Chapter 7

TRANSFORMATION COUPLING

1 Introduction

The last three chapters were concerned with shifting one-sided stochastic processes, $\theta_t Z = (Z_{t+s})_{0 \leqslant s < \infty}$. This chapter extends the view to an abstract setup where general random elements replace stochastic processes and a semigroup of transformations G replaces the shift-maps θ_t, $0 \leqslant t < \infty$.

As mental preparation for this leap, we start off in Section 2 by observing that the shift-coupling theory of Chapter 5 (Sections 2 through 5) extends from one-sided processes to two-sided processes, $Z = (Z_s)_{-\infty < s < \infty}$. The two-sided case is even easier to deal with, since, while the one-sided shifts do not form a group, the two-sided shifts, $\theta_t Z := (Z_{t+s})_{-\infty < s < \infty}$, do: if we shift the origin to t, we have not lost what happened before time t and can shift back, $\theta_{-t}\theta_t Z = Z$. The same observation applies to random fields with the index set $\mathbb{R}^d$ (processes in d-dimensional time).

The main part of the chapter, Sections 3 through 6, then deals with transformation coupling: the generalization of shift-coupling. In order to stress similarities (and dissimilarities), the treatment parallels that of shift-coupling presented in Sections 2 through 5 of Chapter 5: we use analogous section titles and enumerate the theorems in the same way. Several proofs are more or less replicas of the proofs in the shift-coupling case, but we go through all details again to explicate where the abstract conditions enter. One of these conditions is the existence of an invariant measure (an analogue of the Lebesgue measure), which is essential in this theory and is simply assumed. This semigroup theory applies, for instance, to random fields with index set $[0, \infty)^d$.

In Section 7 we spell out the implications of transformation coupling in the special case when G is a locally compact second countable topological group. This has similarities with the step from one-sided processes to two-sided, but this step is even more pleasant because it hands us the existence of an invariant measure, the Haar measure. Section 8 indicates applications: selfsimilarity, exchangeability, rotational invariance, ...

Section 9 rounds off by considering briefly a possible generalization of exact coupling, taking random fields as a specific example.

2 Shift-Coupling Random Fields

In this section we consider shift-coupling of random fields in d dimensions, highlighting aspects that distinguish this case from the case of one-sided stochastic processes. This is in part a preview of what is to come because for several claims we refer to the theory of transformation coupling to be developed in the subsequent sections.

2.1 Preliminaries

Call a stochastic process (see Section 2 of Chapter 4) with the index set $\mathbb{R}^d (d \geqslant 1)$ a *random field* in d dimensions. Thus, in this terminology, a *two-sided* continuous-time stochastic process is a random field in one dimension. Call $\mathbb{R}^d$ the *site set* and a random element in $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ a *random site*. Let

$$Z = (Z_s)_{s \in \mathbb{R}^d} \quad \text{and} \quad Z' = (Z'_s)_{s \in \mathbb{R}^d}$$

be two shift-measurable random fields with a general state space $(E, \mathcal{E})$ and path space $(H, \mathcal{H})$. Define the shift-maps θ_t, $t \in \mathbb{R}^d$, by

$$\theta_t z = (z_{t+s})_{s \in \mathbb{R}^d}, \quad z \in H.$$

Shift-measurability means that $\theta_t H = H, t \in \mathbb{R}^d$, and that the mapping taking (z, t) in $H \times \mathbb{R}^d$ to $\theta_t z$ in H is $\mathcal{H} \otimes \mathcal{B}(\mathbb{R}^d)/\mathcal{H}$ measurable [that is, the mapping taking (z, t) in $H \times \mathbb{R}^d$ to z_t in E is $\mathcal{H} \otimes \mathcal{B}(\mathbb{R}^d)/\mathcal{E}$ measurable]. Shift-measurability holds in the standard settings [when $(E, \mathcal{E})$ is Polish and the processes right-continuous; see Section 2.8 in Chapter 4]. Unlike what was the case in Chapter 5, we need never impose any restrictions beyond shift-measurability in this section.

2.2 Shift-Coupling – Distributional Shift-Coupling

Say that $(\hat{Z}, \hat{Z}', T, T', C)$ is a (nondistributional) *shift-coupling* of Z and Z' if $(\hat{Z}, \hat{Z}')$ is a coupling of Z and Z', T and T' are random sites, and C is an event such that

$$\theta_T \hat{Z} = \theta_{T'} \hat{Z}' \quad \text{on } C. \tag{2.1}$$

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

$$\mathbf{P}(\theta_T \hat{Z} \in \cdot, C) = \mathbf{P}(\theta_{T'} \hat{Z}' \in \cdot, C'). \qquad [\text{thus } \mathbf{P}(C) = \mathbf{P}(C')] \tag{2.2}$$

The shift-coupling (distributional or not) is *successful* if $\mathbf{P}(C) = 1$. In this case we somtimes leave out the events and only write $(\hat{Z}, \hat{Z}', T)$ or $(\hat{Z}, \hat{Z}', T, T')$ for the shift-coupling.

Observe that for one-sided processes we obtain the shift-couplings of Chapter 5 (Section 2) from (2.1) and (2.2) by taking $C = \{T < \infty\}$ and $C' = \{T' < \infty\}$. In the present case it is no longer natural to use $\{T < \infty\}$ and $\{T' < \infty\}$ for the *shift-coupling events* C and C'.

With these definitions the shift-coupling results in Chapter 5 (Sections 2 through 5) still hold. This can be seen either by repeating the arguments in Chapter 5 with straightforward modifications, or by referring to the abstract group theory in Section 7 below. In fact, the present case is easier to deal with, since [unlike the one-sided shifts in Chapter 5] the shift-maps now form a group: if we shift the origin to t, we have not lost a part of the process and can shift back to the initial origin. In particular, this yields [see Theorem 3.2 below] that now

distributional shift-coupling can always be made nondistributional

without assuming [as we had to do in Theorem 2.2 of Chapter 5] that $(E, \mathcal{E})$ is Polish and the processes right-continuous.

2.3 Shift-Coupled Fields Identical, Only with Different Origins

The group property allows us also to simplify the definition of nondistributional shift-coupling. With

$$S := T - T'$$

definition (2.1) can be rewritten as

$$\theta_S \hat{Z} = \hat{Z}' \quad \text{on } C. \tag{2.3}$$

Thus on C the two random fields are really the *same*, only with different origins.

Call $(\hat{Z}, \hat{Z}', S, C)$ a shift-coupling of Z and Z' with *shift* S if (2.3) holds. The distributional version of this is as follows: call $(\hat{Z}, \hat{Z}', S, C, C')$ a distributional shift-coupling of Z and Z' with *shift* S if

$$\mathbf{P}(\theta_S \hat{Z} \in \cdot, C) = \mathbf{P}(\hat{Z}' \in \cdot, C').$$

In the successful case this becomes

$$\theta_S \hat{Z} \stackrel{D}{=} \hat{Z}',$$

and we can immediately turn the distributional shift-coupling $(\hat{Z}, \hat{Z}', S)$ into the nondistributional shift-coupling $(\hat{Z}, \theta_S \hat{Z}, S)$.

2.4 Shift-Coupling Inequality – Følner Averaging Sets

The shift-coupling inequality and the associated Cesaro total variation convergence over intervals $[0, t]$ in Chapter 5 (Section 3) extend naturally to the sets $[0, t]^d$, but also to $[-t, 0]^d$ and to $[-t, t]^d$. In fact the sets can be quite general.

Let λ be the Lebesgue measure on $\mathcal{B}(\mathbb{R}^d)$ and for all $B \in \mathcal{B}(\mathbb{R}^d)$ such that $0 < \lambda(B) < \infty$, let $\lambda(\cdot|B)$ be the uniform distribution on B, that is,

$$\lambda(A|B) := \lambda(A \cap B)/\lambda(B), \quad A \in \mathcal{B}(\mathbb{R}^d).$$

The following *shift-coupling inequality* holds (see Section 7.3 below): for $B \in \mathcal{B}(\mathbb{R}^d)$ such that $0 < \lambda(B) < \infty$,

$$\|\mathbf{P}(\theta_{U_B} Z \in \cdot) - \mathbf{P}(\theta_{U_B} Z' \in \cdot)\| \leqslant 2 - 2\mathbf{E}[\lambda(S + B|B); C], \tag{2.4}$$

where U_B is uniform on B [that is, U_B has the distribution $\lambda(\cdot|B)$] and independent of Z and Z'.

Thus the Cesaro total variation convergence extends to the following general class of averaging sets. Call a family $B_h \in \mathcal{B}(\mathbb{R}^d)$, $0 < h < \infty$, *Følner averaging sets* if

$$\begin{aligned} &0 < \lambda(B_h) < \infty, \\ &\lambda(t + B_h|B_h) \to 1 \text{ as } h \to \infty, \quad t \in \mathbb{R}^d. \end{aligned} \tag{2.5}$$

When $\mathbf{P}(C) = 1$, we obtain from (2.5) [take $B = B_h$ in (2.4)] the Cesaro total variation convergence

$$\|\mathbf{P}(\theta_{U_{B_h}} Z \in \cdot) - \mathbf{P}(\theta_{U_{B_h}} Z' \in \cdot)\| \to 0, \quad h \to \infty. \tag{2.6}$$

This generalization of the Cesaro total variation convergence is not restricted to random fields with index set $\mathbb{R}^d$. It works for one-sided processes and for random fields with index set $[0, \infty)^d$; see Section 4 below.

2.5 The Sets hB Are Følner

We shall now give a nice example of Følner averaging sets, which shows clearly how general they are.

Theorem 2.1. *If $B \in \mathcal{B}(\mathbb{R}^d)$ and $0 < \lambda(B) < \infty$, then the family*

$$hB := \{hs \in \mathbb{R}^d : s \in B\}, \quad 0 < h < \infty,$$

are Følner averaging sets.

PROOF. Note first that

$$\lambda(hB \cap (t + hB))/\lambda(hB) = \lambda(B \cap (t/h + B))/\lambda(B), \quad t \in \mathbb{R}^d, \tag{2.7}$$

and that [with $\|\cdot\|_2$ denoting the L_2 norm with respect to λ]

$$\begin{aligned}\lambda(B)-\lambda(B\cap(t/h+B)) &= 2^{-1}\int(1_B-1_{t/h+B})^2\,d\lambda \\ &= 2^{-1}(\|1_B-1_{t/h+B}\|_2)^2.\end{aligned} \tag{2.8}$$

Let f_n, $n \geqslant 1$, be a sequence of bounded continuous functions such that

$$\|1_B - f_n\|_2 \to 0, \quad n \to \infty, \qquad \text{[see Ash (1972), Theorem 2.4.14]}$$

to obtain [use $\|1_{t/h+B} - f_n(\cdot - t/h)\|_2 = \|1_B - f_n\|_2$ in the first step]

$$\begin{aligned}\|1_B - 1_{t/h+B}\|_2 &\leqslant 2\|1_B - f_n\|_2 + \|f_n - f_n(\cdot - t/h)\|_2 \\ &\to 2\|1_B - f_n\|_2 \qquad \text{as } h \to \infty \\ &\to 0 \qquad \text{as } n \to \infty.\end{aligned}$$

From this and (2.8) we obtain $\lambda(B) - \lambda(B \cap (t/h + B)) \to 0$ as $h \to \infty$, and a reference to (2.7) completes the proof. □

2.6 The Invariant σ-Algebra – Equivalences

Define the *invariant σ-algebra* by

$$\mathcal{I} = \{A \in \mathcal{H} : \theta_t A = A, t \in \mathbb{R}^d\}.$$

Note that $\mathcal{I}$ also equals $\{A \in \mathcal{H} : \theta_t^{-1} A = A, t \in \mathbb{R}^d\}$, since $\theta_t^{-1} = \theta_{-t}$ for $t \in \mathbb{R}^d$. The following claims are equivalent [see the end of Section 7 below]:

(a) There exists a successful distributional shift-coupling of Z and Z'.

(a') There exists a random site T such that $\theta_T Z \overset{D}{=} Z'$.

(b) For some Følner averaging sets B_h, $0 < h < \infty$, (2.6) holds.

(b') For all Følner averaging sets B_h, $0 < h < \infty$, (2.6) holds.

(c) $\mathbf{P}(Z \in \cdot)_{\mathcal{I}} = \mathbf{P}(Z' \in \cdot)_{\mathcal{I}}$.

When Z' is stationary [that is, $Z' \overset{D}{=} \theta_t Z'$ for all $t \in \mathbb{R}^d$], then (2.6) becomes

$$\theta_{U_{B_h}} Z \overset{tv}{\to} Z', \quad h \to \infty.$$

It follows from the equivalence of (c) and (b) that two stationary random fields agree in distribution on $\mathcal{I}$ if and only if they are identically distributed.

This chapter ends our general treatment of coupling. In the second half of the book (the remaining three chapters) the focus will be on other topics, first stationarity and then regeneration, with coupling entering only as a tool. We therefore conclude at this point with some general comments on coupling.

There are many aspects of coupling that have not been treated here, like the domination coupling in partially ordered Polish spaces (see Section 3 in Chapter 1) and the many ingenious coupling tricks that have been devised in particular models. Also, as the last two subsections have demonstrated, there is much yet to be done along the above lines, both in applying the theory to specific problems and in developing new theory.

Finally, the many coupling equivalences encountered in Chapters 1–7 suggest that the following might be a useful guideline:

WORKING HYPOTHESIS. Any meaningful distributional relation should have a coupling counterpart.