A counterexample to the Cantelli conjecture through the Skorokhod embedding problem

Victor A. Kleptsyn, Aline Kurtzmann

A counter-example to the Cantelli conjecture

Victor Kleptsyn, Aline Kurtzmann

February 3, 2012

Abstract

In this paper, we construct a counter-example to a question by Cantelli, asking whether there exists a non-constant positive measurable function φ such that for i.i.d. r.v. X, Y of law $\mathcal{N}(0,1)$, the r.v. $X + \varphi(X) \cdot Y$ is also Gaussian.

For the construction that we propose, we introduce a new tool, the Brownian mass transport: the mass is transported by Brownian particles that are stopped in a specific way. This transport seems to be interesting by itself, turning out to be related to the Skorokhod and Stefan problems.

1 Introduction

1.1 A bit of history

The general thema of this paper is the following:

Cantelli conjecture (1918). *Let X, Y be two real random variables, of standard Gaussian distribution law. Suppose that X and Y are independent. Let φ be a measurable non-negative function. Then the random variable $X + \varphi(X) \cdot Y$ has a Gaussian distribution law if and only if φ is constant.*

Actually, Cantelli has originally mentioned this as a question, asking whether it is possible to have a non-constant function φ, in his paper [2, p.407], but later it became known as Cantelli conjecture. This conjecture has been previously studied by different authors. First, Tortorici [12] has given some restriction on the function φ to satisfy the conjecture. To do

that, he has developed φ in a Hermite series and has approached the solution (via a truncation of the series). Then, Tricomi [13] has used analytical tools in order to describe some properties satisfied by the function φ (through the characteristic function). In the same paper, he has also given a survey on this subject. Later, Dudley [4] has exposed two unsolved problems about finite-dimensional Gaussian measures. One of them was Cantelli conjecture. Dudley said about it "The problem seems to be a mere curiosity, but that will perhaps be unclear until it is solved". Letac has also worked on this problem and has emphasized this question, in his exercise book with Malliavin [6]. Indeed, they have suggested an exercise, showing that the decomposition of φ with respect to the Hermite polynomials, that is $\varphi(x) = \sum_{n\geq 0} \varphi_n \frac{H_n(x)}{n!}$ (in the $L^2(e^{-x^2/2}\frac{\mathrm{d}x}{\sqrt{2\pi}})$ sense) is such that $\varphi_1 = 0$, $-2\varphi_2 = \sum_{n\geq 2} \frac{\varphi_n^2}{n!}$ and $\varphi(x) \leq \varphi_0 + 1$ almost everywhere.

Finally, this striking question has been mentionned by de Meyer, Roynette, Vallois and Yor [3]. Actually, they answered a related question, asked by Tortrat. Consider a standard $(\mathcal{F}_t, t \geq 0)$–Brownian motion, denoted by $(B_t,\, t \geq 0)$. Can one find an a.s. bounded random variable Z, non-constant and $\mathcal{F}_1$–measurable, such that $B_1 + Z(B_2 - B_1)$ has a Gaussian distribution law? De Meyer et al. have proved the existence of a linear standard $(\mathcal{F}_t,\, t \geq 0)$–Brownian motion $(B_t,\, t \geq 0)$, and a stopping time T (w.r.t. $(\mathcal{F}_t,\, t \geq 0)$) which is bounded by 1, non-constant and such that B_T has a Gaussian distribution law. Thanks to this result, they have shown that the random variable $B_1 + \sqrt{T}(B_2 - B_1)$ has a Gaussian distribution law. In their example, $\sqrt{T}$ is $\mathcal{F}_1$–measurable, bounded and non-constant. However, one cannot write $\sqrt{T}$ as a function of B_1. So this construction does not contradict the Cantelli conjecture.

Before turning to the results of this paper, we would like to mention two problems, that turn out to be closely related to the Brownian transport notion, which is an essential part of our proof: the Skorokhod embedding problem and the Stefan problem.

The Skorokhod embedding problem is the following. For a given centered probability measure μ with finite second moment and a Brownian motion B, one looks for an integrable stopping time T such that the distribution law of B_T is μ. Several authors have developed different techniques to solve this problem, which has stimulated research in probability theory since the first formulation of Skorokhod [11]. A survey has been written by Oblój [7] on this subject. Some properties of stopping times or Skorokhod stopping have

also been intensively studied in the 70's and 80's (see for instance [5, 8, 10]). In this paper, we will give a solution to a somehow similar problem, and construct (under some assumptions) a stopping time T which is a function of the Brownian motion: $\exists g$ such that $T = g(B_T)$.

Our problem can actually be reduced to a PDE problem of the Stefan type (see for instance Rubinstein [9]), as we will explain later. Stefan problem is an old question, first considered in 1831 by Lamé and Clapeyron. It is actually a free boundary problem, initially used to predict ice formation/melting. For a historical survey on the problem, we refer the reader to Vuik [14].

1.2 Statement of the results

Our main result will be the following

Theorem 1.1. *There exists a measurable non-constant function $\varphi : \mathbb{R} \to \mathbb{R}_+$ such that for two independent standard Gaussian variables $X, Y \sim \mathcal{N}(0,1)$, the random variable $X + \varphi(X) \cdot Y$ is also Gaussian.*

In fact, as we will see from the construction in Section 2, the function φ can be taken to be a "choice" between two continuous functions:

$$\varphi(x) = \begin{cases} \varphi_0(x), & x \in \mathcal{K}, \\ \varphi_1(x), & x \notin \mathcal{K}, \end{cases}$$

where $\mathcal{K}$ is a Cantor set of positive Lebesgue measure and $\varphi_0, \varphi_1 \in C(\mathbb{R})$. Actually, the function φ we construct here is discontinuous. We believe that Cantelli conjecture is true if we impose the continuity of the function φ, but we have no proof for that.

In the proof of Theorem 1.1, there is a notion that naturally appears, the one of a new mass transport, that we name Brownian transport

Definition 1. Let μ_0, μ_1 be two probability measures, with the same mean and square integrable. We say that there exists a *Brownian transport* from μ_0 to μ_1 if, for a random process $(X_t, t \geq 0)$ such that $X_0 \sim \mu_0$ and $\mathrm{d}X_t = \mathrm{d}B_t$ (where B is a real Brownian motion independent of μ_0), one can find a stopping time T with finite expectation, and a function f such that

i) $X_T \sim \mu_1$,

ii) a.s. $T = f(X_T)$.

We say that f is the *stopping funtion* of this transport.

If moreover the time T is a.s. bounded, we say that there exists a *bounded Brownian transport* from μ_0 to μ_1.

Its study seems highly interesting on its own (see §4.1-4.2). However, in this paper, we will prove only two existence theorems, and their assumptions are clearly far from being optimal: we prove the results that would suffice to conclude the proof of Theorem 1.1.

We point out that such a transport does not always exist. In particular, a Brownian transport cannot exist if the target measure μ_1 is atomic (unless μ_0 has atoms at the same place). Moreover, we cannot for instance transport the uniform measure on $[-1,1]$ to the uniform measure on $[-1/2,1/2]$: the variance cannot be decreased by a Brownian transport. The bounded Brownian transport is even more restrictive: for instance, it cannot create "holes" inside the support of the measure, a necessary condition for its existence is that $\mathrm{Supp}(\mu_0) \subset \mathrm{Supp}(\mu_1)$.

We will discuss the existence of a Brownian transport later, in Section 2. Nevertheless, to construct a counter-example to Cantelli conjecture, we will consider probability measures such that a bounded Brownian transport exists (and the proof of its existence is an essential step in the construction). The second main result here is the following

Theorem 1.2. *Let μ_0, μ_1 be two centered probability measures, square integrable and which support is $\mathbb{R}$. Suppose that, for any R large enough, the troncated probability measures $\tilde{\mu}_0^R = \frac{\mu_0|_{[-R,R]}}{\mu_0([-R,R])}$ and $\tilde{\mu}_1^R = \frac{\mu_1|_{[-R,R]}}{\mu_1([-R,R])}$ satisfy:*

i) $\tilde{\mu}_0^R$ and $\tilde{\mu}_1^R$ are absolutely continuous with respective densities ρ_{μ_0} and ρ_{μ_1},

ii) there exist $a_R, b_R > 0$ such that for all $-R \leq x \leq R$, we have $\rho_{\mu_0}(x) \geq a_R$ and $\rho_{\mu_1}(x) \leq b_R$,

iii) there exists $\alpha_R > 0$ such that for any $J \subset [-R,R]$, we have $\mu_1(J) \geq e^{-\alpha_R/|J|}$.

Assume also that

iv) for any $x \in \mathbb{R}$, we have $\Phi_{\mu_0 \to \mu_1}(x) := \int_{-\infty}^x (\mu_0 - \mu_1)((-\infty, s])\, \mathrm{d}s > 0$,

v) $\limsup_{|x| \to +\infty} \frac{\rho_{\mu_0}(x)}{\rho_{\mu_1}(x)} < 1$.

Then, there exists a Brownian transport from μ_0 to μ_1, with a possibly unbounded stopping time T. Moreover, this Brownian transport is given by the first intersection time with the graph of some continuous function f.

Remark. We can actually suppose in the latter theorem that the measures μ_0, μ_1 have the same mean (instead of being centered).

An analogous question can be asked also for measures supported on an interval. This question, on one hand, turns out to be a bit simpler than the real line one (due to the compactness and lack of effects at infinity). On the other hand, it becomes one of the steps in our proof of Theorem 1.2: the function f is constructed as a limit of a subsequence of functions f_R corresponding to a "cut-off" problem. The corresponding theorem is

Theorem 1.3. *Let μ_0, μ_1 be two probability measures, with the same mean, square integrable and which support is an interval $I \subset \mathbb{R}$. Suppose that they satisfy the hypotheses:*

i) μ_0, μ_1 are absolutely continuous with respective densities ρ_{μ_0}, ρ_{μ_1},

ii) there exist $a, b > 0$ such that for all $x \in I$, we have $\rho_{\mu_0}(x) \geq a$ and $\rho_{\mu_1}(x) \leq b$,

iii) there exists $\alpha > 0$ such that for any interval $J \subset I$, we have $\mu_1(J) \geq e^{-\alpha/|J|}$,

iv) for all $x \in I$, we have $\Phi_{\mu_0 \to \mu_1}(x) := \int_{-\infty}^{x} (\mu_0 - \mu_1)((-\infty, s])\, \mathrm{d}s > 0$,

v) $\rho_{\mu_0} > \rho_{\mu_1}$ in some inner neighborhood $\mathcal{U}_\varepsilon(\partial I) \cap I$.

Then, there exists a bounded Brownian transport from μ_0 to μ_1, given by the first intersection time with the graph of some continuous function f.

The proof of these two results will be done in several steps. First, we will do some a priori estimates and transformations, answering the question "assuming that such a transport exists, how should it look like?". The understanding coming from these steps will leave us with some kind of a PDE problem, of the Stefan type.

However we could not establish the existence theorems for this problem directly, by PDE methods (in fact, it seems to be an interesting question to us), we establish them via a discretization procedure: we solve an analogous discrete problem and pass to the limit as the mesh goes to 0. This part is rather technical and is postponed to §4.2.

Remark. Some assumptions of Theorems 1.2 and 1.3 seem non restrictive, such as the positivity of $\Phi_{\mu_0 \to \mu_1}$ inside I for Theorem 1.3. Indeed, a necessary condition is that the function Φ is non-negative (see Corollary 3.2). Though, in the case of a non-negative function Φ that is not positive everywhere inside I, one can simply split the interval I into the intervals of positivity of Φ (see Lemma 3.4). Other assumptions, such as (iii), seem unavoidable in order to assure the uniform boundedness of the stopping time. Indeed, otherwise, the first intersection measure of the Brownian motion with the graph of an unbounded function f, say $f(x) = \frac{1}{|x|}$, would satisfy the assumptions of the theorem. But it is very likely that it cannot be obtained by means of a bounded Brownian transport. Finally, some assumptions (such as the absolute continuity of μ_0 or the lower bound for its density) surely can be weakened. It is highly interesting to study what are the "correct assumptions" for the existence of a Brownian transport. But we are not doing it in the present work: the statement of Theorem 1.3 suffices for our construction.

1.3 Outline

Let us indicate how the rest of the paper is organized. In Section 2, we present a construction of a counter-example to the Cantelli conjecture, based upon the existence of a certain bounded Brownian transport.

Next, in Section 3, we are doing some a priori estimates: assuming that a Brownian transport exists, we deduce some conclusions describing it. In particular here, the function Φ defined in Theorem 1.2 naturally appears (and becomes one of the main objects of our consideration). We will also explain how it is naturally related to a PDE of the Stefan type. In §3.2, using these estimates, we deduce from Theorem 1.2 that the Brownian transport (that we refer to in our counter-example construction) is indeed a bounded one (checking the behaviour of the corresponding function f at the infinity). This completes the construction of our counter-example (modulo Theorem 1.2).

In Section 4, we are proving Theorems 1.2 and 1.3. First, in §4.1, considering the restriction on the intervals $[-R, R]$ and passing to the limit, we prove the existence of a Brownian transport for measures on the real line, deducing Theorem 1.2 from Theorem 1.3. Then, in §4.2, by means of the discretization technique, we establish the existence of a Brownian transport on an interval, proving Theorem 1.3.

2 Construction

The first step in the proof of Theorem 1.1 is the following idea, close to [3]. Consider the standard Brownian motion $(B_t,\, t \geq 0)$, and let $T = T(\omega)$ be a stopping time (w.r.t. the standard family $(\mathcal{F}_t, t \geq 0)$ of σ-algebras), such that $T < C$ almost surely for some constant C. Then,

$$B_C = B_T + (B_C - B_T) = B_T + \sqrt{C - T} \cdot \xi, \tag{1}$$

where the random variable $\xi := \frac{B_C - B_T}{\sqrt{C-T}}$ is a standard Gaussian variable $\mathcal{N}(0,1)$ and is independent from B_T due to the Markov property.

Now note that B_C is a Gaussian random variable, so

$$B_T + \sqrt{C - T} \cdot \xi \sim \mathcal{N}(0, C), \quad B_T \perp\!\!\!\perp \xi, \quad \xi \sim \mathcal{N}(0,1). \tag{2}$$

Compare it to what we need to prove Theorem 1.1 (and hence to disprove the Cantelli conjecture):

$$X + \varphi(X) \cdot Y \sim \mathcal{N}(0, \cdot), \quad X \perp\!\!\!\perp Y, \quad X, Y \sim \mathcal{N}(0,1). \tag{3}$$

This comparison immediately gives us the following conclusion:

Proposition 2.1. *Let $T = T(\omega)$ be a non-constant stopping time for the standard Brownian motion $(B_t,\, t \geq 0)$, and assume that the following holds:*

i) $\exists C: \quad \forall \omega \quad T(\omega) < C$*;*

ii) The law of B_T is the standard Gaussian law: $B_T \sim \mathcal{N}(0,1)$;

iii) There exists a measurable function $f : \mathbb{R} \to \mathbb{R}_+$, such that almost surely $T = f(B_T)$.

Then, the function $\varphi(x) = \sqrt{C - f(x)}$ provides us a counter-example to the Cantelli conjecture.

Remark. There is one subtlety with the property iii) that we would like to emphasize. While this property *says* that the stopping moment T should be equal to a function of the place B_T where the process was stopped, it *does not* say that we should stop the process immediately once the equality $t = f(B_t)$ is satisfied. Moreover, for the construction in the proof of Theorem 1.1, it is *not* true that $T = \min\{t \colon t = f(B_t)\}$.

This is where the Brownian transport notion naturally appears. The assumption of this proposition can be rephrased as the existence of a Brownian transport from the Dirac measure δ_0 to the Gaussian measure $\mathcal{N}(0,1)$, with a non-constant bounded stopping time T. In fact, the Brownian transport notion will appear further in even more "clear" way, see Problem 1 below.

We can now describe how the stopping time T, satisfying the assumptions of Proposition 2.1, will be constructed. We will fix a moment $t_0 \in (0,1)$ and choose in a small neighborhood of the origin a Cantor set $\mathcal{K} \subset \mathbb{R}$ of positive Lebesgue measure (with some restrictions on its geometry), such that on this set the density of the law $\mathcal{N}(0,1)$ is everywhere upper bounded by the density of the law $\mathcal{N}(0,t_0)$:

$$\rho_{\mathcal{N}(0,t_0)}(x) > \rho_{\mathcal{N}(0,1)}(x) \quad \forall x \in \mathcal{K}.$$

Then, at the moment t_0, for any $x \in \mathcal{K}$, we stop $\frac{\rho_{\mathcal{N}(0,1)}(x)}{\rho_{\mathcal{N}(0,t_0)}(x)}$'s part of all the trajectories passing through x at this moment. To do so, one can either use a probabilistic Markov time, modifying the initial probability space of the Brownian motion by multiplying it by [0,1], or note that the random variable $S_{t_0}(\omega) := \sup_{0 \le t \le t_0} |B_t(\omega)|$ has a continuous conditional distribution w.r.t. any condition $B_{t_0} = x$, and hence, denoting by $\varkappa(\alpha, x)$ the α-quantile of the corresponding conditional distribution, we can put

$$T(\omega) = t_0 \quad \text{if } x := B_{t_0}(\omega) \in \mathcal{K} \text{ and } S_{t_0}(\omega) \le \varkappa\left(\frac{\rho_{\mathcal{N}(0,1)}(x)}{\rho_{\mathcal{N}(0,t_0)}(x)}, x\right). \tag{4}$$

This stopping ensures that the transport time T and the corresponding function f are non-constant: there is something left to transport.

The following problem now remains. At the moment t_0, there is a conditional distribution of not yet stopped trajectories, with the density:

$$\rho_0(x) = \begin{cases} c^{-1}\rho_{\mathcal{N}(0,t_0)}(x), & x \notin \mathcal{K}, \\ c^{-1}\left(\rho_{\mathcal{N}(0,t_0)}(x) - \rho_{\mathcal{N}(0,1)}(x)\right), & x \in \mathcal{K}, \end{cases} \tag{5}$$

where $c = \mathbb{P}(\mathcal{N}(0,1) \notin \mathcal{K})$.

Now, we are left with the following problem. We want to stop these trajectories at a bounded stopping time T, such that

i) $T = f(B_T)$,

ii) the law of B_T conditionally to $T > t_0$ is the restriction (to $\mathbb{R} \setminus \mathcal{K}$) of the standard Gaussian law $\mathcal{N}(0,1)|_{\mathbb{R}\setminus\mathcal{K}}$.

In other words, we are looking for a solution of the following

Problem 1. *Find a bounded Brownian transport from $\mu_0 = \rho_0 \, \mathrm{d}x$, given by* (5)*, to μ_1 which is the conditional distribution of $\mathcal{N}(0,1)$ on $\mathbb{R} \setminus \mathcal{K}$.*

Indeed, once Problem 1 is solved with the bounded stopping time T_1 such that $T_1 = f_1(B_{T_1})$, we can take for the original problem

$$T(\omega) = \begin{cases} t_0, & \text{if } x := B_{t_0}(\omega) \in \mathcal{K} \text{ and } S_{t_0}(\omega) \leq \varkappa(\frac{\rho_{\mathcal{N}(0,1)}(x)}{\rho_{\mathcal{N}(0,t_0)}(x)}, x), \\ t_0 + T_1, & \text{otherwise.} \end{cases} \tag{6}$$

where T_1 is evaluated on the trajectory $X_t = B_{t_0+t}$. We then have

$$f(x) = \begin{cases} t_0, & \text{if } x \in \mathcal{K}, \\ t_0 + f_1(x), & \text{if } x \notin \mathcal{K}. \end{cases} \tag{7}$$

Remark. It is important to note that, due to the choice of the "target measure" μ_1, the stopping point of the process $(X_t, t \geq 0)$ a.s. does not belong to $\mathcal{K}$. Hence, even though in (7), the function f on $\mathcal{K}$ does not coincide with $t_0 + f_1(x)$, the equality $T = f(B_T)$ still a.s. holds for the trajectories not yet stopped at time t_0.

Actually, Problem 1 is a particular case of a wider question, closely related to the Skorokhod embedding problem, that we state below. This question seems very interesting to us, but, at the best of our knowledge, it has not been studied until now.

Problem 2. *Let two probability measures μ_0 and μ_1, with the same mean and finite second moment and which support is $\mathbb{R}$, be given. Find a bounded Brownian transport from μ_0 to μ_1.*

It is obvious that such a Brownian transport does not always exist. Evidently, one has to ask $\mathbb{E}(\mu_0) = \mathbb{E}(\mu_1)$ and $\mathbb{V}\mathrm{ar}(\mu_0) \leq \mathbb{V}\mathrm{ar}(\mu_1)$. However, these two hypotheses are clearly insufficient. We will study this question in the next section and find a stronger necessary condition, and, what is more important, a sufficient one. Moreover, under this sufficient condition, the

function f can be taken to be continuous, with the moment T being the first intersection time of the trajectory $(X_t, t \geq 0)$ with the graph of f:

$$T(\omega) = \inf\{t \geq 0 : \quad t = f(X_t)\}.$$

Problem 1 will then be deduced as a result of an application of this sufficient condition. It will be done below, in Section 3. Namely, we have the following result

Theorem 2.2. *Assume that $\mathcal{K} \subset [-1, 1]$ and that there exists $\alpha > 0$ such that, for any interval $I \subset [-1, 1]$, one has $Leb(I \setminus \mathcal{K}) \geq \exp\{-\alpha/|I|\}$. Then, there exists a solution T_1 to Problem 1 and the corresponding function f_1 is continuous. Moreover, T_1 can be represented as a "first intersection" moment*

$$T_1(\omega) = \inf\{t \geq 0 : \quad t = f_1(X_t)\}.$$

Figure 1 below shows a simulation of the functions f_1 and φ (that one can do thanks to an almost explicite nature of our construction).

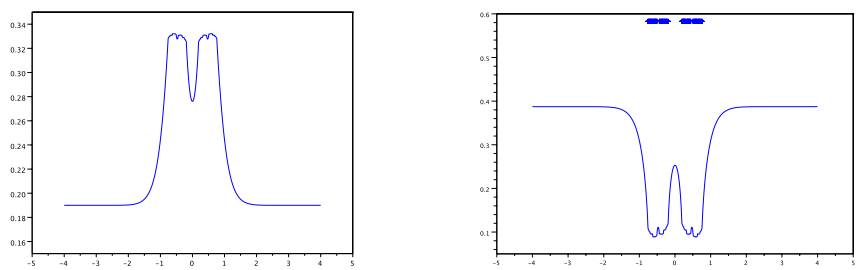

Figure 1: On the left: the graph of the function f_1. On the right: the graph of the resulting function φ

It is not difficult to construct a compact set $\mathcal{K}$ satisfying the assumptions of Theorem 2.2. Actually, if in the standard construction of the Cantor set, one chooses to remove on the n-th step an $\frac{1}{(n+1)^2}$-th part around the middle of the previously constructed intervals, the obtained Cantor set $\mathcal{K}$ satisfies the assumptions of Theorem 2.2. Moreover, for this Cantor set, an even stronger estimate holds: $Leb(I \setminus \mathcal{K}) \geq \alpha|I|^2$ for some universal constant α.

Once such a set $\mathcal{K}$ is constructed, the above arguments allow us to deduce Theorem 1.1 from Theorem 2.2. So the task of disproving the Cantelli conjecture is reduced to proving Theorem 2.2.

3 Ideas: the function Φ and some a priori arguments

In the following, for a regular time-space function $\Phi : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$, $(t, x) \mapsto \Phi_t(x)$, we denote by $\dot{\Phi}_t(x) = \partial_t \Phi_t(x)$ its time-derivative. Moreover, for an absolutely continuous measure μ, we denote by ρ_μ its density distribution function. The ε-neighborhood of a set I is denoted by $\mathcal{U}_\varepsilon(I)$. As all the objects we consider in this section are invariant by a translation, we will suppose that the measures μ_0, μ_1 are centered.

3.1 The transport function Φ and Stefan-type problem

Before going deeper into the proof of the existence theorems (Theorems 2.2, 1.2 and 1.3), let us first do some a priori arguments: assuming that a Brownian transport from some centered measure μ_0 to some other centered measure μ_1 exists (both μ_0, μ_1 having a finite second moment), what could be its properties and how could it be described?

The tool that is very useful for such a description is the following notion

Definition 2. Let μ be a measure on $\mathbb{R}$, with finite second moment. Then, we denote by Φ_μ the primitive of its repartition function $F_\mu(x) := \mu((-\infty, x])$:

$$\Phi_\mu(x) := \int_{-\infty}^{x} \mu((-\infty, s])\, \mathrm{d}s. \tag{8}$$

An easy computation then shows that

$$\begin{aligned} \Phi_\mu(x) &= \int_{-\infty}^{x} (x - y)\, \mathrm{d}\mu(y) = x\mu((-\infty, x]) - \int_{-\infty}^{x} y\, \mathrm{d}\mu(y) \\ &= x - \left(x\mu([x, +\infty)) + \mathbb{E}(\mu) - \int_{x}^{+\infty} y\, \mathrm{d}\mu(y) \right) \\ &= x - \mathbb{E}(\mu) + \int_{x}^{+\infty} \mu([s, +\infty))\, \mathrm{d}s. \end{aligned} \tag{9}$$

In particular, for any two such measures μ_0, μ_1, the difference between the corresponding functions

$$\Phi_{\mu_0 \to \mu_1}(x) := \Phi_{\mu_1}(x) - \Phi_{\mu_0}(x) \tag{10}$$

converges to 0 as x tends to $-\infty$ and as $x \to +\infty$.

The role of Φ is then given by the following result. Let (X_t, T) be a Brownian transport from μ_0 to μ_1. Denote by $\tilde{X}_t := X_{t\wedge T}$ the "stopped" process, by $\tilde{\nu}_t$ its distribution law at time t, and by ν_t the (non-probability) measure given by the "not yet stopped" particles: for any Borel set A, we have

$$\nu_t(A) = \mathbb{P}(X_t \in A, t < T).$$

Lemma 3.1. $\dot{\Phi}_{\tilde{\nu}_t} = \frac{1}{2}\rho_{\nu_t}$.

Proof. Indeed, we have $\mathrm{d}\tilde{X}_t = \mathbb{1}_{t<T}\,\mathrm{d}B_t$ and hence by the heat equation, we have $\dot{\Phi}_{\tilde{\nu}_t} = \frac{1}{2}\rho_{\nu_t}$. □

An immediate corollary to this lemma is the following

Corollary 3.2. *Let μ_0, μ_1 be two centered absolutely continuous probability measures, with finite second moment. Suppose that there exists a Brownian transport from μ_0 to μ_1. Then, for any $x \in \mathbb{R}$, we have $\Phi_{\mu_0\to\mu_1}(x) \geq 0$.*

Proof. It is obvious from Lemma 3.1 that the functions $\Phi_t(x) := \Phi_{\tilde{\nu}_t\to\mu_1}(x) = \Phi_{\mu_1}(x) - \Phi_{\tilde{\nu}_t}(x)$ are monotonically decreasing with t for any fixed x. The only thing we have to check is that $\Phi_t(x)$ converges pointwise to 0 (what is evident in the case of a bounded Brownian transport, but needs to be justified in general). Indeed, $\tilde{X}_t$ is a martingale and its variation

$$\mathbb{V}\mathrm{ar}(\tilde{X}_t) = \mathbb{V}\mathrm{ar}(\tilde{X}_0) + \mathbb{E}(t\wedge T) \leq \mathbb{V}\mathrm{ar}(\mu_0) + \mathbb{E}T < \infty$$

is uniformly bounded. Hence (see for instance Thm 4.3.3 in [1]), we have that $\tilde{X}_t$ converges in L^2 to $\tilde{X}_\infty(\omega) := \lim_{t\to\infty} X_t(\omega)$ and thus

$$\begin{aligned}\Phi_{\tilde{\nu}_t}(x) &= \int_{-\infty}^{x} \mathbb{P}(\tilde{X}_t \leq s)\,\mathrm{d}s = \int_\Omega |\tilde{X}_t(\omega) - s|_-\,\mathrm{d}\mathbb{P}(\omega) \\ &\xrightarrow[t\to\infty]{} \int_\Omega |\tilde{X}_\infty(\omega) - s|_-\,\mathrm{d}\mathbb{P}(\omega) = \Phi_{\mu_1}(x),\end{aligned}$$

where we have denoted $|a|_- := |a| \cdot \mathbb{1}_{a\leq 0}$. □

These proofs, in fact, suggest us a way of constructing the stopping time T. Namely, together with the process $(X_t, t \geq 0)$, we consider an increasing family of closed sets $K_t = \{\Phi_t = 0\}$ (that will be in fact sections of the

supergraph of f: $K_t = \{x \in \mathbb{R} : \ t \geq f(x)\}$, as shown in Figure 2 below). We stop the process once it reaches this family:

$$T = \inf\{t \geq 0 : \quad X_t \in K_t\}.$$

The function f will then be defined as

$$f(x) = \inf\{t \geq 0 : \quad x \in K_t\} = \inf\{t \geq 0 : \quad \Phi_t(x) = 0\}.$$

Roughly speaking, we let the function $\Phi_t = \Phi_{\tilde{\nu}_t \to \mu_1}$ decrease (as $\dot{\Phi}_t \leq 0$), and once it vanishes somewhere, we add this place to the set K_t of "stopped motion". Due to this description, we will call in the future $\Phi_{\mu_0 \to \mu_1}$ the *cost function* of the Brownian transport from μ_0 to μ_1.

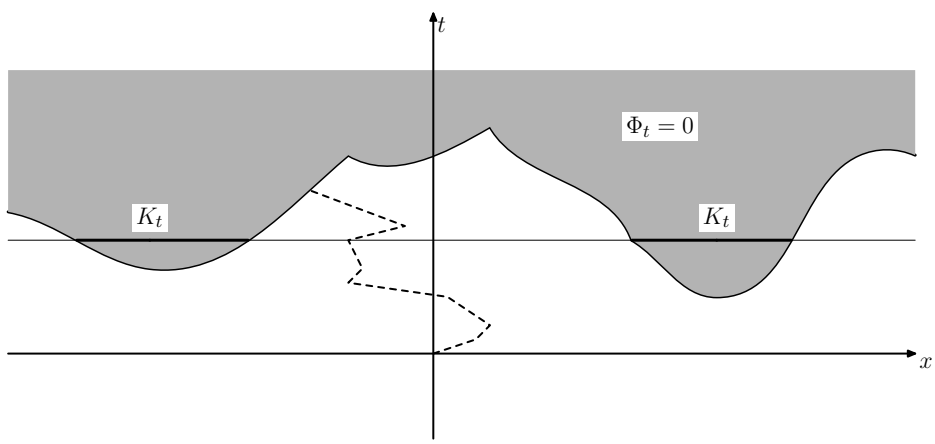

Figure 2: Construction of K_t

We wish to emphasize that the above description is absolutely unrigorous. It cannot be used without proving the corresponding existence theorems that do not seem to have an obvious direct proof. So, we will prove them in Section 4, via the discretization procedure. However, it gives an explanation why Theorems 1.2 and 1.3 should hold.

Moreover, this description can be (for the case of an absolutely continuous measure μ_0) rephrased in terms of Stefan-type problem. Namely, the density $\rho_t = \rho_{\nu_t}$ obeys the heat equation $\dot{\rho}_t = \frac{1}{2}\Delta\rho_t$ with the (moving) Dirichlet boundary condition $\rho_t|_{K_t} = 0$. So, the couple (Φ_t, ρ_t) and the function $f(x)$

obey the system

$$
\begin{cases}
\dot{\Phi}_t = -\frac{1}{2}\rho_t, \\
\dot{\rho}_t = \frac{1}{2}\Delta\rho_t \ \text{ if } \ t < f(x), \\
\Phi_{f(x)}(x) = 0, \\
\rho_{f(x)}(x) = 0,
\end{cases}
\tag{11}
$$

where the third equation defines the function f, while the last one is considered as a boundary condition on ρ.

We will not go deeper into giving fully formal sense to the system (11) (for instance, note that on the graph of f, the derivative $\dot{\Phi}_t$ can be discontinuous and if f is constant on some interval, then at the corresponding points, the density ρ will abruptly go to 0). As we have already mentioned in the Introduction, we could not prove the existence theorem here by PDE methods (though it would be interesting to find such a direct proof). However, we would like to emphasize here that the system (11) seems analogous to the Stefan problem of melting ice (see [9],[14]).

Even though we have not yet established the existence of the process described by the above rules, for the rest of this paragraph, we will –in order to understand the ideas before passing to the technical part– assume that it exists, and will study its behaviour. Note that one of the questions appearing (and that will be answered below) is the following one: does Φ_t vanish everywhere in finite time? To answer this question, it is natural to consider the connected components of $\mathbb{R} \setminus K_t$ and to study their evolution. In fact, to prove Theorem 2.2, we have to show that any of them disappears in a finite time. This will be done in Lemma 3.4. The next result deals with "disconnecting" different intervals from each other, allowing us to study their evolution separately.

Lemma 3.3. *Let $\mu, \tilde{\nu}$ be two centered absolutely continuous probability measures on $\mathbb{R}$, with finite second moment, such that $\Phi_{\tilde{\nu}\to\mu}$ is non-negative. Let $x \in \mathbb{R}$ be such that $\Phi_{\tilde{\nu}\to\mu}(x) = 0$. Then, the measures μ and $\tilde{\nu}$ of the interval $(-\infty, x]$ coincide, as well as the expectations of the conditional measures $\frac{\tilde{\nu}|_{(-\infty,x]}}{\tilde{\nu}((-\infty,x])}$ and $\frac{\mu|_{(-\infty,x]}}{\mu((-\infty,x])}$.*

The same holds for the restrictions on the interval $[x, +\infty)$ and on any interval $[x, y]$ provided that $\Phi_{\tilde{\nu}\to\mu}$ vanishes at both of its endpoints.

Proof. As the measures $\tilde{\nu}$ and μ are non-atomic, the function $\Phi_{\tilde{\nu}\to\mu}$ is of class C^1. But, as $\Phi_{\tilde{\nu}\to\mu}$ is non-negative and $\Phi_{\tilde{\nu}\to\mu}(x) = 0$, the point x is

a minimum of the function $\Phi_{\tilde{\nu}\to\mu}$. Hence $\partial_x \Phi_{\tilde{\nu}\to\mu}(x) = 0$. Noting that $\partial_x \Phi_{\tilde{\nu}\to\mu}(x) = -\mu((-\infty, x]) + \tilde{\nu}((-\infty, x])$, we obtain the first conclusion of the lemma. Now, remember the identity (9):

$$\Phi_\mu(x) = \int_{-\infty}^{x} (x-y)\ \mathrm{d}\mu(y) = x\mu((-\infty, x]) - \int_{-\infty}^{x} y\ \mathrm{d}\mu(y).$$

As $\Phi_{\tilde{\nu}\to\mu}(x) = 0$, and thus $\Phi_\mu(x) = \Phi_{\tilde{\nu}}(x)$, we have

$$x\mu((-\infty, x]) - \int_{-\infty}^{x} y\ \mathrm{d}\mu(y) = x\tilde{\nu}((-\infty, x]) - \int_{-\infty}^{x} y\ \mathrm{d}\tilde{\nu}(y). \tag{12}$$

The equality between the first terms in the left and right hand sides of (12) is already established and thus implies the equality between the last terms.

The other issues of the lemma are direct corollaries of the proved ones. □

3.2 Some a priori arguments and proof of Theorem 2.2

We are now ready to deduce Theorem 2.2 from Theorem 1.2. In other words, still assuming that the description in §3.1 defines us the desired process, we conclude the construction of the counter-example to the Cantelli conjecture. This deduction will be split in several lemmas.

A natural notion that has already inexplicitely appeared is the following

Definition 3. Let μ_0, μ_1 be two centered measures, with finite second moment. We say that there exists a *continuous Brownian transport* from μ_0 to μ_1 if there exists a Brownian transport (X_t, T) from μ_0 to μ_1 such that $T = f(X_T)$ where f is continuous and T is a first intersection time with the graph of f.

The transport we have constructed in Section 2 is actually a continuous one: the transports of Theorems 1.2 and 1.3 are continuous.

A first tool that we need is the following general lemma, that allows to estimate from above the time in which a connected component of $\mathbb{R} \setminus K_t$ "disappears".

Lemma 3.4. *Let $(\tilde{X}_t, K_t)$ be constructed as described above (§3.1) for some probability measures μ_0, μ_1 with the same mean and finite second moment (but perhaps with no time $\bar{t}$ such that $K_{\bar{t}} = \mathbb{R}$). Let I be an interval which is a connected component of $\mathbb{R} \setminus K_t$ (at some time t). Assume that for any interval $J \subset I$, we have $\mu_1(J) \geq \exp\{-\alpha/|J|\}$. Then, there exists a constant θ (which does not depend on I) such that $I \subset K_{t+\theta\alpha|I|}$.*

Proof. We will first prove the following auxiliary statement: there exists a constant θ_0 such that, at the moment $t' := t + \theta_0\alpha|I|$, any connected component of $I \setminus K_{t'}$ is of length less than $|I|/2$. This statement will imply the conclusion of the lemma. Indeed, applying it again to the connected components of $I\setminus K_{t+\theta_0\alpha|I|}$, we see that, at the moment $t'' = t+\theta_0\alpha|I|+\frac{\theta_0}{2}\alpha|I|$, the lengths of connected components of $I\setminus K_{t''}$ do not exceed $\alpha\frac{|I|}{4}$. We repeat this procedure. Thus, at the moment $t + 2\theta_0\alpha|I|$, we have $I \subset K_{t+2\theta_0\alpha|I|}$. This concludes the proof.

Let us now prove the latter statement. Indeed, note that for any interval of complement $J \subset \mathbb{R} \setminus K_t$, the Wiener measure of the trajectories that are still moving inside J at the time t is equal to $\mu_1(J)$. Indeed, as J is a connected component of $\mathbb{R} \setminus K_t$, we have $\Phi_t|_{\partial J} = 0$ and hence Lemma 3.3 can be applied. So to prove that at some moment $t' > t$, the length of any connected component $J \subset I \setminus K_{t'}$ is less than $|I|/2$, it suffices to show that, at this moment, the proportion of trajectories that have not yet intersected the graph of f is at most $\exp(-\frac{\alpha}{|I|/2})$.

To do this, we consider a weaker stopping condition: the trajectory is stopped once it reaches the boundary of I. The density of such a process is given by the heat equation with the Dirichlet boundary conditions on I. The measure of not yet stopped trajectories at the moment $t+\tau$ is then given by the scalar product $\langle\varphi_\tau, 1/|I|\rangle$, where

$$\dot{\varphi}_\tau = \frac{1}{2}\Delta\varphi_\tau, \quad \varphi_\tau|_{\partial I} = 0, \quad \varphi_0 = \rho_t.$$

As the Laplace operator is self-adjoint, this scalar product is equal to $\langle\psi_\tau, \varphi_0\rangle$, where

$$\dot{\psi}_\tau = \frac{1}{2}\Delta\psi_\tau, \quad \psi_\tau|_{\partial I} = 0, \quad \psi_0 = \frac{1}{|I|}.$$

Thus, this scalar product doesn't exceed $|I| \cdot \sup_I \psi_\tau$. Re-scaling the interval I to $[0,1]$ and accordingly multiplying the time by $1/|I|^2$ and the initial function by $|I|$, we obtain an upper bound by

$$\sup_{[0,1]} \sum_n c_{2n+1} \exp\left\{-\frac{\pi^2(2n+1)^2}{2|I|^2}\tau\right\} \sin(\pi(2n+1)x), \tag{13}$$

where $c_{2n+1} = \frac{2}{2n+1}$ are the nonzero Fourier coefficients of the function 1 with respect to the eigenfunctions $\sin(\pi(2n+1)x)$ of the Laplace operator on $[0,1]$.

Estimating c_n by 1 in (13) and the exponents by a geometric series, we see that this supremum does not exceed

$$\exp\{-\frac{\pi^2}{2|I|^2}\tau\} \cdot \frac{1}{1-\exp\{-\frac{\pi^2}{|I|^2}\tau\}}.$$

Now, note that for $\tau = \frac{8}{\pi^2}\alpha|I|$, the first factor is $\exp\{-4\frac{\alpha}{|I|}\} = \Big(\exp\{-\frac{\alpha}{|I|/2}\}\Big)^2$. Thus, the product is at most

$$\exp\{-\frac{\alpha}{|I|/2}\} \cdot \frac{\exp\{-\frac{\alpha}{|I|/2}\}}{1-\exp\{-\frac{\alpha}{|I|/2}\}}. \tag{14}$$

Note finally that $\exp\{-\frac{\alpha}{|I|/2}\}$ is at most $1/2$, as otherwise the μ_1-measures of both left and right halves of I would be greater than $1/2$. Hence, the second factor in (14) is not greater than 1 and we have obtained the desired estimate by $\exp\{-\frac{\alpha}{|I|/2}\}$. □

The next results are for the particular case of the transport in Theorem 2.2, based essentially on the specifics of Gaussian distributions. Namely, let μ_0 and μ_1 be as in Theorem 2.2.

Lemma 3.5. *μ_0 and μ_1 satisfy the assumptions of Theorem 1.2.*

Proof. The conditions i) and v) are obvious and the fact that the measures μ_0, μ_1 have the same mean comes from the fact that we are removing the same part from $\mathcal{N}(0,t_0)$ and $\mathcal{N}(0,1)$. Conditions ii) and iii) are due to the assumptions on $\mathcal{K}$. We only have to prove iv). Indeed, the function Φ_μ depends linearly on μ: $\Phi_{\alpha\mu+\beta\nu} = \alpha\Phi_\mu + \beta\Phi_\nu$. Due to the definition of μ_0 and μ_1, we have that

$$\begin{aligned} \mathrm{Law}(\mathcal{N}(0,t_0)) &= c\mu_0 + (1-c)\mu', \\ \mathrm{Law}(\mathcal{N}(0,1)) &= c\mu_1 + (1-c)\mu' \end{aligned}$$

where μ' is the conditional distribution law of $\mathcal{N}(0,1)$ on $\mathcal{K}$. Hence,

$$\Phi_{\mu_0\to\mu_1}(x) = c^{-1}\Phi_{\mathcal{N}(0,t_0)\to\mathcal{N}(0,1)}(x) = c^{-1}\int_{t_0}^1 \frac{1}{\sqrt{2\pi t}}\exp\{-\frac{x^2}{2t}\}\,\mathrm{d}t > 0.$$

□

Now, let the Brownian transport (X_t, T_1), where $T_1 = f_1(X_{T_1})$, be a continuous Brownian transport of Theorem 1.2. We have to show that the (continuous) function f_1 is bounded. In other words, we have to estimate its behaviour at infinity. Actually, we will prove the stronger statement

Proposition 3.6. $\lim_{x\to\infty} f_1(x) = 1 - t_0$. *Moreover, there exists a constant* $\beta > 0$ *such that for all* $|x|$ *large enough, one has* $1-t_0 \leq f_1(x) \leq 1-t_0+e^{-\beta x^2}$.

A first step in proving this proposition is the following

Lemma 3.7. $\forall x \in \mathbb{R}$, $f_1(x) \geq 1 - t_0$.

Proof. It is here easier to work with the non-normalized measures $\hat{\mu}_0 = c\mu_0$ and $\hat{\mu}_1 = c\mu_1$, and with the corresponding non-normalized cost function

$$\Phi_{\hat{\mu}_0\to\hat{\mu}_1} = c\Phi_{\mu_0\to\mu_1} = \Phi_{\mathcal{N}(0,t_0)\to\mathcal{N}(0,1)}.$$

It is clear that they satisfy the system (11). In fact, one can simply divide everything by c, to pass to the normalized case, but it seems to us that the explanation would be less clear.

If we had not removed, at the initial moment from $\mathcal{N}(0, t_0)$, the particles corresponding to $(1 - c)\mu' = \mathcal{N}(0, 1)|_{\mathcal{K}}$, we would have had

$$\int_0^{1-t_0} \rho_{\mathcal{N}(0,t+t_0)}(x)\ \mathrm{d}t = \Phi_{\hat{\mu}_0\to\hat{\mu}_1}(x).$$

As our initial condition is only a part of $\mathcal{N}(0, t_0)$, we have $\forall t > 0\ \forall x \in \mathbb{R}$ $\rho_t(x) < \rho_{\mathcal{N}(0,t_0)}(x)$, where ρ_t is the density of the process started with $\hat{\mu}_0$ and stopped at the moment of touching the graph of f_1. Hence, we have

$$\forall x \in \mathbb{R} \quad \int_0^{t_0} \rho_t(x)\ \mathrm{d}t < \Phi_{\hat{\mu}_0\to\hat{\mu}_1}(x),$$

and as $\int_0^{f_1(x)} \rho_t\ \mathrm{d}t = \Phi_{\hat{\mu}_0\to\hat{\mu}_1}$, we have proved the result. □

Now, let us consider the density that we obtain at the time $1 - t_0$. The next lemma estimates its behaviour at infinity:

Lemma 3.8. *There exists a constant* $\beta_0 > 0$ *such that for all* $|x|$ *large enough, one has*

$$\rho_{\mathcal{N}(0,1)}(x) \cdot (1 - e^{-\beta_0 x^2}) \leq \rho_{1-t_0}(x) \leq \rho_{\mathcal{N}(0,1)}(x).$$

Proof. The measure ν_{1-t_0} is the convolution of the initial measure $\hat{\mu}_0$ with $\mathcal{N}(0, 1-t_0)$. If, instead of $\hat{\mu}_0$, we had $\mathcal{N}(0, t_0)$, we would obtain exactly $\mathcal{N}(0,1)$. But as $\hat{\mu}_0$ is only a part of $\mathcal{N}(0, 1-t_0)$, we immediately have $\rho_{1-t_0}(x) \leq \rho_{\mathcal{N}(0,1)}(x)$.

The difference $\rho_{\mathcal{N}(0,1)}(x) - \rho_{1-t_0}(x)$ is the part of the density that comes from the removed part $\mathcal{N}(0,1)|_{\mathcal{K}}$ of the initial condition. This part is supported by $[-1,1]$. Hence, the difference

$$\rho_{\mathcal{N}(0,1)}(x) - \rho_{1-t_0}(x) = \rho_{\mathcal{N}(0,1)|_{\mathcal{K}} * \mathcal{N}(0,1-t_0)}(x)$$

can be estimated from above as $u \cdot e^{-\frac{(x-1)^2}{2(1-t_0)}}$, where $u > 0$ is a constant. This is asymptotically less that $e^{-\beta_0 x^2} \cdot \rho_{\mathcal{N}(0,1)}(x)$ for any $\beta_0 < \frac{1}{2}\left(\frac{1}{1-t_0} - 1\right)$. □

From now on, let us fix β_0 as in Lemma 3.8. We can estimate the behaviour of the function Φ at the same moment $1 - t_0$

Lemma 3.9. *For all $|x|$ large enough, we have $\Phi_{1-t_0}(x) \leq e^{-\beta_0 x^2} \rho_{\mathcal{N}(0,1)}$.*

Proof. From the definition of Φ, we indeed have

$$\begin{aligned}
\Phi_{1-t_0}(x) &= \int_{-\infty}^{x} \left(\mu_1((-\infty, s]) - \tilde{\nu}_{1-t_0}((-\infty, s])\right) \, \mathrm{d}s \\
&= \int_{-\infty}^{x} (x-s)(\rho_{\mu_1} - \rho_{\tilde{\nu}_{1-t_0}})(s) \, \mathrm{d}s \\
&= \int_{-\infty}^{x} (x-s)(\rho_{\mathcal{N}(0,1)} - \rho_{\nu_{1-t_0}})(s) \, \mathrm{d}s.
\end{aligned}$$

Applying Lemma 3.8, we have as $x \to -\infty$

$$\begin{aligned}
\Phi_{1-t_0}(x) &= \int_{-\infty}^{x} (x-s) \cdot e^{-\beta_0 s^2} \cdot \rho_{\mathcal{N}(0,1)}(s) \, \mathrm{d}s \\
&\leq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} |s| e^{-(\beta_0 + 1/2)s^2} \, \mathrm{d}s = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} e^{-(\beta_0+1/2)v^2} \, \mathrm{d}(v^2/2) \\
&\leq \frac{1}{\sqrt{2\pi}} e^{-(\beta_0+1/2)x^2} = e^{-\beta_0 x^2} \cdot \rho_{\mathcal{N}(0,1)}(x).
\end{aligned}$$

In the same way, using the integral representation of $\Phi_{\mu \to \nu}$ via the integral (9), one can estimate $\Phi_{1-t_0}(x)$ for any large positive x. □

Having obtained this estimate, we can conclude that the inequality $f_1(x) \leq 1 - t_0 + e^{-\beta_0 x^2/2}$ will be satisfied for a "very dense" at infinity set of points x. Namely, denote $\ell(x) := e^{-\beta_0 x^2/2}$.

Lemma 3.10. *For any $|x|$ large enough, there exist two points $y_+ \in [x, x + \ell(x)]$ and $y_- \in [x - \ell(x), x]$ such that $f_1(y_+) \leq 1 - t_0 + \ell(x)$ and $f_1(y_-) \leq 1 - t_0 + \ell(x)$.*

Proof. Assume the contrary: for instance, that $\forall y \in [x, x + \ell(x)]$, $f_1(y) > (1 - t_0) + \ell(x)$. This implies that the set K_t does not intersect the rectangle $[x, x + \ell(x)] \times [1 - t_0, 1 - t_0 + \ell(x)]$, and for any point of this rectangle, the density $\rho_t(y)$ can be estimated from below via the solution of the heat equation $\dot{u} = \frac{1}{2}\Delta u$ on $[x, x + \ell(x)]$ with the initial conditions $u_{1-t_0} = \rho_{1-t_0}$.

For all $|x|$ large enough, $\rho_{\mathcal{N}(0,1)}$ varies on $[x, x + \ell(x)]$ at most 2 times, and hence we have a lower bound for the initial condition $\forall y \in [x, x + \ell(x)]$

$$\rho_{1-t_0}(y) \geq \frac{1}{3}\rho_{\mathcal{N}(0,1)}(m) \geq \frac{1}{3}\sin\left(\frac{\pi}{\ell(x)} \cdot (y - x)\right) \cdot \rho_{\mathcal{N}(0,1)}(m),$$

where $m = x + \frac{1}{2}\ell(x)$ is the middle of the interval $[x, x + \ell(x)]$. The function $\sin\left(\frac{\pi}{\ell(x)} \cdot (y - x)\right)$ is an eignefunction of the Laplace operator with the eigenvalue $\lambda = \frac{\pi^2}{\ell(x)^2}$ and hence for all $t \in [1 - t_0, 1 - t_0 + \ell(x)]$, we have a lower bound

$$\begin{aligned} \rho_t(y) &\geq \frac{1}{3}\exp\{-\frac{t - (1 - t_0)}{2} \cdot \frac{\pi^2}{\ell(x)^2}\} \cdot \sin\left(\frac{\pi}{\ell(x)} \cdot (y - x)\right) \cdot \rho_{\mathcal{N}(0,1)}(m) \\ &\geq \frac{1}{4}\sin\left(\frac{\pi}{\ell(x)} \cdot (y - x)\right) \cdot \rho_{\mathcal{N}(0,1)}(m). \end{aligned}$$

In particular, for the middle point m of the interval, we have

$$\rho_t(m) \geq \frac{1}{4}\rho_{\mathcal{N}(0,1)}(m).$$

Thus,

$$\int_{1-t_0}^{1-t_0+\ell(x)} \rho_t(m)\ \mathrm{d}t \geq \frac{\ell(x)}{4} \cdot \rho_{\mathcal{N}(0,1)}(m). \tag{15}$$

As $\ell(x) = e^{-\beta_0 x^2/2}$, we have due to Lemma 3.9

$$\Phi_{1-t_0}(m) \leq e^{-\beta_0(x+\ell(x))^2} \cdot \rho_{\mathcal{N}(0,1)}(m).$$

So we have

$$\int_{1-t_0}^{1-t_0+\ell(x)} \rho_t(m)\ \mathrm{d}t > \Phi_{1-t_0}(m).$$

The obtained contradiction proves the lemma. □

We can now conclude the proof of Proposition 3.6.

Proof of Proposition 3.6. Lemma 3.10 implies that for any $|x|$ large enough, either $f_1(x) \leq 1 - t_0 + \ell(x)$ or the connected component I of $\mathbb{R} \setminus K_t$ that contains x is a subset of $[x - \ell(x), x + \ell(x)]$. We are now going to show that then $f_1(x) \leq (1-t_0)+\ell(x)+\theta_1\ell(x)^2$, where the constant θ_1 can be chosen not depending on x. Indeed, due to Lemma 3.3, we can consider the continuous Brownian transport problem from $\nu_{1-t_0+\ell(x)}|_I$ to $\hat{\mu}_1|_I$ independently of the rest of the real line. Let us then rescale I to $[0,1]$, normalizing the measures $\nu_{1-t_0+\ell(x)}|_I$ and $\hat{\mu}_1|_I$ to probability ones, and rescaling the time by the factor $\frac{1}{|I|^2}$.

The density of the new probability measure $\tilde{\mu}_1$ on $\tilde{I} = [0,1]$ takes value on $[1/2, 2]$ (as $\rho_{\mathcal{N}(0,1)}$ varies at most two times on I). Hence, it satisfies the assumptions of Lemma 3.4 with some uniform (not depending on x) constant α. Thus, the rescaled time in which the interval "disappears" is uniformly (for $|x|$ large enough) bounded by some constant θ_3 and hence $x \in I \subset K_{(1-t_0)+\ell(x)+\theta_3\ell(x)^2}$. As $\ell(x) \ll 1$, the latter statement implies the desired upper bound for $f_1(x)$. □

This concludes the proof of Theorem 2.2: the function f_1 is bounded on $\mathbb{R}$.

4 Existence of a Brownian transport

4.1 Brownian transport on the real line

In this paragraph, we will deduce Theorem 1.2 from Theorem 1.3 (which will be proved in the next paragraph). To do so, assume that the measures μ_0, μ_1 satisfy the assumptions of Theorem 1.2. Naturally, the idea here will be to find a family of compactly supported measures μ_0^R and μ_1^R that approximate μ_0 and μ_1 and for which there exist continous Brownian transports. The simplest case is when the measures μ_0, μ_1, in addition to be centered, are symmetric.

We will then consider the sequence of conditional normalized measures

$$\tilde{\mu}_0^R := \frac{\mu_0|_{[-R,R]}}{\mu_0([-R,R])}, \text{ and } \tilde{\mu}_1^R := \frac{\mu_1|_{[-R,R]}}{\mu_1([-R,R])}.$$

For the case of general centered measures μ_0 and μ_1, we will have to modify this construction, as their restrictions on $[-R,R]$ are no longer forced to have the same mean. Namely, denote for any measure μ such that $\mu((-\infty,0)) > 0$ and $\mu((0,\infty)) > 0$ by $\gamma(\mu)$ the measure

$$\gamma(\mu) := c(\mu)\mu|_{(-\infty,0)} + d(\mu)\mu|_{(0,\infty)},$$

where $(c(\mu), d(\mu))$ is the unique solution to the system

$$\begin{cases} c(\mu)\mu((-\infty,0)) + d(\mu)\mu((0,\infty)) = 1, \\ -c(\mu)\int_{-\infty}^0 |x| \, \mathrm{d}\mu + d(\mu)\int_0^\infty x \, \mathrm{d}\mu = 0. \end{cases}$$

It is then easy to see that $\gamma(\mu)$ is always a centered measure and we have $c(\tilde{\mu}_j^R) \xrightarrow[R\to\infty]{} 1$ and $d(\tilde{\mu}_j^R) \xrightarrow[R\to\infty]{} 1$ (as the second equation tends to $c = d$ as $R \to \infty$). Then we can consider the families $\mu_0^R = \gamma(\tilde{\mu}_0^R)$ and $\mu_1^R = \gamma(\tilde{\mu}_1^R)$.

Now we would like to consider continuous Brownian transports from μ_0^R to μ_1^R, then extract a convergent subsequence from the sequence of corresponding functions f_R, and finally show that the limit function f indeed defines a continuous Brownian transport from μ_0 to μ_1. So, a first step in the realization of this scheme is to check that for all R large enough, Theorem 1.3 is indeed applicable for finding a continuous Brownian transport from μ_0^R to μ_1^R.

Lemma 4.1. *For any R large enough, there exists a continuous Brownian transport from μ_0^R to μ_1^R.*

Proof. We have to check that the assumptions of Theorem 1.3 are satisfied for all R large enough. As the conditions i)-iii) are the same in Theorems 1.2 and 1.3, we only have to check the two last ones.

Recall that we have $\lambda := \limsup_{x\to\infty} \frac{\rho_{\mu_0}(x)}{\rho_{\mu_1}(x)} < 1$. Hence, for some constant M, we have $\frac{\rho_{\mu_0}(x)}{\rho_{\mu_1}(x)} < \frac{1+\lambda}{2}$ outside $[-M,M]$. Now, for $x \in (-M,M)$, we have

$$\frac{\rho_{\mu_0^R}(x)}{\rho_{\mu_1^R}(x)} = \frac{\rho_{\mu_0}(x)}{\rho_{\mu_1}(x)} \cdot \frac{\mu_1([-R,R])}{\mu_0([-R,R])} \cdot \left(\frac{c(\tilde{\mu}_0^R)}{c(\tilde{\mu}_1^R)} \cdot \mathbb{1}_{x<0} + \frac{d(\tilde{\mu}_0^R)}{d(\tilde{\mu}_1^R)} \cdot \mathbb{1}_{x\geq 0} \right).$$

Note that the second factor in the right hand side tends (uniformly) to 1 as $R \to \infty$. Thus, for any R large enough, it is less than $\frac{2}{1+\lambda}$ and hence $\exists M : \forall |x| > M$, $\frac{\rho_{\mu_0^R}(x)}{\rho_{\mu_1^R}(x)} < \frac{2}{1+\lambda} \cdot \frac{1+\lambda}{2} = 1$. This proves the desired condition v).

Moreover, note that due to the finiteness of the first moment of μ_0 and μ_1, we have $\Phi_{\mu_0^R \to \mu_1^R}(x) \xrightarrow[R\to\infty]{} \Phi_{\mu_0 \to \mu_1}(x)$ uniformly on $x \in [-M, M]$. Thus, for all R large enough, we have $\Phi_{\mu_0^R \to \mu_1^R} > 0$ on $[-M, M]$.

Next, for all $R > M$ and $x \in (-R, -M]$, we have

$$\Phi_{\mu_0^R \to \mu_1^R}(x) = \int_{-\infty}^{x} (\mu_1^R - \mu_0^R)((-\infty, s])\, \mathrm{d}s = \int_{-\infty}^{x} (x-s)(\rho_{\mu_1^R}(s) - \rho_{\mu_0^R}(s))\, \mathrm{d}s > 0.$$

Finally, if $R > M$ and $x \in [M, R)$, we have

$$\Phi_{\mu_0^R \to \mu_1^R}(x) = \int_{x}^{\infty} (\mu_1^R - \mu_0^R)([s, +\infty))\, \mathrm{d}s = \int_{x}^{\infty} (s-x)(\rho_{\mu_1^R}(s) - \rho_{\mu_0^R}(s))\, \mathrm{d}s > 0.$$

Thus, for all R large enough and all $x \in (-R, R)$, we have $\Phi_{\mu_0^R \to \mu_1^R}(x) > 0$. This proves iv) and thus concludes the proof. □

We will choose and fix a value $R_0 \geq 1$ such that for any $R > R_0$, there exists a continuous Brownian transport from μ_0^R to μ_1^R, and we will consider the corresponding family of stopping functions f_R.

A next step is to assure the possibility of extracting a convergent subsequence from the family of functions f_R.

Proposition 4.2. *The family (f_R) is precompact in the topology of uniform convergence on the compact sets.*

This proposition, due to the Arzelà-Ascoli theorem, is equivalent to the union of the following two results.

Lemma 4.3. *The family of functions (f_R) is locally uniformly bounded: for any interval $I = [-\ell, \ell]$, there exists $C' = C'(\ell)$ such that $\forall R \geq R_0$, we have $f_R|_I \leq C'$.*

Proposition 4.4. *Let μ_0, μ_1 be two probability measures, supported on a finite or infinite interval $I \subset \mathbb{R}$, for which there exists a continuous Brownian transport from μ_0 to μ_1 with some stopping function f. Assume that, for an interval $I' \subset I$ and a constant $C' > 0$, the following holds:*

i) μ_0, μ_1 *satisfy the hypotheses of Theorem 1.3 on* $\mathcal{U}_1(I') \cap I$.

ii) $f|_{\mathcal{U}_1(I')\cap I} \leq C'$.

iii) $\mu_0|_{I'}$ *and* $\mu_1|_{I'}$ *satisfy the hypotheses of Theorem 1.3 for some constants* a', b', α'.

Let $\delta_0 := \min\{\frac{\varepsilon}{3\theta_0\alpha'}, \frac{1}{2}\}$. *Then the inverse of the modulus of continuity of* $f|_{I'}$, *denoted by* $\delta_{f|_{I'}}(\varepsilon)$, *is lower bounded by:*

$$\delta_{f|_{I'}}(\varepsilon) \geq \frac{\varepsilon\pi \cdot a'}{2\delta_0 \cdot b'} \exp\{-\frac{\pi^2 C'}{\delta_0^2}\}. \tag{16}$$

Proof of Lemma 4.3. We will first prove that the functions f_R "take small values somewhere". Namely, that there exist some constants ℓ_1, C'' such that $\forall R \geq R_0$, $\exists x \in [-\ell_1, \ell_1]$: $f_R(x) \leq C''$. Indeed, as we have already mentioned, the functions $\Phi^R := \Phi_{\mu_0^R \to \mu_1^R}$ converge to the function $\Phi := \Phi_{\mu_0 \to \mu_1}$. In particular, the values $\Phi^R(0)$ are uniformly bounded by some constant C_1.

Now, let us consider a Brownian motion started from $\mu_0|_{[-1,1]}$. Its density ρ_{BM} at 0 has an asymptotics of $\frac{1}{\sqrt{t}}$ and thus, its integral diverges. Hence, there exists C'' such that

$$\int_0^{C''} \rho_{BM}(t)\ \mathrm{d}t > C_1. \tag{17}$$

By continuity, (17) holds also in the case of the density ρ of the process starting with an initial measure $\mu_0^R|_{[-1,1]} > \mu_0|_{[-1,1]}$, and which trajectories are stopped outside a large enough interval $[-\ell_1, \ell_1]$. Hence, for any R large enough (so that $\mu_0^R|_{[-1,1]}$ is close enough to $\mu_0|_{[-1,1]}$), there exists $x \in [-\ell_1, \ell_1]$ such that $f(x) \leq C''$. Indeed, otherwise, we would have an inequality

$$\int_0^{C''} \rho_t^R(0)\ \mathrm{d}t > \Phi_{\mu_0^R \to \mu_1^R}(0),$$

which would be a contradiction.

Now, for the Brownian transport from μ_0^R to μ_1^R, let us consider the total measure $\nu_t(\mathbb{R} \setminus K_t)$ of the not yet stopped trajectories at some time t. Note that, due to the recurrence of the Brownian motion on $\mathbb{R}$: $\forall \varepsilon > 0$, $\forall \ell_2$, there exists a time $\bar{t} = \bar{t}(\varepsilon, \ell_2)$ such that for any $x \in [-\ell_2, \ell_2]$, a Brownian trajectory, starting at x, crosses the rectangle $[-\ell_1, \ell_1] \times [C'', \bar{t}]$ left to right with probability at least $1 - \varepsilon$.

Choose now ℓ_2 large enough so that $\forall R \geq R_0$, $\mu_0^R([-\ell_2, \ell_2]) \geq 1 - \varepsilon$. Then, for any $R \geq R_0$, the total measure $\nu_{\bar{t}}(\mathbb{R} \setminus K_{\bar{t}})$ of the not yet stopped trajectories at time $\bar{t}$ will be at most 2ε, as crossing the rectangle implies stopping due to the choice of ℓ_1 and C''. In particular, taking

$$\varepsilon := \frac{1}{4} \min(\mu_0(-\ell - 1, -\ell), \mu_0(\ell, \ell + 1)),$$

we see that

$$\begin{aligned} \nu_{\bar{t}}(\mathbb{R} \setminus K_t) &\leq \frac{1}{2}\mu_0(-\ell - 1, -\ell) \leq \mu_0^R(-\ell - 1, -\ell) \\ \nu_{\bar{t}}(\mathbb{R} \setminus K_t) &\leq \frac{1}{2}\mu_0(\ell, \ell + 1) \leq \mu_0^R(\ell, \ell + 1). \end{aligned}$$

Hence, any connected component of $\mathbb{R} \setminus K_{\bar{t}}$ that intersects $I = (-\ell, \ell)$ is contained in $(-\ell - 1, \ell + 1)$.

Applying now Lemma 3.4 for all the connected components of $\mathbb{R} \setminus K_{\bar{t}}$ that intersect I, we conclude that all of them disappear in at most time $\theta \cdot \alpha_{\ell+1} \cdot |[-\ell - 1, \ell + 1]|$. Hence $\forall R \geq R_0$, $f_R|_{[-\ell,\ell]} \leq \bar{t} + \theta \cdot \alpha_{\ell+1} \cdot (2\ell + 2)$ and we have the desired upper bound. □

We are now ready to prove the uniform continuity for the family f_R, that is Proposition 4.4. A basic idea here is the following one: assume that the function f is smooth and (piecewise) monotonic. Then, considering a point x in a neighborhood of which f is monotonically increasing, we see that between the moments $t = f(x)$ and $t + \Delta t = f(x + \Delta x)$, the left end of the interval of complement to K_t absorbs approximatively the mass $\Delta t \cdot \rho_t'(x)$ of Brownian particles and this should be equal to the mass μ_1 of the interval $[x, x + \Delta x]$. Hence,

$$\Delta t \approx \frac{\mu_1([x, x + \Delta x])}{\rho_t'(x)} \approx \frac{\rho_{\mu_1}(x)}{\rho_t'(x)} \cdot \Delta x.$$

Estimating from above the numerator by b, and from below the denominator (by a comparison with the heat equation on an interval), we obtain the desired bound for $f' = \frac{\Delta t}{\Delta x}$. Let us now make these computations rigorous.

Proof of Proposition 4.4. Note first that Lemma 3.4 guarantees that the functions $f|_I$ cannot have "high thin peaks": if $y, z \in \mathcal{U}_1(I') \cap I$ and $f(y) = f(z)$, then

$$\max_{x \in [y,z]} f(x) \leq f(y) + \theta\alpha' \cdot |[y, z]|.$$

Now, take $\delta = \delta_0 = \min\left(\frac{\varepsilon}{\theta_0\alpha'}, \frac{1}{2}\right)$ and let us show the estimate (16). Namely, assume first that $x, y \in I'$ with the distance between x and y less than the right hand-side of (16). We want to show that $|f(x) - f(y)| \leq \varepsilon$. Without any loss of generality, we can assume that $f(x) < f(y)$. We can also assume that $\forall x' \in [x, y]$, $f(x') > f(x)$ (as otherwise, we can replace x with the rightmost point x' of the level set $f^{-1}(f(x)) \cap [x, y]$).

Consider now the behaviour of f on $[x, x + \delta_0]$. Denote $t_1 = f(x)$ and $t_2 = \min_{[y, x+\delta_0]} f$. Due to Lemma 3.4 and the choice of δ_0, we have

$$f(y) \leq \max(t_1, t_2) + \theta\alpha'\delta_0 \leq \max(t_1, t_2) + \frac{\varepsilon}{2}.$$

Thus, if $t_2 \leq t_1 + \frac{\varepsilon}{2}$, everything is proven. (In particular, this rules out the case of $x + \delta_0$ falling outside I: the lower limit of f at an endpoint of I is zero.)

Thus, we can assume that $t_2 > t_1 + \frac{\varepsilon}{2}$. Consider now the Brownian paths of the process X_t that were not stopped. Note that any such path, starting anywhere in $[x, x + \delta_0]$, stays in this interval till the moment t_1 and then leaves it through the left end before the moment t_2, as shown in Figure 3 below. The first intersection point of such a path with the graph of f is somewhere above $[x, y]$. Hence, the measure $\mu_1([x, y])$ is greater or equal to the measure of such paths.

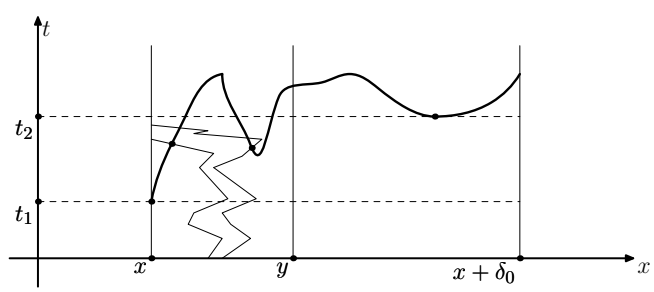

Figure 3: Two Brownian paths crossing the graph of f

Finally, we can easily estimate this measure from below through the heat equation. Namely, the condition $\rho_{\mu_0}|_I \geq a$ allows us to estimate the initial density on $[x, x + \delta]$ from below by an eigenfunction of the Laplace operator, that is $u(z) = a \sin \frac{\pi(z-x)}{\delta_0}$ with the eigenvalue $\lambda = \frac{\pi^2}{\delta_0^2}$. Hence, the density of the trajectories that have never left $[x, x + \delta]$ up to time t is greater

than $e^{-\lambda t} \cdot a' \sin \frac{\pi(z-x)}{\delta_0}$, and thus the density of those who are first-leaving the interval through its left end is at least $a' \frac{\pi}{\delta_0} e^{-\lambda t}$. The total mass of the trajectories leaving between the moments t_1 and t_2 is

$$\int_{t_1}^{t_2} a \frac{\pi}{\delta_0} e^{-\lambda t} \, \mathrm{d}t \geq a' \frac{\pi}{\delta_0} (t_2 - t_1) e^{-\lambda t_2}.$$

As we have $t_2 - t_1 \geq \frac{\varepsilon}{2}$ and $t_2 \leq C'$, we finally have obtained a lower bound for the total mass of such trajectories and thus for $\mu_1([x, y])$. This lower bound is given by

$$a' \frac{\pi \cdot \varepsilon}{\delta_0 \cdot 2} e^{-C' \pi^2 / \delta_0^2}.$$

Though, due to our assumption, $\mu_1([x, y]) \leq b'(y - x)$, and due to our choice of $\delta(\varepsilon)$, this gives us a contradiction. $\square$

Having proved both Lemma 4.3 and Proposition 4.4, we have thus proved Proposition 4.2. We are now ready start concluding the proof of Theorem 1.2. Namely, as the family (f_R) is precompact, there exists a convergent subsequence $f_{R_k} \xrightarrow[k \to \infty]{} f$. A natural conclusion would then be that the first intersection measure with the graph of f for the initial measure $\mu_0 = \lim \mu_0^{R_k}$ is exactly $\mu_1 = \lim \mu_1^{R_k}$. To make this argument work rigorously, we will need the following

Definition 4. Let $f \in C(\mathbb{R}, \mathbb{R}_+)$ be a continuous positive function and $x \in \mathbb{R}$. The *first intersection measure* $m_{x,f}$ is defined as the law of the $x-$coordinate of the first intersection between the graph of f and the trajectory of the Brownian motion started from the point x: $X_t = x + B_t$, $T = \inf\{t \geq 0 : \ t = f(X_t)\}$ and $m_{x,f} = \ \mathrm{Law}(X_T)$. Similarly, we denote by $m_{\mu,f}$ the first intersection measure between the process started from the distribution μ and the graph of the stopping function f.

Proposition 4.5. *The first intersection measure $m_{x,f}$ depends continuously (in the sense of the weak* convergence) on $x \in \mathbb{R}$ and $f \in C(\mathbb{R}, \mathbb{R}_+)$ (where $C(\mathbb{R}, \mathbb{R}_+)$ is equipped with the topology of uniform convergence on compact sets).*

The following lemma is an easy exercise

Lemma 4.6. *Denote by $(X_t, t \geq 0)$ the standard Brownian motion. For all $\varepsilon > 0$, there exists $\delta > 0$ such that, with probability at least $1 - \varepsilon$, there*

exist $t_+, t_- \in [\delta, \varepsilon]$ *such that* $X_{t_+} = \delta$, $X_{t_-} = -\delta$ *and* $\sup_{0 \le t \le \max(t_+, t_-)} |X_t| \le \varepsilon$. *In other words, the Brownian motion crosses horizontally the rectangle* $[-\delta, \delta] \times [\delta, \varepsilon]$, *and before this crossing, it stays inside the strip* $[-\varepsilon, \varepsilon] \times \mathbb{R}_+$ *(see Figure 4 below).*

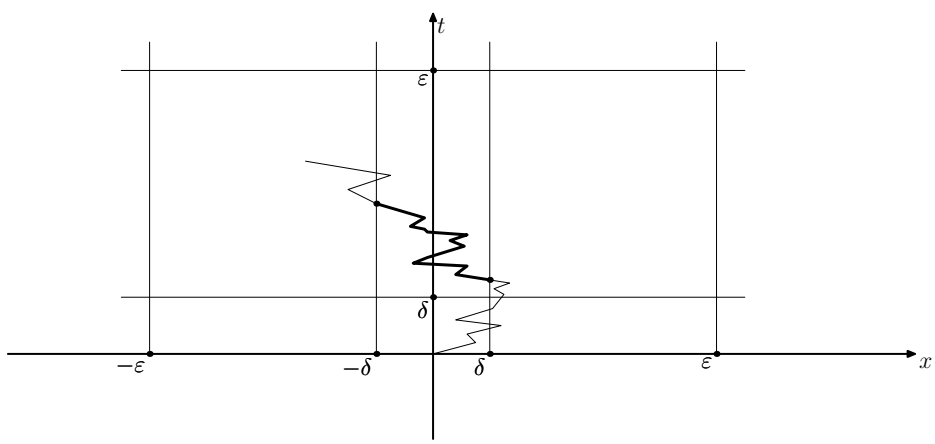

Figure 4: A Brownian path crossing the strip

Proof of Proposition 4.5. Let $f_1 \in C(\mathbb{R}, \mathbb{R}_+)$ and $x_1 \in \mathbb{R}$ be given. Take an arbitrary $\varepsilon > 0$ and let $\delta > 0$ be defined by Lemma 4.6. It is easy to see that, for some $R > 0$, for any initial point $x \in \mathcal{U}_1(x_1)$ and for any f such that $|f(x_1) - f_1(x_1)| \le 1$, the Brownian motion started at x intersects f before leaving the strip $[-R, R] \times \mathbb{R}_+$ with probability at least $1 - \varepsilon$.

Consider now $x_2 \in \mathcal{U}_\delta(x_1)$ and $\|f_2 - f_1\|_{C([-R-\delta, R+\delta])} \le \delta$. We will estimate the difference between m_{x_1, f_1} and m_{x_2, f_2}. To do this, take the trajectory of the same Brownian motion B_t shifted to the initial points x_1 and x_2: $X_t^1 = x_1 + B_t$ and $X_t^2 = x_2 + B_t$.

Consider the moment of the first intersection of these processes with the corresponding graphs. Let $T_j := \inf\{t \ge 0 : \ t = f_j(X_t^j)\}$ for $j = 1, 2$ and $T := \min(T_1, T_2)$. Note that T_1 and T_2 are two Markov hitting times and hence, the conditional behaviour of X_t^j under any condition $T = T_0$ and $X_{T_0}^j = \bar{x}_j$ is simply the Brownian motion shifted to the initial point $(T_0, \bar{x}_j)$. See Figure 5 below.

Now, let us prove that we have $|X_{T_1}^1 - X_{T_2}^2| \le \varepsilon$ with probability at least $(1 - \varepsilon)^2$. To show this, we first note that, due to the choice of R, we

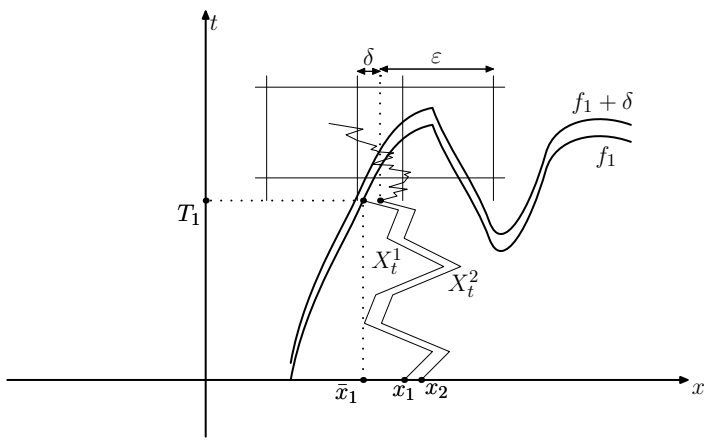

Figure 5: First intersection of X^1, X^2 with the graphs of f_1 and $f_1 + \delta$ respectively

have $X^j_T \in \mathcal{U}_{R+\delta}(x_1)$ with probability at least $1-\varepsilon$. Now, under any "first intersection condition" $T_2 \geq T_1 = \bar{t}$, $X^1_{\bar{t}} = \bar{x}_1 \in \mathcal{U}_R(x_1)$, the trajectory of $X^2_{\bar{t}}$ intersects the graph of f_2 inside $\mathcal{U}_{\varepsilon+\delta}(\bar{x}_1) \times [\bar{t}, \bar{t}+\varepsilon]$ with probability at least $1-\varepsilon$. Indeed, under this condition, the trajectory of $X^2_{\bar{t}}$ is the trajectory of the Brownian motion started from the point $(\bar{t}, X^2_{\bar{t}})$. Meanwhile, we have $|X^2_{\bar{t}} - X^1_{\bar{t}}| = |x_2 - x_1| \leq \delta$. Also, we have $f_2(\bar{x}_1) \leq f_1(\bar{x}_1) + \delta$. Recalling the definition of δ, we obtain the desired estimate on the conditional probability.

In the same way, under any condition $T_1 \geq T_2 = \bar{t}$ and $X^2_{\bar{t}} = \bar{x}_2 \in \mathcal{U}_R(x_1)$, we have $|X^1_{T_1} - X^2_{T_2}| \leq \varepsilon + \delta$ with probability at least $1-\varepsilon$. Considering the first intersection moment, we see that, with probability at least $(1-\varepsilon)$, the corresponding point belongs to $\mathcal{U}_R(x_1)$, and conditionally to it we have $|X^1_{T_1} - X^2_{T_2}| \leq \varepsilon + \delta$ with probability at least $1-\varepsilon$. Hence, we have finally

$$\mathbb{P}(|X^1_{T_1} - X^2_{T_2}| \leq \varepsilon + \delta) \geq (1-\varepsilon)^2. \tag{18}$$

As $m_{x_1,f_1} = \mathrm{Law}(X^1_{T_1})$ and $m_{x_2,f_2} = \mathrm{Law}(X^2_{T_2})$, (18) gives us the desired comparison between these two measures. □

As it can be easily seen from the latter proof, the continuity in Proposition 4.5 is uniform for x belonging to any compact set in $\mathbb{R}$.

For further arguments, it will be useful to consider the following distance between probability measures

Definition 5. Let μ, μ' be two probability measures. We define the *distance* between them as $d(\mu, \mu') := \inf\{\delta > 0 : \ \exists$ random variables U, V : $\mathrm{Law}(U) = \mu,\ \mathrm{Law}(V) = \mu'$ and $\mathbb{P}(|U - V| \le \delta) \ge 1 - \delta\}$.

It is easy to see that this distance defines on the space of probability measures precisely the weak* convergence. In fact, in the proof of Proposition 4.5, we obtain the estimate

$$d(m_{x_1,f_1}, m_{x_2,f_2}) \le \max(1 - (1-\varepsilon)^2, \varepsilon + \delta) \le 2\varepsilon.$$

Now, let us pass to the first intersection measures starting from arbitrary initial distributions

Lemma 4.7. *Let $\mu_0^{(k)} \to \mu_0$ be a weak* convergent sequence of measures, and $f_{(k)}, f \in C(\mathbb{R}, \mathbb{R}_+)$ be such that $f_{(k)} \to f$ uniformly on any compact set. Then, $m_{\mu_0^{(k)}, f_{(k)}} \xrightarrow[k\to\infty]{} m_{\mu_0, f}$.*

If additionally, the corresponding expectations of the first intersection times $T_{(k)}$ are uniformly bounded by some constant C, then the expectation of the first intersection time T is also finite and does not exceed C.

Proof. Indeed, for any $\varepsilon > 0$, there exist $\ell_1, \ell_2, \delta > 0, \delta \le \varepsilon$ such that

i) $\mu_0(-\ell_1, \ell_1) \ge 1 - \varepsilon$

ii) if $|x| \le \ell_1$, $|y - x| \le \delta$ and $\|f - \tilde{f}\|_{C([-\ell_1-\ell_2, \ell_1+\ell_2])} \le \delta$, then we have $d(m_{x,f}, m_{y,\tilde{f}}) \le \varepsilon$.

(The second conclusion comes from the uniform version of Proposition 4.5).

For any k large enough, we have $d(\mu_0, \mu_0^{(k)}) < \delta$. Hence, for any such k, we can choose the processes X^1, X^2 such that $\mathrm{Law}(X_0^1) = \mu_0$, $\mathrm{Law}(X_0^2) = \mu_0^{(k)}$, $\mathrm{d}X_t^1 = \mathrm{d}X_t^2 = \mathrm{d}B_t$ and $\mathbb{P}(|X_0^1 - X_0^2| \le \delta) \ge 1 - \delta$. Then, with probability at least $1 - \delta - \varepsilon$, we have

$$|X_0^1| \le \ell_1 \quad \text{and} \quad |X_0^1 - X_0^2| \le \delta. \tag{19}$$

Due to the property ii), the conditional probability of $|X_T^1 - X_{T_k}^2| \le \varepsilon$ is at least $1 - \varepsilon$ under the condition (19), where $T = \inf\{t \ge 0 : \ t = f(X_t^1)\}$ and $T_{(k)} = \inf\{t \ge 0 : \ t = f_{(k)}(X_t^2)\}$ are first intersection stopping times.

Hence, with probability at least $1 - \delta - 2\varepsilon$, we have $|X_T^1 - X_{T_2}^2| \le \varepsilon$ and hence

$$d(m_{\mu_0,f}, m_{\mu_0^{(k)}, f_{(k)}}) \le \delta + 2\varepsilon \le 3\varepsilon.$$

As ε is arbitrarily chosen, we have $m_{\mu_0^{(k)},f_{(k)}} \xrightarrow[k\to\infty]{} m_{\mu_0,f}$.

Now, let us prove the second statement of the lemma. Actually, for any k large enough, and any realization as before, we have $|T - T_{(k)}| \leq \varepsilon$ with probability at least $1-\delta-2\varepsilon \geq 1-3\varepsilon$. Thus, we have obtained a lower bound for the integral of T over a set of probability $1-3\varepsilon$, which is $\mathbb{E}T_{(k)}+\varepsilon \leq C+\varepsilon$. As $\varepsilon > 0$ is arbitrary, this implies that $\mathbb{E}T \leq C$. □

We can now conclude the proof of Theorem 1.2.

Proof of Theorem 1.2. We have now constructed continuous Brownian transports from $\mu_0^{R_k}$ to $\mu_1^{R_k}$ with stopping functions f_{R_k} converging uniformly on compact sets to some continuous function f. Then, due to the first part of Lemma 4.7, we have

$$m_{\mu_0,f} = \lim_{k\to\infty} m_{\mu_0^{R_k},f_{R_k}} = \lim_{k\to\infty} \mu_1^{R_k} = \mu_1.$$

The expectations of the corresponding passage times $T_{(k)}$ are also equal to

$$\mathbb{E}T_{(k)} = \mathbb{V}\mathrm{ar}\,\mu_1^{R_k} - \mathbb{V}\mathrm{ar}\,\mu_0^{R_k}$$

and thus, due to the choice of $\mu_0^{R_k}, \mu_1^{R_k}$, the latter difference converges to $\mathbb{V}\mathrm{ar}\,\mu_1 - \mathbb{V}\mathrm{ar}\,\mu_0 < \infty$. Hence, these expectations are uniformly bounded and due to the second part of Lemma 4.7, we have $\mathbb{E}T < \infty$. We have finally constructed a continuous Brownian transport from μ_0 to μ_1. □

4.2 Brownian transport on an interval: discretization

4.2.1 Discretization

We are now going prove Theorem 1.3. As we have already mentioned, we will do it by means of a discretization procedure, replacing the Brownian motion by a discrete random walk, and then passing to the limit as the mesh of the lattice goes to zero.

We will first study a discretized version of our problem. Namely, instead of a Brownian motion on $\mathbb{R}$, we consider a random walk on $\mathbb{Z}$:

$$Y_{t+1} = \begin{cases} Y_t + 1, & \text{with probability } 1/2, \\ Y_t - 1, & \text{with probability } 1/2. \end{cases}$$

We have to modify the setting of a continuous Brownian transport in the following way. The stopping time T is now a probabilistic Markov moment, that is related to the new function g in the following way:

$$\begin{cases} \text{if } t > g(Y_t), \text{ then the process is stopped,} \\ \text{if } t = g(Y_t), \text{ then the process is stopped with probability } q(Y_t), \end{cases} \tag{20}$$

where $q : \mathbb{Z} \to [0,1]$ is a new auxiliary function. A Brownian transport in this setting will be called a *discrete Brownian transport*.

The new discrete functions corresponding to Φ are defined as

$$\Phi_\mu^{\mathbb{Z}}(x) = \sum_{y<x} \sum_{z \le y} \mu(z) = \sum_{z<x} (x-z)\mu(z),$$

and $\Phi^{\mathbb{Z}}_{\mu_0 \to \mu_1}(x) := \Phi^{\mathbb{Z}}_{\mu_1}(x) - \Phi^{\mathbb{Z}}_{\mu_0}(x)$. It is then easy to check that for a centered measure μ on $\mathbb{Z}$ and for an integer x, one has $\Phi_\mu(x) = \Phi^{\mathbb{Z}}_\mu(x)$. So, we will in further mostly omit the upper index "$\mathbb{Z}$". The discrete function Φ works in the same way as its continuous analogue: an easy computation shows that

$$\Phi_{\ldots 00100 \ldots \to \ldots 0 \frac{1}{2} 0 \frac{1}{2} 0 \ldots}(x) = \frac{1}{2}\delta_0(x).$$

Hence, we have for any displacement defined by (20)

$$\Phi_{\nu_t \to \nu_{t+1}}(x) = \frac{1}{2} \cdot \begin{cases} \nu_t(x), & \text{if } g(x) > t, \\ 0, & \text{if } g(x) < t, \\ \nu_t(x) \cdot q(x), & \text{if } g(x) = t. \end{cases} \tag{21}$$

This allows us, for two centered measures μ_0, μ_1, to define recursively the transport process in the following way:

i) Initial state: $K_{-1} = \emptyset$.

ii) Evolution: for any $t \ge 0$, any $x \in \mathbb{Z} \setminus K_{t-1}$, if $\Phi_{\nu_t \to \mu_1}(x) > \frac{1}{2}\nu_t(x)$, where ν_t is the occupation measure at time t, there is nothing to be done. Otherwise, take $g(x) := t$ with $q(x) = 2\frac{\Phi_{\nu_t \to \mu_1}(x)}{\nu_t(x)}$ (and 0 if $\Phi_{\nu_t \to \mu_1}(x) = \nu_t(x) = 0$).

Due to (21), we then have

$$\Phi_{\nu_{t+1} \to \mu_1}(x) = \Phi_{\nu_t \to \mu_1}(x) - \min\left(\frac{1}{2}\nu_t(x), \Phi_{\nu_t \to \mu_1}(x)\right).$$

In particular, we can easily see by induction that all the functions $\Phi_t := \Phi_{\nu_t \to \mu_1}$ are non-negative, and the procedure is thus well-defined for all t. Also the latter construction implies the following:

i) if at some time t, at cell x, we have $\Phi_{\nu_t \to \mu_1}(x) = 0$, then the cell (t, x) is frozen and any particle coming to it at this moment (or afterwards) is stopped,

ii) if $\Phi_{\nu_t \to \mu_1}(x) \geq \frac{1}{2}\nu_t(x)$, then the cell (t, x) is fully diffused,

iii) if $0 < \Phi_{\nu_t \to \mu_1}(x) < \frac{1}{2}\nu_t(x)$, then the cell (t, x) is "partially frozen", meaning that a part of the particles of total measure $2\Phi_{\nu_t \to \mu_1}(x)$ is diffused, whereas the others are frozen. In this case, $\Phi_{\nu_{t+1} \to \mu_1}(x) = 0$, so that, starting from the moment $t+1$, the cell x becomes fully frozen.

We have the following

Proposition 4.8. *Let μ_0, μ_1 be two centered measures on $\mathbb{Z}$, both with finite support. Suppose that μ_1 is everywhere positive on the interval $I := [\min \mathrm{Supp}(\mu_0), \max \mathrm{Supp}(\mu_0)]$ and $\Phi_{\mu_0 \to \mu_1} \geq 0$. Then, the procedure* (20) *provides us with everywhere defined functions g, q that define a discrete bounded Brownian transport from μ_0 to μ_1.*

To prove this result, we will first need the following lemma, which is a discrete analogue of Lemma 3.3.

Lemma 4.9. *Let μ, ν be two centered (discrete) measures of finite support. Suppose that $\Phi_{\nu \to \mu} \geq 0$ and $\Phi_{\nu \to \mu}(x) = \Phi_{\nu \to \mu}(y) = 0$ for some $x < y$. Then, we have $\mu([x, y]) \geq \nu([x, y]) \geq \nu([x+1, y-1]) \geq \mu([x+1, y-1])$.*

Proof. Note that $\nu(z) = (\Phi_\nu(z+1) - \Phi_\nu(z)) - (\Phi_\nu(z) - \Phi_\nu(z-1))$. Taking the difference between such representations for $\mu(z)$ and $\nu(z)$, and summing up on $z \in [x+1, y-1]$, we have

$$\begin{aligned} \sum_{z \in [x+1, y-1]} (\mu(z) - \nu(z)) &= (\Phi_{\nu \to \mu}(y) - \Phi_{\nu \to \mu}(y-1)) - (\Phi_{\nu \to \mu}(x+1) - \Phi_{\nu \to \mu}(x)) \\ &= -\Phi_{\nu \to \mu}(y-1) - \Phi_{\nu \to \mu}(x+1). \end{aligned}$$

Hence, we get

$$\nu([x+1, y-1]) - \mu([x+1, y-1]) = \Phi_{\nu \to \mu}(y-1) + \Phi_{\nu \to \mu}(x+1).$$

On the other hand, summing on $z \in [x, y]$, we have

$$\begin{aligned}\sum_{z\in[x,y]} (\mu(z) - \nu(z)) &= (\Phi_{\nu\to\mu}(y+1) - \Phi_{\nu\to\mu}(y)) - (\Phi_{\nu\to\mu}(x) - \Phi_{\nu\to\mu}(x-1)) \\ &= \Phi_{\nu\to\mu}(y+1) + \Phi_{\nu\to\mu}(x-1) \geq 0.\end{aligned}$$

Thus, we conclude that

$$\mu([x, y]) \geq \nu([x, y]) \geq \nu([x+1, y-1]) \geq \mu([x+1, y-1]).$$

□

Proof of Proposition 4.8. Consider the value $m_t := \nu_t(\{x : \ \Phi_{\nu_t\to\mu_1}(x) > 0\})$. On one hand, the sequence (m_t) converges to 0. Indeed, ν_t is a part of the occupation measure of a random walk on $\mathbb{Z}$ with the initial distribution μ_0, that is in particular conditioned to never exit the interval $I := \ \mathrm{Supp}(\mu_1)$. The probability of staying inside I during t steps converges to 0, and thus, so does m_t. On the other hand, Lemma 4.9 implies that

$$\nu_t(\{x : \ \Phi_{\nu_t\to\mu_1}(x) > 0\}) \geq \mu_1(\{x : \ \Phi_{\nu_t\to\mu_1}(x) > 0\})$$

and thus

$$m_t \geq \sharp\{x : \ \Phi_{\nu_t\to\mu_1}(x) > 0\} \cdot \min_{z\in I} \mu_1(z).$$

As $b := \min_{z\in I} \mu_1(z) > 0$ due to the hypothesis of the proposition, once $m_t < b$, we have $\Phi_t \equiv 0$ and hence $\nu_t = \mu_1$. □

4.2.2 Proof of Theorem 1.3

We are now ready to prove Theorem 1.3. Let two centered measures μ_0 and μ_1, supported on some interval $I \subset \mathbb{R}$, be given and assume that, for these measures, the hypotheses i)-v) of the theorem are satisfied. Up to a rescaling of space and time, we can assume that $I = [-1, 1]$.

For any natural n, one can consider the discretized measures $\mu_0^{(n)}$ and $\mu_1^{(n)}$ on $\frac{1}{n}\mathbb{Z}$, defined as

$$\mu_i^{(n)}\left(\frac{k}{n}\right) = n \int_{\frac{k-1}{n}}^{\frac{k+1}{n}} \left(1 - \left|x - \frac{k}{n}\right|\right) \, \mathrm{d}\mu_i(x), \quad i = 0, 1. \tag{22}$$

Note that the measures $\mu_0^{(n)}$ and $\mu_1^{(n)}$ are supported on the sets $\{-1, \frac{-n+1}{n}, \ldots, \frac{n-1}{n}, 1\}$, and have the same mean.

Consider now the corresponding random walks (with the elementary time step $\frac{1}{n^2}$) and the corresponding functions

$$\Phi^{\frac{1}{n}\mathbb{Z}}_{\mu_i^{(n)}}\left(\frac{k}{n}\right) = \sum_{y<x,\ y\in\frac{1}{n}\mathbb{Z}} \left(y - \frac{k}{n}\right) \mu_i^{(n)}(y)$$

which, as earlier for $\mathbb{Z}$, are the restrictions on $\frac{1}{n}\mathbb{Z}$ of the continuous functions $\Phi_{\mu_i^{(n)}}(x)$. A first step in applying the discretization technique is a check that there exists a discrete Brownian transport from $\mu_0^{(n)}$ to $\mu_1^{(n)}$.

Lemma 4.10. *For any n large enough, the measures $\mu_0^{(n)}$ and $\mu_1^{(n)}$ satisfy the hypotheses of Proposition 4.8.*

Proof. Note that the functions $\Phi_{\mu_0^{(n)}\to\mu_1^{(n)}}$ converge uniformly to the function $\Phi_{\mu_0\to\mu_1}$ that is positive inside I. Hence,

$$\forall\delta\ \exists n_0:\ \forall n > n_0 \quad \Phi_{\mu_0^{(n)}\to\mu_1^{(n)}}|_{I\setminus\mathcal{U}_\delta(\partial I)} > 0.$$

On the other hand, due to the assumption v), we have

$$\exists n_1:\ \forall n \geq n_1,\ \forall x \in \mathcal{U}_\varepsilon(\partial I)\cap I\cap\frac{1}{n}\mathbb{Z} \quad \mu_1^{(n)}(x) > \mu_0^{(n)}(x),$$

what assures $\Phi_{\mu_0^{(n)}\to\mu_1^{(n)}}|_{\mathcal{U}_\varepsilon(\partial I)\cap I\cap\frac{1}{n}\mathbb{Z}} \geq 0$. Choosing then $\delta = \varepsilon/2$, we see that $\Phi_{\mu_0^{(n)}\to\mu_1^{(n)}}$ is positive everywhere on I once n is large enough. □

Consider now the corresponding discrete transport functions $g_{(n)}(x)$ that we extend to $[-1,1]$ piecewise linearly. Note that, for these functions, we still have the (uniform in n) estimates, analogous to Lemma 3.4 and Proposition 4.5 (proven by the same methods). Hence, the family of functions $g_{(n)}$ is precompact and we can extract a convergent subsequence $g_{(n_k)} \to f$. On the other hand, discrete random walks tend, as $n\to\infty$, to the Brownian motion. Hence, the same arguments as in Proposition 4.5 and Lemma 4.7 imply that the first intersection measure for the initial distribution $\mu_0 = \lim_{k\to\infty}\mu_0^{(n_k)}$ with the stopping function $f = \lim_{k\to\infty} g_{(n_k)}$ will be $\lim_{k\to\infty}\mu_1^{(n_k)} = \mu_1$. This concludes the proof of Theorem 1.3.

Acknowledgement: V.K. thanks S. Pirogov, A. Sobolevski, B. Gurevich, A. Rybko and S. Shlosman for interesting discussions. A.K. thanks Bernard Roynette for telling her the Cantelli conjecture.

References

[1] Benaïm M. & El Karoui N. [2004], *Promenade aléatoire*, Ed. Ecole Polytechnique.

[2] Cantelli F.P. [1918], Sullo schema lexiano della dispersione ipernormale, *Memorie Accad. Naz. Lincei* (5) **12**, 395-411.

[3] De Meyer B., Roynette B., Vallois P. & Yor M. [2002], On independent times and positions for Brownian motions, *Rev. Mat. Iberoamericana* **18**, no. 3, 541–586.

[4] Dudley R.M. [1974], The Gaussian Process and How to Approach It, *Proceedings of the International Congress of Mathematicians, Vancouver*, vol. 2, 143-146.

[5] Fitzsimmons P.J. [1988], Penetration time and Skorohod stopping, *Sem. Probab.* **22**, 166-174.

[6] Letac G. & Malliavin P. [1995], *Integration and Probability (Exercises and solutions)*, Springer Verlag.

[7] Obłój J. [2004], The Skorokhod embedding problem and its offspring, *Probability Surveys*, vol. 1, 321-392.

[8] Rost H. [1971], The stopping distributions of a Markov process, *Inventiones math*, vol. 14 **1**, 1-16.

[9] Rubinstein L.I. [1994], *The Stefan Problem*, Translations of Mathematical Monographs, 27.

[10] Shih C.T. [1990], Ordered Skorokhod Stopping for a Sequence of Measures, *Ann. Probab.*, vol. 18 **4**, 1623-1634.

[11] Skorokhod A.V. [1961], *Issledovaniya po teorii sluchainykh protsessov (Stokhasticheskie differentsialnye uravneniya i predelnye teoremy dlya protsessov Markova)*, (Studies in the theory of random processes, Translated from the Russian in 1965), Izdat. Kiev. Univ., Kiev.

[12] Tortorici P. [1948], Soluzione approssimata di un'equazione integrale di Cantelli, *Statistics Annali di Matematica Pura ed Applicata* **27**(1), 75-86.

[13] Tricomi F.G. [1967], Einige ungelöste Probleme der klassischen Analysis, *Abh. Math. Sem. Univ. Hamburg* **31**, 25-32.

[14] Vuik C. [1993], Some historical notes about the Stefan problem, *Nieuw Archief voor Wiskunde* **11**, 157-167.

Victor Kleptsyn: IRMAR (UMR 6625 CNRS), Université Rennes 1, Campus de Beaulieu, F-35042 Rennes Cedex, France.
Email: victor.kleptsyn@univ-rennes1.fr

Aline Kurtzmann: Institut Elie Cartan Nancy, Université de Lorraine, B.P. 70239, F-54506 Vandœuvre-lès-Nancy Cedex, France.
Email: aline.kurtzmann@univ-lorraine.fr