Chapter 4

STOCHASTIC PROCESSES

1 Introduction

In this chapter we shall be concerned with coupling general stochastic processes (in one-sided time) in such a way that they ultimately merge. This was our main concern in the first half of Chapter 2 and in Section 5 of Chapter 3. Recall, for instance, the classical coupling: two differently started versions of a Markov chain run independently until they meet, at a time T say, and then run together from time T onward.

For lack of a better term we shall call this kind of coupling *exact coupling*. The qualifier 'exact' here refers to the fact that the processes coincide *exactly* from T onward as opposed to what was the case in the latter part of Chapter 2 and will be the case in Chapter 5, where the processes merge only modulo a random time shift (shift-coupling, epsilon-coupling). Exact coupling is in fact what some writers still call only 'coupling'. The word coupling then refers to the *merging*, and not to *joint construction* in general as in this book and as is becoming more and more common.

Section 2 starts with preliminaries establishing notation and discussing the definition of a stochastic process to make sure that our simple abstract framework and the reasons why it is chosen are understood.

Section 3 then introduces exact coupling and its *distributional* version, which is obtained by replacing pointwise merging by distributional merging. Distributional exact coupling is not as intuitively appealing as its nondistributional counterpart but has several merits. It applies (slightly?) more generally and has the same distributional implications. It is easier to establish and can serve as a first step in constructing a nondistributional exact coupling. Section 4 takes a brief look at distributional coupling in general and establishes the coupling event inequality.

Section 5 presents the coupling time inequality and the resulting limit theory. Section 6 proves the central maximality result, and Section 8 reformulates the proof after the concept of coupling with respect to a sub-σ-algebra has been introduced in Section 7. Section 9 introduces the tail σ-algebra $\mathcal{T}$ and proves a result on maximal coupling with respect to $\mathcal{T}$ that leads to a basic set of equivalences between successful exact coupling, convergence in total variation, and distributional identity on $\mathcal{T}$.

This chapter is followed by Chapter 5, where analogous theory is established for two generalizations of exact coupling: shift-coupling and epsilon couplings. Chapter 6 then considers the implications of these three sets of coupling results in the Markov case, while Chapter 7 extends the view beyond stochastic processes.

2 Preliminaries – What Is a Stochastic Process?

Before turning to the coupling theory we spend a number of pages (a third of the chapter!) on some general but simple aspects of stochastic processes that are basic for our purposes. The impatient reader could skim rapidly through this section. Much of it is a motivation for the technically straightforward property of 'shift-measurability', which we often impose in order to be able to shift our processes randomly, and which is satisfied in the standard settings such as when the state space is Polish and the paths right continuous.

2.1 Classical Definition

The classical definition of a stochastic process goes as follows: a *stochastic process* with *index set* $\mathbb{I}$ and *state space* $(E, \mathcal{E})$ is a family $Z = (Z_s)_{s \in \mathbb{I}}$, where the Z_s are random elements defined on a common probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and all taking values in $(E, \mathcal{E})$.

We shall here think of the index set as time and mostly restrict the use of the term 'stochastic process' to the following four cases:

$\mathbb{I} = \mathbb{R}$	two-sided continuous time,
$\mathbb{I} = [0, \infty)$	one-sided continuous time,
$\mathbb{I} = \mathbb{Z}$	two-sided discrete time,
$\mathbb{I} = \{0, 1, 2, \dots\}$	one-sided discrete time.

In this chapter we shall, in fact, only be concerned with one-sided processes. But in this section the index set $\mathbb{I}$ is kept general for later purposes.

WARNING. To avoid misunderstanding it should be stressed that here the role of the index set $\mathbb{I}$ is different from the role it had in Chapter 3. We are not going to couple the collection of random elements Z_s, $s \in \mathbb{I}$, that is, their joint distribution will be fixed. We shall couple Z and another process Z'; the collection to be coupled will be Z and Z'.

2.2 Stochastic Process as a Random Mapping in $(E^{\mathbb{I}}, \mathcal{E}^{\mathbb{I}})$

Let $(E^{\mathbb{I}}, \mathcal{E}^{\mathbb{I}})$ denote the product space

$$(E^{\mathbb{I}}, \mathcal{E}^{\mathbb{I}}) := \bigotimes_{s \in \mathbb{I}} (E, \mathcal{E}).$$

Rather than regarding Z in the classical way as a family of random elements in $(E, \mathcal{E})$, we can equivalently regard Z as a random mapping, that is, as a single random element in $(E^{\mathbb{I}}, \mathcal{E}^{\mathbb{I}})$ defined by

$$Z(\omega) = (Z_s(\omega))_{s \in \mathbb{I}}, \quad \omega \in \Omega.$$

The two points of view are equivalent because each Z_s is a measurable mapping from $(\Omega, \mathcal{F})$ to $(E, \mathcal{E})$ if and only if Z is a measurable mapping from $(\Omega, \mathcal{F})$ to $(E^{\mathbb{I}}, \mathcal{E}^{\mathbb{I}})$.

The *distribution* of a stochastic process Z is the distribution of Z as a random element in $(E^{\mathbb{I}}, \mathcal{E}^{\mathbb{I}})$. The distribution of Z is uniquely determined by the *finite-dimensional distributions*, that is, by the distributions of the $(E^n, \mathcal{E}^n)$ valued random elements $(Z_{t_1}, \dots, Z_{t_n})$, $t_1, \dots, t_n \in \mathbb{I}$, $n \geqslant 1$. These finite-dimensional distributions are in turn determined by

$$\mathbf{P}(Z_{t_1} \in A_1, \dots, Z_{t_n} \in A_1), \;\; A_1, \dots, A_n \in \mathcal{E}, \;\; t_1, \dots, t_n \in \mathbb{I}, \;\; n \geqslant 1.$$

In particular, when Z is *real valued*, that is, $(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B})$, then the finite-dimensional distributions are determined by the *finite-dimensional distribution functions*

$$\mathbf{P}(Z_{t_1} \leqslant x_1, \dots, Z_{t_n} \leqslant x_n), \;\; x_1, \dots, x_n \in \mathbb{R}, \;\; t_1, \dots, t_n \in \mathbb{I}, \;\; n \geqslant 1.$$

Finally, for Polish $(E, \mathcal{E})$, Kolmogorov's consistency theorem (Fact 3.2 in Chapter 3) states that if a consistent collection of finite-dimensional distributions is given, then there always exists a stochastic process having these finite-dimensional distributions.

2.3 Path Space $(H, \mathcal{H})$ – Standard Settings

The *paths* of Z are the realizations $Z(\omega)$, $\omega \in \Omega$, of the random mapping Z. Sometimes restrictions are put on the paths, for instance in the continuous-time case that they are continuous, or right-continuous, or right-continuous with left-hand limits, or more generally that they lie in some subset H of $E^{\mathbb{I}}$. In this case it is natural to consider Z as a random element not in $(E^{\mathbb{I}}, \mathcal{E}^{\mathbb{I}})$ but in $(H, \mathcal{H})$, where $\mathcal{H}$ is the σ-algebra on H generated by the projection mappings taking z in H to z_t in E, $t \in \mathbb{I}$. Note that $\mathcal{H}$ is the *trace* of H on $\mathcal{E}^{\mathbb{I}}$, that is,

$$\mathcal{H} := \mathcal{E}^{\mathbb{I}} \cap H := \{A \cap H : A \in \mathcal{E}^{\mathbb{I}}\}.$$

Again the two points of view are equivalent: Z has H valued paths and each Z_t is a measurable mapping from $(\Omega, \mathcal{F})$ to $(E, \mathcal{E})$ if and only if Z is a measurable mapping from $(\Omega, \mathcal{F})$ to $(H, \mathcal{H})$. When a particular *path set* H is given, call $(H, \mathcal{H})$ the *path space* of Z. (One could conceive of allowing $\mathcal{H}$ to be more general than just the trace of H on $\mathcal{E}^{\mathbb{I}}$, but we shall not do so here.)

Note that H need not be an element of $\mathcal{E}^{\mathbb{I}}$. In particular, H is not an element of $\mathcal{E}^{\mathbb{I}}$ in the *standard settings* in continuous time when $(E, \mathcal{E})$ is Polish, $\mathbb{I} = \mathbb{R}$ or $\mathbb{I} = [0, \infty)$, and H is one of the sets

$C_E(\mathbb{I})$ = continuous maps from $\mathbb{I}$ to E,

$D_E(\mathbb{I})$ = right-continuous maps from $\mathbb{I}$ to E with left-hand limits,

$R_E(\mathbb{I})$ = right-continuous maps from $\mathbb{I}$ to E.

Both $C_E(\mathbb{I})$ and $D_E(\mathbb{I})$ can be metrized such that the Borel σ-algebras, $\mathcal{C}_E(\mathbb{I})$ and $\mathcal{D}_E(\mathbb{I})$, are the traces of the respective path sets $C_E(\mathbb{I})$ and $D_E(\mathbb{I})$ on $\mathcal{E}^{\mathbb{I}}$ (see Ethier and Kurtz (1986)), but we shall not need this fact here [except for three isolated results, Theorem 7.4 in Chapter 5, Theorem 5.4 in Chapter 8, and the (v)-part of Theorem 3.4(b) in Chapter 10].

The *distribution* of a stochastic process Z with path space $(H, \mathcal{H})$ is the distribution of Z as a random element in $(H, \mathcal{H})$. The distribution is again uniquely determined by the finite-dimensional distributions.

However, even if $(E, \mathcal{E})$ is Polish and a consistent collection of finite-dimensional distributions is given, there need *not* be a stochastic process with path space $(H, \mathcal{H})$ having these finite-dimensional distributions. This has to be checked in each individual case. One example where it can be established is for the Wiener process: there is a one-sided continuous-time real-valued process with path space $(C_E(\mathbb{I}), \mathcal{C}_E(\mathbb{I}))$ having independent stationary increments (see Billingsley (1986) or Kallenberg (1997), we will only use this fact for an isolated example, Section 8.1 in Chapter 7).

2.4 Observing a Process at a Random Time

We would like to be able to observe a stochastic process at a random time. There is no complication with this in discrete time, but in continuous time there is. So let Z be a one-sided continuous-time stochastic process with state space $(E, \mathcal{E})$ and let T be a random time in $[0, \infty)$. By Z_T we mean the E valued mapping defined on Ω in the obvious way:

$$Z_T(\omega) := Z_{T(\omega)}(\omega), \quad \omega \in \Omega.$$