Orbit coupling

Hans-Otto Georgii
Mathematisches Institut der Universität München
Theresienstr. 39, D–80333 München
E–mail: georgii@rz.mathematik.uni–muenchen.de

Abstract

We consider the action of a semigroup S on a standard space E. An orbit coupling of two probability measures on E is a coupling of these measures giving the largest possible weight to the event that the orbits of the two coordinates meet each other. We establish the existence of such orbit couplings for a large class of semigroups S. We also discuss some applications.

Résumé

On considère l'action d' un semigroupe S sur une espace standard E. Un couplage orbital de deux mesures de probabilité sur E est un couplage de ces mesures qui maximise la probabilité de rencontre des orbites des deux coordonnés. On établit l' existence d' un tel couplage orbital pour une grande classe de semigroupes S. On discute aussi quelques applications.

1 Introduction and result

Let $(E, \mathcal{B})$ be a standard measurable space (i.e., $\mathcal{B}$ is the Borel σ–algebra for some Polish topology on E) and μ, ν two probability measures on $(E, \mathcal{B})$. A *coupling* of μ and ν is a probability measure P on $(E \times E, \mathcal{B} \otimes \mathcal{B})$ with marginals μ and ν, which means that $P \circ X^{-1} = \mu$ and $P \circ Y^{-1} = \nu$ for the two projections X, Y of $E \times E$ onto E. Of course, one is primarily interested in couplings with some additional nice properties, and it is well- known that the construction of suitable couplings is a particularly useful technique of modern probability theory. This paper is devoted to the construction of couplings which reflect the ergodic properties of μ and ν with repect to a given class of transformations of $(E, \mathcal{B})$.

To motivate this coupling we consider first the well–known tail coupling of (distributions of) stochastic processes. Suppose $(E, \mathcal{B}) = (F, \mathcal{F})^{\mathbb{N}}$, the countably infinite product of some standard space $(F, \mathcal{F})$. Let $\theta : E \to E$ be the left–shift and $\mathcal{T}$ the tail σ-field. Then there exists a coupling P of μ and ν such that

$$P(\theta^n X = \theta^n Y \text{ for some } n \in \mathbb{N}) = 1 - \tfrac{1}{2}\|\mu - \nu\|_{\mathcal{T}}. \tag{1.1}$$

Here and below, we use the notation

$$\|\mu - \nu\|_{\mathcal{A}} = 2 \sup_{A \in \mathcal{A}} |\mu(A) - \nu(A)|$$

for the total variation distance of two probability measures which are defined on, or restricted to, a σ–algebra $\mathcal{A}$. A detailed account of tail couplings can be found in [8]. (In traditional terminology, tail couplings P as in (1.1) are referred to as *maximal couplings*. This is because each coupling P of μ and ν satisfies (1.1) with "$\leq$" instead of "$=$". In our context, however, this terminology is misleading: we need to distinguish between couplings of different nature rather than to stress the maximality which is a common feature of all couplings considered here.) An existence proof for tail couplings will also come out in Section 3 below.

An analogous result involving the ergodic rather than the tail behaviour of μ and ν was recently obtained by Aldous and Thorisson [1] and Thorisson [12]: Let $S = \mathbb{N}$ or $\mathbb{Z}$, $(E, \mathcal{B}) = (F, \mathcal{F})^S$ for some standard space $(F, \mathcal{F})$, θ the left–shift of E, and $\mathcal{I}$ the σ-algebra of all θ–invariant events in $\mathcal{B}$. Then any two probability measures μ, ν on $(E, \mathcal{B})$ admit a coupling P such that

$$P(\theta^m X = \theta^n Y \ \text{ for some } \ m, n \in S) = 1 - \tfrac{1}{2}\|\mu - \nu\|_{\mathcal{I}}. \tag{1.2}$$

Aldous and Thorisson call such a coupling a *shift coupling*. For actions of $\mathbb{R}_+$, a similar result was derived in [11]. It is the purpose of this note to establish the existence of such couplings for actions of more general semigroups or groups. (After completion of this paper, I learned of Thorisson's independent work [13] which deals with actions of locally compact second countable groups. This should be read as a companion paper.)

Let $(S, \mathcal{S})$ be a standard measurable semigroup, that is, a standard measurable space with an associative measurable multiplication. Next, let $(E, \mathcal{B})$ be any standard measurable space and $T : S \times E \to E, (s, x) \to T_s x$, a measurable left action of S on E, i.e., an $\mathcal{S} \otimes \mathcal{B}$–$\mathcal{B}$–measurable mapping such that $T_{st}x = T_s T_t x$ for all $s, t \in S$ and $x \in E$. If S has an identity e, we also require that $T_e x = x$ for all $x \in E$. Finally, we let $\mathcal{I} = \mathcal{I}(S)$ be the σ-algebra of all invariant events, viz.

$$\mathcal{I} = \{A \in \mathcal{B} : T_s^{-1} A = A \ \text{ for all } \ s \in S\}. \tag{1.3}$$

Definition. An *orbit coupling* of two probability measures μ, ν on $(E, \mathcal{B})$ is a coupling P of μ and ν such that

$$P(T_s X = T_t Y \ \text{ for suitable } \ s, t \in S) = 1 - \tfrac{1}{2}\|\mu - \nu\|_{\mathcal{I}}. \tag{1.4}$$

Note that the coupling event

$$C = \{T_s X = T_t Y \ \text{ for some } \ s, t \in S\} \tag{1.5}$$

on the left-hand side of (1.4) can be rewritten as $\{\mathcal{O}(X) \cap \mathcal{O}(Y) \neq \emptyset\}$, where $\mathcal{O}(x) = \{T_s x : s \in S\}$ is the orbit of $x \in E$. In the case when S is a group, this is equal to $\{\mathcal{O}(X) = \mathcal{O}(Y)\}$. This explains our choice of terminology.

Remarks. (a) The coupling event C in (1.5) is measurable whenever S is countable. This is because — $(E, \mathcal{B})$ being standard — the diagonal Δ in $E \times E$ is measurable, and C is the (then countable) union of the sets $(T_s \times T_t)^{-1}(\Delta) \in \mathcal{B} \otimes \mathcal{B}$. In the general case, C is the projection onto $E \times E$ of the measurable set $(T \times T)^{-1}(\Delta)$ in $(S \times E) \times (S \times E)$. This shows that C is analytic and therefore universally measurable, in that $C \in (\mathcal{B} \otimes \mathcal{B})_*$, the universal completion of $\mathcal{B} \otimes \mathcal{B}$; see Section 8.4 of Cohn [3]. Moreover, the cross section theorem (Theorem 8.5.3 of [3]) implies the existence of two $(\mathcal{B} \otimes \mathcal{B})_*$–$\mathcal{S}$ –measurable mappings $\sigma, \tau : E \times E \to S$ such that

$$C = \{T_\sigma X = T_\tau Y\} \equiv \{(x, y) \in E \times E : T_{\sigma(x,y)}x = T_{\tau(x,y)}y\}. \tag{1.6}$$

This representation of C is particularly convenient.

(b) For an arbitrary coupling P of μ and ν and all $A \in \mathcal{I}$ we have $1_A \circ X = 1_A \circ T_\sigma X = 1_A \circ T_\tau Y = 1_A \circ Y$ on C and therefore

$$|\mu(A) - \nu(A)| \leq \int_{C^c} |1_A \circ X - 1_A \circ Y| \, dP \leq P(C^c).$$

This gives (1.4) with "$\leq$" instead of "$=$" and reveals the maximality of orbit couplings. However, as explained above in connection with (1.1), we prefer to skip the epithet "maximal" to avoid confusion.

(c) In addition to the apparent similarity of (1.1) and (1.4), there is also a deeper connection between tail couplings and orbit couplings. Suppose F is a finite set and $E = F^{\mathbb{N}}$. Let $\Gamma = \mathrm{Perm}(F)^{\mathbb{N}}$ be the countable group of all transformations of E which act coordinatewise on finitely many coordinates by permutations of F. Then $\mathcal{T} = \mathcal{I}(\Gamma)$, and the orbit coupling event C in (1.5) with $S = \Gamma$ is equal to the tail coupling event on the left-hand side of (1.1). The existence of a tail coupling (for finite F) is therefore a special case of our theorem below on the existence of orbit couplings. Conversely, as pointed out to me by B. Weiss, a celebrated theorem of Dye (and others, see [6]) asserts that every measure preserving action of a countable amenable group on a standard probability space is orbit equivalent to the action of Γ on E (which preserves the Bernoulli measure). That is, there exists a measure space isomorphism which maps orbits onto orbits. As a consequence, the existence of orbit couplings for actions of countable amenable groups can be deduced from the existence of tail couplings, at least when μ and ν admit a dominating invariant probability measure. As a matter of fact, our construction of orbit couplings for countable S completely parallels the construction of tail couplings.

(d) For the trivial action of the identity group $S = \{e\}$, (1.4) takes the form

$$P(X = Y) = 1 - \tfrac{1}{2}\|\mu - \nu\|_{\mathcal{B}}. \tag{1.7}$$

The existence of such a "diagonal coupling" (called γ–coupling in [8]) is well-known and easy to derive by putting $\mu \wedge \nu$ on the diagonal Δ and adding a suitable multiple of $(\mu - \mu \wedge \nu) \otimes (\nu - \mu \wedge \nu)$.

(e) In the case when S acts transitively on E, in that $\mathcal{O}(x) = E$ for all $x \in E$, we have $\mathcal{I} = \{E, \emptyset\}$ and $C = E \times E$, so that every coupling is an orbit coupling. More generally, it is easy to see that the existence of orbit couplings is trivial whenever E splits into at most countably many orbits. $\square$

Our main result will state that an orbit coupling exists whenever S is either a countable normal semigroup, or a compact metric group, or when S is composed of finitely many such building blocks.

To describe this in detail we need the counterparts for semigroups of some basic concepts for groups. Let $(S, \mathcal{S})$ be a measurable semigroup. A subsemigroup R of S is called *normal* (in S) if $sR = Rs$ for all $s \in S$. In particular, S is called *normal* if it is normal in itself. [Any group and any abelian semigroup is normal.] A normal subsemigroup R of S defines an equivalence relation $\sim_R$ on S by $s \sim_R s'$ iff $sR \cap s'R \neq \emptyset$. The quotient $S/R \equiv S/\sim_R$ with its natural multiplication is then a semigroup, the *factor semigroup.*

Definition. Suppose $(S, \mathcal{S})$ is a measurable semigroup and $R \in \mathcal{S}$ a normal subsemigroup. We will say

(a) S is a *countable extension* of R if S/R is countable and normal; and
(b) S is a *compact group extension* of R if S/R is a group and, for a suitable topology on S/R with Borel σ–algebra $\mathcal{B}(S/R)$,
    (i) S/R is a compact metrizable topological group, and
    (ii) the canonical projection $\gamma : S \to S/R$ admits a measurable section, that is, a $\mathcal{B}(S/R)$–$\mathcal{S}$–measurable mapping $\phi : S/R \to S$ with $\gamma \circ \phi = \text{id}$.

[Sufficient conditions for the existence of a measurable section can be found in Exercise 4 on p. 287 of [3], for example. These conditions are satisfied if S is Polish, the $\sim_R$ equivalence classes are closed, and γ is open. In specific examples, the existence of a measurable section can often be seen directly.]

We write $\mathbf{S}$ for the class of all standard measurable semigroups $(S, \mathcal{S})$ which admit a finite increasing sequence $R_1 \subset \ldots \subset R_n = S$ of measurable subsemigroups such that, for all $1 \leq k < n$, R_{k+1} is either a countable extension or a compact group extension of R_k. $\mathbf{S}$ includes

(1) the countable semigroups resp. groups $\mathbb{N}^d$, $\mathbb{Z}^d$ and S_∞ (the group of all finite permutations of $\mathbb{N}$);
(2) the compact metric groups $\mathbb{R}^d/\mathbb{Z}^d$, $\text{SO}(d)$, and $\{-1, 1\}^{\mathbb{N}}$; and therefore
(3) $\mathbb{R}^d$ and $\mathbb{R}^d_+$ (because $\mathbb{R}^d_+/\mathbb{N}^d$ is isomorphic to $\mathbb{R}^d/\mathbb{Z}^d$);
(4) the semidirect products $\mathbb{R}^d \odot \text{SO}(d)$ (the group of Euclidean motions) and $\{-1, 1\}^{\mathbb{N}} \odot \mathsf{S}_\infty$; and so on.

Here is our main result.

Theorem. For any (left) action of a semigroup $S \in \mathbf{S}$ on a standard space $(E, \mathcal{B})$, any two probability measures on $(E, \mathcal{B})$ admit an orbit coupling.

The proof of this theorem will be given in Section 3. The next section is devoted to a few elementary applications which illustrate the possible use of orbit couplings.

2 Some applications

In the setting of (1.1), it is well-known that $\mu = \nu$ on $\mathcal{T}$ if and only if $\|\mu \circ \theta^{-n} - \nu \circ \theta^{-n}\|_{\mathcal{B}} \to 0$ as $n \to \infty$. Our first goal in this section is an analogue of this statement which refers to the ergodic rather than the tail behaviour. We need some definitions.

Consider a standard measurable semigroup $(S, \mathcal{S})$. The right translation of S by some $t \in S$ is defined by $r_t : s \to st$. A net $(\rho_i)_{i \in D}$ of probability measures on $(S, \mathcal{S})$ is called *right–ergodic* if $\lim_{i\in D}\|\rho_i\circ r_t^{-1}-\rho_i\|_{\mathcal{S}}=0$ for all $t\in S$. Left translations and left–ergodic nets are defined similarly. The existence of a right–ergodic net of probability measures on $(S,\mathcal{S})$ is equivalent to the right–amenability of S, which means that there exists a right–invariant mean on S [4]. If S is a locally compact group with a countable basis, this is further equivalent to the existence of a right–ergodic sequence of the form $\rho_n=m(\cdot\,|\Lambda_n)$, where m is right Haar measure and (Λ_n) a Følner sequence of compact sets. See Section 3 of the Appendix of [10] for references and further discussion. The equivalence (b) $\Leftrightarrow$ (c) in the following corollary is an extension of a result of Berbee [2] for $S=\mathbb{N}$; see also [1, 12, 13].

Corollary 2.1 For any action of a semigroup $S\in\mathbf{S}$ on a standard space $(E,\mathcal{B})$ and any two probability measures μ,ν on $(E,\mathcal{B})$, the following statements are equivalent:

(a) $\mu=\nu$ on $\mathcal{I}=\mathcal{I}(S)$

(b) There exists a successful orbit coupling of μ and ν, i.e., a coupling P with $P(T_\sigma X=T_\tau Y)=1$.

If S is right–amenable, these statements are equivalent to

(c) For any right–ergodic net $(\rho_i)_{i\in D}$ on $(S,\mathcal{S})$,

$$\lim_{i\in D}\|\rho_i(\mu)-\rho_i(\nu)\|_{\mathcal{B}}=0,$$

where $\rho_i(\mu)=\int\rho_i(ds)\mu\circ T_s^{-1}$.

Proof. (a) $\Rightarrow$ (b): This is obvious from the theorem.

(b) $\Rightarrow$ (a): This follows from Remark (b) in Section 1.

(b) $\Rightarrow$ (c): Let P be a successful orbit coupling of μ and ν. Then for all $B\in\mathcal{B}$ and $i\in D$ we have

$$\begin{aligned}
|\rho_i(\mu)(B)-\rho_i(\nu)(B)| &\le \int dP\left|\int\rho_i(ds)(1_B\circ T_sX-1_B\circ T_sY)\right|\\
&\le \int dP\Big[\Big|\int\rho_i(ds)(1_B\circ T_sX-1_B\circ T_{s\sigma}X)\Big|\\
&\qquad +\Big|\int\rho_i(ds)(1_B\circ T_{s\tau}Y-1_B\circ T_sY)\Big|\Big]\\
&= \int dP\left[|(\rho_i-\rho_i\circ r_\sigma^{-1})(\Lambda_X)|+|(\rho_i\circ r_\tau^{-1}-\rho_i)(\Lambda_Y)|\right],
\end{aligned}$$

where $\Lambda_x=\{s\in S:T_sx\in B\}\in\mathcal{S}$ for $x\in E$. In the second step we have used that $T_{s\sigma}X=T_sT_\sigma X=T_{s\tau}Y$ for all $s\in S$ P–almost surely. Taking the supremum over $B\in\mathcal{B}$ in the preceding estimate we find

$$\|\rho_i(\mu)-\rho_i(\nu)\|_{\mathcal{B}}\le\int dP\Big[\|\rho_i-\rho_i\circ r_\sigma^{-1}\|_{\mathcal{S}}+\|\rho_i\circ r_\tau^{-1}-\rho_i\|_{\mathcal{S}}\Big].$$

The result thus follows from the dominated convergence theorem.

(c) $\Rightarrow$ (a): This is obvious. □

Assertion (c) above is dual to a mean ergodic theorem, as is specified in the following example.

Example 2.2 Suppose $S \in \mathbf{S}$ is an amenable group, $(\lambda_i)_{i \in D}$ a left–ergodic net and T a measure preserving left action of S on a standard probability space $(E, \mathcal{B}, \mu)$. Then for all $f \in L^1(\mu)$,

$$\lim_{i \in D} \int d\mu \left| \int \lambda_i(ds) f \circ T_s - \bar{f} \right| = 0, \tag{2.1}$$

where $\bar{f}$ is the conditional expectation of f relative to $\mathcal{I}$ and μ. (We refer to [10] for a general account of this and other ergodic theorems.)

To verify this we can assume without loss that $f \geq 0$ and $\int f \, d\mu = 1$. Then $f\mu$ and $\bar{f}\mu$ are probability measures that agree on $\mathcal{I}$, and $\bar{f}\mu$ is invariant. Corollary 2.1 therefore implies that

$$\lim_{i \in D} \|\rho_i(f\mu) - \bar{f}\mu\|_{\mathcal{B}} = 0, \tag{2.2}$$

where ρ_i is the image of λ_i under the inversion $s \to s^{-1}$. But $\rho_i(f\mu)$ has the μ–density $\int \lambda_i(ds) f \circ T_s$, so that the left–hand sides of (2.1) and (2.2) coincide.

In the non–invertible case when S is only a right–amenable semigroup, $\rho_i(f\mu)$ has the μ–density $\int \rho_i(ds) P_s f$, where P_s is the Perron–Frobenius operator corresponding to T_s. So, in this case we obtain that for any right–ergodic net (ρ_i), $\int \rho_i(ds) P_s f$ converges to $\bar{f}$ in $L^1(\mu)$–norm. $\square$

We continue with two further elementary examples which illustrate the use of successful orbit couplings. A less immediate application will appear in a separate paper [5].

Example 2.3 Let T be a measure preserving transformation of a standard probability space $(E, \mathcal{B}, \mu)$, $A \in \mathcal{B}$ a set of positive measure, and T_A the induced transformation on A. It is well–known that the conditional probability $\mu_A = \mu(\cdot\,|A)$ is T_A–invariant. It is also well–known that T_A is ergodic whenever T is ergodic. Let us derive this fact using an orbit coupling.

Suppose $B \subset A$ with $\mu(B) > 0$ is T_A–invariant. Since μ is T–ergodic, μ_A and μ_B coincide on $\mathcal{I}(T)$, so that there exists a successful orbit coupling P of μ_A and μ_B relative to T. In fact, P is even a successful orbit coupling relative to T_A. For, if $T^m X = T^n Y$ for suitable $m, n \in \mathbb{N}$ then $T^{m+\ell} X = T^{n+\ell} Y$ for all $\ell \geq 0$. Since $X, Y \in A$ with P–probability 1, by Poincaré recurrence we can find an ℓ such that $m + \ell$ is a return time of X to A, and thus $n + \ell$ is a return time of Y to A. Hence $T_A^j X = T_A^k Y$ for suitable $j, k \in \mathbb{N}$ with P–probability 1. By Corollary 2.1, $\mu_A = \mu_B$ on $\mathcal{I}(T_A)$ and therefore $\mu_A(B) = \mu_B(B) = 1$. This proves the ergodicity of T_A.

An analogous argument works in a continuous time setting. For example, it can be used to show that a flow under an integrable function (cf. [9], p.11) is ergodic if and only if the underlying discrete dynamical system is ergodic. In particular, a stationary marked point process on $\mathbb{R}$ of finite intensity is ergodic if and only if, under the associated normalized Palm measure, the stationary sequence of point marks and spacings is ergodic. We leave this to the reader. Further applications of orbit couplings to point processes on $\mathbb{R}^d$ can be found in [12, 14]. $\square$

Example 2.4 Let G be a countable abelian (additive) group and p a probability measure on G which is strongly aperiodic, in that G is generated by the set $\{u - v : p(u)p(v) > 0\}$. A variant of an argument of Ornstein (see [7], pp. 68–70, or [5]) then shows that for each $z \in G$ there exist two processes $(Z_n)_{n\geq 0}$ and $(Z'_n)_{n\geq 0}$ on a common probability space such that

(i) $(Z_n)_{n\geq 0}$ and $(Z'_n)_{n\geq 0}$ are random walks on G with jump distribution p and start at 0;

(ii) $Z'_n = Z_n + z$ eventually with probability 1.

In fact, one can achieve that each increment of each random walk is independent from the previous increments of the other walk.

Here we will show how one can use an orbit coupling to construct such a pair of processes. This construction, however, does not produce the additional independence property.

Let $E = G^{\mathbb{N}}$ (with its natural σ–algebra) and $S = \mathsf{S}_\infty$ the countable group of all finite permutations of $\mathbb{N}$, i.e., of all bijections $s : \mathbb{N} \to \mathbb{N}$ with $s(n) = n$ eventually. S acts on E in a natural way by interchanges of coordinates. Let $\mu = p^{\mathbb{N}}$ be the product measure on E. By the Hewitt–Savage zero–one law, μ is trivial on $\mathcal{I}(S)$. Since p is strongly aperiodic, we can find $z_1, \ldots, z_k, z'_1, \ldots, z'_k \in G$ with $p(z_j)p(z'_j) > 0$ for all $1 \leq j \leq k$ and $z = \sum_{j=1}^k (z_j - z'_j)$. Let $A = [z_1, \ldots, z_k]$ and $A' = [z'_1, \ldots, z'_k]$ be the cylinder events in E which fix the first k coordinates as indicated. The conditional probabilities μ_A and $\mu_{A'}$ are then well–defined and agree on $\mathcal{I}(S)$. By the theorem, they admit a successful S–orbit coupling P. Under P, the shifted processes $\theta^k X = (X_j)_{j>k}$ and $\theta^k Y = (Y_j)_{j>k}$ have distribution μ. In other words, the sum processes $Z_n = \sum_{j=1}^n X_{k+j}$ and $Z'_n = \sum_{j=1}^n Y_{k+j}$ satisfy property (i). On the other hand, we know that, with P–probability 1, the sequence $(z_1, \ldots, z_k, X_{k+1}, X_{k+2}, \ldots)$ is a finite permutation of $(z'_1, \ldots, z'_k, Y_{k+1}, Y_{k+2}, \ldots)$, so that $\sum_{j=1}^k z_j + Z_n = \sum_{j=1}^k z'_j + Z'_n$ for sufficiently large n. This gives (ii).

The construction above can easily be extended to give the following result: For any standard measurable abelian group $(G, \mathcal{G})$, any probability measure p on $(G, \mathcal{G})$, and any $U, U' \in \mathcal{G}$ such that $p^{*k}(U)p^{*k}(U') > 0$ for some $k \geq 1$, there exists a pair of processes (Z_n) and (Z'_n) satisfying (i) and

(ii') $Z_n - Z'_n = Z_{n-1} - Z'_{n-1} \in U - U'$ eventually, with probability 1.

One may use this coupling for a proof of the Choquet–Deny theorem and of Blackwell's renewal theorem, for example. □

3 Proof

The theorem consists of two parts which correspond to countable extensions and compact group extensions. These two parts are stated explicitly in Propositions 3.1 and 3.5 below. For convenience we will call a measurable semigroup $(S, \mathcal{S})$ an *orbit coupling semigroup* if for any left action of S on a standard measurable space, any two probability measures on this space admit an orbit coupling.

Proposition 3.1 Suppose $(S, \mathcal{S})$ is a standard measurable semigroup and $R \in \mathcal{S}$ a normal subsemigroup such that S/R is countable and normal. If R is an orbit coupling semigroup, then so is S.

This proposition implies in particular that any countable normal semigroup S is an orbit coupling semigroup. This follows from Remark (d) in Section 1 because we can assume without loss that S has an identity.

Turning to the proof we let S and R be as in the hypothesis and assume that R is an orbit coupling semigroup. We then have the following lemma.

Lemma 3.2 Let $(E, \mathcal{B})$ and $(E', \mathcal{B}')$ be two standard spaces, $\varphi, \psi : E \to E'$ two measurable mappings, T an action of R on E', and μ, ν two finite measures on E of equal mass. Then there exists a coupling P of μ and ν such that

$$2\, P(T_r \circ \varphi \circ X \neq T_s \circ \psi \circ Y \text{ for all } r, s \in R) = \|\mu \circ \varphi^{-1} - \nu \circ \psi^{-1}\|_{\mathcal{I}(R)}. \tag{3.1}$$

Proof. Let us note first that the event on the left–hand side is universally measurable. This follows from Remark (a) and Lemma 8.4.6 of [3].

Next we assume without loss that μ and ν have mass 1. Let $\mu' = \mu \circ \varphi^{-1}$, $\nu' = \nu \circ \psi^{-1}$ and P' an R–orbit coupling of μ' and ν'. Define

$$P = \int P'(dx', dy')\ \mu^{x'} \otimes \nu^{y'},$$

where $\mu^{x'}$ and $\nu^{y'}$ are regular versions of the conditional probabilities $\mu(\cdot\,|\varphi = x')$ and $\nu(\cdot\,|\psi = y')$, respectively. P is clearly a coupling of μ and ν. Its image under $\varphi \times \psi$ is P' because $\mu^{x'}(\varphi^{-1}A') = 1_{A'}(x')$ for μ'–almost all x' and all $A' \in \mathcal{B}'$, and therefore $P(\varphi^{-1}A' \times \psi^{-1}B') = P'(A' \times B')$ for all $A', B' \in \mathcal{B}'$. The conclusion is thus equivalent to the R–orbit coupling identity (1.4) for P'. □

Remark 3.3 Let D denote the decoupling event on the left–hand side of (3.1), and let P be as in the preceding lemma. Then

$$P(D^c \cap \{\varphi \circ X \in \cdot\,\}) = P(D^c \cap \{\psi \circ Y \in \cdot\,\}) \text{ on } \mathcal{I}(R)$$

and therefore

$$\|\mu \circ \varphi^{-1} - \nu \circ \psi^{-1}\|_{\mathcal{I}(R)} = \|P(D \cap \{\varphi \circ X \in \cdot\,\}) - P(D \cap \{\psi \circ Y \in \cdot\,\}\|_{\mathcal{I}(R)}.$$

Equation (3.1) is thus equivalent to the statement that the measures $P(D \cap \{\varphi \circ X \in \cdot\,\})$ and $P(D \cap \{\psi \circ Y \in \cdot\,\})$ are mutually singular on $\mathcal{I}(R)$. □

To show that S is an orbit coupling semigroup we fix a standard space $(E, \mathcal{B})$ and two probability measures μ and ν on $(E, \mathcal{B})$. We need to construct an S–orbit coupling of μ and ν. This construction, which consists in an iterated application of Lemma 3.2, is the common core of the construction of orbit couplings and tail couplings. Under different wrappings, it appears both in Goldstein's proof of the tail coupling (1.1) (see [8]), and in Aldous and Thorisson's proof of the shift coupling (1.2) [1, 12]. To make this evident we consider a generalized setting as follows.

Suppose we are given two sequences φ_n, ψ_n of measurable mappings from $(E, \mathcal{B})$ into itself, and $(E, \mathcal{B})$ is equipped with an action of R. Consider the coupling events

$$C_n = \{T_r \circ \varphi_n \circ X = T_s \circ \psi_n \circ Y \text{ for some } r, s \in R\} \tag{3.2}$$

in $E \times E$ and the decoupling events $D_n = C_n^c$, and define $C = \bigcup_{n \geq 1} C_n$ and $D = C^c = \bigcap_{n \geq 1} D_n$

Recursive construction of a coupling P with coupling event C. Let P_1 be a coupling of μ and ν such that

$$P_1(D_1 \cap \{\varphi_1 \circ X \in \cdot\}) \perp P_1(D_1 \cap \{\psi_1 \circ Y \in \cdot\}) \text{ on } \mathcal{I}(R).$$

Such a P_1 exists by Lemma 3.2 and Remark 3.3. If P_n is already defined, we let μ_{n+1} and ν_{n+1} be the two marginals of $1_{D_n} P_n$ and use Lemma 3.2 to find a coupling P_{n+1} of μ_{n+1} and ν_{n+1} such that

$$P_{n+1}(D_{n+1} \cap \{\varphi_{n+1} \circ X \in \cdot\}) \perp P_{n+1}(D_{n+1} \cap \{\psi_{n+1} \circ Y \in \cdot\}) \text{ on } \mathcal{I}(R). \tag{3.3}$$

Then we can write

$$\begin{aligned} \mu &= (1_{C_1} P_1) \circ X^{-1} + \mu_2 \\ &= (1_{C_1} P_1) \circ X^{-1} + (1_{C_2} P_2) \circ X^{-1} + \mu_3 = \dots \\ &= \sum_{k=1}^{n} (1_{C_k} P_k) \circ X^{-1} + \mu_{n+1} \end{aligned} \tag{3.4}$$

so that

$$\sum_{k \geq 1} (1_{C_k} P_k) \circ X^{-1} \leq \mu$$

and similarly

$$\sum_{k \geq 1} (1_{C_k} P_k) \circ Y^{-1} \leq \nu.$$

We can therefore pick a measure P_∞ on $E \times E$ with marginals $\mu - \sum_{k \geq 1} (1_{C_k} P_k) \circ X^{-1}$ and $\nu - \sum_{k \geq 1} (1_{C_k} P_k) \circ Y^{-1}$. Setting $C_\infty = E \times E$, we define

$$P = \sum_{1 \leq k \leq \infty} 1_{C_k} P_k \,.$$

It is then evident that P is a coupling of μ and ν. The essential feature of this coupling is that

$$P(D \cap \{\varphi_n \circ X \in \cdot\}) \perp P(D \cap \{\psi_n \circ Y \in \cdot\}) \text{ on } \mathcal{I}(R) \text{ for all } n \geq 1. \tag{3.5}$$

To verify this property we note that for all $n \geq 1$

$$\begin{aligned} (1_D P) \circ X^{-1} &= \mu - (1_C P) \circ X^{-1} \\ &= \sum_{k=1}^{n} (1_{C_k} P_k) \circ X^{-1} + \mu_{n+1} - (1_C P) \circ X^{-1} \\ &\leq \mu_{n+1} = (1_{D_n} P_n) \circ X^{-1} \,. \end{aligned} \tag{3.6}$$

The second equality comes from (3.4), and the inequality from the definition of P. Similarly, $(1_D P) \circ Y^{-1} \leq (1_{D_n} P_n) \circ Y^{-1}$. Assertion (3.5) thus follows from (3.3).

When specialized to the cases of orbit couplings and tail couplings, assertion (3.5) implies the required maximality of the coupling P. Let us demonstrate this first in the case of tail couplings.

Let $(E, \mathcal{B}) = (F, \mathcal{F})^{\mathbb{N}}$ for some standard space $(F, \mathcal{F})$, $\varphi_n = \psi_n = \theta^n$ the n'th iterate of the shift, and $R = \{e\}$ the trivial group. The coupling event C defined below (3.2) is then equal to the tail coupling event on the left–hand side of (1.1), and property (3.5) means that $P(D \cap \{X \in \cdot\}) \perp P(D \cap \{Y \in \cdot\})$ on $\theta^{-n}\mathcal{B}$ for all n, and therefore on $\mathcal{T}$. An analogue of Remark 3.3 thus implies (1.1). In fact, an obvious refinement of (3.6) even shows that for each n

$$2\, P(\theta^n X \neq \theta^n Y) = \|\mu - \nu\|_{\theta^{-n}\mathcal{B}} .$$

This is the usual finite version of (1.1).

We now return to the proof of Proposition 3.1. Let T be an action of S on $(E, \mathcal{B})$, and suppose S/R is countable and normal. Then we can find a countable section $\Sigma \subset S$ (i.e., a countable set of representatives for the equivalence classes in S/R) and an enumeration $n \to (s(n), t(n))$ of $\Sigma \times \Sigma$. We set $\varphi_n = T_{s(n)}$, $\psi_n = T_{t(n)}$ and let P be constructed as above relative to these mappings. Then it is easily checked that the set C defined below (3.2) is given by

$$\begin{aligned} C &= \Big\{\exists\, n \in \mathbb{N}\ \exists\, q, r \in R : T_{qs(n)}X = T_{rt(n)}Y\Big\} \\ &= \Big\{\exists\, s, t \in S : T_s X = T_t Y\Big\}; \end{aligned}$$

that is, C is the S–coupling event (1.5). On the other hand, property (3.5) of P can be exploited as follows.

Lemma 3.4 Under the conditions above, $(1_D P) \circ X^{-1} \perp (1_D P) \circ Y^{-1}$ on $\mathcal{I}(S)$.

Proof. For brevity we write $\pi = (1_D P) \circ X^{-1}$ and $\rho = (1_D P) \circ Y^{-1}$. Assertion (3.5) then reads

$$\pi \circ T_s^{-1} \perp \rho \circ T_t^{-1} \ \text{ on } \mathcal{I}(R) \ \text{ for all } \ s, t \in \Sigma. \tag{3.7}$$

In a first step, we construct a set $B \in \mathcal{I}(R)$ such that $\pi \circ T_s^{-1}(B) = 0$ and $\rho \circ T_t^{-1}(B^c) = 0$ for all $s, t \in \Sigma$. Let $s \to n(s)$ be a bijection from Σ to $\mathbb{N}$ and

$$\lambda = \sum_{s \in \Sigma} 2^{-n(s)} \left[\pi \circ T_s^{-1} + \rho \circ T_s^{-1}\right] .$$

λ is finite, and for any $s \in \Sigma$ there exist $\mathcal{I}(R)$–measurable functions f_s, $g_s \geq 0$ such that $\pi \circ T_s^{-1} = f_s \lambda$ on $\mathcal{I}(R)$ and $\rho \circ T_s^{-1} = g_s \lambda$ on $\mathcal{I}(R)$. By (3.7), $f_s \wedge g_t = 0$ λ–almost surely for all $s, t \in \Sigma$. We define $B = \bigcap_{s \in \Sigma}\{f_s = 0\}$ and $B' = \bigcap_{t \in \Sigma}\{g_t = 0\}$. Then $\pi \circ T_s^{-1}(B) = 0$ for all $s \in \Sigma$. Also, for all $t \in \Sigma$ we have $\rho \circ T_t^{-1}(B') = 0$,

$$\lambda(B^c \setminus B') = \lambda\Big(\bigcup_{s,t \in \Sigma} \{f_s \wedge g_t > 0\}\Big) = 0$$

and therefore $\rho \circ T_t^{-1}(B^c) = 0$.

We will now use the set B to construct a set $A \in \mathcal{I}(S)$ which separates π and ρ. Since $B \in \mathcal{I}(R)$, the mapping $\beta : t \to T_t^{-1} B$ from S to $\mathcal{I}(R)$ is constant on the $\sim_R$

equivalence classes and can therefore be considered as a map from the countable normal semigroup $Q \equiv S/R$ to $\mathcal{I}(R)$. We thus can define

$$A = \bigcup_{a \in Q} \bigcap_{b \in aQ} \beta(b) \ .$$

Then $A \in \mathcal{I}(R)$, so that also the map $\alpha : t \to T_t^{-1} A$ on S can be viewed as a mapping on Q. For all $c \in Q$ we have

$$\alpha(c) = \bigcup_{a \in Q} \bigcap_{b \in aQ} \beta(bc) = \bigcup_{a \in Q} \bigcap_{b \in aQc} \beta(b).$$

Since $Qc \subset Q$, $aQc \subset aQ$ and therefore $\alpha(c) \supset A$. On the other hand, since Q is normal we have $aQc = Qac$, whence

$$\alpha(c) = \bigcup_{a \in Qc} \bigcap_{b \in Qa} \beta(b) \subset A.$$

Hence $\alpha(c) = A$ for all $c \in Q$, which means that $A \in \mathcal{I}(S)$. Finally, it is evident that

$$A \subset \bigcup_{a \in Q} \beta(a) \ \text{ and } \ A^c \subset \bigcup_{b \in Q} \beta(b)^c,$$

and from the first step we know that $\pi(\beta(a)) = 0$ and $\rho(\beta(b)^c) = 0$ for all $a, b \in Q$. This shows that $\pi(A) = 0$ and $\rho(A^c) = 0$. $\Box$

To complete the proof of Proposition 3.1 we now only need to note that eq. (1.4) follows from Lemma 3.4 by an analogue of Remark 3.3.

We now turn to the proof of the second half of the theorem, which can be stated as follows.

Proposition 3.5 Let $(S, \mathcal{S})$ be a standard measurable semigroup, and suppose $(S, \mathcal{S})$ is a compact group extension of a normal subsemigroup $R \in \mathcal{S}$. If R is an orbit coupling semigroup, then so is S.

Together with Remark (d), the proposition shows in particular that each compact metric group is an orbit coupling group.

Let S and R be as in the hypothesis of the proposition, T a left action of S on a standard measurable space $(E, \mathcal{B})$, and μ, ν two probability measures on $(E, \mathcal{B})$. By assumption, the factor semigroup $Q \equiv S/R$ is a compact metric group, which gives us the existence of the normalized Haar measure m on Q. We consider the probability measures $\tilde{\mu} = m \otimes \mu$ and $\tilde{\nu} = m \otimes \nu$ on the standard space $(\tilde{E}, \tilde{\mathcal{B}}) \equiv (Q \times E, \mathcal{B}(Q) \otimes \mathcal{B})$. Writing $\phi : Q \to S$ for the measurable section which exists by assumption, we also consider the measurable mapping $\Phi : (a, x) \to T_{\phi(a)} x$ from $(\tilde{E}, \tilde{\mathcal{B}})$ to $(E, \mathcal{B})$. By Lemma 3.2, there exists a coupling $\tilde{P}$ of $\tilde{\mu}$ and $\tilde{\nu}$ such that

$$\tilde{P}\Big(T_r \circ \Phi \circ \tilde{X} = T_s \circ \Phi \circ \tilde{Y} \ \text{ for some } \ r, s \in R\Big) = 1 - \tfrac{1}{2} \|\tilde{\mu} \circ \Phi^{-1} - \tilde{\nu} \circ \Phi^{-1}\|_{\mathcal{I}(R)} \ . \quad (3.8)$$

In the above, $\tilde{X}, \tilde{Y}$ are the two projections from $\tilde{E} \times \tilde{E}$ onto $\tilde{E}$. Finally, we introduce the projection ξ from $\tilde{E} \times \tilde{E}$ onto $E \times E$ and define $P = \tilde{P} \circ \xi^{-1}$. By construction, P is a coupling of μ and ν. We claim that P is an S–orbit coupling.

Consider first the S–coupling event C in (1.5). Since ϕ is a section of $Q = S/R$,

$$\begin{aligned} C &= \Big\{\exists\, a, b \in Q\ \exists\, r, s \in R : T_{r\phi(a)}X = T_{s\phi(b)}Y\Big\} \\ &= \xi\Big\{\exists\, r, s \in R : T_r \circ \Phi \circ \tilde{X} = T_s \circ \Phi \circ \tilde{Y}\Big\} . \end{aligned}$$

Together with (3.8), this shows that

$$P(C) = \tilde{P}(\xi^{-1}C) \geq 1 - \tfrac{1}{2}\|\tilde{\mu} \circ \Phi^{-1} - \tilde{\nu} \circ \Phi^{-1}\|_{\mathcal{I}(R)} \ . \tag{3.9}$$

To estimate the variation distance on the right–hand side we fix any $A \in \mathcal{I}(R)$ and consider the function

$$f_A(x) = \int m(da)\ 1_A(T_{\phi(a)}x)$$

on E. Since $T_s^{-1}A$ only depends on the equivalence class $[s]$ of $s \in S$, we have

$$f_A(T_s x) = \int m(da)\ 1_A(T_{\phi(a[s])}x) = f_A(x)$$

for all $s \in S$. The second equality comes from the invariance of m under translations. This shows that f_A is $\mathcal{I}(S)$–measurable. It follows that

$$\begin{aligned} |\tilde{\mu} \circ \Phi^{-1}(A) - \tilde{\nu} \circ \Phi^{-1}(A)| &= |\int f_A\ d(\mu - \nu)| \\ &= |\int (f_A - \tfrac{1}{2}) d(\mu - \nu)| \\ &\leq \|f_A - \tfrac{1}{2}\|\ \|\mu - \nu\|_{\mathcal{I}(S)} \end{aligned}$$

and therefore

$$\|\tilde{\mu} \circ \Phi^{-1} - \tilde{\nu} \circ \Phi^{-1}\|_{\mathcal{I}(R)} \leq \|\mu - \nu\|_{\mathcal{I}(S)} .$$

Combining this estimate with (3.9) and Remark (b) we obtain (1.4). This completes the proof of Proposition 3.5 and the theorem.

Acknowledgement

I am indebted to an anonymous referee for bringing references [13, 14] to my attention.

References

[1] D. Aldous and H. Thorisson, Shift–coupling, *Stochastic Proc. Appl.* **44**, 1993, pp. 1–14.

[2] H. C. P. Berbee, *Random Walks with Stationary Increments and Renewal Theory*, Math. Centre Tract No. 112, Mathematisch Centrum, Amsterdam, 1979.

[3] D.L. Cohn, *Measure Theory*, Birkhäuser, Boston, 1980.

[4] M.M. Day, Semigroups and amenability, K. W. Folley, Ed., *Semigroups, Proc. Symp. Wayne State Univ. 1968*, pp. 5–53, Academic Press, New York London, 1969.

[5] H.O. Georgii, Mixing properties of induced random transformations, *Ergodic Theory & Dyn. Systems*, 1996, to appear.

[6] Y. Katznelson and B. Weiss, The classification of non–singular actions, revisited. *Ergodic Th. & Dyn. Systems* **11**, 1991, pp. 333–348.

[7] Th.M. Liggett, *Interacting Particle Systems*, Springer Verlag, New York, 1985.

[8] T. Lindvall, *Lectures on the Coupling Method*, Wiley, New York, 1992.

[9] K. Petersen, *Ergodic Theory*, Cambridge University Press, 1983.

[10] A. Tempelman, *Ergodic Theorems for Group Actions*, Kluwer, Dordrecht, 1992.

[11] H. Thorisson, Shift coupling in continuous time, *Probab. Theory Rel. Fields* **99**, 1994, pp. 477–483.

[12] H. Thorisson, On time– and cycle–stationarity, *Stochastic Proc. Appl.* **55**, 1995, pp. 183–209.

[13] H. Thorisson, Transforming random elements and shifting random fields, *Ann. Probab.*, 1996, to appear.

[14] H. Thorisson, Point-stationarity in d dimensions and Palm theory, preprint, 1995.