S-E A Mohammed

Stochastic functional differential equations

S-E A Mohammed

University of Khartoum

Stochastic functional differential equations

π

Pitman Advanced Publishing Program

BOSTON · LONDON · MELBOURNE

PITMAN PUBLISHING LIMITED
128 Long Acre, London WC2E 9AN

PITMAN PUBLISHING INC
1020 Plain Street, Marshfield, Massachusetts 02050

Associated Companies
Pitman Publishing Pty Ltd, Melbourne
Pitman Publishing New Zealand Ltd, Wellington
Copp Clark Pitman, Toronto

First published 1984

AMS Subject Classifications: (main) 60H05, 60H10, 60H99, 60J25, 60J60
(subsidiary) 35K15, 93E15, 93E20

Library of Congress Cataloging in Publication Data

Mohammed, S. E. A.
Stochastic functional differential equations.

(Research notes in mathematics; 99)
Bibliography: p.
Includes index.
1. Stochastic differential equations. 2. Functional differential equations. I. Title. II. Series.
QA274.23.M64 1984 519.2 83-24973
ISBN 0-273-08593-X

British Library Cataloguing in Publication Data

Mohammed, S. E. A.
Stochastic functional differential equations.—
(Research notes in mathematics; 99)
1. Stochastic differential equations
I. Title II. Series
519.2 QA274.23

ISBN 0-273-08593-X

Reproduced and printed by photolithography
in Great Britain by Biddles Ltd, Guildford

Preface

Many physical phenomena can be modelled by stochastic dynamical systems whose evolution in time is governed by random forces as well as intrinsic dependence of the state on a finite part of its past history. Such models may be identified as stochastic (retarded) functional differential equations (stochastic FDE's).

Our main concern in this book is to elucidate a general theory of stochastic FDE's on Euclidean space. In order to have an idea about what is actually going on in such differential systems, let us consider the simplest stochastic delay differential equation. For a non-negative number $r > 0$, this looks like

$$dx(t) = x(t-r)\ dw(t) \qquad \text{(SDDE)}$$

where w is a one-dimensional Brownian motion on a probability space (Ω, F, P) and the solution x is a real-valued stochastic process. It is interesting to compare (SDDE) with the corresponding deterministic delay equation

$$dy(t) = y(t-r)dt. \qquad \text{(DDE)}$$

One can then immediately draw the following analogies between (SDDE) and (DDE):

(a) If both equations are to be integrated forward in time and starting from zero, then it is necessary to specify *a priori* an *initial process* $\{\theta(s): -r \leq s \leq 0\}$ for (SDDE) and a *deterministic fucntion* $\eta:[-r,0] \to \mathbb{R}$ for (DDE). In the ordinary case $(r = 0)$, a simple application of the (Itô) calculus gives the following particular solutions of (SDDE) and (DDE) in closed form:

$$x(t) = e^{w(t)-\frac{1}{2}t},\ y(t) = e^{t},\ t \in \mathbb{R}.$$

For positive delays $(r > 0)$, no simple closed-form solution of (SDDE) is known to me. On the other hand, (DDE) admits exponential solutions $y(t) = e^{\lambda t}$, $t \in \mathbb{R}$, where $\lambda \in \mathbb{C}$ solves the characteristic equation

$$\lambda - e^{-\lambda r} = 0.$$

(b) When $r > 0$, both (SDDE) and (DDE) can be solved uniquely through θ, η, respectively just by integrating forward over steps of size r; e.g.

$$x(t) = \begin{cases} \theta(0) + \int_0^t \theta(u-r)\,dw(u) & 0 \leq t \leq r \\ \theta(t) & -r \leq t < 0 \end{cases}$$

and

$$y(t) = \begin{cases} \eta(0) + \int_0^t \eta(u-r)\,du & 0 \leq t \leq r \\ \eta(t) & -r \leq t < 0 \end{cases}$$

Similar formulae hold over the successive intervals $[r,2r]$, $[2r,3r]$, etc.

(c) In just the same way as the continuation property fails for actual solutions $y:[-r,\infty) \to \mathbb{R}$ of (DDE), it is clear that a Markov property cannot hold for solutions $\{x(t): t \geq -r\}$ of (SDDE) in $\mathbb{R}$ when $r > 0$. Heuristically speaking, a positive delay upsets Markov behaviour in a stochastic delay equation.

(d) To overcome the difficulty in (c), we let $C \equiv C([-r,0],\mathbb{R})$ denote the Banach space of all continuous real functions on $[-r,0]$ given the supremum norm $\|\cdot\|_C$. For each $t \geq 0$, pick the *slice* of the solution paths over the interval $[t-r,t]$ and so obtain trajectories $\{x_t: t \geq 0\}$, $\{y_t: t \geq 0\}$ of (SDDE) and (DDE) traced out in C. It now follows from our results in Chapter III (Theorems (III.1.1), (III.2.1), (III.3.1)) that trajectories $\{x_t: t \geq 0\}$ describe a time-homogeneous continuous Feller process on C.

(e) As functions of the initial path $\eta \in C$, trajectories $\{{}^{\eta}x_t: t \geq 0\}$, $\{{}^{\eta}y_t: t \geq 0\}$ of (SDDE) and (DDE) through η define a *trajectory field*

$$T_t^s: C \to \mathcal{L}^2(\Omega, C), \quad t \geq 0,$$
$$\eta \mapsto {}^{\eta}x_t$$

and a *semi-flow*

$$T_t^d: C \to C, \quad t \geq 0$$
$$\eta \mapsto {}^{\eta}y_t$$

respectively.

Both T_t^s and T_t^d are continuous linear, where $\mathcal{L}^2(\Omega,C)$ is the complete space of all F-measurable $\theta:\Omega\to C$ such that $\int_\Omega \|\theta(\omega)\|_C^2 \, dP(\omega) < \infty$, furnished with the $\mathcal{L}^2$ semi-norm

$$\|\theta\|_{\mathcal{L}^2} = \left[\int_\Omega \|\theta(\omega)\|_C^2 \, dP(\omega)\right]^{1/2}.$$

(Cf. Theorem (II.3.1)). However, in Chapter V §3, we show that the trajectory field T_t^s, $t \geq 0$, does *not* admit 'good' sample function behaviour. Thus, despite the fact that Borel measurable versions always exist, no such version of the trajectory field has almost all sample functions locally bounded (or even linear) on C (cf. Corollary (V.4.7.1), V §3, VI §3). It is intriguing to observe here that this type of erratic behaviour is peculiar to *delayed diffusions* (SDDE) with $r > 0$. Indeed for the ordinary case $r = 0$ it is well-known that the trajectory field has sufficiently smooth versions with almost all sample functions diffeomorphisms of Euclidean space onto itself (Kunita [45], Ikeda and Watanabe [35], Malliavin [51], Elworthy [19], Bismut [5]).

(f) At times $t \geq r$, the deterministic semi-flow T_t^d maps continuous paths into C^1 paths, while the corresponding trajectory field T_t^s takes continuous paths into α-Hölder continuous ones with $0 < \alpha < \frac{1}{2}$ (cf. Theorem (V. 4.4)).

Now our discussion has so far been with reference to the rather special examples of stochastic and deterministic delay equations (SDDE) and (DDE) above. However, this is indeed no serious restriction; it is one of our main contentions in this book that the observations (a) - (f) are essentially valid for a much wider class of stochastic FDE's than just (SDDE). Thus in Chapter II we establish existence, uniqueness and continuous dependence on the initial process for solutions to general stochastic FDE's of the form

$$dx(t) = g(t,x_t)dz(t), \quad t > 0$$

$$x_0 = \theta \in \mathcal{L}^2(\Omega,C)$$

where the *coefficient process* $g:\mathbb{R}^{\geq 0} \times \mathcal{L}^2(\Omega,C) \to \mathcal{L}^2(\Omega,\mathbb{R}^n)$ and the *initial process* $\theta \in \mathcal{L}^2(\Omega,C)$ are given, with z a McShane-type noise on a filtered probability space $(\Omega,F,(F_t)_{t\geq 0},P)$. (Refer to Conditions (E)(i) of Chapter II).

Chapter III essentially says that for systems of type

$$dx(t) = H(t,x_t)dt + G(t,x_t)dw(t),\ t > 0,\ x_0 = \eta \in C$$

the trajectory field $\{^{\eta}x_t : t \geq 0\}$ describes a Feller process on the state space C. Here the *drift coefficient* $H:\mathbb{R}^{\geq 0} \times C \to \mathbb{R}^n$ takes values in $\mathbb{R}^n$ and the *diffusion coefficient* $G:\mathbb{R}^{\geq 0} \times C \to L(\mathbb{R}^m,\mathbb{R}^n)$ has values in the space $L(\mathbb{R}^m,\mathbb{R}^n)$ of all linear maps $\mathbb{R}^m \to \mathbb{R}^n$. The noise w is a (standard) m-dimensional Brownian motion. If the stochastic FDE is autonomous, the trajectory field is a *time-homogeneous diffusion* on C.

In Chapter IV we look at autonomous stochastic FDE's

$$dx(t) = H(x_t)dt + G(x_t)dw(t)$$

and investigate the structure of the associated one-parameter semi-group $\{P_t\}_{t\geq 0}$ given by the time-homogeneous diffusion on C. A novel feature of such diffusions *when* $r > 0$ is that the semi-group $\{P_t\}_{t\geq 0}$ is *never strongly continuous* on the Banach space $C_b(C,\mathbb{R}) \equiv C_b$ of all bounded uniformly continuous real-valued functions on C endowed with the supremum norm (Theorem (IV. 2.2)). Hence a weak generator A of $\{P_t\}_{t\geq 0}$ can be defined on the latter's *domain of strong continuity* $C_b^0 \subset C_b$ and a general formula for A is established in Theorem (IV. 3.2). Due to the absence of non-trivial differentiable functions on C having bounded supports, we are only able to define a *weakly dense* class of smooth functions on C which is rich enough to generate the Borel σ-algebra of C. These are what we call *quasi-tame functions* (IV §4). On such functions the weak generator assumes a particularly simple and concrete form (Theorem (IV. 4.3)).

Distributional and sample regularity properties for trajectory fields of autonomous stochastic FDE's are explored in Chapter V. We look at two extreme examples: the highly erratic delayed diffusions mentioned above, and the case of stochastic FDE's with *ordinary diffusion coefficients* viz.

$$dx(t) = H(x_t)dt + g(x(t))dw(t),\ t > 0.$$

If g satisfies a *Frobenius condition*, the trajectory field of the latter class admits sufficiently smooth and locally compactifying versions for $t \geq r$ (Theorem (V. 2.1), Corollaries V (2.1.1) - V (2.1.4)). In general, the compactifying nature of the trajectory field for $t \geq r$ is shown to persist in a *distributional sense* for autonomous stochastic FDE's with arbitrary Lipschitz coefficients (Theorems (V. 4.6), (V. 4.7)).

There are many examples of stochastic FDE's. In Chapters VI and VII we highlight only a few. Among these are stochastic ODE's ($r = 0$, VI §2), stochastic delay equations VI §3), linear FDE's forced by white noise (VI §4), a model for physical Brownian motion (VII §2), stochastic FDE's with discontinuous initial data (VII §3), stochastic integro-differential equations (VII §4), and stochastic FDE's with an infinite memory ($r = \infty$, VII §5). Chapter VII contains also some open problems and conjectures with a view to future developments.

From a historical point of view, equations with zero diffusions (RFDE's) or zero retardation (stochastic ODE's) have been the scene of intensive study during the past few decades. There is indeed a vast amount of literature on RFDE's e.g. Hale [26], [27], [28], Krasovskii [43], El'sgol'tz [18], Mishkis [56], Jones [42], Banks [3], Bellman and Cooke [4], Halanay [25], Nussbaum [62], [63], Mallet-Paret [49], [50], Oliva [64], [65], Mohammed [57], and others. On stochastic ODE's, one could point out the outstanding works of Itô [36], [37], [38], Itô and McKean [40], McKean [52], Malliavin [51], McShane [53], Gihman and Skorohod [24], Friedman [22], Stroock and Varadhan [73], Kunita [45], Ikeda and Watanabe [35], and Elworthy [19]. However, general stochastic FDE's have so far received very little attention from stochastic analysts and probabilists. In fact a surprisingly small amount of literature is available to us at present on the theory of stochastic FDE's. The first work that we are aware of goes back to an extended article of Itô and Nisio [41] in 1964 on stationary solutions of stochastic FDE's with infinite memory ($r = \infty$). The existence of invariant measures for non-linear FDE's with white noise and a finite memory was considered by M. Scheutzow in [69], [70].

Apart from Section VII §5 and except when otherwise stated, all the results in Chapters II-VII are new. Certain parts of Chapters II, III and IV were included in preprints [58], [59], [60], by the author during the period 1978-1980. Section VI §4 is joint work of S.E.A. Mohammed, M. Scheutzow and H.v. Weizsäcker.

The author wishes to express his deep gratitude to K.D. Elworthy, K.R. Parthasarathy, P. Baxendale, R.J. Elliott, H.v. Weizsäcker, M. Scheutzow and S.A. Elsanousi for many inspiring conversations and helpful suggestions.

For financial support during the writing of this book I am indebted to the British Science and Engineering Research Council (SERC), the University of Khartoum and the British Council.

Finally, many thanks go to Terri Moss who managed to turn often illegible scribble into beautiful typescript.

Salah Mohammed

Khartoum 1983.

Contents

I Preliminary background

§1. Introduction

In this chapter we give an assortment of basic ideas and results from Probability Theory and Linear Analysis which make necessary prerequisites for reading the subsequent chapters. Due to limitations of space, almost all proofs have been omitted. However, we hope that the referencing is adequate.

§2. Measure and Probability

A *measurable space* (Ω, F) is a pair consisting of a non-empty set Ω and a σ-algebra F of subsets of Ω. If E is a real Banach space, an *E-valued measure* μ on (Ω, F) is a map $\mu: F \to E$ such that (i) $\mu(\emptyset) = 0$, (ii) μ is *σ-additive* i.e. for any disjoint countable family of sets $\{A_k\}_{k=1}^{\infty}$ in F the series $\sum_{k=1}^{\infty} \mu(A_k)$ converges in E and $\mu(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \mu(A_k)$. When $E = \mathbb{R}$, μ is called a *signed measure* and if $\mu(F) \subseteq \mathbb{R}^{\geq 0}$ it is called a *positive measure*. If $\sup\{|\mu(A)| : A \in F\} < \infty$, μ is a *finite measure*. A positive finite measure P on (Ω, F) such that $P(\Omega) = 1$ is a *probability measure* on Ω; the triple (Ω, F, P) is then a *probability space*. The set of all finite real-valued measures on Ω is denoted by $M(\Omega)$ and the subset of all probability measures by $M_p(\Omega)$.

A probability space (Ω, F, P) is *complete* if every subset of a set of P-measure zero belongs to F i.e. whenever $B \in F$, $P(B) = 0$, $A \subseteq B$, then $A \in F$. In general any probability space can be completed with respect to its underlying probability measure. Indeed let (Ω, F, P) be an arbitrary probability space and take

$$\bar{F}^P = \{A \cup \Delta : A \in F, \Delta \subset \Delta_0 \in F, P(\Delta_0) = 0\}$$

to be the *completion* of F *under* P. Extend P to $\bar{F}^P$ by setting $P(A \cup \Delta) = P(A)$, $\Delta \subset \Delta_0$, $P(\Delta_0) = 0$. Then $(\Omega, \bar{F}^P, P)$ is the *smallest* P-complete probability space with $F \subset \bar{F}^P$. Because of this property, we often find it technically simpler to assume from the outset that our underlying probability space is complete.

When Ω is a Hausdorff topological space and F is its Borel σ-algebra, Borel Ω, generated by all open (or closed) sets, a measure μ on Ω is *regular* if

$$\mu(B) = \sup \{\mu(C) : C \subseteq B, C \text{ closed}\}$$
$$= \inf \{\mu(U) : B \subseteq U, U \text{ open}\}.$$

If Ω is metrizable, every $\mu \in M(\Omega)$ is regular and hence completely determined by its values on the open (or closed) sets in Ω. (Parthasarathy [66], Chapter II, Theorem 1.2, p. 27).

Let Ω be a Hausdorff space and E a real Banach space. An E-valued measure μ on (Ω, Borel Ω) is *tight* if (i) $\sup \{|\mu(B)|: B \in \text{Borel } \Omega\} < \infty$, where $|\cdot|$ denotes the norm on E; and (ii) for each $\varepsilon > 0$, there is a compact set K_ε in Ω such that $|\mu(\Omega \setminus K_\varepsilon)| < \varepsilon$

Theorem (2.1): Let Ω be a *Polish space* i.e. a complete separable metrizable space. Then every finite real Borel measure on Ω is tight.

Proof: Parthasarathy ([66], Chapter II, Theorem 3.2, p. 29); Stroock and Varadhan ([73], Chapter 1, Theorem (1.1.3), pp. 9-10). □

Let Ω be a separable metric space and F = Borel Ω. Denote by $C_b(\Omega,\mathbb{R})$ the Banach space of all bounded uniformly continuous functions $\phi:\Omega \to \mathbb{R}$ given the supremum norm

$$||\phi||_{C_b} = \sup \{|\phi(\eta)|: \eta \in \Omega\}.$$

The natural bilinear pairing

$$\langle\cdot,\cdot\rangle : C_b(\Omega,\mathbb{R}) \times M(\Omega) \to \mathbb{R}$$
$$\langle\phi,\mu\rangle = \int_{\eta\in\Omega} \phi(\eta)\, d\mu(\eta), \ \phi \in C_b(\Omega,\mathbb{R}),\ \mu \in M(\Omega),$$

induces an embedding

$$M(\Omega) \longrightarrow C_b(\Omega,\mathbb{R})^*$$
$$\mu \longmapsto \langle\cdot,\mu\rangle$$

where $C_b(\Omega,\mathbb{R})^*$ is the strong dual of $C_b(\Omega,\mathbb{R})$. Indeed each $\mu \in M(\Omega)$ corresponds to the continuous linear functional

$$C_b(\Omega,\mathbb{R}) \longrightarrow \mathbb{R}$$
$$\phi \longmapsto \int_{\eta\in\Omega} \phi(\eta)\, d\mu(\eta)$$

because for every $\phi \in C_b(\Omega,\mathbb{R})$

$$\left|\int_{\eta\in\Omega} \phi(\eta)\, d\mu(\eta)\right| \leq \|\phi\|_{C_b}\, v(\mu)(\Omega)$$

where $v(\mu)(\Omega) = \sup\{\sum_{k=1}^{p} |\mu(A_k)| : A_k \in F,\ k = 1,\ldots,p \text{ disjoint},\ \Omega = \bigcup_{k=1}^{p} A_k,$ $p < \infty\}$ is the *total variation of* μ on Ω (Dunford and Schwartz [15], Chapter III, pp. 95-155). As a subset of $C_b(\Omega,\mathbb{R})^*$ give $M(\Omega)$ the induced weak * topology. Now this turns out to be the same as the *weak toplogy* or *vague topology* on measures because of the following characterizations.

Theorem (2.2): Let Ω be a metric space and μ, $\mu_k \in M(\Omega)$ for $k = 1,2,3,\ldots$. Then the following statements are all equivalent:

(i) $\mu_k \to \mu$ as $k \to \infty$ in the weak * topology of $M(\Omega)$;

(ii) $\lim_{k\to\infty} \int_{\eta\in\Omega} \phi(\eta)d\mu_k(\eta) = \int_{\eta\in\Omega} \phi(\eta)d\mu(\eta)$, for every $\phi \in C_b(\Omega,\mathbb{R})$;

(iii) $\limsup_k \mu_k(C) \leq \mu(C)$ for every closed set C in Ω;

(iv) $\liminf_k \mu_k(U) \geq \mu(U)$ for every open set U in Ω;

(v) $\lim_{k\to\infty} \mu_k(B) = \mu(B)$ for every $B \in$ Borel Ω such that $\mu(\partial B) = 0$.

For proofs of the above theorem, see Parthasarathy ([66] Chapter II, Theorem 6.1, pp. 40-42) or Stroock and Varadhan ([73], Theorem 1.1.1, pp. 7-9).

The weak topology on $M(\Omega)$, when Ω is a separable metric space, can be alternatively described in the following two equivalent ways:

(a) Define a base of open neighbourhoods of any $\mu \in M(\Omega)$ by

$$U_\mu(\phi_1,\ldots,\phi_p;\ \varepsilon_1,\ldots,\varepsilon_p) = \{\nu : \nu \in M(\Omega),\ \left|\int \phi_k\, d\nu - \int \phi_k\, d\mu\right| < \varepsilon_k,\ k = 1,2,\ldots,p\}$$

where $\phi_1,\ldots,\phi_p \in C_b(\Omega,\mathbb{R})$, $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_p > 0$.

(b) Furnish $M(\Omega)$ with a metric ρ in the following manner. Compactify the separable metric space Ω to obtain $\tilde{\Omega}$. Then $C_b(\Omega,\mathbb{R})$ is Banach space-isomorphic, to the separable Banach space $C(\tilde{\Omega},\mathbb{R})$ of all continuous real functions on $\tilde{\Omega}$, given the supremum norm. Pick a countable dense sequence $\{\phi_k\}_{k=1}^{\infty}$ in $C_b(\Omega,\mathbb{R})$ and define the metric ρ on $M(\Omega)$ by

$$\rho(\mu,\nu) = \sum_{k=1}^{\infty} \frac{1}{2^k \|\phi_k\|_{C_b}} \left| \int_\Omega \phi_k \, d\mu - \int_\Omega \phi_k \, d\nu \right| ,$$

$\mu,\nu \in M(\Omega)$ (Stroock and Varadhan [73], Theorem 1.1.2, p. 9; Parthasarathy [66], pp. 39-52). Note that $M(\Omega)$ is complete if and only if Ω is so. Similarly, $M_p(\Omega)$ is compact if and only if Ω is compact. More generally compact subsets of $M(\Omega)$ are characterized by the well-known theorem of Prohorov given in Chapter V (Theorem (V.4.5)).

There is a theory of (Bochner) integration for maps $X:\Omega \to E$ where E is a real Banach space and (Ω,F,μ) a real measure space (Dunford and Schwartz [15], Chapter III §1-6).

On a probability space (Ω,F,P) an (F, Borel E)-measurable map $X:\Omega \to E$ is called an E-valued *random variable*. Such a map is P-*integrable* if there is a sequence $X_n:\Omega \to E$, $n = 1,2,\ldots$, of simple (F, Borel E)-measurable maps so that $X_n(\omega) \to X(\omega)$ as $n \to \infty$ for a.a. $\omega \in \Omega$ and $\lim_{m,n\to\infty} \int_{\omega\in\Omega} |X_n(\omega)-X_m(\omega)|_E dP(\omega) = 0.$

Define the *expectation of* an integrable random variable $X:\Omega \to E$ by

$$EX = \int_\Omega X(\omega)dP(\omega) = \lim_{n\to\infty} \int_\Omega X_n(\omega)dP(\omega) \in E.$$

This definition is independent of the choice of sequence $\{X_n\}_{n=1}^{\infty}$ converging a.s. to X (Rao [67], Chapter I, §1.4; Yosida [78], Chapter V §5, pp. 132-136).

For a separable Banach space E, $X: \Omega \to E$ is a random variable if and only if one of the following conditions holds:

(i) There is a sequence $X_n:\Omega \to E$, $n = 1,2,\ldots$, of simple (F, Borel E)-measurable maps such that $X_n(\omega) \to X(\omega)$ as $n \to \infty$ for a.a. $\omega \in \Omega$;

(ii) X is *weakly measurable* i.e. for each $\lambda \in E^*$, $\lambda \circ X:\Omega \to \mathbb{R}$ is (F, Borel $\mathbb{R}$)-measurable.

(Elworthy [19], Chapter I, §1(C) pp. 2-4; Rao [67], Chapter I §1.4).

Denote by $\mathcal{L}^0(\Omega,E;F)$ the real vector space of all E-valued random variables $X:\Omega \to E$ on the probability space (Ω,F,P). The space $\mathcal{L}^0(\Omega,E;F)$ is a complete

TVS under the complete pseudo-metric

$$d(X_1,X_2) = \inf[\varepsilon+P\{\omega:\omega \in \Omega, |X_1(\omega)-X_2(\omega)|_E > \varepsilon\} : \varepsilon > 0]$$

for $X_1, X_2 \in \mathcal{L}^0(\Omega,E;F)$. The norm in our real Banach space is always denoted by $|.|_E$ or sometimes just $|.|$. A sequence $\{X_n\}_{n=1}^{\infty}$ of random variables $X_n:\Omega \to E$ *converges in probability* to $X \in \mathcal{L}^0(\Omega,E;F)$ if for every $\varepsilon > 0$ $\lim_{n\to\infty} P\{\omega:\omega \in \Omega, |X_n(\omega) - X(\omega)|_E > \varepsilon\} = 0$. A random variable $X:\Omega \to E$ is (Bochner) integrable if and only if the function

$$|X(\cdot)|_E : \Omega \longrightarrow \mathbb{R}^{\geq 0}$$

$$\omega \longmapsto |X(\omega)|_E$$

is P-integrable, in which case $|\int_\Omega X(\omega)dP(\omega)|_E \leq \int_\Omega |X(\omega)|_E \, dP(\omega)$ i.e. $|EX|_E \leq E|X(\cdot)|_E$. The space $\mathcal{L}^1(\Omega,E;F)$ of all integrable random variables is a complete real TVS with respect to the $\mathcal{L}^1$-semi-norm

$$||X||_{\mathcal{L}^1} = \int_\Omega |X(\omega)|_E \, dP(\omega), \quad X \in \mathcal{L}^1(\Omega,E;F).$$

Similarly for any integer $k > 1$ define the complete space $\mathcal{L}^k(\Omega,E;F)$ of all F-measurable maps $X:\Omega \to E$ such that $\int_\Omega |X(\omega)|^k \, dP(\omega) < \infty$, endowed with the semi-norm

$$||X||_{\mathcal{L}^k} = [\int_\Omega |X(\omega)|_E^k \, dP(\omega)]^{1/k} .$$

Note that the spaces $\mathcal{L}^k(\Omega;E;F)$ become the real Banach spaces $\mathbb{L}^k(\Omega,E;F)$ if we identify random variables which agree on a set of full P-measure. If $X_n, X \in \mathcal{L}^0(\Omega,E;F)$, $n = 1,2,\ldots$, say $X_n \to X$ as $n \to \infty$ *a.s.* or $X_n(\omega) \to X(\omega)$ as $n \to \infty$ *for a.a.* $\omega \in \Omega$ if there exist a set $\Omega_0 \in F$ of full P-measure such that $X_n(\omega) \to X(\omega)$ as $n \to \infty$ for all $\omega \in \Omega_0$.

The various notions of convergence for E-valued random variables are linked by the following theorem which is proved in the same way as for real random variables (cf. Dunford and Schwartz [15], Chapter III §2-4; Rao [67], Chapter I §1.4 pp. 16-29; Halmos [30]).

<u>Theorem (2.3)</u>: Let $X_n, X:\Omega \to E$ be random variables for $n = 1,2,\ldots$, and let E be a real separable Banach space.

(i) If $X_n \to X$ as $n \to \infty$ a.s., then $X_n \to X$ as $n \to \infty$ in probability.

(ii) If $X_n, X \in \mathcal{L}^k(\Omega,E;F)$, $n = 1,2,\ldots$ and $X_n \to X$ as $n \to \infty$ in $\mathcal{L}^k$ $(k > 0)$, then $X_n \to X$ as $n \to \infty$ in probability.

(iii) If $X_n \to X$ as $n \to \infty$ in probability, then there is a subsequence $\{X_{n_i}\}_{i=1}^{\infty}$ of $\{X_n\}_{n=1}^{\infty}$ such that $X_{n_i} \to X$ as $i \to \infty$ a.s.

(iv) Dominated Convergence: Let $X_n \in \mathcal{L}^1(\Omega,E;F)$, $n = 1,2,\ldots$ and $X \in \mathcal{L}^0(\Omega,E;F)$ be such that $X_n \to X$ as $n \to \infty$ in probability. Suppose there exists $Y \in \mathcal{L}^1(\Omega,\mathbb{R}^{\geq 0};F)$ such that, for a.a. $\omega \in \Omega$, $|X_n(\omega)|_E \leq Y(\omega)$ for all $n \geq 1$. Then $X \in \mathcal{L}^1(\Omega,E;F)$ and $\int_\Omega X(\omega)dP(\omega) = \lim_{n\to\infty} \int_\Omega X_n(\omega)dP(\omega)$.

Chebyshev's inequality also holds for Banach-space-valued random variables X. It follows trivially by applying its classical version to the real-valued random variable $|X(\cdot)|_E$ (Chung [7], p. 48):

Theorem (2.4) (*Chebyshev's Inequality*): If E is a Banach space and $X \in \mathcal{L}^k(\Omega,E;F)$, $k \geq 1$, then for every $\varepsilon > 0$

$$P\{\omega:\omega \in \Omega,\ |X(\omega)|_E \geq \varepsilon\} \leq \frac{1}{\varepsilon^k}\int_\Omega |X(\omega)|_E^k \, dP(\omega).$$

In particular the map

$$\begin{aligned} \mathcal{L}^k(\Omega,E;F) &\longrightarrow M_p(E) \\ X &\longmapsto P \circ X^{-1} \end{aligned}$$

is continuous, for each $k \geq 1$. If E is separable, then the above map is continuous also for $k = 0$.

In a probability space (Ω,F,P), two events $A,B \in F$ are *independent* (*under* P) if $P(A \cap B) = P(A)P(B)$; two sub-σ-algebras g_1, g_2 of F are *independent* (*under* P) if $P(A \cap B) = P(A)P(B)$ for all $A \in g_1$ and all $B \in g_2$; two random variables $X, Y : \Omega \to E$ are *independent* (*under* P) if the σ-algebras $\sigma(X)$, $\sigma(Y)$ generated by X, Y respectively are independent under P.

Theorem (2.5) (*Borel-Cantelli Lemma*) Let (Ω,F,P) be a probability space and $\{\Omega^k\}_{k=1}^{\infty} \subset F$.

(i) If $\sum_{k=1}^{\infty} P(\Omega^k)$ converges, then $P(\limsup_k \Omega^k) = 0$ i.e. $P(\liminf_k \Omega \sim \Omega^k) = 1$;