal{F}(\bullet; m_{x,k}, m_{y,k}, h_k), \quad \text{with } p_n \geq 0, \quad n = 1, \ldots, N, \quad \sum_{n} p_n = 1 \]  

(53)

where every \( \mathcal{F}(\bullet; m_{x,k}, m_{y,k}, h_k) \) \( (k = 1, \ldots, n) \) is a Gaussian kernel characteristic functional.

Based on the above assumptions we can conclude that the set of equations (42) will take the form

\[ \mathcal{T}\left( \tilde{L} \left[ \mathcal{F}_{xy}^k \right], \mathcal{F}_{xy}^k ; C \right)(u,v) = 0, \quad k = 1, \ldots, n \]  

(54)

or,

\[ \tilde{L}\left[ \mathcal{F}_{xy}^k \right](u,v) = 0, \quad u \in C_c^{\infty,N}, \quad v \in C_c^{\infty,M}, \quad k = 1, \ldots, n \]  

(55)

Note that for the case described we need general amplitudes \( 0 \leq p_k \leq 1, \sum_{k=1}^{n} p_k = 1 \) (instead of equal with each other, assumed at the previous sections) for every kernel characteristic functional. These amplitudes will be assigned from the approximation algorithm, every time that a concentration parameter exceeds the threshold value. Now let us derive the ordinary
differential equations associated with every Gaussian kernel, for the specific case of the one dimensional system with a linear and cubic term. Equations (55) will take the form
\[
\frac{d}{dt} \delta \mathcal{F}_w^k(u,v) + k \frac{\delta \mathcal{F}_w^k}{\delta u(t)}(u,v) - a \frac{\delta^3 \mathcal{F}_w^k}{\delta u(t)^3}(u,v) = \delta \mathcal{F}_w^k(u,v)
\] (56)
for every \( u \in C^\infty_c, v \in C^\infty_c (k = 1, \ldots, n) \). We observe that every characteristic functional controls a specific region of the probability space and evolves independently from the others. Note also, that the interchange of probability between the kernels characteristic functionals is carried on completely from the approximation algorithm and only at the case when some concentration parameter exceeds the threshold value. Although this assumption is very loose, since the probability must interchange continuously (with respect to time), between the kernels, it gives us a first picture of the method. We emphasize that with the analytical calculation of the functional integrals the interchange of probability occurs continuously since the system of ordinary differential equations governing the kernel parameters is coupled. Hence, for this case the interchange of probability is introduced from the nondiagonal terms of equation (42).

An important feature of the discussed method is the ability for parallelized computations. Parallelization techniques can be physically applied both to the dynamical evolution of the kernels and the optimization algorithm. For the first case the algorithm can take advantage of the independent evolution of every kernel. Additionally for the optimization algorithm we can approximate the given density (since is already splitted into Gaussian kernels) by independent group of kernels with small variance. Hence, in this way we can succeed very fast computations for system of higher dimensions, with various excitations.

To derive a set of ordinary differential equations for every FDE (56) we will get moment equations using the result of Section 3 of the present chapter. Hence, since every characteristic functional \( \delta \mathcal{F}_w^k \) is Gaussian, the functional differential equation (56) will be fulfilled if the three first moment equations are satisfied. By direct application of Section 5.3.2 of the present Chapter and Appendix I lead us to the following subset of ordinary differential equations for every kernel \( (k = 1, \ldots, n) \)
\[
m^k_{X,t}(t) + k m^k_X(t) + 3am^k_X(t)B^k_{XX,t}(t,t) + d[m^k_X(t)]^3 = m^k_X(t)
\]
\[
B^k_{XX,t}(t,s) + kB^k_{XX,t}(t,s) + 3am^k_X(t)B^k_{XX,t}(t,s)m^k_X(t) + 3aB^k_{XX,t}(t,t)B^k_{XX,t}(t,s) = B^k_{XY,t}(s,t)
\]
\[
B^k_{XY,t}(t,s) + kB^k_{XY,t}(t,s) + 3aB^k_{XY,t}(t,t)B^k_{XY,t}(t,s) + 3am^k_X(t)B^k_{XY,t}(t,s)m^k_X(t) = B^k_{YY,t}(s,t)
\]
for all \((t,s) \in I \times I\). We will not come into details concerning the numerical solution of the above equations. We just emphasize that special treatment was needed, since as we note, there is no participation of the \( s \) variable to the dynamics of the equations. Hence we have to solve the above system for every value of \( s \).

Let us now give a precise description of the optimization algorithm used for the rebuild of a measure (more precisely density) with kernels with smaller variance. As we mention we will have to approximate a given superposition of Gaussian densities by another superposition with the restriction of a maximum variance. We thus have the following optimization problem,
\[
\min_{\{A_i, m_i, \sigma_i\}_{i=1}^M} \left\{ \int_{-\infty}^{+\infty} \sum_{i=1}^N \frac{A_i}{\sigma_i \sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{x - \bar{m}_i}{\bar{\sigma}_i} \right)^2 \right] - \sum_{j=1}^M \frac{A_j}{\sigma_0 \sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{x - m_j}{\sigma_0} \right)^2 \right] \right\}^2 \, dx \right\}
\]

where \( \{\bar{A}_i, \bar{m}_i, \bar{\sigma}_i\}_{i=1}^N \) are the amplitudes, mean values, variances of the given superposed kernels and \( \{A_i, m_i, \sigma_0\}_{i=1}^M \) are the parameters of the kernels that approximates the given density. We also restrict ourselves, for simplicity, to the case of a common variance \( \sigma_0 \) for all the Gaussian kernels. More details for the optimization algorithm can be found at Appendix II.

In what follows we will examine the case of the above system with two different set of parameters. Hence a white noise excitation case will be studied with the system parameters described by the table

<table>
<thead>
<tr>
<th>Non-dimensional parameters</th>
<th>Case I</th>
<th>Case II</th>
</tr>
</thead>
<tbody>
<tr>
<td>( k )</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>( a )</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>White noise intensity</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

We note that for the first case the nonlinear system has one stable fixed point located at zero. As the parameter \( a \) decreases a pitchfork bifurcation occurs at zero and two symmetric stable points appear at \( \pm \sqrt{\frac{k}{a}} \) and the fixed point at zero becomes unstable. Hence we have the bifurcation diagram

The above analysis will also be confirmed from the resulted densities. Thus we expect that for the case II, the probability will concentrate around the fixed points and hence we will have a bimodal distribution. For the first case we expect a unimodal distribution with the probability concentrated at zero. Since the above fixed points are global attractors we expect to have these
results after some time, independent of the initial density. For the numerical results presented above we are using two cases of initial distributions. The first case is consisting of a Gaussian distribution centered at zero with variance $\sigma = 0.1$. The second case is a convex superposition of two Gaussian distributions with characteristics $m_1 = 0, m_2 = 0.6, \sigma_1 = 0.1, \sigma_2 = 0.6$ and amplitudes $a_1 = 0.4$ and $a_2 = 0.6$ respectively.

The first plot shows with solid lines the mean values of the kernel probability measures. Their shape verifies the stable property of the fixed point at 0. We can see that the initial distribution is spreading and approaches very the final steady state distribution. A very interesting point is that the concentration of probability is not stopping as can be seen from the outer kernels. This behaviour reveals the necessity for the study of the tail behaviour, since a representation which predicts the tail revolution will overcome the mentioned problem.
$f_{x}(t, x)$ for the case of Gaussian White Noise excitation with one stable fixed point at 0 and initial conditions of the second type.

Again we can see from the first plot that the kernels mean values are attracted from the stable fixed point (global attractor). In this case, since the initial distribution is more complex, more kernels are used for the efficiently representation of the solution. Notice that the interchange of probability between kernels takes place every 0.2 seconds, which is every time we have discontinuity of the mean value curves.

Notice that since the steady state probability distribution has an asymptotic behaviour of $o(e^{-x})$ and simultaneously we are using Gaussian representations, we will constantly have the previously described behaviour of the outer kernels. Hence, it is very important to use suitable kernel representations with identical tail behaviour. However, we can see that the above representation produces satisfactory results for the main probability mass.
$f_\epsilon(t,x)$ for the case of Gaussian White Noise excitation with two stable fixed points at $\pm 1$ and initial conditions of the first type.

In this case we examine the behaviour of the system for different system parameters. Hence, we have for the underlying Hamiltonian system two stable fixed points at $\pm 1$ while the fixed point at 0 has become unstable. For the first case of initial conditions we have full symmetry with respect to the spatial 0. The symmetry is maintained for the kernel mean values as well. Note also that, despite the discontinuities of the kernels mean values the probability distribution is continuous (second plot).

After a short time we can see that the probability is concentrated near the stable fixed points and a steady state is approached. Again we can see the higher concentration of outer kernels, something that can be explained from the tail evolution problems. It must also be noted that the rapid spread of probability is explained from the chosen white noise intensity.
This is the most interest case. We start with a bimodal initial condition and we are approaching a bimodal steady state distribution. From the first plot we note that until the time of 0.4 seconds the starting bimodal distribution is transforming to a unimodal one. As the system evolves, the probability mass is transferred from the stable fixed point located at -1 to the stable fixed point at 1. This is a very interesting phenomenon since between the two stable fixed points is an unstable one which divides the phase space of the deterministic problem and does not allow the transition of orbits from the one sine to another. This is not the case for the probability mass, were we know that after a sufficient time period the probability of the two stable fixed points regions will be equal.
5.5. References


PROHOROV, YU.V. & SAZONOV, V.V., 1961, Some Results Associated with Bochner’s Thoerem. *Theory of Probability and Its Applications*, VI, 82-87.


Subjects for Future Research

As it has been seen in the text there are numerous threads for future investigations. One very interesting direction would be the application of the results presented for the numerical solution of stochastic differential equations describing complex dynamical systems. Some topics on this direction of future research should be

- Implementation of numerical method for the solution of the partial differential equations derived at Section 5.3.4 through the reduction of the Functional differential equation. In this way we would have results for the joint characteristic function of the response, its derivatives and the excitation at various time instants. Probably the most convenient way should be the use of kernel characteristic functions representations with given marginals.

- Direct application of the proposed method of Section 5.4.4 without any simplifications. In this way we could solve the full statistical problem and get the complete probability structure through the characteristic functional. This direction should require analytical computation of infinite dimensional integrals by applying the theorems presented at Section 5.4.1.

- Calculation of the steady-state probability measures associated with a dynamical system, by means of suitable kernel characteristic functional representation. This representation should describe a stationary measure and its application for the numerical computation of its parameters might use a variation of the proposed method of Kernel Characteristic Functionals.

- Improvement of the simplified method, presented at Section 5.4.5, and especially of the optimization algorithm described at Appendix III for faster computation of the probability densities as well as for numerical solution of higher dimensional problems. In this direction should be take in account the embedded characteristic of the method for parallel computation, discussed at Section 5.4.5.

Another research thread of great importance should be the analytical investigation of dynamical systems by means of the kernel probability measure representations. More specifically some topics on this direction could be

- Study of the evolution of the asymptotic behaviour (tail) of the response, directly thorough the analytical study of the characteristic functional and more specifically thought the study of the asymptotic behaviour of the characteristic functional at the neighbourhood of zero. In this direction a useful tool might be the Tatarskii characteristic functional. The above topic is of great importance since the knowledge of the tail behaviour will allow solving the problem with few kernels.

- Analytical study of the behaviour of the probability structure of the response by studying analytically the parameters of the kernel characteristic functionals in relation with critical system parameters. An interesting point on this would be how the bifurcation analysis of the underlying deterministic system would/could predict the probability structure of the response through the kernels. In this way we could study more clearly the effects of nonlinearity and stochasticity simultaneously.
Finally an important topic concerning the study of kernel characteristic functionals is the proof of theorems analogous to theorem 4.13 of Section 5.4.3 due to D.G. Konstant and V.I. Piterbarg, concerning the probability structure of the extremes of stochastic processes with known probability measure. Hence in this way we would have necessary conditions for a probability measure/characteristic functional to be admissible for the solution of the statistical problems.
Appendix I

Volterra Derivatives of Gaussian Characteristic Functional

For the sake of convenience we shall now compute, the essential derivatives of Gaussian functional for our analysis. Let,

$$\mathcal{F}(u,v) = \exp\left(\frac{i}{\hbar}(m_x u + i (m_y v) - \frac{1}{2} \langle C_{xx} u, u \rangle - \frac{1}{2} \langle C_{yy} v, v \rangle - \langle C_{xy} u, v \rangle)\right)$$

Then we have the Volterra derivatives computed at \((0,0)\).

$$\frac{\delta \mathcal{F}(0,0)}{\delta u_i} = m_x(t_i)$$

$$\frac{\delta \mathcal{F}(0,0)}{\delta v_i} = m_y(t_i)$$

$$\frac{\delta^2 \mathcal{F}(0,0)}{\delta u_i \delta u_j} = B_{XX}(t_i, t_j) + m_x(t_i) m_x(t_j)$$

$$\frac{\delta^2 \mathcal{F}(0,0)}{\delta v_i \delta v_j} = B_{YY}(t_i, t_j) + m_y(t_i) m_y(t_j)$$

$$\frac{\delta^3 \mathcal{F}(0,0)}{\delta u_i \delta u_j \delta u_k} = B_{XX}(t_i, t_j) B_{XX}(t_k, t_j) + m_x(t_i) m_x(t_j) m_x(t_k)$$

$$\frac{\delta^3 \mathcal{F}(0,0)}{\delta u_i \delta u_j \delta v_k} = m_x(t_i) B_{XY}(t_j, t_k) + m_x(t_j) B_{XY}(t_i, t_k) + m_y(t_k) B_{XY}(t_i, t_j) + m_x(t_i) m_x(t_j) m_y(t_k)$$
\[
\frac{\delta^4 \mathcal{F}^I (0,0)}{\delta u_1 \delta u_2 \delta u_3 \delta u_4} \frac{1}{i^4} = B_{XX} (t_1,t_2) B_{XX} (t_2,t_3) B_{XX} (t_3,t_4) + B_{XX} (t_2,t_3) B_{XX} (t_3,t_4) \]

\[
+ m_x (t_1) B_{XX} (t_2,t_3) m_x (t_4) + m_x (t_2) B_{XX} (t_1,t_3) m_x (t_4) +
\]

\[
+ m_x (t_1) B_{XX} (t_2,t_4) m_x (t_3) + m_x (t_2) B_{XX} (t_1,t_4) m_x (t_3) +
\]

\[
+ m_x (t_1) B_{XX} (t_3,t_4) m_x (t_2) + m_x (t_3) m_x (t_2) m_x (t_3) m_x (t_4) +
\]

\[
+ B_{XX} (t_1,t_2) B_{XX} (t_3,t_4) + m_x (t_1) B_{XX} (t_3,t_4) m_x (t_2) +
\]

\[
\frac{\delta^4 \mathcal{F}^I (0,0)}{\delta u_1 \delta u_2 \delta u_3 \delta u_4} \frac{1}{i^4} = B_{XX} (t_1,t_3) B_{XY} (t_2,t_4) + B_{XX} (t_2,t_3) B_{XY} (t_1,t_4) +
\]

\[
+ m_x (t_1) B_{XX} (t_2,t_3) m_y (t_4) + m_x (t_2) B_{XX} (t_1,t_3) m_y (t_4) +
\]

\[
+ m_x (t_1) B_{XY} (t_2,t_4) m_x (t_3) + m_x (t_2) B_{XY} (t_1,t_4) m_x (t_3) +
\]

\[
+ m_x (t_1) B_{XX} (t_3,t_4) m_y (t_2) + m_x (t_3) m_x (t_2) m_y (t_3) m_y (t_4) +
\]

\[
+ B_{XX} (t_1,t_2) B_{XY} (t_3,t_4) + m_x (t_1) B_{XY} (t_3,t_4) m_x (t_2) +
\]

The above calculations will for the derivation of moment equations from the FDE. These equations correspond to the Gaussian closure technique. In Section 5.4.5 of the Chapter 5 they will be used for the application of the Gaussian closure technique locally in the phase space and will allows us to get a first picture of the proposed method.
Appendix II

Optimization Algorithm

We have the following optimization problem,

\[
\min_{\{A, m, \sigma_0\}_{i=1}^M} \left\{ \int_{-\infty}^{+\infty} \sum_{i=1}^{N} \frac{A_i}{\tilde{\sigma}_i \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left( \frac{x - \tilde{m}_i}{\tilde{\sigma}_i} \right)^2 \right\} dx \right\} - \sum_{j=1}^{M} A_j \exp \left\{ -\frac{1}{2} \left( \frac{x - \tilde{m}_j}{\tilde{\sigma}_j} \right)^2 \right\} dx \right\}^2 \right\} \]

where \( \{A_i, \tilde{m}_i, \tilde{\sigma}_i\}_{i=1}^N \) are the amplitudes, mean values, variances of the given superposed kernels and \( \{A, m, \sigma_0\}_{i=1}^M \) are the parameters of the kernels that approximates the given density. We also restrict ourselves, for simplicity, to the case of a common variance \( \sigma_0 \) for all the Gaussian kernels. Hence we will have the restrictions,

\[
LB < \sigma_0 < UB \quad \text{and} \quad \sum_{j=1}^{M} A_j = 1
\]

Expanding the above, we have the function for minimization

\[
F(\{A\}, \sigma_0, M) = \int_{-\infty}^{+\infty} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{A_i \tilde{A}_j}{\tilde{\sigma}_i \tilde{\sigma}_j 2\pi} \exp \left\{ -\frac{1}{2} \left( \frac{x - \tilde{m}_i}{\tilde{\sigma}_i} \right)^2 - \frac{1}{2} \left( \frac{x - \tilde{m}_j}{\tilde{\sigma}_j} \right)^2 \right\} dx -
\]

\[
-2 \int_{-\infty}^{+\infty} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{A_i \tilde{A}_j}{\tilde{\sigma}_i \sigma_0 2\pi} \exp \left\{ -\frac{1}{2} \left( \frac{x - \tilde{m}_i}{\tilde{\sigma}_i} \right)^2 - \frac{1}{2} \left( \frac{x - m_j}{\sigma_0} \right)^2 \right\} dx +
\]

\[
+ \int_{-\infty}^{+\infty} \sum_{i=1}^{M} \sum_{j=1}^{M} \frac{A_i \tilde{A}_j}{\sigma_0 \sigma_0 2\pi} \exp \left\{ -\frac{1}{2} \left( \frac{x - m_i}{\sigma_0} \right)^2 - \frac{1}{2} \left( \frac{x - m_j}{\sigma_0} \right)^2 \right\} dx.
\]

The integrations can be carried out analytically. Hence we will have,

\[
F(\{A\}, \sigma_0, M) = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{A_i \tilde{A}_j}{2 \sqrt{\left( \tilde{\sigma}_i^2 + \tilde{\sigma}_j^2 \right) \pi}} \exp \left\{ -\frac{\left( \tilde{m}_i - \tilde{m}_j \right)^2}{\tilde{\sigma}_i^2 + \tilde{\sigma}_j^2} \right\} -
\]

\[
-2 \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{A_i \tilde{A}_j}{2 \sqrt{\left( \sigma_0^2 + \tilde{\sigma}_i^2 \right) \pi}} \exp \left\{ -\frac{\left( m_i - \tilde{m}_j \right)^2}{\sigma_0^2 + \tilde{\sigma}_i^2} \right\} +
\]

\[
+ \sum_{i=1}^{M} \sum_{j=1}^{M} \frac{A_i \tilde{A}_j}{2 \sqrt{\left( \sigma_0^2 + \sigma_0^2 \right) \pi}} \exp \left\{ -\frac{\left( m_i - m_j \right)^2}{\sigma_0^2 + \sigma_0^2} \right\}.
\]

To minimize the above quantity we fixed the common variance \( \sigma_0 \) and the mean values \( m_i, i = 1, 2, \ldots M \) defined by

\[
m_i = \left[ \min (\tilde{m}_0) - \frac{3}{2} \max (\sigma_0) \right] + \frac{1}{M - 1} \left[ \max (\tilde{m}_0) - \min (\tilde{m}_0) + 3 \max (\sigma_0) \right]
\]
Thus the minimization problem is equivalent with the set of equations,

\[ B_{ij} A_j = C_i \]

where,

\[ B_{ij} = \frac{1}{2\sigma_0^2} \exp \left\{ -\frac{(m_j - m_i)^2}{2\sigma_0^2} \right\} \quad i = 1, \ldots, M \quad j = 1, \ldots, M \]

\[ C_i = \sum_{k=1}^{N} \frac{\tilde{A}_k}{2\sigma_0^2} \exp \left\{ -\frac{(\tilde{m}_k - m_i)^2}{2\sigma_0^2} \right\} \quad i = 1, \ldots, M \]

And the value of the minimizing function will be,

\[ F(\{A_i\}, \sigma_0, M) = \tilde{A}_j D_{ij} \tilde{A}_j - 2C_i A_i + A_i B_{ij} A_j \]

To the linear system of equations we must add the condition

\[ \sum_{j=1}^{M} A_j = 1 \]

Thus we have a least-square problem, with \( M + 1 \) equations.

For our case we make the function

\[ G(\sigma_0, M) = F(\{A_{i, \text{opt}}\}, \sigma_0, M) \]

with the amplitudes \( \{A_{i, \text{opt}}\} \) defined by the least-square problem. Thus, for every \( M \) we minimize the function \( G(\sigma_0, M) \) with respect to \( \sigma_0 \) and then we choose the optimum \( \sigma_{0, \text{opt}}, M_{\text{opt}}, \{A_{i, \text{opt}}\} \).
Appendix III

Application on Ship Stability and Capsizing on Random Sea

The Probabilistic Model

The current regulations and criteria for assuring the stability of a ship and preventing it from capsizing (see the Codes of the International Maritime Organization (IMO) and of the Germanischer Lloyd) are empirical and are based on the properties of the righting lever of the ship, taking only hydrostatic forces into account. For details and criticism see Kreuzer and Wendt (p. 1836). Those static criteria which neglect the motion of the ship as well as the impact of seaway and wind do obviously not guarantee total stability as they cannot prevent at least 30 ships yearly of tonnage greater than 500 GT from getting lost due to severe weather conditions. Hence researchers agree that those criteria have to be modified by using hydrodynamic models of the ship–sea system, by describing the sea as a random field and by analyzing the ship as a rigid body with 6 degrees of freedom using methods of nonlinear dynamics and the theory of random dynamical systems.

The present state of (deterministic as well as stochastic) ship dynamics research has been recently documented in the Theme Issue of the *Phil. Trans. Royal Soc. London (Series A)* entitled “The nonlinear dynamics of ships”, edited by Spyrou and Thompson. The volume includes the extended overview by Spyrou and Thompson, presenting also the historical development of the field as well as future directions of research, the prime one being the “integration of the probabilistic character of the seaway into the study of nonlinear dynamics” – which we will try to address.

In the following analysis we will not consider the linear model for rolling equation since it is the result of the assumptions that the wave is small. Hence the linearized model describes small ship motions. Transition to large roll angles requires the introduction of nonlinear terms into the ship motion equations. Unfortunately, there is no theory available to derive ship motions equations in a consistent manner. So we used a physics based approach, considering forces of a different nature separately and trying to combine them into one model of nonlinear rolling, suitable for capsizing study.

The forces acting on a ship in a seaway can be classified in the following manner BELENKY L.V. AND SEVASIANOV N.B., (Stability and Safety of Ships):

- Inertial hydrodynamic forces (e.g. added mass)
- Wave damping hydrodynamic forces
- Viscous damping forces
- Hydrostatic forces
- Wave excitation forces
- Other forces, including
- Aerodynamic forces
- Appendage forces (including dumping form rollers and anti-rolling fins)
- Forces due to fluid motion in internal tanks including anti-roll tanks
- Hull “manoeuvring” forces, propeller thrust, etc.

There are other systems to classify these forces. This classification assigns forces of a different nature to the same category. However, it reflects a role that force plays in ship motions.

For our analysis we will use the simplified, but nonlinear, rolling equation

\[(I_{xx} + a_{44})\ddot{\phi} + 2N_\phi \dot{\phi} + mg \cdot GZ (\phi) = mg \cdot GM \cdot a_E (t)\]  

(1)

where

- We have neglect nonlinear damping forces
- We have neglect the effect of bulwark entering into water during heeling
- The change of GZ curve due to irregular seas
- We have neglect effects of special possibly anti-rolling devices
- \(N_\phi\) is a linear dumping coefficient
- \(a_E (t)\) describes the excitation due to wave forces

To “randomize” a capsizing model we need to:

- Introduce a model for the stochastic excitation due to wave forces
- Calculate probabilistic characteristics of rolling, roll velocities and other processes included in the model of capsizing
- Calculate probabilistic characteristics of the “carrier process” – the process, by which crossing of the boundary is associated with the capsizing
- Calculate the probability of crossing

In the direction described above we will express the stochastic forces using the well-known model of Pierson – Longuet Higgins Model (see Chapter 2). Hence we assume that \(a_E (t)\) is a normal stationery stochastic process with given spectrum or correlation function. Moreover we assume that all the necessary ship’s characteristics (such as GZ curve, GM, etc.) are known.

For the probabilistic analysis we will first write eq. (1) in a non-dimensional, state-space form, i.e.

\[\dot{\phi} = \theta\]

\[\dot{\theta} + 2b\theta + V'(\phi) = F_w (t)\]

(2)

where,

- \(b = N_\phi / (I_{xx} + a_{44})\)
- \(V'(\phi) = \frac{mg}{(I_{xx} + a_{44})} \int GZ (\varphi) d\varphi\)
- \(F_w (t) = mg \cdot GM \cdot a_E (t) / (I_{xx} + a_{44})\)
Functional Differential Equation for the Rolling Equation

We shall derive the corresponding functional differential equation for the stochastic system (2). We assume that

a) $F_w(\bullet, \bullet) : I \times \Omega \to \mathbb{R}$ is a normal, stationary, mean square continuous stochastic process defined onto the probability space $(\Omega, \mathcal{B}(\Omega), \mathbb{P})$ with the associated probability space $(C_c^\infty, \mathcal{B}(C_c^\infty), \mathbb{P}_F)$, and with characteristic functional $F_F$.

b) $(\theta_0(\bullet), \phi_0(\bullet)) : \Omega \to \mathbb{R}^2$ are random variables with the associated probability space $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mathbb{P}_{\theta_0, \phi_0})$ and characteristic functions $\Psi_{\theta_0, \phi_0}$.

c) $V'(\bullet) : I \times \mathbb{R} \to \mathbb{R}$ is a measurable, deterministic function that can be expanded into a series of the form

$$V'(\phi) = a_1\phi + a_3\phi^3$$ (3)

Applying Theorem 2.1 of Chapter 5 we will have the functional differential equation

$$\frac{d}{dt} \frac{\delta F_{\theta_0, \phi}(u_\phi, u_\theta, v)}{\delta u_\phi(t)} = \frac{\delta F_{\theta_0, \phi}(u_\phi, u_\theta, v)}{\delta u_\phi(t)}$$ (4a)

$$\frac{d}{dt} \frac{\delta F_{\theta_0, \phi}(u_\phi, u_\theta, v)}{\delta u_\phi(t)} + 2b \frac{\delta F_{\theta_0, \phi}(u_\phi, u_\theta, v)}{\delta u_\phi(t)} + a_t \frac{\delta F_{\theta_0, \phi}(u_\phi, u_\theta, v)}{\delta u_\phi(t)} +$$

$$- a_3 \frac{\delta^3 F_{\theta_0, \phi}(u_\phi, u_\theta, v)}{\delta u_\phi(t)^3} = \frac{\delta F_{\theta_0, \phi}(u_\phi, u_\theta, v)}{\delta v(t)}$$ (4b)

Where $F_{\theta_0, \phi}(u_\phi, u_\theta, v)$ is the joint characteristic functional of the stochastic processes $F_w, \phi, \theta$.

Concerning the initial conditions of the problem we have

$$F_{\theta_0, \phi}(v_\phi \delta_{t_0}(\bullet), v_\theta \delta_{t_0}(\bullet), 0) =$$

$$= \int \int \exp \left[ i \left\langle \phi, v_\phi \delta_{t_0}(\bullet) \right\rangle + i \left\langle \theta, v_\theta \delta_{t_0}(\bullet) \right\rangle + i \left\langle F, 0 \right\rangle \right] F_{\theta_0, \phi} \left( d\phi, d\theta, dF \right) =$$

$$= \int \exp i \left( v_\phi \phi(t_0) + v_\theta \theta(t_0) \right) \mathcal{F}_{\theta_0, \phi} \left( d\phi, d\theta, dF \right) = \mathcal{F}_{\theta_0, \phi} \left( v_\phi, v_\theta \right)$$ (6)
**PDE for the joint probability density function of the r.v.’s \( F_W(t), \phi(t), \theta(t) \)**

Following the analysis of Chapter 5/Section 5.3.4 we have the reduction of the Functional Differential Equation (4a) and (4b) to the PDE

\[
\frac{\partial \phi_{OF}(t,t,s)}{\partial t} - v_\phi \frac{\partial \phi_{OF}(t,t,t)}{\partial v_\phi} - 2bv_\theta \frac{\partial \phi_{OF}(t,t,t)}{\partial v_\theta} - a_3v_\theta \frac{\partial^3 \phi_{OF}(t,t,t)}{\partial v^3_\phi} = v_\theta \frac{\partial \phi_{OF}(t,t,t)}{\partial v}, \quad v_\theta, v_\phi, v \in \mathbb{R}, \quad t \in I
\]

where \( \phi_{OF}(t_1,t_2,t_3) = \phi_{OF}(v_\phi,t_1;v_\theta,t_2;v,t_3) \) is the joint characteristic function of the random variable (vector) \((\phi(t_1),\theta(t_2),F_W(t_3))\).

Moreover the additional condition holds

\[
\phi_{OF}(0,0,0,t_2;v,t) = \phi_{\mathcal{F}}(0,0,t_2;v,t) = \mathcal{F}_F(v \cdot \delta(\cdot - t)), \quad v \in \mathbb{R}
\]

and the initial condition

\[
\phi_{OF}(v_\phi,0;v_\theta,0,t_3) = \phi_{\mathcal{F}}(v_\phi,0;v_\theta,0) = \mathcal{F}_{\mathcal{F}_{\theta}}(v_\phi,v_\theta), \quad v_\phi, v_\theta \in \mathbb{R}
\]

Hence we have derived a partial differential equation for the joint characteristic function \((\phi(t),\theta(t),F_W(t))\).

An alternative form of eq. (7)-(9) can be derived, by taking their Fourier transform. In this case we will have

\[
\frac{\partial f_{OF}(t,t,s)}{\partial t} + \frac{\partial}{\partial \phi} \left[ \theta f_{OF}(t,t,t) \right] + \frac{\partial}{\partial \theta} \left[ (2b\theta + a_3\phi + a_5\phi^3 + F) f_{OF}(t,t,t) \right], \quad t \in I
\]

where \( f_{OF}(t_1,t_2,t_3) = f_{OF}(\phi(t_1),\theta(t_2);F(t_3)) \) is the joint probability density function of the random variable (vector) \((\phi(t_1),\theta(t_2),F_W(t_3))\).

Moreover the additional condition holds

\[
\int \int f_{OF}(\phi,t_1;\theta,t_2;F,t) d\phi d\theta = f_{\mathcal{F}}(F,t) = \int_{-\infty}^{\infty} e^{i\nu t} \mathcal{F}_F(\nu \cdot \delta(\cdot - t)) d\nu, \quad F \in \mathbb{R}
\]

and the initial condition

\[
\int f_{OF}(\phi,0;\theta,0;F,t) dF = f_{\mathcal{F}}(\phi,0;\theta,0) = \mathcal{F}^{-1} \left[ \mathcal{F}_{\mathcal{F}_{\theta}}(v_\phi,v_\theta) \right], \quad \phi, \theta \in \mathbb{R}
\]

(12)

Using this information, we can additionally determine the probability structure of the second derivative \( \dot{\phi}(t) \) by means of a simple calculation using eq. (4b). Having this information we can calculate the probability of capsizing, thus, determining the ship stability and safety.
References