MAXIMAL COUPLING OF EUCLIDEAN BROWNIAN MOTIONS

ELTON P. HSU AND KARL-THEODOR STURM

Abstract. We prove that the mirror coupling is the unique maximal Markovian coupling in the class of Markovian couplings of two Euclidean Brownian motions starting from single points.

1. Introduction

Let $(E_1, \mathscr{B}_1, \mu_1)$ and $(E_2, \mathscr{B}_2, \mu_2)$ be two probability spaces. A coupling of the probability measures μ_1 and μ_2 is a probability measure μ on the product measurable space $(E_1 \times E_2, \mathscr{B}_1 \times \mathscr{B}_1)$ whose marginal probabilities are μ_1 and μ_2, respectively. We denote the set of coupling of μ_1 and μ_2 by $\mathscr{C}(\mu_1, \mu_2)$. Thus, loosely speaking, a coupling of two Euclidean Brownian motions on $\mathbb{R}^n$ starting from x_1 and x_2, respectively, is a $C(\mathbb{R}_+, \mathbb{R}^n \times \mathbb{R}^n)$-valued random variable (X_1, X_2) on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ such that the components X_1 and X_2 have the law of Brownian motion starting from x_1 and x_2, respectively. In this case, we say simply that (X_1, X_2) is a coupling of Brownian motions from (x_1, x_2).

In the present work we discuss the uniqueness problem of maximal coupling of Euclidean Brownian motion. We first use the coupling inequality to motivate our definition of maximal coupling and show that the mirror coupling is a maximal coupling. We then show by an example that in general the maximal coupling is not unique. Second, we define the concept of Markovian couplings and show that the mirror coupling is the unique maximal coupling in the class of Markovian couplings.

We thank Pat Fitzsimmons and Wilfrid Kendall for their private communications on maximal couplings which are not the mirror coupling. We also thank Mu-Fa Chen for his useful comments.

2. Maximal coupling

Let

$$p(t, x, y) = \left(\frac{1}{2\pi t}\right)^{n/2} e^{-|x-y|^2/2t}$$

1991 *Mathematics Subject Classification.* Primary-60D58; secondary-28D05.

Key words and phrases. Euclidean Brownian motion, mirror coupling, maximal coupling.

This research was supported in part by the NSF grant DMS-010479 (Hsu) and by the SFB 611 at the University of Bonn (Sturm).

be the Gaussian heat kernel on $\mathbb{R}^n$. Here

$$|x|^2 := \sum_{i=1}^n |x^{(i)}|^2, \quad x = (x^{(1)}, \dots, x^{(n)}) \in \mathbb{R}^n.$$

Define the function

$$\phi_t(r) = \frac{2}{\sqrt{2\pi t}} \int_0^{r/2} e^{-\rho^2/2t}\, d\rho.$$

When $t = 0$, we let

$$\phi_0(r) = \begin{cases} 0, & r = 0, \\ 1, & r > 0. \end{cases}$$

The following facts are easy to verify:

(1) $r \mapsto \phi_t(r)$ is strictly concave on $\mathbb{R}_+ = [0, \infty)$ for $t > 0$;

(2) $t \mapsto \phi_t(r)$ is a tail distribution function on $\mathbb{R}_+$ for $r > 0$; in fact it is the tail distribution of the first passage time of a one dimensional Brownian motion from 0 to $r/2$:

$$\mathbb{P}\left\{\tau_{r/2} \geq t\right\} = \phi_t(r);$$

(3) we have

$$\phi_t(|x_1 - x_2|) = \frac{1}{2} \int_{\mathbb{R}^n} |p(t, x_1, y) - p(t, x_2, y)|\, dy. \tag{2.1}$$

Fix two distinct points x_1 and x_2 in $\mathbb{R}^n$. Let $X = (X_1, X_2)$ be a coupling of Euclidean Brownian motions from (x_1, x_2). The coupling time $T(X_1, X_2)$ is the earliest time at which the two Brownian motions coincide afterwards:

$$T(X_1, X_2) = \inf\left\{t > 0 : X_1(s) = X_2(s) \text{ for all } s \geq t\right\}.$$

The following coupling inequality gives a lower bound for the tail probability of the coupling time.

Proposition 2.1. *Let (X_1, X_2) be a coupling of Brownian motions from (x_1, x_2). Then*

$$\mathbb{P}\left\{T(X_1, X_2) \geq t\right\} \geq \phi_t(|x_1 - x_2|).$$

Proof. For any $A \in \mathscr{B}(\mathbb{R}^n)$, we have

$$\begin{aligned} \mathbb{P}\left\{T(X_1, X_2) > t\right\} &\geq \mathbb{P}\left\{X_1(t) \neq X_2(t)\right\} \\ &\geq \mathbb{P}\left\{X_1(t) \in A,\ X_2(t) \notin A\right\} \\ &\geq \mathbb{P}\left\{X_1(t) \in A\right\} - \mathbb{P}\left\{X_2(t) \in A\right\}. \end{aligned}$$

Hence,

$$\begin{aligned} \mathbb{P}\left\{T(X_1, X_2) > t\right\} &\geq \sup_{A \in \mathscr{B}(\mathbb{R}^n)} \int_A \left\{p(t, x_1, y) - p(t, x_2, y\right\} dy \\ &= \frac{1}{2} \int_{\mathbb{R}^n} |p(t, x_1, y) - p(t, x_2, y)|\, dy \\ &= \phi_t(|x_1 - x_2|). \end{aligned}$$

In the last step we have used (2.1) □

The coupling inequality

$$\mathbb{P}\left\{T(X_1, X_2) > t\right\} \ge \frac{1}{2}\int_{\mathbb{R}^n} \left|p\left(t, x_1, y\right) - p\left(t, x_2, y\right)\right| dy$$

holds under a much more general setting, see Lindvall[5].

In view of the coupling inequality, the following definition is natural.

Definition 2.2. *A coupling (X_1, X_2) of Brownian motions from (x_1, x_2) is called maximal at time t_0 if*

$$\mathbb{P}\left\{T(X_1, X_2) \ge t_0\right\} = \phi_{t_0}(|x_1 - x_2|).$$

It is called maximal if this holds for all $t_0 \ge 0$.

Throughout this work the starting points (x_1, x_2) are fixed unless otherwise stated. H will denote the hyperplane bisecting the segment $[x_1, x_2]$:

$$H = \left\{x \in \mathbb{R}^n : \langle x - x_0, n\rangle = 0\right\},$$

where

$$x_0 = \frac{x_1 + x_2}{2}$$

is the middle point and

$$n = \frac{x_1 - x_2}{|x_1 - x_2|}$$

is the unit vector in the direction of the segment (pointing from x_2 to x_1). We use $R : \mathbb{R}^n \to \mathbb{R}^n$ to denote the mirror reflection with respect to the hyperplane H:

$$Rx = x - 2\langle x - x_0, n\rangle n.$$

We now describe the mirror coupling. Let

$$\tau = \inf\left\{t \ge 0 : X_1(t) \in H\right\}$$

be the first hitting time of H by X_1. It is well known that

$$\mathbb{P}\left\{\tau \ge t\right\} = \phi_t(|x_1 - x_2|).$$

A coupling (X_1, X_2) of Brownian motions from (x_1, x_2) is a mirror coupling or X_2 is the mirror coupling of X_1 if X_2 is the mirror reflection of X_1 with respect to H before time τ and coincides with X_1 afterwards; namely,

$$X_2(t) = \begin{cases} RX_1(t), & t \in [0, \tau]; \\ X_1(t), & t \in [\tau, \infty). \end{cases}$$

In this case the coupling time $T(X_1, X_2) = \tau$. By definition the mirror coupling is a maximal coupling.

We say that a coupling (X_1, X_2) of Brownian motions is a mirror coupling up to time t_0 if the above relation between X_1 and X_2 holds for all $t \le t_0$.

The mirror coupling is not the only maximal coupling. Both Pat Fitzsimmons and Wilfrid Kendall communicated to us other maximal couplings. We describe the former in dimension 1. Let

$$l = \sup\{t \le \tau :\ X_1(t) = x_1\}$$

be the last time the first Brownian motion X_1 is at x_1 before τ_{x_0}. Note that X_1 starts from x_1. We define X_2 to be the time reversal of X_1 (shifted to x_2) before time l, the mirror reflection of X_1 (with respect to x_0) between l and τ, and X_2 after τ; namely,

$$X_2(t) = \begin{cases} x_2 - x_1 + X_1(l-t), & t \in [0, l]; \\ x_1 + x_2 - X_1(t), & t \in [l, \tau]; \\ X_1(t), & t \in [\tau, \infty). \end{cases}$$

Of course X_2 is not the mirror coupling of X_1. On the other hand, by Williams' decomposition of Brownian path $\{X_1(t), 0 \le t \le \tau\}$ (see Revuz and Yor[6], 244–245 and 304–305), X_2 is a Brownian motion starting from x_2. The coupling time for (X_1, X_2) is again equal to τ, which implies that the coupling is indeed a maximal coupling.

We now define a but natural subclass of couplings of Brownian motions.

Definition 2.3. *Let $X = (X_1, X_2)$ be a coupling of Brownian motions from $x = (x_1, x_2)$. Let $\mathscr{F}_*^X = \left\{\mathscr{F}_t^X\right\}$ be the filtration of σ-fields generated by the process X:*

$$\mathscr{F}_s^X = \sigma\left\{(X_1(u), X_2(u)) :\ u \le s\right\}.$$

We say that X is a Markovian coupling if for each $s \ge 0$, conditioned on the σ-field $\mathscr{F}_s^X$, the shifted process

$$\{(X_1(t+s), X_2(t+s)),\ t \ge 0\}$$

is still a coupling of Brownian motions from $(X_1(s), X_2(s))$.

Remark 2.4. The condition that (X_1, X_2) is a Markovian coupling only requires that, conditioned on $\mathscr{F}_s^{X_1, X_2}$, each time-shifted component is a Brownian motion. Thus it is a condition imposed solely on the joint distribution of (X_1, X_2). In particular, (X_1, X_2) is a Markovian coupling as soon as each component is a Brownian motion with respect to the common filtration $\mathscr{F}_*^X$; for instance if $\mathscr{F}_*^{X_1} = \mathscr{F}_*^{X_2}$. Note that the definition does not imply automatically that (X_1, X_2) is a Markov process. It does include, but not limited to, the case where (X_1, X_2) is a (possibly time-nonhomogeneous) Markov process whose generator L satisfies the conditions

$$Lf_1(z_1) = \frac{1}{2}\Delta_1 f_1(z_1), \quad Lf_2(z_2) = \frac{1}{2}\Delta_2 f_2(z_2),$$

where f_1 and f_2 are, respectively, functions of z_1 and z_2 alone.

The main result of this paper is the following.

Theorem 2.5. *Let $x_1, x_2 \in \mathbb{R}^n$ and $t_0 \geq 0$. A Markovian coupling of n-dimensional Brownian motions (X_1, X_2) starting from (x_1, x_2) which is maximal at time t_0 must be the mirror coupling up to time t_0.*

Corollary 2.6. *Let $x_1, x_2 \in \mathbb{R}^n$. The mirror coupling is the only maximal Markovian coupling of n-dimensional Brownian motions starting from (x_1, x_2).*

Before coming to the proof of this uniqueness result, we need to study optimal couplings of two Gaussian distributions of the same variance.

3. Optimal Coupling of Gaussian distributions

Given $t \geq 0$ and $x \in \mathbb{R}^n$ we use $N(x,t)$ to denote the n-dimensional Gaussian distribution with density function $z \mapsto p\,(t, x, z)$.

Definition 3.1. *Let $x_1, x_2 \in \mathbb{R}^n$ and $t \geq 0$, the mirror coupling $N(x_1, x_2, t)$ of the two Gaussian distributions $N(x_1, t)$ and $N(x_2, t)$ is the probability measure on $\mathbb{R}^n \times \mathbb{R}^n$ defined by*

$$N(x_1, x_2, t)(dy_1, dy_2) = \delta_{y_1}(dy_2)h_0(y_1)dy_1 + \delta_{Ry_1}(dy_2)h_1(y_1)dy_1$$

where R is the mirror reflection with respect to the hyperplane bisecting the segment $[x_1, x_2]$,

$$h_0(z) = p\,(t, x_1, z) \wedge p\,(t, x_2, z),$$

and

$$h_1(z) = p\,(t, x_1, z) - h_0(z).$$

The mirror coupling $N = N(x_1, x_2, t)$ can also be characterized uniquely by either one of the following two equivalent relations:

$$\begin{aligned} \int_{\mathbb{R}^n \times \mathbb{R}^n} f(z_1, z_2)N(dz_1 dz_2) &= \int_{\mathbb{R}^n} f(z, z)h_0(z)dz + \int_{\mathbb{R}^n} f(z, Rz)h_1(z)dz, \\ \int_{\mathbb{R}^n \times \mathbb{R}^n} f(z_1, z_2)N(dz_1 dz_2) &= \int_{\mathbb{R}^n} f(z, z)h_0(z)dz + \int_{\mathbb{R}^n} f(Rz, z)h_2(z)dz, \end{aligned}$$

where

$$h_2(z) = p\,(t, x_2, z) - h_0(z).$$

It is easy to verify that the marginal probabilities of $N(x_1, x_2, t)$ are $N(x_1, t)$ and $N(x_2, t)$, respectively; namely,

$$N(x_1, x_2, t) \in \mathscr{C}(N(x_1, t), N(x_2, t)).$$

It is concentrated on the following two perpendicular n-dimensional hyperplanes containing the point $(x_0, x_0) \in \mathbb{R}^n \times \mathbb{R}^n$:

$$D = \{(z, z) :\, z \in \mathbb{R}^n\}, \quad L = \{(z, Rz) :\, z \in \mathbb{R}^n\}.$$

The n-dimensional densities of $N(x_1, x_2, t)$ on these two hyperplanes are $h_0(z)$ and $h_1(z)$, respectively.

If an $\mathbb{R}^n \times \mathbb{R}^n$-valued random variable (ξ_1, ξ_2) is distributed as $N(x_1, x_2, t)$, we say that the Gaussian random variables ξ_1 and ξ_2 are mirror coupled. In this case, conditioned on ξ_1, the random variable ξ_2 takes only two possible values, ξ_1 and $R\xi_1$.

The significance of the mirror coupling $N(x_1, x_2, t)$ is that in a very general sense, it is the unique optimal coupling of the distributions $N(x_1, t)$ and $N(x_2, t)$. We clarify the meaning of this statement in THEOREMS 3.2 and 3.3 below. For this purpose, we introduce the Wasserstein (or Kantorovich-Rubinstein) distance of two probability measures μ_1 and μ_2 on $\mathbb{R}$ with respect to a nonnegative (cost) function ϕ:

$$D^W_\phi(\mu_1, \mu_2) = \inf_{\mu \in \mathscr{C}(\mu_1,\mu_2)} \int_{\mathbb{R}\times\mathbb{R}} \phi(|x-y|)\, \mu(dxdy).$$

As in SECTION 1, let

$$\phi_t(r) = \frac{2}{\sqrt{2\pi t}} \int_0^{r/2} e^{-\rho^2/2t}\, d\rho$$

for $t > 0$ with the obvious proviso for $t = 0$, and recall that $r \mapsto \phi_t(r)$ is strictly concave on $\mathbb{R}_+ = [0, \infty)$ for $t > 0$.

Theorem 3.2. *Let $(s, t) \in \mathbb{R}_+ \times \mathbb{R}_+$ and $(x_1, x_2) \in \mathbb{R} \times \mathbb{R}$. Then*

$$D^W_{\phi_s}(N(x_1, t), N(x_2, t)) = \phi_{s+t}(|x_1 - x_2|). \tag{3.1}$$

Proof. The hyperplane H bisecting the segment $[x_1, x_2]$ divides $\mathbb{R}^n$ into two half spaces. Let

$$u_0(x) = \frac{1}{2}\, \mathrm{sgn}\langle x - x_0, x_1 - x_2\rangle,$$

i.e., $u_0(x) = 1/2$ on the half space containing x_1 and $u_0(x) = -1/2$ on the half space containing x_2. Define

$$u_t(x) = \int_{\mathbb{R}^n} p\,(t, x, y) u_0(y)\, dy.$$

The following assertions hold:

$$u_t(z) = \frac{1}{2}\, \phi_t(2|z - x_0|)\mathrm{sgn}\langle z - x_0, x_1 - x_2\rangle, \qquad z \in \mathbb{R}^n; \tag{3.2}$$

$$\phi_t(|z_1 - z_2|) \ge u_t(z_1) - u_t(z_2), \qquad (z_1, z_2) \in \mathbb{R}^n \times \mathbb{R}^n; \tag{3.3}$$

$$\phi_t(|x_1 - x_2|) = u_t(x_1) - u_t(x_2); \tag{3.4}$$

$$u_{t+s}(x) = \int_{\mathbb{R}^n} p\,(t, x, z) u_s(z)\, dz, \qquad x \in \mathbb{R}^n. \tag{3.5}$$

(3.2) can be verified by a direct computation; (3.3) follows from the identity

$$u_t(z_1) - u_t(z_2) = \int_{\mathbb{R}^n} \{p\,(t, z_1, z) - p\,(t, z_2, z)\}\, u_0(z)\, dz$$

and (2.1) (note that $|u_0(z)| \le 1/2$); (3.4) follows from setting $z = x_1$ and $z = x_2$ in (3.2); and (3.5) follows from the semigroup property of the Gaussian heat kernel $p\,(t, z_1, z_2)$.

Now suppose that (ξ_1, ξ_2) is a coupling of the distributions $N(x_1, t)$ and $N(x_2, t)$. We have

$$\begin{aligned}\phi_{s+t}(|x_1 - x_2|) &= u_{s+t}(x_1) - u_{s+t}(x_2)\\ &= \mathbb{E}\left[u_s(\xi_1) - u_s(\xi_2)\right]\\ &\le \mathbb{E}\,\phi_s(|\xi_1 - \xi_2|).\end{aligned}$$

This shows that

$$D^W_{\phi_s}(N(x_1, t), N(x_2, t)) \ge \phi_{s+t}(|x_1 - x_2|). \tag{3.6}$$

On the other hand, if (ξ_1, ξ_2) has the distribution $N = N(x_1, x_2, t)$ (mirror coupling), then we have

$$\begin{aligned}\mathbb{E}\,\phi_s(|\xi_1 - \xi_2|) &= \int_{\mathbb{R}^n\times\mathbb{R}^n} \phi_s(|z_1 - z_2|)N(dz_1dz_2)\\ &= \int_{\mathbb{R}^n} \phi_s(2|\langle z - x_0, n\rangle|)h_1(z)\,dz\\ &= 2\int_{\langle z-x_0,n\rangle\ge 0}^{\infty} u_s(z)\left\{p\,(t, x_1, z) - p\,(t, x_2, z)\right\}dz\\ &= \int_{\mathbb{R}^n} u_s(z)\left\{p\,(t, x_1, z) - p\,(t, x_2, z)\right\}dz\\ &= u_{s+t}(x_1) - u_{s+t}(x_2)\\ &= \phi_{t+s}(|x_1 - x_2|).\end{aligned}$$

Here $n = (x_1 - x_2)/|x_1 - x_2|$ is the unique vector pointing from x_2 to x_1 and $x_0 = (x_1 + x_2)/2$ is the middle point. Note that both functions $u_s(z)$ and $p\,(t, x_1, z) - p\,(t, x_2, z)$ are odd with respect to the hyperplane H bisecting the segment $[x_1, x_2]$. The above sequence of equalities shows that the distance $D^W_{\phi_s}(N(x_1, t), N(x_2, t))$ is attained at the mirror coupling. □

For calculations related to the above proof, see Sturm[7], Example 4.6.

We now show that the Wasserstein distance $D^W_{\phi_s}(N(x_1, t), N(x_2, t))$ between the Gaussian distributions $N(x_1, t)$ and $N(x_2, t)$ is attained only at the mirror coupling. This fact is a consequence of the following general uniqueness theorem for the transport problem arising from a strictly concave cost function.

Theorem 3.3. *Let $s > 0$. Then the mirror coupling $N(x_1, x_2, t)$ is the unique coupling of $N(x_1, t)$ and $N(x_2, t)$ which realizes the Wasserstein distance $D^W_{\phi_s}(N(x_1, t), N(x_2, t))$.*

Proof. We have shown that the mirror coupling $N(x_1, x_2, t)$ indeed attains the Wasserstein distance, i.e.

$$\int_{\mathbb{R}^n\times\mathbb{R}^n} \phi_s(|y_1 - y_2|)N(x_1, x_2, t)(dy_1dy_2) = D^W_{\phi_s}(N(x_1, t), N(x_2, t)).$$

It is enough to prove the uniqueness.

The argument below will work if ϕ_s is replaced by any strictly concave cost function ϕ. For this reason we will drop the subscript s in what follows. We regard $\phi(|y_1 - y_2|)$ as the cost of transporting a unit mass from y_1 to y_2. Each $\mu \in \mathscr{C}(\mu_1, \mu_2)$ represents a way of transporting μ_1 to μ_2 and the total cost is the integral

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} \phi(|y_1 - y_2|)\mu(dy_1 dy_2).$$

Thus the Wasserstein distance $D^W_\phi(\mu_1, \mu_2)$ represents the minimal cost of transporting μ_1 to μ_2 and an optimal probability measure in $\mathscr{C}(\mu_1, \mu_2)$ is a transport which realizes the minimal cost. In our case,

$$\mu_1 = N(x_1, t) = p\,(t, x_1, z)dz, \quad \mu_2 = N(x_2, t) = p\,(t, x_2, z)dz.$$

Let μ be an optimal transport in this case. We have to show that $\mu = N(x_1, x_2, t)$, the mirror coupling.

Let

$$D = \{(x, x) : x \in \mathbb{R}^n\}$$

be the diagonal in $\mathbb{R}^n \times \mathbb{R}^n$. We first show that the restriction of μ to D is

$$\mu\Big|_D(dz) = \nu_0(dz) := h_0(z)dz, \tag{3.7}$$

where

$$h_0(z) := p\,(t, x_1, z) \wedge p\,(t, x_2, z).$$

First of all, since the marginal distributions of μ are

$$N(x_1, t) = p\,(t, x_1, z)dz, \quad N(x_2, t) = p\,(t, x_2, z)dz,$$

respectively, it is clear that $\mu|_D \le \nu_0$. The problem is to show that the equality holds.

We first explain the basic idea by giving an intuitive argument. Suppose that the strict inequality holds at a point y_0. From the fact that the first marginal distribution of μ is $p\,(t, x_1, z)dz$ we argue that there must be a point $y_2 \ne y_0$ such that (y_0, y_2) is in the support of μ. Similarly there must be a point $y_1 \ne y_0$ such that (y_1, y_0) is in the support of μ. This means that a positive mass is transported from y_1 to y_0 and then from y_0 to y_2. But then μ cannot be optimal because from the inequality

$$\phi(|y_1 - y_0|) + \phi(|y_0 - y_2|) > \phi(|y_1 - y_2|)$$

it is more efficient to transport the mass directly from y_1 to y_2.

For a rigorous argument, we write μ in the following forms:

$$\mu(dy_1 dy_2) = k_1(y_1, dy_2)\mu_1(dy_1) = k_2(y_2, dy_1)\mu_2(dy_2) \tag{3.8}$$

where k_1 and k_2 are appropriate Markov kernels on $\mathbb{R}^n$. Define two sub-probability measures

$$\nu_1 = \mu_1 - \nu_0, \quad \nu_2 = \mu_2 - \nu_0$$

on $\mathbb{R}^n$, and a probability measure on $\mathbb{R}^n \times \mathbb{R}^n$:

$$\nu(dy_1 dy_2) = \frac{1}{2}\delta_{y_1}(dy_2)\nu_0(dy_1) + \frac{1}{2}\int_{\mathbb{R}^n} k_2(y_0, dy_1)\, k_1(y_0, dy_2)\, \nu_0(dy_0) + \frac{1}{2}k_1(y_1, dy_2)\, \nu_1(dy_1) + \frac{1}{2}k_2(y_2, dy_1)\, \nu_2(dy_2).$$

Then a straightforward calculation shows that ν is also a coupling of μ_1 and μ_2. Comparing the transportation costs of ν with that of μ yields

$$\begin{aligned}\int_{\mathbb{R}^n\times\mathbb{R}^n} &\phi(|y_1 - y_2|)\nu(dy_1dy_2) - \int_{\mathbb{R}^n\times\mathbb{R}^n} \phi(|y_1 - y_2|)\mu(dy_1dy_2)\\ &= \frac{1}{2}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \{\phi(|y_1 - y_2|) - \phi(|y_1 - y_0|) - \phi(|y_2 - y_0|)\} \times \\ &\qquad\qquad k_2(y_0, dy_2)\, k_1(y_0, dy_1)\, \nu_0(dy_0).\end{aligned}$$

By the strict concavity of ϕ, the right-hand side does not exceed zero and is equal to zero only if y_0 is equal to either y_1 or y_2 almost surely with respect to the measure in the last line of the above display. This fact together with (3.8) imply (3.7).

We now investigate μ off diagonal. Recall that

$$\mu_1 = \nu_0 + \nu_1, \qquad \mu_2 = \nu_0 + \nu_2.$$

We have known already that each optimal transport leaves the part ν_0 unchanged. Moreover, the measures ν_1 and ν_2 are supported, respectively, on the two half spaces S_1 and S_2 separated by the hyperplane

$$H = \{z \in \mathbb{R}^n : \langle z - x_0, n\rangle = 0\}.$$

The idea is that transporting a mass from a point $y_1 \in S_1$ to $y_2 \in S_2$ costs the same as transporting the same mass from Ry_1 to Ry_2, but the two transports together are more expensive than the transports of y_1 to Ry_1 and of y_2 to Ry_2.

To make this argument rigorous, we first note that with the notation established in the first part of the proof μ can be written as

$$\mu(dy_1, dy_2) = \delta_{y_1}(dy_2)\nu_0(dy_1) + k_1(y_1, dy_2)\, \nu_1(dy_1).$$

Here $y_2 \in S_2$ almost surely with respect to $k_1(y_1, .)$ whenever $y_1 \in S_1$. Comparing the transportation costs for μ with those for the mirror coupling $N(dy_1, dy_2) = \delta_{y_1}(dy_2)\nu_0(dy_1) + \delta_{Ry_1}(dy_2)\nu_1(dy_1)$ yields

$$\begin{aligned}\int_{\mathbb{R}^n\times\mathbb{R}^n} &\phi(|y_1 - y_2|)N(dy_1dy_2) - \int_{\mathbb{R}^n\times\mathbb{R}^n} \phi(|y_1 - y_2|)\mu(dy_1dy_2)\\ &= \frac{1}{2}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \{\phi(|y_1 + Ry_1|) + \phi(|y_2 + Ry_2|) - 2\phi(|y_1 + y_2|)\} \times \\ &\qquad\qquad k_1(y_1, dy_2)\nu_1(dy_1).\end{aligned}$$

Again, by the strict concavity of ϕ, the right-hand side does not exceed zero and is equal to zero only if $y_2 = Ry_1$ almost surely with respect to the measure in the last line of the above display. This means that that μ

is the mirror coupling of the two Gaussian distributions $\mu_1 = N(x_1, t)$ and $\mu_2 = N(x_2, t)$. □

The above uniqueness result holds under much more general setting, see Gangbo and McCann[3], THEOREM 1.4. See also Villani[8], SECTION 4.3, THEOREM 3.

Corollary 3.4. *Let $s > 0$. If $\mu \in \mathscr{C}(N(x_1, t), N(x_2, t))$ (i.e., μ is a coupling of the Gaussian measures $N(x_1, t)$ and $N(x_2, t)$) such that*

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} \phi_s(|z_1 - z_2|)\mu(dz_1 dz_2) = \phi_{s+t}(|x_1 - x_2|),$$

then μ must be the mirror coupling.

Proof. This follows from THEOREMS 3.2 and 3.3. □

Remark 3.5. For the Gaussian measures $N(x_1, t)$ and $N(x_2, t)$, the coupling which achieves the distance $D^W_\phi(\mu_1, \mu_2)$ is independent of ϕ; see EXAMPLE 1.5 of Gangbo and McCann[3] for more general examples of this kind.

4. PROOF OF THE MAIN THEOREM

We now come to the proof of the main result THEOREM 2.5. Suppose that (X_1, X_2) is a Markovian coupling of Brownian motions from (x_1, x_2) which is maximal at time t_0. We have

$$\begin{aligned} \phi_{t_0}(|x_1 - x_2|) &\geq \mathbb{P}\{X_1(t_0) \neq X_2(t_0)\} \\ &= \mathbb{E}\,\phi_0(|X_1(t_0) - X_2(t_0)|). \end{aligned}$$

We fix a $t \in [0, t_0)$ and condition on $\mathscr{F}^X_t$. Since the coupling is Markovian, under this conditioning $(X_1(t_0), X_2(t_0))$ is a coupling of the Gaussian distributions $N(X_1(t), t_0 - t)$ and $N(X_2(t), t_0 - t)$. Hence by PROPOSITION 3.2,

$$\mathbb{E}\left\{\phi_0(|X_1(t_0) - X_2(t_0)|)|\mathscr{F}^X_t\right\} \geq \phi_{t_0-t}(|X_1(t) - X_2(t)|).$$

It follows that

$$\phi_{t_0}(|x_1 - x_2|) \geq \mathbb{E}\,\phi_{t_0-t}(|X_1(t) - X_2(t)|).$$

By PROPOSITION 3.2 again, this is bounded from below by $\phi_{t_0}(|x_1 - x_2|)$. Therefore the equality must hold and we have

$$\phi_{t_0}(|x_1 - x_2|) = \mathbb{E}\,\phi_{t_0-t}(|X_1(t) - X_2(t)|).$$

Using COROLLARY 3.4, we see that as Gaussian random variables, $X_2(t)$ must be mirror coupled with $X_1(t)$.

The fact that $X_1(t)$ and $X_2(t)$ are mirror coupled implies that with probability one, $X_2(t)$ is equal to either $X_1(t)$ or its mirror reflection $R(X_1(t))$. By the sample path continuity, we conclude that

$$\mathbb{P}\{X_2(t) = X_1(t) \text{ or } R(X_1(t)) \text{ for all } t \in [0, t_0]\} = 1.$$

It is now clear that before t_0 and before the first hitting time

$$\tau = \inf\{t \geq 0 : X_1(t) \in H\}$$

of the hyperplane H bisecting the segment $[x_1, x_2]$ we must have $X_2(t) = R(X_1(t))$. To show that X_2 is the mirror coupling of X_1 all the way up to time t_0, it is enough to show that $X_1(t) = X_2(t)$ for $t \in [\tau, t_0]$ if $\tau < t_0$.

We have shown that $(X_1(t), X_2(t))$ is mirror coupled for each $t < t_0$. Hence,

$$\mathbb{P}\{X_1(t) = X_2(t)\} = \int_{\mathbb{R}^n} \{p(t, x_1, z) \wedge p(t, x_2, z)\}\, dz.$$

Of course this equality should holds if X_2 is the mirror coupling of X_1, in which case the probability on the left-hand side is equal to the probability $\mathbb{P}\{\tau \leq t\}$, i.e.,

$$\mathbb{P}\{\tau \leq t\} = \int_{\mathbb{R}^n} \{p(t, x_1, z) \wedge p(t, x_2, z)\}\, dz.$$

Hence we have

$$\mathbb{P}\{X_1(t) = X_2(t)\} = \mathbb{P}\{\tau \leq t\}.$$

On the other hand, we know that $X_1(t) = RX_2(t)$ if $t < \tau$, or equivalently $X_1(t) \neq X_2(t)$ if $t < \tau$. This implies that

$$\{X_1(t) = X_2(t)\} \subseteq \{\tau \leq t\}.$$

Since the two events have the same probability, it follows that with probability one $X_1(t) = X_2(t)$ if $\tau \leq t$.

5. More general initial distributions

Given two probability measures μ_1, μ_2, it is not clear that a maximal Markovian coupling always exists. The problem is that the probability $\mathbb{P}\{T(X_1, X_2) \geq t\}$ can be minimized for each fixed t but not at the same coupling for all t. In this respect, we can obtain some positive results by taking advantage of certain situations in which the unique minimizers of the Wasserstein distance $D^W_\phi(\mu_1, \mu_2)$ are independent of the choice of strictly concave function ϕ. This is the case, for example, if $(\mu_1 - \mu_2)_+$ is supported on a half space and $(\mu_1 - \mu_2)_-$ is the reflection of $(\mu_1 - \mu_2)_+$ in the other half space, or if $(\mu_1 - \mu_2)_+$ is supported on an open ball and $(\mu_1 - \mu_2)_-$ is the spherical image of $(\mu_1 - \mu_2)_-$ (see Remark 3.5).

Let (X_1, X_2) be a maximal Markovian coupling of Brownian motions starting from (μ_1, μ_2). We expect that conditioned on $\mathscr{F}_0^{X_1, X_2}$, the coupling (X_1, X_2) is a maximal Markov coupling from $(X_1(0), X_2(0))$. This means that

$$\mathbb{P}\left\{T(X_1, X_2) \geq t \middle| \mathscr{F}_0^{X_1, X_2}\right\} = \phi_t(|X_1(0) - X_2(0)|)$$

and

$$\mathbb{P}\{T(X_1, X_2) \geq t\} = \mathbb{E}\,\phi_t(|X_1(0) - X_2(0)|).$$

Since the distribution of $(X_1(0), X_2(0))$ is a coupling of (μ_1, μ_2), the right side is bounded from below by $D^W_{\phi_t}(\mu_1, \mu_2)$. The function ϕ_t is strictly concave for strictly positive t. By our choice of (μ_1, μ_2), there is a unique minimizer $M_{\mu_1,\mu_2} \in \mathscr{C}(\mu_1, \mu_2)$ which realizes the distance simultaneously for all $t > 0$. From this fact it is now easy to conclude that the unique maximal Markovian coupling of Brownian motions starting from the initial distributions (μ_1, μ_2) and it is given by

$$\int_{\mathbb{R}^n} \mathscr{N}^{x_1,x_2} \, M_{\mu_1,\mu_2}(dx_1 dx_2),$$

where $\mathscr{N}^{x_1,x_2}$ denotes the law of the mirror coupling of Brownian motions from (x_1, x_2).

6. Another uniqueness theorem

There are other, arguably less interesting, conditions under which we can prove that the mirror coupling is the only maximal coupling.

Theorem 6.1. *If (X_1, X_2) is a maximal coupling of Euclidean Brownian motions from (x_1, x_2) and at the same time a martingale with respect to a filtration $\mathscr{F}_* = \{\mathscr{F}_t\}$, then they must be mirror coupled.*

Proof. We sketch a proof in dimension 1. The difference $X_1 - X_2$ is a continuous martingale. Let

$$\sigma_t = \frac{1}{4} \langle X_1 - X_2 \rangle(t).$$

There is a Brownian motion W such that

$$X_1(t) - X_2(t) = 2W(\sigma_t).$$

By the Kunita-Watanabe inequality, we have

$$d\langle X_1, X_2 \rangle(t) \le \sqrt{d\langle X_1 \rangle d\langle X_2 \rangle} = dt. \tag{6.1}$$

Hence $\sigma_t \le t$. Now let

$$\tau_1 = \inf \{ t \ge 0 : X_1(t) = X_2(t) \} .$$

We have $T(X_1, X_2) \ge \tau_1$. Let

$$\tau_2 = \inf \{ t \ge 0 : W(t) = 0 \} .$$

Then τ_2 is the first passage time of a Brownian motion from $|x_1 - x_2|/2$ to 0. The maximality of the coupling (X_1, X_2) means that $T(X_1, X_2)$ and τ_2 have the same distribution. On the other hand, we have $\sigma_{\tau_1} = \tau_2$, hence

$$T(X_1, X_2) \ge \tau_1 \ge \sigma_{\tau_1} = \tau_2.$$

It follows that $T(X_1, X_2) = \tau_2$ and $\sigma_{\tau_1} = \tau_2 = \tau_1$. Therefore the coupling time coincides with the first meeting time of X_1 and X_2, and before they meet the equality must hold in the Kunita-Watanabe inequality (6.1). This latter fact forces

$$X_2(t) = X_2(0) + X_1(0) - X_1(t) = 2x_0 - X_1(t), \quad 0 \le t \le T(X_1, X_2),$$

which just means that X_2 is the mirror coupling of X_1. □

7. Concluding remarks

It is clear that the results proved in this paper depend heavily on certain symmetry properties of the transition density function (the heat kernel) of Euclidean Brownian motion. As such it does not lend itself in any way to an immediate extension to a general Riemannian manifold. However, the argument is valid (at least for point mass starting distributions) for certain symmetric spaces, such as complete, simply connected spaces of constant curvature (space forms). The proof of this extension follows verbatim the proof we have given above for Euclidean space, and the only point which may not be immediately clear is the following fact.

Proposition 7.1. *Suppose that M is a complete, simply connected Riemannian manifold of constant curvature (space form) and x_1 and x_2 two distinct points on M. Let N be the totally geodesic hypersurface perpendicular to the minimal geodesic joining x_1 and x_2. Let*

$$\tau_N = \inf\{t \geq 0 : X_1(t) \in N\}$$

be the first hitting time of N. Then

$$\mathbb{P}\{\tau_N \geq t\} = \frac{1}{2}\int_M |p_M(t, x_1, y) - p_M(t, x_2, y)|\, dy.$$

Proof. Let M_1 be the half space of M containing x_1. Then the Dirichlet heat kernel of M_1 is

$$p_{M_1}(t, x, y) = p_M(t, x, y) - p_M(t, \tilde{x}, y),$$

where $\tilde{x}$ is the mirror reflection of x with respect to the hypersurface $N = \partial M_1$. Of course $x_2 = \tilde{x}_1$; hence,

$$\begin{aligned}\mathbb{P}\{\tau_N \geq t\} &= \int_{M_1} p_{M_1}(t, x_1, y)\, dy \\ &= \int_{M_1} \{p_M(t, x_1, y) - p_M(t, x_2, y)\}\, dy.\end{aligned}$$

By symmetry, we have

$$\begin{aligned}&\int_{M_1} \{p_M(t, x_1, y) - p_M(t, x_2, y)\}\, dy \\ &\qquad = \int_{M_2} \{p_M(t, x_2, y) - p_M(t, x_1, y)\}\, dy.\end{aligned}$$

The desired relation follows immediately. □

It should be pointed out that if the curvature is negative, coupling may fail with positive probability.

How to construct a maximal coupling for Brownian motions on a general Riemannian manifold is an interesting problem. The works of Griffeath[4] seems to indicate that a maximal coupling always exists as long as we do not restrict ourselves to Markovian couplings, and it is generally believed (Griffeath[4], Chen[2], Burdzy and Kendall[1], Lindvall[5]) that a maximal coupling is in general non-Markovian. However, Mu-Fa Chen communicated to us recently that the work of Zhang[9] implies a construction of a maximal Markovian coupling for a general discrete-time Markov chain. Whether such a construction can be carried out for a continuous-time Markov process remains to be seen.

References

[1] Burdzy, K. and Kendall, W. S., Efficient Markovian couplings: examples and counterexamples, *Ann. Appl. Prob.*, **10** (2000) 362-409

[2] Chen, M. F., Optimal Markovianian couplings and applications, *Acta Math. Sinica* (New Series), **10**, no. 3 (1994), 260–275.

[3] Gangbo, M. and McCann, R. J., The geometry of optimal transportation, *Acta Math.*, **177** (1996), 113-161.

[4] Griffeath, D., A maximal coupling for Markov chains, *Z. Wahr.*, **31** (1975), 94–106.

[5] Lindvall, T., *Lectures on the Coupling Method*, John Wiley & Sons, Chichester and New York, 1992.

[6] Revuz, D. and Yor, M., *Continuous Martingales and Brownian Motion*, 2nd edition, Springer-Verlag (1994).

[7] Sturm. K. T., A semigroup approach to harmonic maps, preprint SFB 611 Bonn, **39** (2003).

[8] Villani, C., *Topics in mass transportation*, Graduate Studies in Mathematics, vol. 57, AMS, Providence, RI (2003).

[9] Zhang, S. Y., Existence and application of optimal Markovian coupling with respect to nonnegative lower semi-continuous functions, *Acta Math. Sin. Eng. Ser.,* **16**, no. 2, 261–270 (2000).

Department of Mathematics, Northwestern University, Evanston, IL 60208
E-mail address: `elton@math.northwestern.edu`

Institute for Applied Mathematics, University of Bonn, Wegelerstrasse 6, 53115 Bonn
E-mail address: `sturm@uni-bonn.de`