Chapter 5

SHIFT-COUPLING

1 Introduction

The previous chapter dealt with coupling one-sided stochastic processes in such a way that their paths eventually merge. This we called exact coupling to distinguish it from the more general *shift-coupling* to be considered in this chapter. Shift-coupling means that the paths eventually do not merge 'exactly' but only modulo a random time shift. In this chapter we shall also consider an issue that arises only in continuous time: what happens when the random time shift can be made arbitrarily small, that is, when *epsilon couplings* exist.

It turns out that both shift-coupling and epsilon-couplings have a theory paralleling that of exact coupling: they can be linked to a mode of convergence (*Cesaro* and *smooth* total variation convergence, respectively) and to a σ-algebra (the *invariant* and the *smooth tail* σ-algebra, respectively) in the same way as exact coupling is linked to plain total variation convergence and to the tail σ-algebra. For both shift-coupling and epsilon-coupling we introduce inequalities that play the same key role as the coupling time inequality in the exact coupling case.

In order to stress the similarities (and the dissimilarities) between these three types of coupling and to make comparison easier, the treatment of first shift-coupling (Sections 2 through 5) and then epsilon-couplings (Sections 6 through 9) is organized in the same way as that of exact coupling (Sections 3, 5, 6, and 9 in Chapter 4): the sections have analogous titles, and the subsections and theorems are enumerated in the same way (when possible). We start with a section defining the concept and its distributional version, continue with a section presenting the inequality and the resulting limit theory, then move on to a section discussing the question of maximality, and finish with a section introducing the σ-algebra and the basic set of equivalences between the coupling, the total variation result, and the σ-algebra.

Throughout this chapter U is a random variable that is uniform on $[0, 1]$ and independent of the processes and the shift-coupling (epsilon couplings). And note that in the continuous-time case we now impose the shift-measurability condition throughout.

2 Shift-Coupling – Distributional Shift-Coupling

This section introduces shift-coupling and its distributional version.

2.1 Shift-Coupling – Definition

Let Z and Z' be one-sided discrete-time or continuous-time shift-measurable stochastic processes with general state space $(E, \mathcal{E})$ and path space $(H, \mathcal{H})$; see Section 2 in Chapter 4. We shall use notation in accordance with continuous time, but all we need to switch to discrete time is to substitute, for instance, t by n and s by k and to introduce the following convention: in the discrete-time case extend the definition of the shift-maps to noninteger times by

$$\theta_t z = \theta_{[t]} z, \qquad t \in [0, \infty), \quad z = (z_0, z_1, \dots) \in E^{\{0,1,\dots\}}.$$

A *shift-coupling* of Z and Z' is a quadruple $(\hat{Z}, \hat{Z}', T, T')$ where $(\hat{Z}, \hat{Z}')$ is a coupling of Z and Z' and T and T' are two random times (integer-valued in the discrete-time case) such that

$$\begin{aligned} &\theta_T \hat{Z} = \theta_{T'} \hat{Z}' \text{ on } \{T < \infty\} \\ &\text{and } \{T < \infty\} = \{T' < \infty\}. \end{aligned} \tag{2.1}$$

Using the convention that a process is absorbed in the cemetery state when shifted to infinity we can rewrite (2.1) simply as

$$\theta_T \hat{Z} = \theta_{T'} \hat{Z}' \qquad \text{(see Figure 2.1).}$$

The times T and T' are the *shift-coupling times*, $\mathbf{P}(T < \infty)$ is the *success probability*, and the shift-coupling is *successful* if $\mathbf{P}(T < \infty) = 1$. When $T < \infty$, then $T - T'$ is the *shift*. There is no shift if $T \equiv T'$, and then the shift-coupling is an exact coupling.

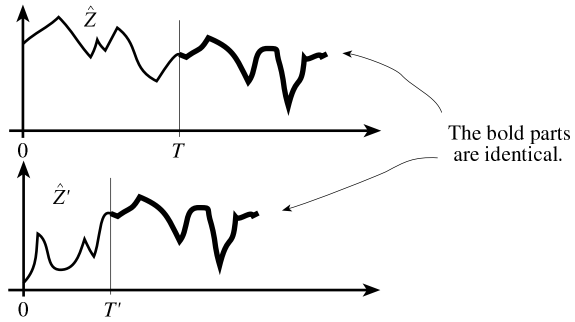

FIGURE 2.1. Nondistributional shift-coupling: merging modulo a time shift.

2.2 Distributional Shift-Coupling – Definition

Say that $(\hat{Z}, \hat{Z}', T, T')$ is a *distributional shift-coupling* of Z and Z' if $(\hat{Z}, \hat{Z}')$ is a coupling of Z and Z', and T and T' are nonnegative random times such that

$$\theta_T \hat{Z} \stackrel{D}{=} \theta_{T'} \hat{Z}'; \tag{2.2}$$

here we again use the convention that the shifted processes are absorbed in the cemetery state when T and T' are infinite.

Certainly a shift-coupling is also a *distributional* shift-coupling. We shall use the word *nondistributional* to distinguish a shift-coupling from a distributional one. Otherwise, we use the same terminology in both cases.

For an example of a successful distributional shift-coupling consider two independent differently started versions, Z and Z', of a countable state space irreducible recurrent Markov chain and let T and T' be the times when Z and Z', respectively, first hit a fixed state. Then (Z, Z', T, T') is a distributional shift-coupling of Z and Z'. A nondistributional shift-coupling is obtained by letting the chains continue in the same way after hitting the state.

Note that (2.2) implies nothing about T and T' except that $\mathbf{P}(T < \infty) = \mathbf{P}(T' < \infty)$. It is an interesting observation, however, that a distributional shift-coupling of the *space-time* processes $(Z_s, s)_{s\in[0,\infty)}$ and $(Z'_s, s)_{s\in[0,\infty)}$ is a distributional *exact* coupling of Z and Z', since

$$\theta_T (\hat{Z}_s, s)_{s\in[0,\infty)} \stackrel{D}{=} \theta_{T'} (\hat{Z}'_s, s)_{s\in[0,\infty)}$$

is equivalent to $(\theta_T \hat{Z}, T) \stackrel{D}{=} (\theta_{T'} \hat{Z}', T')$. And in the nondistributional case a shift-coupling of the space-time processes is equivalent to $T = T'$ and $\theta_T \hat{Z} = \theta_T \hat{Z}'$.

2.3 The Hats May Be Dropped in the Distributional Case

If we have a distributional shift-coupling of Z and Z', then we can take $\hat{Z}$ and $\hat{Z}'$ to be the original processes Z and Z'.

Theorem 2.1. *Suppose $(\hat{Z}, \hat{Z}', \hat{T}, \hat{T}')$ is a distributional shift-coupling of Z and Z'. Then the underlying probability space can be extended to support random times T and T' such that*

$$(Z, T) \stackrel{D}{=} (\hat{Z}, \hat{T}) \quad \textit{and} \quad (Z', T') \stackrel{D}{=} (\hat{Z}', \hat{T}').$$

In particular,

$$\theta_T Z \stackrel{D}{=} \theta_{T'} Z'. \tag{2.3}$$

Proof. This follows from the transfer extension in Section 4.5 of Chapter 3. In order to obtain T take $Y_1 := Z$ and $(Y_1', Y_2') := (\hat{Z}, \hat{T})$ and define $T := Y_2$. Similarly, in order to obtain T' take $Y_1 := Z'$ and $(Y_1', Y_2') := (\hat{Z}', \hat{T}')$ and define $T' := Y_2$. □

This theorem motivates again dropping the hats when discussing distributional shift-coupling, when there is no danger of confusion. Say that T and T' are distributional shift-coupling times *of* Z and Z' if (2.3) holds.

2.4 Turning Distributional into Nondistributional

In the standard settings a distributional shift-coupling can always be turned into a nondistributional one.

Theorem 2.2. *Let $(\hat{Z}, \hat{Z}', \hat{T}, \hat{T}')$ be a distributional shift-coupling of Z and Z'. Suppose there exists a weak-sense-regular conditional distribution of $\hat{Z}'$ given $\theta_{\hat{T}'} \hat{Z}'$ [this holds in discrete time when the state space is Polish and in continuous time when the state space is Polish and the paths are right-continuous]. Then the underlying probability space $(\Omega, \mathcal{F}, \mathbf{P})$ can be extended to support T, Z'', and T'' such that*

$$(Z, T) \stackrel{D}{=} (\hat{Z}, \hat{T}) \quad \textit{and} \quad (Z'', T'') \stackrel{D}{=} (\hat{Z}', \hat{T}')$$

and (Z, Z'', T, T'') is a nondistributional shift-coupling of Z and Z'.

Proof. Let T be as in Theorem 2.1. To obtain (Z'', T'') use the transfer extension in Section 2.12 of Chapter 4 as follows. Take $Y_1 := \theta_T Z$ and $(Y_1', Y_2') := (\theta_{\hat{T}'} \hat{Z}', \kappa_{\hat{T}'} \hat{Z}')$ to obtain Y_2 such that [see Section 2.9 of Chapter 4 for the definition of the killing maps κ_t]

$$(\theta_T Z, Y_2) \stackrel{D}{=} (\theta_{\hat{T}'} \hat{Z}', \kappa_{\hat{T}'} \hat{Z}').$$

Define (Z'', T'') by

$$(\theta_{T''} Z'', \kappa_{T''} Z'') := (\theta_T Z, Y_2) \qquad (\text{thus } \theta_T Z = \theta_{T''} Z'').$$

Since $(\theta_{T''}Z'', \kappa_{T''}Z'')$ is a copy of $(\theta_{\hat{T}'}\hat{Z}', \kappa_{\hat{T}'}\hat{Z}')$, it follows that (Z'', T'') is a copy of $(\hat{Z}', \hat{T}')$ because (Z'', T'') is determined in the same measurable way by $(\theta_{T''}Z'', \kappa_{T''}Z'')$ as $(\hat{Z}', \hat{T}')$ is by $(\theta_{\hat{T}'}\hat{Z}', \kappa_{\hat{T}'}\hat{Z}')$. □

3 Shift-Coupling – Inequality and Asymptotics

The last section was devoted to the definition of shift-coupling and its distributional version. We shall now go on to the limit implications.

3.1 Shift-Coupling Inequality – And Its Reformulations

Rather than shifting Z to a nonrandom t as in the coupling time inequality we now shift to a point picked uniformly at random in $[0, t]$.

Theorem 3.1. *Let Z and Z' be one-sided discrete-time or continuous-time shift-measurable stochastic processes with a general state space $(E, \mathcal{E})$ and path space $(H, \mathcal{H})$. If there is a distributional shift-coupling of Z and Z' with times T and T', then the underlying probability space can be extended to support a copy R of T' such that $\{T < \infty\} = \{R < \infty\}$ and, for $0 \leqslant t < \infty$,*

$$\begin{aligned} &\|\mathbf{P}(\theta_{Ut}Z \in \cdot) - \mathbf{P}(\theta_{Ut}Z' \in \cdot)\| && \text{SHIFT-COUPLING} \\ &\quad \leqslant 2\mathbf{P}(T \vee R > Ut), && \text{INEQUALITY} \end{aligned}$$

where U is uniform on $[0, 1]$ and independent of Z, Z', T, and R. In the nondistributional case we can take $R := T'$ and the shift-coupling inequality becomes

$$\|\mathbf{P}(\theta_{Ut}Z \in \cdot) - \mathbf{P}(\theta_{Ut}Z' \in \cdot)\| \leqslant 2\mathbf{P}(T \vee T' > Ut).$$

Proof. With U independent of the shift-coupling $(\hat{Z}, \hat{Z}', T, T')$, note that the remainder when $T + Ut$ is divided by t,

$$(T + Ut) \bmod t := (T/t + U - [T/t + U])t,$$

is uniform on $[0, t]$ and independent of $\hat{Z}$. Therefore, $\theta_{(T+Ut) \bmod t}\hat{Z}$ is a copy of $\theta_{Ut}Z$. Similarly, $\theta_{(T'+Ut) \bmod t}\hat{Z}'$ is a copy of $\theta_{Ut}Z'$. Thus

$$\begin{aligned} &(\theta_{(T+Ut) \bmod t}\hat{Z}, \theta_{(T'+Ut) \bmod t}\hat{Z}') \\ &\text{is a coupling of } \theta_{Ut}Z \text{ and } \theta_{Ut}Z'. \end{aligned} \tag{3.1}$$

In the nondistributional case $\theta_T\hat{Z} = \theta_{T'}\hat{Z}'$ yields the second identity in

$$\begin{aligned} \theta_{(T+Ut) \bmod t}\hat{Z} &= \theta_{Ut}\theta_T\hat{Z} = \theta_{Ut}\theta_{T'}\hat{Z}' \\ &= \theta_{(T'+Ut) \bmod t}\hat{Z}' \quad \text{on} \quad \{Ut \leqslant t - T \vee T'\}, \end{aligned}$$