

# Stochastic calculus and martingales on trees

(Calcul stochastique et martingales sur les arbres)

Jean Picard

Laboratoire de Mathématiques Appliquées (CNRS UMR 6620)

Université Blaise Pascal

63177 Aubière Cedex, France

`Jean.Picard@math.univ-bpclermont.fr`

## Abstract

Considering trees as simple examples of singular metric spaces, we work out a stochastic calculus for tree-valued processes. We study successively continuous processes and processes with jumps, and define notions of semimartingales and martingales. We show that martingales of class (D) converge almost surely as time tends to infinity, and prove on some probability spaces the existence and uniqueness of a martingale of class (D) with a prescribed integrable limit; to this end, we use either a coupling method or an energy method. This problem is related with tree-valued harmonic maps and with the heat semigroup for tree-valued maps.

## Résumé

Considérant que les arbres sont des exemples simples d'espaces métriques singuliers, nous développons un calcul stochastique pour les processus à valeurs dans les arbres. Nous étudions successivement les processus continus et avec sauts, et définissons les notions de semimartingales et martingales. Nous montrons que les martingales de classe (D) convergent presque sûrement quand le temps tend vers l'infini, et établissons sur certains espaces de probabilité l'existence et l'unicité d'une martingale de classe (D) avec limite intégrable fixée; pour cela, nous utilisons soit une méthode de couplage, soit une méthode d'énergie. Ce problème a des liens avec les applications harmoniques à valeurs dans les arbres, et avec le semi-groupe de la chaleur pour les applications à valeurs dans les arbres.

**Mathematics Subject Classification (2000).** 60G07 60G48 58J65 58E20 47H20

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Geometric preliminaries</b>	<b>7</b>
2.1	Geodesics and convex functions . . . . .	7
2.2	Barycentre . . . . .	8
2.3	Properties of the barycentre . . . . .	10
<b>3</b>	<b>Stochastic calculus on a star</b>	<b>13</b>
3.1	Semimartingales . . . . .	13
3.2	Quasimartingales . . . . .	15
3.3	Continuous martingales . . . . .	17
<b>4</b>	<b>Martingales and coupling</b>	<b>23</b>
4.1	The main result . . . . .	23
4.2	Feller property of the semigroup . . . . .	28
4.3	Stars and hyperbolic geometry . . . . .	29
<b>5</b>	<b>Martingales and energy minimisation</b>	<b>32</b>
5.1	The Dirichlet space . . . . .	32
5.2	Energy minimising maps . . . . .	36
5.3	Martingales for symmetric diffusions . . . . .	40
<b>6</b>	<b>Generalisation to trees</b>	<b>44</b>
6.1	The geometry of trees . . . . .	44
6.2	Stochastic calculus on a tree . . . . .	48
6.3	Quasimartingales . . . . .	52
6.4	Continuous martingales . . . . .	53
6.5	Martingales with prescribed limit . . . . .	55
<b>7</b>	<b>Coupling of diffusions on trees</b>	<b>59</b>
7.1	Coupling of spiders . . . . .	59
7.2	Coupling of snakes . . . . .	63
<b>8</b>	<b>Stochastic calculus with jumps</b>	<b>69</b>
8.1	Martingales with jumps . . . . .	69
8.2	Martingales with prescribed limit . . . . .	73
	<b>References</b>	<b>78</b>

# 1 Introduction

The relationship between manifold-valued harmonic maps and manifold-valued continuous martingales have been investigated in several works in the two last decades, see for instance [18, 19, 25, 28] for the stochastic construction of harmonic maps. In this type of problem, one considers two manifolds  $M$  and  $N$ . On  $M$ , one is given a second-order differential operator  $L$ , or equivalently a diffusion  $X_t$ ; for instance, if  $M$  is Riemannian, one can consider the Laplace-Beltrami operator  $L$ , or equivalently the Brownian motion  $X_t$  on  $M$ . Then one can associate to  $L$  the notions of heat semigroup and harmonic functions on  $M$ , and these notions have stochastic counterparts; for instance, it is well known that a harmonic function  $h$  transforms the diffusion  $X_t$  into a real local martingale  $h(X_t)$ . On the other hand, on the second manifold  $N$  (the target), one is given a connection (more precisely a linear connection on the tangent bundle  $T(N)$ ); for instance, if  $N$  is Riemannian, one can consider the Levi-Civita connection. The operator  $L$  acts on functions  $f : M \rightarrow \mathbb{R}$ , but the connection enables to also define it on functions  $f : M \rightarrow N$ , and one obtains a function  $L_N f : M \rightarrow T(N)$  (called the tension field). Then it is again possible to consider the notions of heat semigroup and harmonic maps; for instance, a smooth map  $h : M \rightarrow N$  is harmonic if  $L_N h = 0$  (see [16]). These notions have a stochastic interpretation; the connection enables to consider continuous martingales in  $N$  (see [24, 7, 10]) which are transformed into submartingales by convex functions, and  $h$  is harmonic if it transforms the diffusion  $X_t$  into a martingale  $h(X_t)$ . This is the stochastic analogue of the analytical property stating that a harmonic map composed with a convex function is subharmonic. In particular, the Dirichlet problem or the heat equation with values in  $N$  are strongly related to the problem of finding a continuous martingale on  $N$  with a prescribed final value. Thus

- the stochastic calculus for the diffusion  $X_t$  and the  $N$ -valued martingales can be applied to the construction and the properties of harmonic maps and of the heat semigroup; in particular, coupling properties of  $X_t$  are very useful for this purpose, see for instance [18, 19];
- conversely, a functional analytic construction of harmonic maps (such as energy minimisation when  $L$  is a symmetric operator) can be applied to the construction of a family of martingales, see [28].

A basic tool in all these studies is Itô's stochastic calculus involving smooth (at least  $C^2$ ) functions.

However, it would be interesting to consider more singular spaces  $M$  and  $N$ . In particular, an analytical theory for energy minimising maps has been

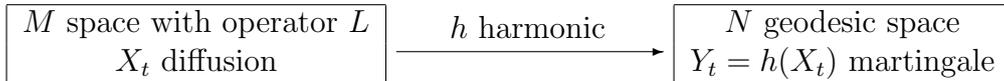


Figure 1: Stochastic interpretation of harmonic maps

worked out in [20] (see also [9] for the case of Riemannian polyhedra, and [15] for another method); a functional analytic approach to the heat semigroup is also given by [31]; it would be desirable to obtain a stochastic interpretation of these theories. For  $M$ , the analytical theory requires a harmonic structure, and the stochastic theory requires a diffusion (a continuous Markov process); the relationship between these two notions has been extensively studied for a long time (see for instance the link between regular Dirichlet forms and symmetric Hunt processes in [14]); we will not insist on it and only consider some properties of these diffusions which will be useful to us, namely their coupling properties; in particular, since this article focusses on trees, we will study the coupling properties of some classical diffusions on trees.

In this article, we will be mainly concerned by the singularity of  $N$ . In the analytical theory, the main assumption on  $N$  is that it is a metric space which is geodesic (the distance between two points is given by the minimal length of a curve joining these two points) and which has nonpositive (or at least bounded above) curvature in the sense of Alexandrov; our aim is therefore to construct a theory of martingales and semimartingales on these spaces, and to explore the links between analytical and probabilistic theories. On a geodesic space, one can consider the notion of convex functions (which are convex on geodesics parameterised by arc length), so the idea is to use stochastic calculus for convex rather than  $C^2$  functions. This point of view is already used on smooth manifolds; a continuous process is a martingale if (at least locally) any convex function maps it to a submartingale, so extending this definition to singular geodesic spaces is tempting. However, there are some difficulties with too general spaces, so here, we only consider a simple type of such spaces, namely trees. Multidimensional generalisations such as Riemannian polyhedra would of course be interesting, but are postponed to future work. In a large part of this article, we will focus on a toy example of tree, namely a star  $\mathbb{Y}^\ell$  with  $\ell$  rays  $R_i$  and a common origin  $O$ .

It appears that martingales of class (D) in  $N$  converge almost surely as in the real case, and the basic problem for the interpretation of harmonic maps is the existence and uniqueness of a martingale in  $N$  with a prescribed limit. In particular, given a diffusion  $X_t$  on a space  $M$ , we first consider limits of type  $g(X_1)$  or  $g(X_\tau)$  for a first exit time  $\tau$ ; then we are able to consider general functionals of the diffusion. Our main result will be to

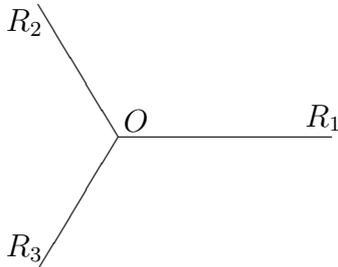


Figure 2: The space  $\mathbb{Y}^3$ .

prove the existence and uniqueness of such a martingale under two different frameworks (coupling method, energy method), and to relate it to analytical problems (heat semigroup, energy minimisation). The difference between the two frameworks lies in the assumptions on the  $M$ -valued diffusion; either it will satisfy some coupling properties, or it will be symmetric and associated to a regular Dirichlet form. The advantage of the coupling method (which is also used in [31]) is that it also yields smoothness properties on the heat semigroup and that it does not require the symmetry of the diffusion; its disadvantage is that the coupling property is not always easy to check; on the other hand, the energy method has been successfully applied in [20, 9]. As a particular framework, we will consider the case where both  $M$  and  $N$  are trees.

Another problem is to extend this theory to non continuous Markov processes on  $M$  (which are associated to non local operators  $L$ ) and non continuous martingales on  $N$ . This has been considered in [26] when  $N$  is smooth; in this case, the connection (which is a local object) is not sufficient, and one needs a global object, namely a notion of barycentre. Fortunately, it is known that barycentres can be constructed on geodesic spaces with nonpositive curvature, so in particular on trees. Thus we want to use them and deduce a notion of martingale with jumps. The definitions which were used in [26] in the smooth case cannot be handled in the case of trees, so we suggest a new definition (which probably is also useful in the smooth case). Then we can extend the results of the continuous case to this setting.

Let us outline the contents of this article. We begin with some geometric preliminaries such as the construction and properties of barycentres in Section 2. Then we work out in Section 3 a stochastic calculus in  $N = \mathbb{Y}^\ell$  (involving semimartingales), and define a notion of continuous martingale; we will see that we also have a notion of quasimartingale. The existence of a martingale with prescribed final value is proved either from the coupling

properties of the diffusion  $X_t$  in Section 4 (Theorem 4.1.4 and Corollary 4.1.13), or by energy minimisation (when  $X_t$  is a symmetric diffusion) in Section 5 (Theorems 5.2.8 and 5.3.1). The techniques which are used for  $N = \mathbb{Y}^\ell$  are extended to a more general class of trees in Section 6 (Theorems 6.5.2 and 6.5.6).

In order to apply the results of Section 4, we will give in Section 7 examples of trees  $M$  and of diffusions on them satisfying the coupling property. For instance, if  $M$  is itself a star, we will see that the Walsh process introduced in [34] satisfies it. We will also consider some other examples such as the Evans process ([13]) and the Brownian snake ([22]).

Finally, we define in Section 8 a notion of martingale with jumps, and extend the above theory to this case (Theorems 8.2.5 and 8.2.9).

## 2 Geometric preliminaries

Let us consider the metric space  $(N, \delta)$  where  $N = \mathbb{Y}^\ell$  is the star with  $\ell$  rays (for  $\ell \geq 3$ ) and  $\delta$  is the tree distance (Figure 2). More precisely, we first consider the disjoint union of  $\ell$  rays  $(R_i, \delta_i)$ , each of them being isometric to  $\mathbb{R}_+$ ; then we glue their origins into a single point  $O$ , and we consider the distance

$$\delta(A, B) = \begin{cases} \delta_i(A, B) & \text{if } A, B \in R_i, \\ \delta_i(O, A) + \delta_j(O, B) & \text{if } A \in R_i, B \in R_j \text{ for } i \neq j. \end{cases}$$

Such a space can be embedded in an Euclidean space by choosing  $\ell$  different unit vectors  $e_i$ , and by putting

$$\mathbb{Y}^\ell = \bigcup_i R_i, \quad R_i = \{r e_i; r \geq 0\}. \quad (2.0.1)$$

Different choices for  $e_i$  lead to different isometric embeddings of the same metric space  $\mathbb{Y}^\ell$  (“isometric” means that the length of a curve in  $\mathbb{Y}^\ell$  is the same when computed for the metric of  $\mathbb{Y}^\ell$  or the Euclidean metric). One can for instance embed  $\mathbb{Y}^\ell$  in  $\mathbb{R}^2$ , but we will generally embed it in  $\mathbb{R}^\ell$  and choose  $(e_i; 1 \leq i \leq \ell)$  as the canonical basis of  $\mathbb{R}^\ell$ ; this will be called the standard embedding of  $N$  into  $\mathbb{R}^\ell$ ; then the distance in  $\mathbb{Y}^\ell$  is equal to the distance of  $\mathbb{R}^\ell$  induced by the norm  $|y| = \sum |y_i|$ ; in particular,  $|y|$  is the distance of  $y$  to  $O$ , and the coordinate  $y_i$  of  $y$  is  $|y|$  if  $y$  is in  $R_i$ , and 0 otherwise. We put  $R_i^* = R_i \setminus \{O\}$ .

### 2.1 Geodesics and convex functions

The singular manifold  $N = \mathbb{Y}^\ell$  is an example of a metric space which is a geodesic space; this means that locally (and here also globally), the distance between two points of  $N$  is the length of an arc with minimal length joining these two points; this arc is called a geodesic. Here, all these arcs are parts of the  $\ell(\ell - 1)/2$  infinite geodesics  $R_i \cup R_j$ ,  $i \neq j$ . Moreover,  $N$  has nonpositive curvature in the sense of Alexandrov (see for instance [9]; this property is crucial for our study, but we will not need the precise definition of curvature). A property which is more specific to trees is that the connected subsets of  $N$  are also its convex subsets (a subset is convex if the geodesic arc linking two points of the subset is included in the subset). Convex functions can be defined similarly to smooth manifolds.

**Definition 2.1.1.** *A real function  $f$  defined on  $N$  is said to be convex if it is convex on the geodesics of  $N$  when they are parameterised by arc length.*

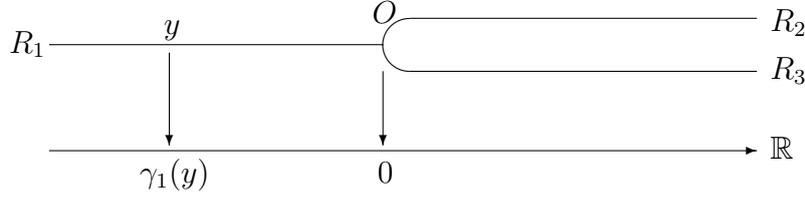


Figure 3: The Busemann function  $\gamma_1$  on  $\mathbb{Y}^3$ .

Consider the restriction  $f_i : r \mapsto f(re_i)$  of  $f$  to the ray  $R_i$ ; it is not difficult to check that  $f$  is convex if it is continuous, and the functions  $f_i$  are convex on  $\mathbb{R}_+$  and satisfy

$$f'_i(0) + f'_j(0) \geq 0 \quad \text{for } i \neq j. \quad (2.1.2)$$

Notice in particular that all but at most one of the functions  $f_i$  are non decreasing.

*Example 2.1.3.* The function  $y \mapsto |y|$ , and more generally the distance functions  $\delta(y_0, \cdot)$  are convex.

*Example 2.1.4.* The distance to a convex subset, for instance the component function  $y \mapsto y_i$ , is convex.

*Example 2.1.5.* The  $\ell$  Busemann functions  $\gamma_i$  associated to  $R_i$  are convex; these functions are defined (see Figure 3) by

$$\gamma_i(y) = \lim_{r \rightarrow \infty} (\delta(y, re_i) - r) = \sum_{j \neq i} y_j - y_i. \quad (2.1.6)$$

Thus the absolute value of  $\gamma_i(y)$  is  $|y|$ , and its sign is negative on  $R_i^*$ , positive on the other rays.

## 2.2 Barycentre

We now introduce the notion of barycentre which replaces the notion of expectation on the real line (see also [32] for more general spaces). Let  $Y$  be a square integrable variable (this means that  $|Y|$  is square integrable, or equivalently that  $\delta(y_0, Y)$  is square integrable for any  $y_0$ ). The function  $y \mapsto \delta^2(y, z)$  is strictly convex on  $N$  (this again means that it is strictly convex on the geodesics), so

$$\phi_Y(y) = \mathbb{E}\delta^2(y, Y)/2$$

is also strictly convex. It tends to  $+\infty$  at infinity and is therefore minimal at a unique point.

**Definition 2.2.1.** *The barycentre of a square integrable  $N$ -valued variable  $Y$  is defined as*

$$\mathbb{B}[Y] = \operatorname{argmin} \phi_Y.$$

The barycentre is computed by solving the variational problem. If  $\phi_i(r) = \phi_Y(r e_i)$ , then its derivative can be written with the Busemann function  $\gamma_i$  of (2.1.6) as

$$\phi'_i(r) = \mathbb{E}\gamma_i(Y) + r.$$

Thus, if there exists a  $i$  such that  $\mathbb{E}\gamma_i(Y) < 0$ , then

$$\mathbb{B}[Y] = -\mathbb{E}\gamma_i(Y) e_i.$$

On the other hand, if  $\mathbb{E}\gamma_i(Y) \geq 0$  for any  $i$ , then  $\phi_i$  is non decreasing for any  $i$ , so  $\mathbb{B}[Y] = O$ . By using the standard embedding  $N \subset \mathbb{R}^\ell$  and the linear extensions (2.1.6) of the functions  $\gamma_i$  to  $\mathbb{R}^\ell$ , we deduce that

$$\mathbb{B}[Y] = \Pi(\mathbb{E}[Y]) \tag{2.2.2}$$

with

$$\begin{aligned} \Pi : \mathbb{R}_+^\ell &\longrightarrow \mathbb{Y}^\ell \\ z &\longmapsto \Pi(z) = \sum \gamma_i(z)^- e_i. \end{aligned}$$

One can compare this result with the case of smooth Riemannian manifolds; if one uses an isometric embedding into a Euclidean space, then the barycentre is approximately (for variables with small support) the orthogonal projection of the expectation on the manifold. Here, the function  $\Pi$  can also be viewed as a projection. The inverse image of  $O$  is a cone  $C_O$ , and each of the  $\ell$  connected components  $C_i$  of  $\mathbb{R}_+^\ell \setminus C_O$  is projected onto a different ray  $R_i^*$  (see Figure 4).

A consequence of (2.2.2) is that the barycentre can be extended to integrable variables.

An equivalent way of characterising the barycentre (which will be useful for more general trees) is as follows. If  $y_0 \in N$ , then  $N \setminus \{y_0\}$  has two or  $\ell$  connected components which are denoted by  $y_0^\alpha$  for  $\alpha$  in  $\{-, +\}$  or  $\{1, \dots, \ell\}$ . The derivative at  $y_0$  of a function  $f$  in the direction of  $y_0^\alpha$  is denoted by  $\partial_\alpha f(y_0)$ . Consider (as in Section 7 of [31]) the oriented distance function

$$\psi(y_0^\alpha, y) = \delta(y_0, y) (1_{\{y \in y_0^\alpha\}} - 1_{\{y \notin y_0^\alpha\}}). \tag{2.2.3}$$

These functions are the functions  $-\gamma_i$ ,  $1 \leq i \leq \ell$  if  $y_0 = O$ , and the functions  $\pm(\gamma_j(y) - \gamma_j(y_0))$  if  $y_0 \in R_j^*$ .

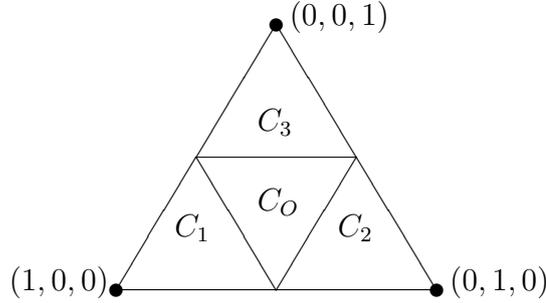


Figure 4: Inverse images of  $R_i^*$  and  $O$  for the projection  $\Pi : \mathbb{R}_+^3 \rightarrow \mathbb{Y}^3$  restricted to the triangle  $\{z_1 + z_2 + z_3 = 1\}$ .

**Proposition 2.2.4.** *If  $Y$  is a square integrable variable, then  $\mathbb{B}[Y] = y_0$  if and only if*

$$\mathbb{E}\psi(y_0^\alpha, Y) \leq 0 \tag{2.2.5}$$

for all the connected components  $y_0^\alpha$  of  $N \setminus \{y_0\}$ .

*Proof.* The derivatives of  $\phi_Y$  at  $y_0$  are

$$\partial_\alpha \phi_Y(y_0) = -\mathbb{E}\psi(y_0^\alpha, Y),$$

and  $\mathbb{B}[Y] = y_0$  if and only if all these derivatives are nonnegative.  $\square$

In particular, we have

$$\mathbb{B}[Y] = O \iff \forall i \quad \mathbb{E}\gamma_i(Y) \geq 0. \tag{2.2.6}$$

## 2.3 Properties of the barycentre

The barycentre satisfies the Jensen inequality

$$f(\mathbb{B}[Y]) \leq \mathbb{E}f(Y) \tag{2.3.1}$$

for convex Lipschitz functions  $f$  (and also for non Lipschitz functions if  $f(Y)$  is integrable); this inequality has been proved for more general metric spaces with nonpositive curvature in [9] (Proposition 12.3); for  $N = \mathbb{Y}^\ell$ , we will actually prove a generalised form (see Proposition 2.3.5 below).

*Remark 2.3.2.* Contrary to the Euclidean case, the barycentre is not characterised by the Jensen inequality; the set of points  $y_0$  satisfying  $f(y_0) \leq \mathbb{E}f(Y)$  for any convex Lipschitz function  $f$  is called the convex barycentre of  $Y$ , see [11]; the Jensen inequality says that the convex barycentre contains the

barycentre, but it generally contain other points. For instance, if  $Y$  is uniformly distributed on  $\{|y| \leq 1\}$  in  $\mathbb{Y}^3$ , its barycentre is  $O$  and its convex barycentre is  $\{|y| \leq 1/6\}$ . On smooth manifolds also, the convex barycentre is not a singleton; however, an important difference is in its size. For small enough smooth manifolds, the diameter of the convex barycentre is dominated by the third order moment of the law, see [2]; this means that

$$\delta(z_1, z_2) \leq C \inf_y \mathbb{E} \delta^q(y, Y) \quad (2.3.3)$$

for  $q = 3$  and for any  $z_1$  and  $z_2$  in the convex barycentre of  $Y$ . On our space  $N = \mathbb{Y}^\ell$ , notice that if  $z_1$  and  $z_2$  are in the convex barycentre, then

$$\delta(y, z_i) \leq \mathbb{E} \delta(y, Y)$$

for any  $y$ , because  $\delta(y, \cdot)$  is convex. Thus (2.3.3) holds with  $C = 2$  and  $q = 1$ . Actually, it is not possible to obtain a higher value for  $q$ ; for instance, if  $Y$  is uniformly distributed on  $\{|y| \leq \varepsilon\}$  in  $\mathbb{Y}^3$ , then its convex barycentre is  $\{|y| \leq \varepsilon/6\}$ , so the left and right sides of (2.3.3) are respectively of order  $\varepsilon$  and  $\varepsilon^q$ .

*Remark 2.3.4.* The functions  $\gamma_i$  are convex, so we see from (2.2.6) that  $O$  is in the convex barycentre if and only if it is the barycentre; actually, it can be seen that the convex barycentre is a closed convex subset of  $N$ , and the barycentre is the point of this subset which is the closest to the origin (this property is particular to our baby tree and cannot be extended to other trees).

**Proposition 2.3.5.** *Let  $Y$  be a square integrable variable and let  $f$  be a Lipschitz function  $f$  which is convex on all the geodesics containing  $\mathbb{B}[Y]$ . Then (2.3.1) holds true.*

*Proof.* Put  $y_0 = \mathbb{B}[Y]$ . The inequality (2.3.1) is evident if  $f$  is minimal at  $y_0$ . Otherwise, there is exactly one connected component  $y_0^\alpha$  of  $N \setminus \{y_0\}$  such that the derivative  $\partial_\alpha f(y_0)$  is negative; for all the other connected components  $y_0^\beta$ , we have

$$\partial_\beta f(y_0) \geq -\partial_\alpha f(y_0).$$

This inequality is indeed (2.1.2) if  $y_0 = O$ , and follows easily from the convexity of  $f$  on  $R_i$  if  $y_0$  is in  $R_i^*$ . If  $y$  is in  $y_0^\alpha$ , the fact that  $f$  is convex on the arc  $[y_0, y]$  implies that

$$f(y) \geq f(y_0) + \partial_\alpha f(y_0) \delta(y_0, y)$$

and if  $y$  is in another  $y_0^\beta$ , we have

$$f(y) \geq f(y_0) + \partial_\beta f(y_0) \delta(y_0, y) \geq f(y_0) - \partial_\alpha f(y_0) \delta(y_0, y).$$

Thus, in both cases,

$$f(y) \geq f(y_0) + \partial_\alpha f(y_0) \psi(y_0^\alpha, y),$$

where  $\psi$  was defined in (2.2.3). We put  $y = Y$ , take the expectation and use (2.2.5) to conclude.  $\square$

Proposition 2.3.5 is a semi-localised version of Jensen's inequality (we use the word "semi-localised" because the function satisfies a global condition on a set of geodesics which can be called "local" at  $\mathbb{B}[Y]$ ). If the barycentre is  $O$ , there is no gain with respect to the classical Jensen inequality, but otherwise  $f$  is only required to be convex on  $(\ell - 1)$  of the  $\ell(\ell - 1)/2$  geodesics. As an example, the functions  $\psi(y_0^\alpha, \cdot)$  of (2.2.3) are convex on the geodesics intersecting  $y_0^\alpha$ , and concave on the geodesics intersecting its complement, so we have

$$\mathbb{B}[Y] \in y_0^\alpha \implies \psi(y_0^\alpha, \mathbb{B}[Y]) \leq \mathbb{E}\psi(y_0^\alpha, Y), \quad (2.3.6a)$$

$$\mathbb{B}[Y] \notin y_0^\alpha \implies \psi(y_0^\alpha, \mathbb{B}[Y]) \geq \mathbb{E}\psi(y_0^\alpha, Y). \quad (2.3.6b)$$

In particular, by taking  $y_0 = \mathbb{B}[Y]$ , we find again (2.2.5), so the semi-localised Jensen inequality is a characterisation of the barycentre (this will be in particular useful for the definition of martingales with jumps).

**Proposition 2.3.7.** *For any square integrable variables  $Y$  and  $Z$ , one has*

$$\delta(\mathbb{B}[Y], \mathbb{B}[Z]) \leq \mathbb{E}\delta(Y, Z).$$

*Proof.* Suppose  $y_0 = \mathbb{B}[Y] \neq \mathbb{B}[Z]$ , let  $y_0^\alpha$  be the connected component of  $N \setminus \{y_0\}$  containing  $\mathbb{B}[Z]$ . Then the relations (2.3.6) imply that

$$\begin{aligned} \delta(\mathbb{B}[Y], \mathbb{B}[Z]) &= \psi(y_0^\alpha, \mathbb{B}[Z]) - \psi(y_0^\alpha, \mathbb{B}[Y]) \\ &\leq \mathbb{E}\psi(y_0^\alpha, Z) - \mathbb{E}\psi(y_0^\alpha, Y) \\ &\leq \mathbb{E}\delta(Y, Z) \end{aligned}$$

because  $\psi(y_0^\alpha, \cdot)$  is non expanding.  $\square$

Proposition 2.3.7 also holds for more general spaces with nonpositive curvature, see [32]; it says that the barycentre is a non expanding operator. Like (2.2.2), this property can be used to extend barycentres to integrable variables. Then the result of Proposition 2.3.5 is also extended to integrable variables.

### 3 Stochastic calculus on a star

We now want to study stochastic calculus on  $N = \mathbb{Y}^\ell$ . To this end, we suppose given a probability space  $\Omega$  with a filtration  $(\mathcal{F}_t)$ . The expression “càdlàg process” will designate a right continuous process with left limits. We assume that  $\Omega$  is a Lusin space; in particular, conditional probabilities exist. We refer to [8] for the classical real stochastic calculus. Let us first give an example of  $N$ -valued process which can be considered as the standard diffusion on  $N$ .

*Example 3.0.1.* Given  $\ell$  parameters  $p_i \geq 0$  such that  $\sum p_i = 1$ , the Walsh process (or spider)  $X_t$  constructed in [34, 4] is the continuous Markov process which is a standard Brownian motion on each ray  $R_i$  and which, when hitting  $O$ , immediately quits it and chooses one of the rays according to the probabilities  $p_i$ ; thus  $\mathbb{P}^O[X_t \in R_i]$  is  $p_i$  for  $t > 0$ . In particular, the isotropic Walsh process corresponds to  $p_i = 1/\ell$ . This process can be defined rigorously by using excursion theory, from its semigroup, or from its Dirichlet form. It has been studied in the last years because its filtration has interesting properties; in particular, it is not the filtration of a Euclidean Brownian motion (see [33]).

#### 3.1 Semimartingales

**Definition 3.1.1.** *An adapted càdlàg process  $Y_t$  on  $N$  is said to be a semimartingale if  $f(Y_t)$  is a semimartingale for any real convex function  $f$ .*

In smooth manifolds, continuous semimartingales are defined by means of  $C^2$  functions, and they are transformed into real semimartingales by (not necessarily  $C^2$ ) convex functions (see [12]); however, non constant convex functions may not exist, so one generally needs a localisation in order to characterise manifold-valued semimartingales. Here, convex functions are numerous enough.

By taking into account the fact that  $N$  is piecewise smooth and by using (3.1.3), we can replace in Definition 3.1.1 convex functions by continuous functions  $f$  which are  $C^2$  on each ray (but this notion will have no sense on general trees).

**Proposition 3.1.2.** *By using the standard embedding of  $N = \mathbb{Y}^\ell$  in  $\mathbb{R}^\ell$ , an adapted càdlàg process  $Y_t$  with values in  $N$  is a semimartingale if and only if its components  $Y_t^i$  are real semimartingales for any  $i$ .*

*Proof.* The component functions are convex, so it is clear that the condition is necessary. Conversely, if  $Y_t^i$  are real semimartingales and if  $f$  is a convex

function, then  $f(Y_t)$  can be written as

$$f(Y_t) = f(O) + \sum_i (f_i(Y_t^i) - f(O)) \quad (3.1.3)$$

with  $f_i$  convex, so  $f(Y_t)$  is a semimartingale.  $\square$

In the continuous case, we only need a single function  $f(y) = |y|$  to test the semimartingale property.

**Proposition 3.1.4.** *An adapted continuous process  $Y_t$  with values in  $N$  is a semimartingale if and only if  $|Y_t|$  is a real semimartingale.*

This result can be deduced from [30]. We give a proof for completeness and because we use it several times. It is actually sufficient to apply the following lemma to  $U_t = |Y_t|$  and  $V_t = Y_t^i$ .

**Lemma 3.1.5.** *Let  $U_t$  be a real continuous semimartingale, and let  $V_t$  be a continuous nonnegative adapted process such that  $dV = dU$  on  $\{V > 0\}$ ; this means that  $V_t - V_s = U_t - U_s$  as soon as  $V_r > 0$  for  $s \leq r \leq t$ . Then  $V_t$  is a semimartingale which can be written as*

$$V_t = V_0 + \int_0^t 1_{\{V_s > 0\}} dU_s + \frac{1}{2} L_t \quad (3.1.6)$$

for a nondecreasing process  $L_t$  which is the local time of  $V$  at 0.

*Proof.* Let  $\varepsilon > 0$ , let  $\tau'_0 = 0$  and consider the sequences of stopping times

$$\tau_k = \inf \left\{ t \geq \tau'_{k-1}; V_t = 0 \right\}, \quad \tau'_k = \inf \left\{ t \geq \tau_k; V_t \geq \varepsilon \right\}$$

which increase to infinity. The process  $U_t$  is a semimartingale, so  $V_t$  is a semimartingale on the intervals  $[\tau'_{k-1}, \tau_k]$  with  $dV = dU$ . Thus Tanaka's formula yields

$$d(V \vee \varepsilon) = 1_{\{V > \varepsilon\}} dU + \frac{1}{2} dL^\varepsilon \quad (3.1.7)$$

on these time intervals, for a local time  $L_t^\varepsilon$ . On the time intervals  $[\tau_k, \tau'_k]$ , then  $V \leq \varepsilon$  so  $V \vee \varepsilon = \varepsilon$  is again a semimartingale and (3.1.7) again holds with  $dL^\varepsilon = 0$ . Thus, by pasting the intervals, we deduce that  $V_t \vee \varepsilon$  is a semimartingale on the whole time interval satisfying

$$V_t \vee \varepsilon = V_0 \vee \varepsilon + \int_0^t 1_{\{V_s > \varepsilon\}} dU_s + \frac{1}{2} L_t^\varepsilon.$$

By taking the limit as  $\varepsilon \downarrow 0$ , the stochastic integral converges to the integral of (3.1.6), so we deduce that  $L_t^\varepsilon$  also converges to a non decreasing process, and the proof is complete.  $\square$

In our case, (3.1.6) can be written as

$$Y_t^i = Y_0^i + \int_0^t 1_{R_i^*}(Y_s) d|Y_s| + \frac{1}{2} L_t^i,$$

where  $L_t^i$  is the local time at  $O$  on  $R_i^*$ ; if  $f$  is a convex function, then  $f(Y_t)$  can be written from (3.1.3) as

$$f(Y_t) = f(Y_0) + \sum_i \int_0^t 1_{\{Y_s^i > 0\}} df_i(Y_s^i) + \frac{1}{2} \sum_i f_i'(0) L_t^i, \quad (3.1.8)$$

where  $df_i(Y_s^i)$  can be written with the classical Itô-Tanaka formula. Subsequently, we will also consider the total local time  $L_t = \sum L_t^i$ .

It is not difficult to check that one can replace in Proposition 3.1.4 the function  $y \mapsto |y|$  by another one such as a Busemann function, or the distance to a fixed point.

*Remark 3.1.9.* If  $Y_t$  is not continuous but càdlàg, the semimartingale property of  $|Y_t|$  is no more sufficient; it is indeed not difficult to construct a deterministic path  $y_t$  such that  $|y_t| = t$  but  $y_t$  has not finite variation (let the ray change at each time  $t = 1/n$ ).

*Example 3.1.10.* A Walsh process (Example 3.0.1) is a continuous semimartingale since  $|X_t|$  is a reflected Brownian motion, and its local time at  $O$  satisfies

$$L_t^i = p_i L_t. \quad (3.1.11)$$

## 3.2 Quasimartingales

As soon as a metric space is endowed with a notion of barycentre, one can also consider conditional barycentres (recall that conditional probabilities exist on  $\Omega$ ) and define a notion of quasimartingale similarly to the real case (see quasimartingales up to infinity of [8]).

**Definition 3.2.1.** *An integrable adapted process  $(Y_t; 0 \leq t < \infty)$  with values in  $N$  is said to be a quasimartingale if*

$$\sup \mathbb{E} \sum_k \delta(Y_{t_k}, \mathbb{B}[Y_{t_{k+1}} | \mathcal{F}_{t_k}]) < \infty, \quad (3.2.2)$$

where the supremum is taken over all the subdivisions  $(t_k)$  of  $[0, \infty]$ , and where  $Y_\infty = O$ .

One can replace  $O$  by another point (or an integrable variable). Definition 3.2.1 is also equivalent to the finiteness of (3.2.2) for subdivisions of compact intervals of  $\mathbb{R}_+$ , and the boundedness of  $|Y_t|$  in  $L^1$  (the boundedness in  $L^1$  follows from (3.2.2) by considering the subdivisions  $\{0, t, \infty\}$ ).

**Proposition 3.2.3.** *If  $Y_t$  is an adapted process in  $N = \mathbb{Y}^\ell$ , then the three following conditions are equivalent.*

1. *The process  $Y_t$  is a quasimartingale in  $N$ .*
2. *The process  $f(Y_t)$  is a real quasimartingale for any Lipschitz convex real function  $f$ .*
3. *The components  $Y_t^i$  are real quasimartingales.*

*Proof.* Let  $Y_t$  be a quasimartingale and  $f$  be a convex Lipschitz function. We deduce from the Jensen inequality (2.3.1) that

$$\begin{aligned} f(Y_{t_k}) - \mathbb{E}[f(Y_{t_{k+1}}) \mid \mathcal{F}_{t_k}] &\leq f(Y_{t_k}) - f(\mathbb{B}[Y_{t_{k+1}} \mid \mathcal{F}_{t_k}]) \\ &\leq C \delta(Y_{t_k}, \mathbb{B}[Y_{t_{k+1}} \mid \mathcal{F}_{t_k}]). \end{aligned}$$

One can replace the left-hand side by its positive part and deduce from (3.2.2) that

$$\sup_k \sum_k \mathbb{E} \left( f(Y_{t_k}) - \mathbb{E}[f(Y_{t_{k+1}}) \mid \mathcal{F}_{t_k}] \right)^+ < \infty.$$

On the other hand,

$$\sum_k \mathbb{E} \left( f(Y_{t_k}) - \mathbb{E}[f(Y_{t_{k+1}}) \mid \mathcal{F}_{t_k}] \right) = \mathbb{E}f(Y_0) - f(O),$$

so

$$\sup_k \sum_k \mathbb{E} \left| f(Y_{t_k}) - \mathbb{E}[f(Y_{t_{k+1}}) \mid \mathcal{F}_{t_k}] \right| < \infty.$$

Thus  $f(Y_t)$  is a quasimartingale, and the first condition of the proposition implies the second one. The fact that the second condition implies the third one is trivial. Finally, we assume that  $Y_t^i$  are quasimartingales; the processes  $\gamma_i(Y_t)$  are quasimartingales, and

$$\begin{aligned} \delta(Y_{t_k}, \mathbb{B}[Y_{t_{k+1}} \mid \mathcal{F}_{t_k}]) &= \max_i \left( \gamma_i(\mathbb{B}[Y_{t_{k+1}} \mid \mathcal{F}_{t_k}]) - \gamma_i(Y_{t_k}) \right) \\ &\leq \max_i \left( \mathbb{E}[\gamma_i(Y_{t_{k+1}}) \mid \mathcal{F}_{t_k}] - \gamma_i(Y_{t_k}) \right) \\ &\leq \sum_i \left| \mathbb{E}[\gamma_i(Y_{t_{k+1}}) \mid \mathcal{F}_{t_k}] - \gamma_i(Y_{t_k}) \right|, \end{aligned}$$

where we have used (2.3.1) in the second line. We deduce (3.2.2), so  $Y_t$  is a quasimartingale on  $N$ .  $\square$

In particular, càdlàg quasimartingales are semimartingales. In the continuous case, by applying the result of [30], it is actually sufficient to suppose that  $|Y_t|$  is a quasimartingale.

### 3.3 Continuous martingales

Our aim is now to define a notion of continuous martingale in  $N = \mathbb{Y}^\ell$  (càdlàg martingales will be considered in Section 8). We first define the class  $\Sigma^+$  as in [33]; it consists of the nonnegative local submartingales  $Y_t$ , the non decreasing part of which increases only on  $\{Y_t = 0\}$ . This implies that  $\int 1_{\{Y>0\}} dY$  is a local martingale. Now consider a continuous real semimartingale  $Y_t$ ; it can be decomposed as

$$Y_t = Y_0 + \int_0^t 1_{\mathbb{R}^+}(Y_s) dY_s + \frac{1}{2}(L_t^+ - L_t^-)$$

for the local times  $L_t^\pm$  at 0 on the two rays  $\mathbb{R}_\pm^*$ , and

$$Y_t^\pm = Y_0^\pm + \int_0^t 1_{\mathbb{R}_\pm^*}(Y_s) dY_s + \frac{1}{2}L_t^\pm.$$

Then  $Y_t$  is a local martingale if and only if  $Y_t^\pm$  are in  $\Sigma^+$  and  $L_t^+ = L_t^-$ . If now  $Y_t$  is  $N$ -valued, one can ask for the same properties, so that  $Y_t^i$  is in  $\Sigma^+$  and  $L_t^i = L_t^j$  (in particular  $Y_t^i$  is a local submartingale). These processes have been called spider martingales in [35]. For instance, the Walsh process is a spider martingale if and only if it is isotropic (recall (3.1.11)).

However, this notion suffers an important limitation with respect to the problem of finding a martingale with prescribed final value. If for instance  $\mathcal{F}_0$  is trivial and  $F$  is a bounded variable on  $N = \mathbb{Y}^3$  such that

$$\mathbb{P}[F \in R_1^*] > 0, \quad \mathbb{P}[F \in R_2^*] > 0, \quad \mathbb{P}[F \in R_3^*] = 0, \quad (3.3.1)$$

then there is no bounded spider martingale converging to  $F$ . One should indeed have  $Y_t^3 = 0$  (because it is a bounded nonnegative submartingale converging to 0), so  $L_t^3 = 0$ ; thus the condition  $L_t^i = L_t^j$  implies that all the local times are 0, so the process cannot quit  $O$  when it has hit it; such a process cannot satisfy (3.3.1).

Let us give another annoying property of spider martingales; Theorem 6.1 of [33] says that for a Brownian filtration, one has

$$dL_t^1 \wedge dL_t^2 \wedge dL_t^3 = 0 \quad (3.3.2)$$

for any  $N$ -valued process such that  $Y_t^i$  is in  $\Sigma^+$ . Thus if  $Y_t$  is a spider martingale for a Brownian filtration, then the condition  $L_t^i = L_t^j$  again implies that the local times are 0 and that  $Y_t$  cannot quit  $O$ .

Our aim is therefore to find another notion of continuous martingale. As in the smooth case (see [7] or Theorem 4.39 of [10]), martingales will be defined by means of convex functions; moreover, we only define local martingales (and as in the manifold-valued case they are simply called martingales).

**Definition 3.3.3.** *A continuous adapted process  $Y_t$  in  $N$  is said to be a martingale if  $f(Y_t)$  is a local submartingale for any Lipschitz convex function  $f$ .*

In the real valued case, this definition corresponds to the notion of local martingale. On the other hand, following another terminology used for real processes (Definition VI.20 of [8]), we say that  $Y_t$  is of class (D) if the family of variables  $|Y_\tau|$ , for  $\tau$  finite stopping time, is uniformly integrable. Then  $Y_t$  is a martingale of class (D) if  $f(Y_t)$  is a submartingale of class (D) for any Lipschitz convex function  $f$ . In this case  $Y_t$  has almost surely a limit  $Y_\infty$  in  $N$  (because the components  $Y_t^i$  are submartingales of class (D) and therefore have limits), and  $f(Y_t)$  should be a submartingale on the compact time interval  $[0, \infty]$ .

**Proposition 3.3.4.** *Let  $Y_t$  be a continuous adapted process in  $N$ . The following conditions are equivalent.*

1. *The process  $Y_t$  is a martingale.*
2. *The processes  $\gamma_i(Y_t)$  are local submartingales (where  $\gamma_i$  are the Busemann functions of (2.1.6)).*
3. *The components  $Y_t^i$  are in the class  $\Sigma^+$  (local nonnegative submartingales, the finite variation parts of which increase only on  $\{Y^i = 0\}$ ), and the local times  $L_t^i$  and total local time  $L_t = \sum L_t^i$  satisfy*

$$dL_t^i/dL_t \leq 1/2. \tag{3.3.5}$$

*Proof.* It is clear that the first condition implies the second one. Let us prove that the second condition implies the third one. If  $\gamma_i(Y_t)$  are local submartingales, then  $Y_t$  is in particular a semimartingale, and (3.1.8) for  $\gamma_i$  is written as

$$\gamma_i(Y_t) = \gamma_i(Y_0) - \int_0^t 1_{\{Y_s^i > 0\}} dY_s^i + \sum_{j \neq i} \int_0^t 1_{\{Y_s^j > 0\}} dY_s^j - \frac{1}{2} L_t^i + \frac{1}{2} \sum_{j \neq i} L_t^j.$$

These processes should be local submartingales for all  $i$ . We deduce that

$$\int_0^t 1_{R_j^*}(Y_s) d\gamma_i(Y_s) = (1_{\{j \neq i\}} - 1_{\{j=i\}}) \int_0^t 1_{R_j^*}(Y_s) dY_s^j$$

are local submartingales for all  $i$  and  $j$ , so the integrals of the right-hand side are actually local martingales; this means that  $Y_t^i$  is in the class  $\Sigma^+$ . Moreover, the finite variation part of  $\gamma_i(Y_t)$  should be non decreasing on  $\{Y = O\}$ , so

$$\sum_{j \neq i} dL_t^j - dL_t^i = dL_t - 2dL_t^i \geq 0$$

and (3.3.5) holds. The only thing which has still to be proved is that the third condition of the proposition implies the first one. Let us write (3.1.8) for a convex Lipschitz function  $f$ , and let us prove that  $f(Y_t)$  is a local submartingale. The property  $Y^i \in \Sigma^+$  and the convexity of  $f_i$  implies that the stochastic integrals are local submartingales, so it is sufficient to check that  $\sum f'_i(0)L_t^i$  is non decreasing; this is evident if the values of  $f'_i(0)$  are nonnegative, and if one of them, say  $f'_1(0)$ , is negative, then  $f'_i(0) \geq |f'_1(0)|$  for  $i \neq 1$  (see (2.1.2)), so

$$\sum f'_i(0)dL_t^i \geq |f'_1(0)| \left( \sum_{i \neq 1} dL_t^i - dL_t^1 \right) \geq 0$$

from (3.3.5). □

*Remark 3.3.6.* Look at Picture 4 about barycentres (for  $\ell = 3$ ), and notice that the vector  $(dL_t^i/dL_t)$  is necessarily in the triangle  $\{z_1 + z_2 + z_3 = 1\}$  of  $\mathbb{R}_+^3$ . Then (3.3.5) says that for martingales, this vector should lie in  $C_O$ . For a Brownian filtration, (3.3.2) says that it is necessarily on the boundary of the triangle, so for Brownian martingales, it can only take three values.

*Example 3.3.7.* Consider a continuous local martingale on a geodesic, say  $R_1 \cup R_2$ , which is isometric to  $\mathbb{R}$ ; it is a  $N$ -valued martingale. In this case, one has  $L_t^1 = L_t^2$  and  $L_t^i = 0$  for  $i \geq 3$ ; in particular, a martingale for a Brownian filtration does not necessarily stop at  $O$  (contrary to spider martingales).

*Example 3.3.8.* The spider martingales of [35] are martingales; the property  $L_t^i = L_t^j$  easily implies (3.3.5).

*Example 3.3.9.* From (3.1.11), a Walsh process is a martingale if and only if  $p_i \leq 1/2$  for any  $i$  (no ray should have a probability greater than  $1/2$ ).

As it is the case for smooth manifolds (Theorem 4.43 of [10]), the class of martingales is stable with respect to uniform convergence in probability.

**Proposition 3.3.10.** *Let  $Y_t^n$  be a sequence of continuous  $N$ -valued martingales and let  $Y_t$  be a continuous process such that*

$$\limsup_n \sup_{t \leq T} \delta(Y_t^n, Y_t) = 0$$

*in probability for any  $T$ . Then  $Y_t$  is a martingale.*

*Proof.* It is sufficient to prove the martingale property for the process  $Y_t$  stopped at the first time at which  $|Y_t| \geq C$ , for  $C > 0$ . This stopped process is the limit of the processes  $Y^n$  stopped at  $\tau \wedge \tau_n$ , where  $\tau_n$  is the first time at which  $|Y_t^n| \geq 2C$ . Thus we are reduced to prove the proposition for uniformly bounded processes. In this case, the submartingale property of  $f(Y_t^n)$  in Definition 3.3.3 is easily transferred to  $f(Y_t)$ .  $\square$

**Proposition 3.3.11.** *If  $Y_t$  and  $Z_t$  are continuous  $N$ -valued martingales, then the distance  $D_t = \delta(Y_t, Z_t)$  is a local submartingale.*

*Proof.* It is sufficient to prove that  $D_t \vee \varepsilon$  is a local submartingale for any  $\varepsilon > 0$ . Let  $\tau_0 = 0$  and consider the sequence of stopping times

$$\tau_{k+1} = \inf \left\{ t \geq \tau_k; \delta(Y_{\tau_k}, Y_t) \vee \delta(Z_{\tau_k}, Z_t) \geq \varepsilon/5 \right\}$$

which tends to infinity. We want to prove that  $D_t \vee \varepsilon$  is a local submartingale on each time interval  $I_k = [\tau_k, \tau_{k+1}]$ . If  $D_{\tau_k} \leq \varepsilon/2$ , then  $D_t \leq \varepsilon$  on  $I_k$  so  $D_t \vee \varepsilon = \varepsilon$  is constant. Otherwise, let  $A$  be the midpoint on the arc linking  $Y_{\tau_k}$  and  $Z_{\tau_k}$ . Then  $Y_t$  and  $Z_t$  do not cross  $A$  on  $I_k$ , so

$$D_t = \delta(A, Y_t) + \delta(A, Z_t).$$

The function  $\delta(A, y)$  is convex, so  $\delta(A, Y_t)$  and  $\delta(A, Z_t)$  are local submartingales, and  $D_t$  is therefore a local submartingale on  $I_k$ . This completes the proof.  $\square$

**Corollary 3.3.12.** *If  $Y$  and  $Z$  are continuous martingales of class (D) such that  $Y_\infty = Z_\infty$ , then  $Y_t = Z_t$  for any  $t$ .*

This property immediately follows from Proposition 3.3.11 since  $D_t$  is a nonnegative submartingale of class (D) converging to 0. It is called the non confluence property. This is the uniqueness to the problem of constructing a martingale with prescribed limit.

**Corollary 3.3.13.** *Let  $(Y_t^n)$  be a sequence of continuous martingales of class (D) with limits  $Y_\infty^n$  and suppose that  $Y_\infty^n$  converges in  $L^1$  to a variable  $Y_\infty$ . Then there exists a continuous process  $Y_t$  with limit  $Y_\infty$ , such that*

$$\lim_n \sup_{0 \leq t \leq \infty} \delta(Y_t^n, Y_t) = 0 \tag{3.3.14}$$

*in probability, and  $Y_t$  is a martingale of class (D).*

*Proof.* It follows from Proposition 3.3.11 that  $\delta(Y_t^n, Y_t^m)$  are submartingales of class (D). In particular,

$$\mathbb{P}\left[\sup_{0 \leq t \leq \infty} \delta(Y_t^n, Y_t^m) \geq C\right] \leq \frac{1}{C} \mathbb{E} \delta(Y_\infty^n, Y_\infty^m)$$

converges to 0 as  $m, n \rightarrow \infty$ , so  $(Y_t^n)$  converges in the sense of (3.3.14) to a continuous process  $Y_t$ . If  $f$  is a convex Lipschitz function, then  $(f(Y_t^n); 0 \leq t \leq \infty)$  is a submartingale, and this property is transferred to  $f(Y_t)$  by means of

$$\delta(Y_t^n, Y_t) \leq \mathbb{E}[\delta(Y_\infty^n, Y_\infty) \mid \mathcal{F}_t].$$

This means that  $Y_t$  is a martingale of class (D).  $\square$

Corollary 3.3.13 means that the set of variables which are limits of martingales of class (D) is closed in the space  $L^1(N)$  of integrable  $N$ -valued variables. The aim of subsequent sections is to find conditions ensuring that this set is the whole space  $L^1(N)$ . Before considering this question, let us give some remarks comparing  $N$ -valued martingales with the real and the manifold-valued cases.

*Remark 3.3.15.* If  $Y_\infty$  is integrable, we can consider its conditional barycentres  $\mathbb{B}[Y_\infty \mid \mathcal{F}_t]$ . On  $\mathbb{Y}^\ell$ , contrary to the Euclidean case, this is generally not a martingale. Let us give an example; let  $Y_\infty = X_\tau$  where  $X_t$  is the isotropic Walsh process (Example 3.0.1) with  $X_0 = O$ , and let  $\tau$  be the first time at which  $|X_t| \geq 1$ . On  $\{t \leq \tau\}$ , the conditional law of  $X_\tau$  given  $\mathcal{F}_t$  is

$$\mathbb{P}[X_\tau = e_i \mid \mathcal{F}_t] = \begin{cases} ((\ell - 1)|X_t| + 1)/\ell & \text{if } X_t \in R_i, \\ (1 - |X_t|)/\ell & \text{otherwise.} \end{cases}$$

After some calculation, we can deduce that

$$\mathbb{B}[X_\tau \mid \mathcal{F}_t] = \begin{cases} \frac{2(\ell-1)X_t^j + 2 - \ell}{\ell} e_j & \text{if } X_t^j > \frac{\ell-2}{2(\ell-1)}, \\ O & \text{if } |X_t| \leq \frac{\ell-2}{2(\ell-1)}. \end{cases}$$

This is not a martingale; when it quits the point  $O$ , it visits for some time only one ray, so that no more than one local time can increase and this is in contradiction with (3.3.5). On the other hand, it is clear that the martingale of class (D) converging to  $X_\tau$  is  $X_{t \wedge \tau}$ .

*Remark 3.3.16.* Remark 3.3.15 is not surprising since the situation is similar for smooth manifolds. Let us now notice a more surprising fact. On small

enough Riemannian manifolds, one can check that a continuous semimartingale  $(Y_t; 0 \leq t \leq 1)$  is a martingale if and only if

$$\lim \sum_k \delta(Y_{t_k}, \mathbb{B}[Y_{t_{k+1}} | \mathcal{F}_{t_k}]) = 0$$

in probability as the mesh of the subdivision  $(t_k)$  of  $[0, 1]$  tends to 0 (Theorem 4.5 of [26]); actually, if  $Y_t$  is a martingale with bounded quadratic variation, this expression converges to 0 in  $L^1$  (Lemma 5.5 of [26]). Here, there are martingales in  $\mathbb{Y}^\ell$  which do not satisfy this condition. Consider the case of the isotropic Walsh process  $X_t$ . For  $s \leq t$ , if  $X_s \in R_i^*$ , then the conditional law of  $X_t$  given  $\mathcal{F}_s$  gives more mass to  $R_i$  than other rays, and we can check that  $\mathbb{E}[\gamma_j(X_t) | \mathcal{F}_s] \geq 0$  for any  $j \neq i$ . Thus, by applying (2.2.6),

$$\mathbb{E}[\gamma_i(X_t) | \mathcal{F}_s] \geq 0 \implies \mathbb{B}[X_t | \mathcal{F}_s] = O \quad \text{on } \{X_s \in R_i^*\}.$$

We can check that the variable  $\mathbb{E}[\gamma_i(X_t) | \mathcal{F}_s]$  tends to  $+\infty$  as  $t \uparrow \infty$ , so this condition holds if  $t$  is large enough; more precisely, by using the scaling property of the process, it holds if  $t - s \geq c|X_s|^2$ . By using the subdivision  $t_k = k/K$  of  $[0, 1]$ , we deduce

$$\sum_k \delta(X_{t_k}, \mathbb{B}[X_{t_{k+1}} | \mathcal{F}_{t_k}]) \geq \sum_k |X_{t_k}| 1_{\{|X_{t_k}| \leq (cK)^{-1/2}\}}.$$

The right-hand side does not converge to 0 in  $L^1$ . This difference with respect to the case of smooth manifolds is essentially due to the difference in the size of convex barycentres (see Remark 2.3.2); if  $Y_t$  is a martingale, both variables  $Y_{t_k}$  and  $\mathbb{B}[Y_{t_{k+1}} | \mathcal{F}_{t_k}]$  are in the conditional convex barycentre of  $Y_{t_{k+1}}$  given  $\mathcal{F}_{t_k}$ , so they are closer to each other in the manifold-valued case than in the tree-valued case.

## 4 Martingales and coupling

Let us now return to the problem of constructing a continuous martingale  $Y_t$  of class (D) with prescribed limit  $Y_\infty$ . It is a property of the probability space  $\Omega$  and its filtration  $(\mathcal{F}_t)$ . For smooth Cartan-Hadamard manifolds, it is known from [3] that these martingales exist as soon as all the  $\mathcal{F}_t$  real martingales are continuous; we can conjecture that the same result holds here, but it seems difficult to adapt the proof. A general exact formula is unlikely to exist (see Remark 3.3.15), and we will limit ourselves to filtrations generated by some diffusion processes. Two techniques can be used (as in the smooth case), namely coupling of the diffusion (this is the aim of this section), or energy minimisation when the diffusion is symmetric (see next section). This leads to the existence in two frameworks.

In this section, we work out the coupling method which is classically used in the manifold-valued case, see for instance [18] where coupling of the Euclidean Brownian motion is applied, see also [31] for an application in the singular case using a more functional analytic approach. Approximation of  $\mathbb{Y}^\ell$  by hyperbolic planes with highly negative curvature is also discussed in this section (see Subsection 4.3).

Let us fix a bounded final variable  $Y$  on  $N = \mathbb{Y}^\ell$  which is  $\mathcal{F}_1$  measurable, and consider a discretization  $\Delta = (t_k)$ ,  $0 \leq k \leq K$  of the time interval  $[0, 1]$ . The idea is to define  $(Y_k)$  by  $Y_K = Y$  and

$$Y_k = \mathbb{B}[Y_{k+1} | \mathcal{F}_{t_k}]. \quad (4.0.1)$$

This sequence can be viewed as a discrete martingale (and this is actually compatible with the definition of martingales with jumps which will be given in Section 8). We deduce from Jensen's inequality (2.3.1) that the sequence  $f(Y_k)$  is a discrete submartingale for any convex Lipschitz function  $f$ . Moreover, Proposition 2.3.7 says that if  $Y$  and  $Y'$  are two final values, then  $\delta(Y_k, Y'_k)$  is a submartingale, so

$$\delta(Y_k, Y'_k) \leq \mathbb{E}[\delta(Y, Y') | \mathcal{F}_{t_k}]. \quad (4.0.2)$$

We are looking for a condition ensuring the convergence of  $(Y_k)$  as the discretization mesh  $\max(t_{k+1} - t_k)$  tends to 0.

### 4.1 The main result

Let  $M$  be a separable metric space with distance  $d$ , let  $\Omega$  be the space of continuous functions  $\omega : \mathbb{R}_+ \rightarrow M$ , and let  $X_t$  be the canonical process  $X_t(\omega) = \omega(t)$  with its natural filtration  $(\mathcal{F}_t)$ . If moreover  $M$  is complete,

then the usual topology of  $\Omega$  (uniform convergence on compact subsets) can be defined by a separable complete distance, and  $\Omega$  is a Lusin space. We consider on  $\Omega$  a family of probability measures  $(\mathbb{P}^x; x \in M)$  under which  $X_t$  is a homogeneous Markov process with initial value  $X_0 = x$ ; let  $P_t$  be its semigroup. As usually, we also denote by  $\mathbb{P}^\nu$  the law of the process with initial law  $\nu$ .

**Definition 4.1.1.** *An admissible coupling of  $X_t$  with itself is a family  $(\mathbb{P}^{x,x'})$ ,  $(x, x') \in M \times M$ , of probability measures on  $\Omega \times \Omega$  with canonical process  $(X_t, X'_t)$ , filtration  $(\mathcal{F}_t'')$ , such that*

$$\mathbb{E}^{x,x'} [f(X_t) \mid \mathcal{F}_s''] = P_{t-s}f(X_s), \quad \mathbb{E}^{x,x'} [f(X'_t) \mid \mathcal{F}_s''] = P_{t-s}f(X'_s)$$

for any bounded Borel function  $f$ .

This means that the laws of  $(X_t)$  and  $(X'_t)$  are respectively  $\mathbb{P}^x$  and  $\mathbb{P}^{x'}$  under  $\mathbb{P}^{x,x'}$ , and that the processes  $X_t$  and  $X'_t$  are Markovian for the filtration of  $(X, X')$ .

This binary coupling is a simplification of the notion of stochastic flow, since we do not consider simultaneously all the initial conditions, but only two of them. We will assume the existence of a good coupling for which  $X_t$  and  $X'_t$  are close to each other when  $x$  and  $x'$  are close.

*Remark 4.1.2.* We do not suppose that the coupling is fully Markovian, since we do not require  $(X, X')$  to be Markovian. Notice that another notion of coupling is used in [31].

A particular class of coupling is the class of coalescent couplings for which the processes  $X_t$  and  $X'_t$  try to meet and are equal after their first meeting time

$$\sigma = \inf \left\{ t \geq 0; X_t = X'_t \right\}.$$

In this case, the coupling is good if

$$\lim_{d(x,x') \rightarrow 0} \mathbb{P}^{x,x'}[\sigma > t] = 0 \tag{4.1.3}$$

for  $t > 0$  fixed. The following main result gives the existence in  $N = \mathbb{Y}^\ell$  of a martingale with final value  $g(X_1)$ ; recall that the uniqueness was proved in Corollary 3.3.12.

**Theorem 4.1.4.** *We suppose that*

$$\mathbb{E}^x [d(x, X_t) \wedge 1] \leq \phi_1(t) \tag{4.1.5}$$

for some function  $\phi_1$  satisfying  $\lim_0 \phi_1 = 0$ , and that there exists an admissible coupling  $(\mathbb{P}^{x,x'})$  such that

$$\mathbb{E}^{x,x'} [d(X_t, X'_t) \wedge 1] \leq \phi_2(d(x, x')) \quad (4.1.6)$$

for some function  $\phi_2$  satisfying  $\lim_0 \phi_2 = 0$ . Then for any uniformly continuous bounded map  $g : M \rightarrow N$ , there exists a uniformly continuous map  $h : [0, 1] \times M \rightarrow N$  such that  $Y_t = h(t, X_t)$ ,  $0 \leq t \leq 1$ , is under  $\mathbb{P}^x$  the bounded martingale with final value  $Y_1 = g(X_1)$ .

*Proof.* On  $\Omega \times \Omega$  with its natural filtration  $(\mathcal{F}_t'')$ , we first construct a coupling  $\mathbb{P}^{(s,x),(s',x')}$  for nonnegative  $s$  and  $s'$  as follows. Suppose for instance that  $s \leq s'$ .

- On the time interval  $[0, s]$ , we put  $(X_t, X'_t) = (x, x')$ .
- On the time interval  $[s, s']$ , we put  $X'_t = x'$  and  $(X_{s+u}; 0 \leq u \leq s' - s)$  evolves according to  $\mathbb{P}^x$ .
- After time  $s'$ , conditionally on  $\mathcal{F}_{s'}''$ , the process  $(X_{s'+u}, X'_{s'+u}; u \geq 0)$  evolves according to  $\mathbb{P}^{x'',x'}$  for  $x'' = X_{s'}$ .

Then  $X$  and  $X'$  are Markovian for the filtration of  $(X, X')$ , and after  $s$ , respectively  $s'$ , they evolve according to  $\mathbb{P}^x$ , respectively  $\mathbb{P}^{x'}$ . We can suppose without loss of generality that  $\phi_2$  is bounded and non decreasing; then, for  $s \leq s' \leq t$ , from (4.1.6),

$$\begin{aligned} \mathbb{E}^{(s,x),(s',x')} [d(X_t, X'_t) \wedge 1 \mid \mathcal{F}_{s'}''] &\leq \phi_2(d(X_{s'}, x')) \\ &\leq \phi_2(d(x, x') + d(x, X_{s'})). \end{aligned}$$

A similar inequality can of course be written for  $s' \leq s \leq t$ . By taking the expectation and applying (4.1.5), we obtain an expression which converges to 0 as  $|s' - s|$  and  $d(x, x')$  tend to 0, so

$$\mathbb{E}^{(s,x),(s',x')} [d(X_t, X'_t) \wedge 1] \leq \phi_3(|s' - s| + d(x, x')) \quad (4.1.7)$$

with  $\lim \phi_3 = 0$ , for  $t \geq s \vee s'$ . Now consider the variable  $Y = g(X_1)$  of the theorem, and a subdivision  $\Delta = (t_k)$  of  $[0, 1]$ . It follows from the Markov property of  $X$  that the discrete martingale (4.0.1) has the form  $Y_k = h^\Delta(t_k, X_{t_k})$  for a bounded function  $h^\Delta$  defined on  $\Delta \times M$ . Moreover, for  $s$  and  $s'$  in  $\Delta$ , it follows from the Markov properties of Definition 4.1.1 that, on  $\Omega \times \Omega$  and under  $\mathbb{P}^{(s,x),(s',x')}$ , the sequences  $h^\Delta(t_k \vee s, X_{t_k})$  and  $h^\Delta(t_k \vee s', X'_{t_k})$

are the discrete martingales with final values  $g(X_1)$  and  $g(X'_1)$ . By applying (4.0.2), we obtain

$$\delta(h^\Delta(s, x), h^\Delta(s', x')) \leq \mathbb{E}^{(s, x), (s', x')} [\delta(g(X_1), g(X'_1))].$$

We deduce from (4.1.7) that  $h^\Delta(t, x)$  is uniformly continuous on  $\Delta \times M$ , and this is uniform in  $\Delta$ . Since  $M$  is separable, there exists a sequence of dyadic subdivisions  $\Delta_n$  such that  $h^{\Delta_n}(t, x)$  converges for  $t$  dyadic and  $x \in M$ ; moreover, the limit  $h(t, x)$  is uniformly continuous and can therefore be extended to  $[0, 1] \times M$ . The process  $Y_t = h(t, X_t)$  is continuous, with value  $g(X_1)$  at time 1; it is transformed into a submartingale by any convex function  $f$  because the process  $f(Y_t)$  is the limit of the uniformly bounded discrete submartingales  $f(Y_k)$ , so  $Y_t$  is a bounded martingale.  $\square$

**Corollary 4.1.8.** *Suppose that (4.1.5) holds true, and consider a coalescent coupling satisfying (4.1.3). Then (4.1.6) holds true, and consequently, the conclusion of Theorem 4.1.4 is valid.*

*Proof.* By applying (4.1.5) and the triangle inequality on one hand, and the coalescence on the other hand, we obtain

$$\mathbb{E}^{x, x'} [d(X_t, X'_t) \wedge 1] \leq \min(d(x, x') + 2\phi_1(t), \mathbb{P}^{x, x'}[\sigma > t]).$$

Fix some  $t_0 > 0$ ; in the right hand side, we use the first term if  $t < t_0$ , and the second one if  $t \geq t_0$ , so

$$\sup_t \mathbb{E}^{x, x'} [d(X_t, X'_t) \wedge 1] \leq d(x, x') + 2 \sup_{t < t_0} \phi_1(t) + \mathbb{P}^{x, x'}[\sigma > t_0].$$

Consequently, from (4.1.3),

$$\limsup_{d(x, x') \rightarrow 0} \sup_t \mathbb{E}^{x, x'} [d(X_t, X'_t) \wedge 1] \leq 2 \sup_{t < t_0} \phi_1(t)$$

for any  $t_0 > 0$ , and is therefore 0. Thus (4.1.6) is satisfied.  $\square$

*Remark 4.1.9.* In Theorem 4.1.4 and Corollary 4.1.8, the conditions were uniform with respect to  $x$ ; however, what we need is that each point of  $M$  has a neighbourhood satisfying (4.1.7). The functions  $h^\Delta$  are indeed uniformly continuous on these neighbourhoods, and we again deduce the convergence and the continuity of the limit.

*Example 4.1.10.* For the real Wiener process, there are two classical couplings satisfying the assumptions of Theorem 4.1.4. Firstly, the two processes can stay at a fixed distance from each other ( $X'_t - X_t = x' - x$ ). Secondly, we can

consider the coalescent coupling for which  $X_t + X'_t = x + x'$  up to the first meeting time (the process  $X'$  is obtained from  $X$  by reflection). Actually, for any coupling for which the processes coincide after their first meeting time, the process  $|X'_t - X_t|$  is a martingale, so all these couplings are similar for the estimation of  $\mathbb{E}|X'_t - X_t| = |x' - x|$  (this is of course false in dimension greater than 1).

*Example 4.1.11.* If  $X$  is the solution of a stochastic differential equation driven by a Wiener process, we can consider the flow associated to this equation, and let  $X$  and  $X'$  be the images of  $x$  and  $x'$  by this flow. We obtain a non coalescent coupling. This can be applied to Brownian motions on Riemannian manifolds; it is also possible but more technical to construct a coalescent coupling, see [19] where a stronger coupling property is actually proved.

*Example 4.1.12.* We will see in Section 7 that a coalescent coupling satisfying (4.1.6) can be constructed for Walsh processes on  $\mathbb{Y}^\ell$ ; more general trees and graphs can also be considered. Of course, the one-dimensional structure of these spaces makes the construction much easier.

**Corollary 4.1.13.** *Consider a diffusion  $X_t$  satisfying the assumptions of Theorem 4.1.4, and a probability  $\mathbb{P}^\nu$  for some initial law  $\nu$ . Then any  $N$ -valued integrable variable is the limit of a unique martingale of class (D).*

*Proof.* Corollary 3.3.13 is used at each of the following steps (except the second one). The existence of the martingale  $Y_t$  with final value  $g(X_1)$  is first extended to any bounded Borel map  $g$  by approximating them by uniformly continuous functions (real functions can be approximated by uniformly continuous functions from a functional monotone class theorem, and the result is easily extended to  $N$ -valued maps). In a second step, we consider final variables of type  $g(X_{t_1}, \dots, X_{t_k})$  and construct the martingale on each time interval  $[t_j, t_{j+1}]$ . In a third step, by approximating general variables by such variables, we deduce that any bounded variable is the limit of a unique bounded martingale; the result is then easily extended to integrable variables and martingales of class (D).  $\square$

*Remark 4.1.14.* Under the conditions of Theorem 4.1.4, it is clear that continuous martingales of class (D) with prescribed integrable limit exist in particular on  $\mathbb{R}$ . This means that real martingales are continuous. Thus we are in a case where the existence result of [3] concerning smooth manifolds can be applied.

## 4.2 Feller property of the semigroup

We have solved the existence problem for the filtration of  $X_t$ ; if  $g$  is bounded, the initial value of the bounded martingale with final value  $g(X_t)$  is denoted by  $Q_t g(x)$ . It is classical to check that  $Q_t$  is a semigroup; this is the (non linear) heat semigroup  $Q_t$  acting on bounded  $N$ -valued maps  $g$ . Then the martingale with final value  $g(X_t)$  is  $(Q_{t-s}g(X_s); s \leq t)$ . The coupling method also implies a Feller property on this semigroup.

**Proposition 4.2.1.** *Under the assumptions of Theorem 4.1.4, the semigroup  $Q_t$  is Feller continuous in the sense that if  $g : M \rightarrow N$  is a bounded continuous function, then  $Q_t g$  is continuous for any  $t$  and  $Q_t g(x)$  converges to  $g(x)$  as  $t \downarrow 0$ . If the coupling is coalescent and satisfies (4.1.3), then  $Q_t$  is regularising, or strongly Feller; this means that for  $t > 0$ , it maps bounded Borel functions to bounded continuous functions.*

*Proof.* From Proposition 3.3.11, one has

$$\delta(Q_t g(x), Q_t g(x')) \leq \mathbb{E}^{x, x'} [\delta(g(X_t), g(X'_t))],$$

and from (4.1.6),  $(X_t, X'_t)$  converges in law to  $(X_t, X_t)$  as  $x' \rightarrow x$ . We deduce that  $Q_t g$  is continuous if  $g$  is continuous. Similarly

$$\delta(g(x), Q_t g(x)) \leq \mathbb{E}^x [\delta(g(x), g(X_t))]$$

tends to 0 as  $t \downarrow 0$  from (4.1.5). In the coalescent case, we use

$$\delta(Q_t g(x), Q_t g(x')) \leq 2 \sup_y |g(y)| \mathbb{P}^{x, x'} [\sigma > t].$$

□

In the coalescent case, if for instance the probability of non coupling (4.1.3) is dominated by  $d(x, x')$ , then  $Q_t g$  is Lipschitz.

Corollary 4.1.13 also enables to solve the Dirichlet problem; let  $M_0$  be an open subset of  $M$  and let  $\tau$  be the first exit time of  $M_0$  for  $X_t$ . If  $\tau$  is finite and if  $g : M \rightarrow N$  is bounded, we can consider the bounded martingale with limit  $g(X_\tau)$  under  $\mathbb{P}^x$ , and let  $h(x)$  be its initial value; then the martingale is  $h(X_{t \wedge \tau})$ . We say that  $h$  is harmonic on  $M_0$ .

**Proposition 4.2.2.** *Under the assumptions of Theorem 4.1.4, if the coupling is coalescent and satisfies (4.1.3), and if  $h : M \rightarrow N$  is bounded and harmonic on an open subset  $M_0$ , then  $h$  is continuous on  $M_0$ .*

*Proof.* If  $x$  and  $x'$  are in  $M_0$ , we consider the coupled process  $(X_t, X'_t)$  and the exit times  $\tau$  and  $\tau'$  of  $M_0$ ; from Proposition 3.3.11,

$$\begin{aligned} \delta(h(x), h(x')) &\leq \mathbb{E}^{x, x'} [h(X_{t \wedge \tau}), h(X'_{t \wedge \tau'})] \\ &\leq 2 \sup_y |h(y)| \mathbb{P}^{x, x'} [\sigma > t \wedge \tau \wedge \tau'] \\ &\leq 2 \sup_y |h(y)| \left( \mathbb{P}^{x, x'} [\sigma > t] + \mathbb{P}^{x, x'} [\tau \wedge \tau' < t] \right) \end{aligned} \quad (4.2.3)$$

for any  $t$ . Fix  $x$  in  $M_0$  and consider a neighbourhood  $V$  of  $x$  which is at a positive distance from the complement of  $M_0$ . The first probability in (4.2.3) tends to 0 as  $x' \rightarrow x$  for  $t > 0$  fixed, and the second one tends to 0 as  $t \downarrow 0$  uniformly for  $x'$  in  $V$  (apply (4.1.5)). Thus  $h(x')$  tends to  $h(x)$ .  $\square$

### 4.3 Stars and hyperbolic geometry

Another result can be worked out under the framework of Theorem 4.1.4; this result says that  $N$ -valued martingales can be approximated by hyperbolic martingales, and it comes from the fact that a hyperbolic plane with highly negative curvature looks like a star.

Consider the plane  $\mathbb{R}^2$ ; it can be endowed with a hyperbolic metric  $|\cdot|_\kappa$  with curvature  $-\kappa$  by putting

$$|u|_\kappa^2 = |u_1|^2 + \frac{\sinh(\sqrt{\kappa}|y|)^2}{\kappa|y|^2} |u_2|^2, \quad (4.3.1)$$

where  $u$  is a vector based at  $y \in \mathbb{R}^2$ , and  $u_1$  and  $u_2$  are its radial and angular parts ( $|u|_\kappa = |u|$  if  $y = 0$ ). We denote by  $\delta_\kappa$  the corresponding distance so that  $\mathbb{H}_\kappa^2 = (\mathbb{R}^2, \delta_\kappa)$  becomes a hyperbolic plane with curvature  $-\kappa$ . Notice that the Euclidean distance  $\delta_0$  is dominated by the hyperbolic distance  $\delta_\kappa$ . Notice also that we can construct in  $\mathbb{H}_\kappa^2$  continuous martingales with prescribed final value (see Remark 4.1.14). Now choose an embedding (2.0.1) of  $N = \mathbb{Y}^\ell$  in  $\mathbb{R}^2$ ; this is also an isometric embedding into  $\mathbb{H}_\kappa^2$ .

**Proposition 4.3.2.** *Under the assumptions of Theorem 4.1.4, let  $(Y_t^\kappa; 0 \leq t \leq 1)$  be the bounded martingale in  $\mathbb{H}_\kappa^2 = (\mathbb{R}^2, \delta_\kappa)$  with final value  $g(X_1)$ . Then*

$$\lim_{\kappa \uparrow \infty} \sup_{0 \leq t \leq 1} \delta_0(Y_t^\kappa, Y_t) = 0 \quad (4.3.3)$$

*in probability for a process  $Y_t$ , and  $Y_t$  is the bounded  $N$ -valued martingale with final value  $g(X_1)$ .*

*Proof.* It is sufficient to prove that for any sequence of curvatures tending to infinity, there exists a subsequence such that (4.3.3) holds almost surely, and  $Y_t$  is a martingale. The martingale  $Y_t^\kappa$  has the form  $h_\kappa(t, X_t)$ . The technique used in Theorem 4.1.4 shows the uniform continuity of  $h_\kappa$  for the hyperbolic metric, uniformly in  $\kappa$ , and therefore also for the Euclidean metric. Thus we can consider a subsequence converging to a  $\mathbb{R}^2$ -valued function  $h$ , so that  $Y_t^\kappa$  converges to  $Y_t = h(t, X_t)$ . Moreover, the convergence is uniform on compact subsets of  $[0, 1] \times M$ , so on  $\{(t, X_t(\omega)); 0 \leq t \leq 1\}$ , and (4.3.3) holds almost surely. On the other hand,  $Y_t^\kappa$  lives in the convex hull  $N_\kappa$  of  $N$  in  $\mathbb{H}_\kappa^2$ ; we see from (4.3.1) that

$$\delta_\kappa(y/\sqrt{\kappa}, z/\sqrt{\kappa}) = \delta_1(y, z)/\sqrt{\kappa},$$

so  $y \mapsto y/\sqrt{\kappa}$  is an isometry from  $(\mathbb{R}^2, \delta_1/\sqrt{\kappa})$  onto  $\mathbb{H}_\kappa^2 = (\mathbb{R}^2, \delta_\kappa)$ . The star  $N$  is invariant for this isometry, so  $N_\kappa = N_1/\sqrt{\kappa}$ . This implies that

$$\sup\{\delta_\kappa(y, N); y \in N_\kappa\} = \frac{1}{\sqrt{\kappa}} \sup\{\delta_1(y, N); y \in N_1\} = \frac{C_1}{\sqrt{\kappa}}$$

for a finite  $C_1$  ( $C_1$  is also the distance of  $O$  to the complement of  $N_1$ ). In particular,

$$\delta_0(Y_t^\kappa, N) \leq \delta_\kappa(Y_t^\kappa, N) \leq C_1/\sqrt{\kappa}, \quad (4.3.4)$$

and at the limit,  $Y_t$  is in  $N$ . We have to prove that it is a martingale. Let  $C_{ij}$  be the scalar product of  $e_i$  and  $e_j$  (the vectors describing the embedding (2.0.1)), and let  $C$  be the maximal value of  $C_{ij}$  for  $j \neq i$ , so that  $C < 1$ . If  $y$  and  $z$  are in two different rays  $R_i$  and  $R_j$  of  $N$ , then we can write in  $\mathbb{R}^2$

$$\begin{aligned} |z - y|^2 &= |y|^2 + |z|^2 - 2C_{ij}|y||z| \\ &\geq |y|^2 + |z|^2 - 2C|y||z| \\ &\geq (1 - C)(|y|^2 + |z|^2) \\ &\geq (1 - C)(|y| + |z|)^2 / 2 \\ &\geq (1 - C)\delta_\kappa(y, z)^2 / 2 \end{aligned}$$

so

$$\delta_\kappa(y, z) \leq C'\delta_0(y, z) \quad (4.3.5)$$

for some  $C' \geq 1$ . This inequality also holds when  $y$  and  $z$  are in the same ray, so it holds on  $N \times N$ . Now let  $Z_t^\kappa$  be a point in  $N$  which minimises the

hyperbolic distance to  $Y_t^\kappa$ . By using (4.3.4) and (4.3.5), one has that

$$\begin{aligned}
\delta_\kappa(Y_t^\kappa, Y_t) &\leq \delta_\kappa(Y_t^\kappa, Z_t^\kappa) + \delta_\kappa(Z_t^\kappa, Y_t) \\
&\leq \delta_\kappa(Y_t^\kappa, Z_t^\kappa) + C' \delta_0(Z_t^\kappa, Y_t) \\
&\leq \delta_\kappa(Y_t^\kappa, Z_t^\kappa) + C' \delta_0(Z_t^\kappa, Y_t^\kappa) + C' \delta_0(Y_t^\kappa, Y_t) \\
&\leq (1 + C') \delta_\kappa(Y_t^\kappa, N) + C' \delta_0(Y_t^\kappa, Y_t) \\
&\leq (1 + C') C_1 / \sqrt{\kappa} + C' \delta_0(Y_t^\kappa, Y_t),
\end{aligned}$$

which converges to 0. Finally, for each ray  $R_i$  consider the hyperbolic Busemann function

$$\gamma_i^\kappa(y) = \lim_{r \rightarrow \infty} (\delta_\kappa(y, r e_i) - r).$$

It is convex so  $\gamma_i^\kappa(Y_t^\kappa)$  is a (bounded) submartingale. The convergence of  $\delta_\kappa(Y_t^\kappa, Y_t)$  to 0 implies the convergence of  $\gamma_i^\kappa(Y_t^\kappa) - \gamma_i^\kappa(Y_t)$  to 0, and  $\gamma_i^\kappa$  converges to  $\gamma_i$  on  $N$ . Thus  $\gamma_i^\kappa(Y_t^\kappa)$  converges to  $\gamma_i(Y_t)$ , and  $\gamma_i(Y_t)$  is therefore a submartingale. We conclude with Proposition 3.3.4.  $\square$

## 5 Martingales and energy minimisation

Let us now describe another framework in which one can prove the existence of the martingale with prescribed limit; this will relate our problem (as in [28] for the smooth case) with a variational problem, namely energy minimisation; this technique is particularly useful for symmetric diffusions for which a coupling seems difficult to construct.

### 5.1 The Dirichlet space

The aim of this subsection is to define and study the notion of Dirichlet space for tree-valued maps. In the case of more general spaces with nonpositive curvature, a definition using the heat semigroup is proposed in [17]; here, we propose another one for the particular case of trees.

On a separable locally compact space  $M$  endowed with a Radon measure  $\mu$ , consider a symmetric diffusion  $(X_t, 0 \leq t < \zeta)$  with lifetime  $\zeta$  associated to a regular strongly local Dirichlet form  $\mathcal{E}$  on  $L^2(\mu)$ , under the law  $\mathbb{P}^\mu$ . The strong locality means that  $X_t$  is continuous and is not killed inside  $M$ . The domain of  $\mathcal{E}$  is the Dirichlet space  $\mathbb{D}$ , and  $\mathcal{E}(f) = \mathcal{E}(f, f)$  is a semi-norm on it; its elements can be chosen quasicontinuous. We refer to [14] for definitions and properties of these spaces and diffusions.

For some purposes, the space  $\mathbb{D}$  is too restrictive and we have to enlarge it; for instance, the space  $\mathbb{D}$  is stable with respect to Lipschitz transformations  $\phi$  such that  $\phi(0) = 0$ , but generally not with respect to all Lipschitz transformations (constant functions are not always in  $\mathbb{D}$ ); this causes some trouble because for tree-valued functions, there is generally not a canonical point which could replace the role of the point 0 of  $\mathbb{R}$ . For this reason, we are going to consider the space  $\mathbb{D}^{loc}$  of functions which are locally in  $\mathbb{D}$  (on each relatively compact open subset of  $M$  there exists a function of  $\mathbb{D}$  which coincides with  $f$ ). In  $\mathbb{D}$ , we can consider energy measures  $\mu_{\langle f, g \rangle}$  and  $\mu_{\langle f \rangle} = \mu_{\langle f, f \rangle}$  so that  $\mathcal{E}(f)$  is the total mass  $\mu_{\langle f \rangle}(M)$ . These measures can also be defined for functions of  $\mathbb{D}^{loc}$ ; one indeed deduces from the locality that if  $f$  and  $g$  coincide on an open set, then  $\mu_{\langle f \rangle}$  and  $\mu_{\langle g \rangle}$  also coincide on this set. Thus one can define the energy  $\mathcal{E}(f) = \mu_{\langle f \rangle}(M)$  (finite or infinite) on  $\mathbb{D}^{loc}$ . We will be particularly interested by the subspace  $\mathbb{D}^b$  consisting of bounded functions of  $\mathbb{D}^{loc}$  with finite energy. In the transient case, this space coincides with the space of bounded functions of the reflected Dirichlet space, see [6].

Constant functions are in  $\mathbb{D}^b$  (and have zero energy), so  $\mathbb{D}^b$  is stable with respect to all Lipschitz functions. Let us give some other useful facts.

**Lemma 5.1.1.** *For any function  $f$  of  $\mathbb{D}^b$ , one has*

$$\mu_{\langle f \rangle} \{f = 0\} = 0. \quad (5.1.2)$$

*For any  $f$  and  $g$  in  $\mathbb{D}^b$  such that  $fg = 0$ , one has*

$$\mu_{\langle f+g \rangle} = \mu_{\langle f \rangle} + \mu_{\langle g \rangle}, \quad (5.1.3)$$

so

$$\mathcal{E}(f + g) = \mathcal{E}(f) + \mathcal{E}(g). \quad (5.1.4)$$

*For any  $f$  and  $g$  in  $\mathbb{D}^b$ , the energy measures  $\mu_{\langle f \rangle}$  and  $\mu_{\langle g \rangle}$  coincide on  $\{f = g\}$ .*

*Proof.* These properties can be localised so it is sufficient to prove them for functions of  $\mathbb{D}$ . One has

$$\mu_{\langle \Phi \circ f \rangle}(dx) = (\Phi' \circ f)(x)^2 \mu_{\langle f \rangle}(dx)$$

for any function  $\Phi$  of class  $C_b^1$ , so if  $\Phi(0) = 0$  and  $\Phi'(0) = 1$ ,

$$\mu_{\langle f \rangle} \{f = 0\} \leq \int (\Phi' \circ f)(x)^2 \mu_{\langle f \rangle}(dx) = \mathcal{E}(\Phi \circ f) \quad (5.1.5)$$

We apply this relation to

$$\Phi_n(z) = \arctan(nz)/n.$$

Then  $\Phi'_n - \Phi'_m$  tends to 0 as  $m$  and  $n$  tend to infinity, so  $(\Phi_n \circ f)$  is a  $\mathcal{E}$ -Cauchy sequence; moreover it converges to 0 in  $L^2$ , so a standard argument shows that  $\mathcal{E}(\Phi_n \circ f)$  converges to 0. Thus (5.1.2) follows from (5.1.5). On the other hand, we deduce from the non negativity of  $\mu_{\langle f \rangle}$  that

$$\left| \mu_{\langle f, g \rangle}(A) \right|^2 \leq \mu_{\langle f \rangle}(A) \mu_{\langle g \rangle}(A),$$

so  $\mu_{\langle f, g \rangle} = 0$  on  $\{f = 0\} \cup \{g = 0\}$ . Thus, if  $fg = 0$ , then  $\mu_{\langle f, g \rangle}$  is 0 and consequently (5.1.3) holds. We also have that

$$\begin{aligned} \left| \mu_{\langle f \rangle}(A) - \mu_{\langle g \rangle}(A) \right| &= \left| \mu_{\langle f-g, f+g \rangle}(A) \right| \\ &\leq \mu_{\langle f-g \rangle}(A)^{1/2} \mu_{\langle f+g \rangle}(A)^{1/2} \\ &= 0 \end{aligned}$$

if  $A \subset \{f = g\}$ , so  $\mu_{\langle f \rangle} = \mu_{\langle g \rangle}$  on  $\{f = g\}$ . □

*Remark 5.1.6.* One can replace  $\{f = 0\}$  by  $\{f = c\}$  in (5.1.2). Similarly, (5.1.3) and (5.1.4) hold true if  $(f - c)(g - c') = 0$ .

Let  $M_0$  be a relatively compact open subset of  $M$ , and let

$$\tau = \inf \left\{ t > 0; X_t \notin M_0 \right\}$$

be the first exit time of  $M_0$ . We suppose that  $\tau < \zeta$  ( $\zeta$  is the lifetime of  $X$ )  $\mathbb{P}^\mu$ -almost surely.

Let  $\mathbb{D}_0$ ,  $\mathbb{D}_0^{loc}$  and  $\mathbb{D}_0^b$  be the spaces of functions of  $\mathbb{D}$ ,  $\mathbb{D}^{loc}$  and  $\mathbb{D}^b$  having a quasicontinuous modification  $f$  such that  $f = 0$  quasi everywhere outside  $M_0$ . Then  $(\mathbb{D}_0, \mathcal{E})$  is a regular Dirichlet form on  $M_0$  (called the part of  $\mathcal{E}$  on  $M_0$ ) and it is associated to the process  $X$  killed at  $\tau$ . The condition  $\tau < \zeta$  implies that it is transient, so (see [14])

$$\int |f(x)| \nu(dx) \leq \mathcal{E}(f)^{1/2} \quad (5.1.7)$$

for  $f$  in  $\mathbb{D}_0$  and  $\nu$  a measure such that  $\mu$  and  $\nu$  are mutually absolutely continuous. Since  $M_0$  is relatively compact in  $M$ , the space  $\mathbb{D}_0^{loc}$  is equal to  $\mathbb{D}_0$ , so  $\mathbb{D}_0^b$  is the space of bounded functions of  $\mathbb{D}_0$ .

If  $g$  is a quasicontinuous function of  $\mathbb{D}^b$ , we let  $\mathbb{D}_g^b$  be the set of functions of  $\mathbb{D}^b$  having a quasicontinuous modification  $f$  such that  $f = g$  quasi everywhere outside  $M_0$ .

**Lemma 5.1.8.** *Let  $g$  be a quasicontinuous function of  $\mathbb{D}^b$  satisfying  $\mathcal{E}(f, g) \geq 0$  for any nonnegative function  $f$  of  $\mathbb{D}_0^b$ . Then  $g(X_{t \wedge \tau})$  is a  $\mathbb{P}^\mu$ -supermartingale.*

*Proof.* The function

$$h(x) = \mathbb{E}^x[g(X_\tau)]$$

is quasicontinuous and is  $\mathcal{E}$ -orthogonal to  $\mathbb{D}_0$ , so that  $\mathcal{E}(f, h) = 0$  for any  $f$  of  $\mathbb{D}_0^b$ ; this was proved in Section 4.3 of [14] when  $g$  is in the Dirichlet space, but can be extended to  $g$  in  $\mathbb{D}^b$  by modifying  $g$  outside the closure of  $M_0$ . The process  $h(X_{t \wedge \tau})$  is the bounded real martingale with final value  $g(X_\tau)$ . The function  $g - h$  is in  $\mathbb{D}_0$  and our assumption implies that  $\mathcal{E}(g - h, f)$  is nonnegative for any nonnegative  $f$  of  $\mathbb{D}_0$ ; thus  $g - h$  is superharmonic for the process killed at  $\tau$ , and  $(g - h)(X_{t \wedge \tau})$  is a  $\mathbb{P}^\mu$ -supermartingale.  $\square$

Let us now extend the notion of Dirichlet space to maps with values in  $N = \mathbb{Y}^\ell$ .

**Definition 5.1.9.** The set  $\mathbb{D}^b(N)$  is the space of  $N$ -valued functions  $f$  such that  $\phi \circ f$  is in  $\mathbb{D}^b$  for any Lipschitz function  $\phi : N \rightarrow \mathbb{R}$ . For  $f$  in this space, we put

$$\mathcal{E}(f) = \sup \left\{ \mathcal{E}(\phi \circ f); \phi \text{ non expanding} \right\}.$$

**Lemma 5.1.10.** A function  $f$  is in  $\mathbb{D}^b(N)$  if and only if its components  $f_i$  are in  $\mathbb{D}^b$ . In this case, one has

$$\mathcal{E}(f) = \mathcal{E}(|f|) = \sum \mathcal{E}(f_i). \quad (5.1.11)$$

*Proof.* It is clear that  $f_i \in \mathbb{D}^b$  is necessary for  $f \in \mathbb{D}^b(N)$ . Conversely, we check that it is sufficient by using the decomposition

$$\phi(y) = \phi(O) + \sum_i (\phi_i(y_i) - \phi(O)) \quad (5.1.12)$$

of any Lipschitz function  $\phi$  into Lipschitz functions  $\phi_i(r) = \phi(r e_i)$  on each ray  $R_i$ . The second equality in (5.1.11) follows from (5.1.4). We have to prove that  $\mathcal{E}(f) = \mathcal{E}(|f|)$ . The inequality  $\mathcal{E}(f) \geq \mathcal{E}(|f|)$  follows easily from Definition 5.1.9. On the other hand, if  $\phi$  is non expanding, we can suppose  $\phi(O) = 0$ , we use (5.1.12) and again (5.1.4) to obtain

$$\mathcal{E}(\phi \circ f) = \sum_i \mathcal{E}(\phi_i \circ f_i) \leq \sum_i \mathcal{E}(f_i),$$

so  $\mathcal{E}(f) \leq \mathcal{E}(|f|)$ . □

*Remark 5.1.13.* One also has  $\mathcal{E}(f) = \mathcal{E}(\gamma_i \circ f)$  for any Busemann function  $\gamma_i$ .

One can define the energy measure of  $f$  by

$$\mu_{\langle f \rangle} = \mu_{\langle |f| \rangle} = \sum \mu_{\langle f_i \rangle},$$

where the second equality follows from (5.1.3). Thus  $\mathcal{E}(f)$  is the total mass of  $\mu_{\langle f \rangle}$ .

**Lemma 5.1.14.** Let  $f_n$  be a sequence in  $\mathbb{D}^b(N)$  which converges almost everywhere to a function  $f$ . Suppose that  $f_n$  and  $\mathcal{E}(f_n)$  are uniformly bounded. Then  $f$  is in  $\mathbb{D}^b(N)$  and

$$\mathcal{E}(f) \leq \liminf \mathcal{E}(f_n).$$

*Proof.* This is deduced from Lemma 5.1.10 and the similar property for  $\mathbb{D}^b$ , which itself is deduced from the property for  $\mathbb{D}$  and a localisation. □

## 5.2 Energy minimising maps

Let  $g$  be a quasicontinuous map of  $\mathbb{D}^b(N)$ ; we are going to prove the existence of the bounded continuous martingale with final value  $g(X_\tau)$ . The space  $\mathbb{D}_g^b(N)$  is defined in an evident way as in the real case, and the martingale will have the form  $h(X_{t \wedge \tau})$  for  $h$  energy minimising in this space.

The existence of  $h$  can be worked out with the method used for general smooth manifolds in [28], but we will take advantage of the nonpositive curvature of our space to apply a more elementary method with slightly weaker assumptions. The following result is an adaptation of a general analytical theory to our framework, see Theorem 2.2 of [20]. Notice that the Poincaré inequality which is classically used in this method is here replaced by the transience of the killed process and (5.1.7). We begin with a preliminary result.

**Lemma 5.2.1.** *Let  $u$  and  $v$  be quasicontinuous functions of  $\mathbb{D}^b(N)$ , and let  $w(x)$  be the midpoint between  $u(x)$  and  $v(x)$ . Then  $w$  is in  $\mathbb{D}^b(N)$ , and*

$$\mathcal{E}(w) \leq \frac{1}{2}\mathcal{E}(u) + \frac{1}{2}\mathcal{E}(v) - \frac{1}{4}\mathcal{E}(\delta(u, v)). \quad (5.2.2)$$

*Proof.* We deduce from

$$w_i(x) = \frac{1}{2}(\gamma_i \circ u(x) + \gamma_i \circ v(x))^- \quad (5.2.3)$$

that  $w$  is a quasicontinuous function of  $\mathbb{D}^b(N)$ . Its energy measure satisfies

$$\begin{aligned} \mu_{\langle w_i \rangle} &\leq \frac{1}{4}\mu_{\langle \gamma_i \circ u + \gamma_i \circ v \rangle} \\ &= \frac{1}{2}\mu_{\langle \gamma_i \circ u \rangle} + \frac{1}{2}\mu_{\langle \gamma_i \circ v \rangle} - \frac{1}{4}\mu_{\langle \gamma_i \circ u - \gamma_i \circ v \rangle}. \end{aligned}$$

On the set  $A_i = \{u \in R_i\} \cup \{v \in R_i\}$ , one has

$$\delta(u, v) = |\gamma_i \circ u - \gamma_i \circ v|,$$

so

$$\mu_{\langle w_i \rangle} \leq \frac{1}{2}\mu_{\langle u \rangle} + \frac{1}{2}\mu_{\langle v \rangle} - \frac{1}{4}\mu_{\langle \delta(u, v) \rangle} \quad (5.2.4)$$

on  $A_i$ . In particular, since the union of the subsets  $A_i$  is  $M$ , the right-hand side is a nonnegative measure on  $M$ . Moreover, the measures  $\mu_{\langle w_i \rangle}$  are supported by the disjoint sets  $\{w_i > 0\}$ , so

$$\mu_{\langle w \rangle} = \sum_i \mu_{\langle w_i \rangle} \leq \frac{1}{2}\mu_{\langle u \rangle} + \frac{1}{2}\mu_{\langle v \rangle} - \frac{1}{4}\mu_{\langle \delta(u, v) \rangle} \quad (5.2.5)$$

by applying (5.2.4) on  $\{w_i > 0\} \subset A_i$ . Then (5.2.2) follows by integration.  $\square$

**Proposition 5.2.6.** *Consider the above framework with a strongly local regular Dirichlet form  $(\mathbb{D}, \mathcal{E})$  on the locally compact separable space  $M$  with diffusion  $X_t$ , a relatively compact open subset  $M_0$  such that  $X_t$  quits  $M_0$  during its lifetime, and a quasicontinuous function  $g$  of  $\mathbb{D}^b(N)$ . Then there exists a unique (within a modification)  $h$  which minimises the energy  $\mathcal{E}(h)$  among functions of  $\mathbb{D}_g^b(N)$ .*

*Proof.* Choose a minimising sequence consisting of quasicontinuous functions  $h^n$  of  $\mathbb{D}_g^b(N)$ ; one applies (5.2.2) to two elements  $h^m$  and  $h^n$  of the sequence; the midpoint  $h^{m,n}$  is again in  $\mathbb{D}_g^b(N)$ , so

$$\inf_{\mathbb{D}_g^b(N)} \mathcal{E} \leq \mathcal{E}(h^{m,n}) \leq \frac{1}{2}\mathcal{E}(h^m) + \frac{1}{2}\mathcal{E}(h^n) - \frac{1}{4}\mathcal{E}(\delta(h^m, h^n)).$$

By letting  $m$  and  $n$  tend to infinity, one deduces that  $\mathcal{E}(\delta(h^n, h^m))$  converges to 0. The functions  $\delta(h^m, h^n)$  are in  $\mathbb{D}_0$ , and by applying (5.1.7), there exists a subsequence of  $(h^n)$  which converges almost everywhere to a function  $h$ . The subsequence will again be denoted by  $(h_n)$ , and

$$\mathcal{E}(\delta(h^n, h)) \leq \liminf_m \mathcal{E}(\delta(h^n, h^m))$$

converges to 0. By applying the quasicontinuity (see Theorem 2.1.4 of [14]), there exists a subsequence converging quasi everywhere, and  $h$  has therefore a modification which is equal to  $g$  quasi everywhere outside  $M_0$ . Then we deduce from Lemma 5.1.14 that  $h$  is energy minimising in  $\mathbb{D}_g^b(N)$ . For the uniqueness, we see from (5.2.2) that two energy minimising maps  $h^1$  and  $h^2$  should satisfy  $\mathcal{E}(\delta(h^1, h^2)) = 0$ , so  $h^1 = h^2$  from (5.1.7).  $\square$

Now, we want to prove that  $h(X_t)$  is a  $\mathbb{P}^\mu$  martingale. In the case of smooth Riemannian manifolds studied in [28], one notices that since  $h$  is solution of a minimisation problem, then

$$\left. \frac{d}{d\varepsilon} \mathcal{E}(h^\varepsilon) \right|_{\varepsilon=0} = 0 \tag{5.2.7}$$

for any smooth family  $h^\varepsilon$  such that  $h^0 = h$ . Actually, when one is given  $f$ , one can use the perturbation  $h^\varepsilon = \mathcal{T}^\varepsilon \circ h$ , where  $(\mathcal{T}^\varepsilon)$  is the flow of diffeomorphisms of  $N$  defined by the ordinary equation

$$\mathcal{T}^0(y) = y, \quad \frac{d}{d\varepsilon} \mathcal{T}^\varepsilon(y) = f(\mathcal{T}^\varepsilon(y)).$$

The relation (5.2.7) can be written (roughly speaking) as

$$\int_M (f(x), L_N h(x)) d\mu(x) = 0,$$

where  $L_N h$  is the tension field of  $h$ . Since this can be obtained for a large class of functions  $f$ , we can deduce that  $L_N h = 0$  (in a weak sense) and that  $h(X_t)$  is a martingale. Here, this method cannot be immediately applied because one cannot define flows  $\mathcal{T}^\varepsilon$  for all real  $\varepsilon$  except if the point  $O$  is fixed (roughly speaking  $f(O) = 0$ ); this is because all homeomorphisms of  $N$  must let  $O$  fixed. Thus we will only consider a semi-flow ( $\mathcal{T}^\varepsilon, \varepsilon \geq 0$ ) of transformations, and the fact that  $h$  is energy minimising will imply that the derivative of (5.2.7) is nonnegative.

**Theorem 5.2.8.** *Under the assumptions of Proposition 5.2.6, if  $h$  is the energy minimising quasicontinuous map of  $\mathbb{D}_g^b(N)$ , then  $h(X_{t \wedge \tau})$  is a  $\mathbb{P}^\mu$  martingale with limit  $g(X_\tau)$ .*

*Proof.* Fix a ray  $R_i$ , its associated Busemann function  $\gamma_i$ , let  $\rho$  be a nonnegative function of  $\mathbb{D}_0^b$  and let  $\mathcal{T}_x^\varepsilon : N \rightarrow N$  be the translation of step  $\varepsilon\rho(x)$  in the direction of  $R_i$ . The perturbation of  $h(x)$  is defined as

$$h^\varepsilon(x) = \mathcal{T}_x^\varepsilon(h(x)).$$

One has

$$(\gamma_i \circ h^\varepsilon)(x) = (\gamma_i \circ h)(x) - \varepsilon\rho(x),$$

so

$$\left. \frac{\partial(\gamma_i \circ h^\varepsilon)}{\partial \varepsilon} \right|_{\varepsilon=0} = -\rho.$$

The function  $h^\varepsilon$  is in the Dirichlet space  $\mathbb{D}_g^b(N)$  and satisfies

$$\mathcal{E}(\gamma_i \circ h^\varepsilon) = \mathcal{E}(h^\varepsilon) \geq \mathcal{E}(h) = \mathcal{E}(\gamma_i \circ h).$$

By differentiating at  $\varepsilon = 0$ , we obtain that  $\mathcal{E}(\gamma_i \circ h, \rho) \leq 0$ , so, from Lemma 5.1.8,  $\gamma_i(h(X_{t \wedge \tau}))$  is a  $\mathbb{P}^\mu$ -submartingale (for any  $i$ ). Thus  $h(X_{t \wedge \tau})$  is a martingale.  $\square$

Corollaries 5.2.9 and 5.2.11 are similar to [27]. If  $M$  is a Riemannian manifold (or a Riemannian polyhedron), an analytical technique can actually provide the Hölder continuity of  $h$ , see [20, 9, 31].

**Corollary 5.2.9.** *Assume the absolute continuity condition*

$$\forall x \in M \quad \forall t > 0 \quad \mathbb{P}^x[X_t \in dz] \ll \mu(dz). \quad (5.2.10)$$

*Then one can choose the modification of  $h$  so that for any  $x$ , the process  $h(X_{t \wedge \tau})$  is under  $\mathbb{P}^x$  the bounded martingale with limit  $g(X_\tau)$ .*

*Proof.* We obtain from Theorem 5.2.8 and the condition (5.2.10) a  $\mathbb{P}^x$  martingale  $h(X_{t \wedge \tau})$  indexed by  $t > 0$ ; it has a limit as  $t \downarrow 0$  because  $\gamma_i(h(X_{t \wedge \tau}))$  are bounded submartingales and have therefore limits. Let  $h_0(x)$  be the limit. If  $\sigma$  is any stopping time, the quasicontinuity of  $h$  shows that  $h(X_\sigma)$  is the limit of  $h(X_{\sigma+t})$ , so we can deduce that  $h_0 = h$  outside a polar set. Then  $h_0$  satisfies our requirements.  $\square$

**Corollary 5.2.11.** *Assume the absolute continuity condition (5.2.10) and suppose moreover that bounded real functions which are harmonic on an open subset of  $M$  are continuous on this subset. Then the function  $h$  of Corollary 5.2.9 is continuous on  $M_0$ .*

*Proof.* Fix a point  $x$  of  $M_0$  and a non decreasing family  $(V_r)_{r>0}$  of open neighbourhoods of  $x$  such that  $\bigcap V_r = \{x\}$ . Let  $\sigma_r$  be the first exit time of  $V_r$  for  $X_t$ . Then  $h(X_{\sigma_r})$  converges  $\mathbb{P}^x$  almost surely to  $h(x)$  as  $r \downarrow 0$ , so for  $\varepsilon > 0$ , we can choose  $r$  so that

$$\mathbb{E}^x [\delta(h(x), h(X_{\sigma_r}))] \leq \varepsilon.$$

On the other hand, the function

$$x' \mapsto \mathbb{E}^{x'} [\delta(h(x), h(X_{\sigma_r}))]$$

is harmonic on  $V_r$ ; it is continuous from our assumption, so

$$\delta(h(x), h(x')) \leq \mathbb{E}^{x'} [\delta(h(x), h(X_{\sigma_r}))] \leq 2\varepsilon$$

if  $x'$  is close to  $x$ .  $\square$

**Corollary 5.2.12.** *Assume the conditions of Proposition 5.2.6 except that  $M_0$  is not supposed to be relatively compact. Then again there exists a function  $h$  of  $\mathbb{D}^b(N)$  such that  $h(X_{t \wedge \tau})$  is a  $\mathbb{P}^\mu$  martingale with limit  $g(X_\tau)$ , and one has  $\mathcal{E}(h) \leq \mathcal{E}(g)$ . The results of Corollaries 5.2.9 and 5.2.11 can also be extended.*

*Proof.* Let  $M_n$  be the intersection of  $M_0$  with a sequence of relatively compact open subsets of  $M$  which increases to  $M$ . Then we obtain a minimising map  $h^n$ . If  $\tau^n$  is the first exit time of  $M_n$ , then  $h^n(X_{t \wedge \tau^n})$  is the bounded martingale with limit  $g(X_{\tau^n})$ . The sequence  $\mathcal{E}(h^n)$  is non increasing and therefore converges; if  $h^{m,n}$  is the midpoint between  $h^m$  and  $h^n$ , one has

$$\mathcal{E}(h^{m,n}) \geq \min(\mathcal{E}(h^m), \mathcal{E}(h^n))$$

because  $h^{m,n}$  coincides with  $g$  outside  $M_{m \vee n}$ , and  $\mathcal{E}(h^{m,n})$  is dominated with (5.2.2). Thus we can use the technique of Proposition 5.2.6, deduce the

existence of a subsequence converging almost everywhere to a function  $h$  satisfying  $\mathcal{E}(h) \leq \mathcal{E}(g)$ , and  $h(X_{t \wedge \tau})$  is a martingale as a limit of martingales. By using a probability  $\mathbb{P}^\nu$  for  $\nu$  a probability equivalent to  $\mu$ , the probability of  $\{\tau^n \neq \tau\}$  tends to 0, so  $g(X_{\tau^n})$  converges to  $g(X_\tau)$ . Thus the process  $h(X_{t \wedge \tau})$  is the martingale with limit  $g(X_\tau)$ .  $\square$

### 5.3 Martingales for symmetric diffusions

We are now going to prove the existence of martingales for probability spaces generated by symmetric diffusions.

**Theorem 5.3.1.** *Consider like previously a strongly local regular Dirichlet form on a separable locally compact space  $M$ , associated to a diffusion  $X_t$  defined on a canonical probability space  $\Omega$  with measure  $\mathbb{P}^\mu$ . Suppose that the form is conservative so that the lifetime  $\zeta$  of the process is infinite. Then, for any integrable  $N$ -valued variable, there exists a unique  $N$ -valued martingale of class (D) converging to this variable.*

*Proof.* As in Corollary 4.1.13, it is sufficient to consider final variables of the type  $g(X_1)$ . We will suppose that  $g$  is continuous and in  $\mathbb{D}^b(N)$  (this is possible since the form is regular). By taking the product of  $\Omega$  with a Wiener space  $\Omega_W$ , we introduce an independent real Wiener process  $W$ ; on this space, we also consider a nonnegative process satisfying

$$dU_t^\varepsilon = -dt + \sqrt{\varepsilon}dW_t + dA_t^\varepsilon,$$

where  $A_t^\varepsilon$  is the reflection term at 0. The diffusion  $U_t^\varepsilon$  is the symmetric process which is associated to the Dirichlet space  $(\mathbb{D}_U^\varepsilon, \mathcal{E}_U^\varepsilon)$  on  $(\mathbb{R}_+, \mu_U^\varepsilon)$ , where

$$\mu_U^\varepsilon(du) = e^{-u/\varepsilon} du / \varepsilon, \quad \mathcal{E}_U^\varepsilon(f) = \frac{\varepsilon}{2} \int_{\mathbb{R}_+} f'(u)^2 \mu_U^\varepsilon(du),$$

and  $\mathbb{D}_U^\varepsilon$  is the completion for  $\mathcal{E}_U^\varepsilon + |\cdot|_{L^2(\mu_U^\varepsilon)}^2$  of the space of smooth functions on  $\mathbb{R}_+$  with compact support. The process  $(X_t, U_t^\varepsilon)$ , which can be defined on  $\Omega \times \Omega_W$ , is then associated to the product Dirichlet space  $(\mathbb{D}_*^\varepsilon, \mathcal{E}_*^\varepsilon)$  on  $(M \times \mathbb{R}_+, \mu \otimes \mu_U^\varepsilon)$ , with

$$\mathcal{E}_*^\varepsilon(f) = \int \mathcal{E}(f(\cdot, u)) \mu_U^\varepsilon(du) + \int \mathcal{E}_U^\varepsilon(f(x, \cdot)) \mu(dx).$$

This is a regular form (this is because  $\mathbb{D} \otimes \mathbb{D}_U^\varepsilon$  is dense in  $\mathbb{D}_*^\varepsilon$ , see Section V.2.1 of [5]). Let  $\tau(\varepsilon)$  be the first exit time of  $M \times (0, \infty)$ , so

$$\tau(\varepsilon) = \inf \left\{ t \geq 0; U_t^\varepsilon = 0 \right\}.$$

Consider the function  $g(x, u) = g(x)$ , so that  $\mathcal{E}_*^\varepsilon(g) = \mathcal{E}(g)$ . We deduce from Corollary 5.2.12 the existence of a martingale  $Y_t^\varepsilon$  with final value  $g(X_{\tau(\varepsilon)})$ . By proceeding as in Corollary 5.2.9, we can take  $U_0^\varepsilon = 1$  as an initial condition since  $(U_t^\varepsilon)$  satisfies the absolute continuity condition (we let  $X_0$  be distributed according to  $\mu$ ). When  $\varepsilon \downarrow 0$ , the exit time  $\tau(\varepsilon)$  and the final variable  $g(X_{\tau(\varepsilon)})$  converge respectively to 1 and  $g(X_1)$ ; thus, from Corollary 3.3.13, the martingales  $Y_t^\varepsilon$  must converge to the martingale  $Y_t$  with final value  $g(X_1)$ . However, this martingale is defined on the enlarged space  $\Omega \times \Omega_W$ , and we want a martingale on  $\Omega$ . We can apply the Yamada-Watanabe method of which is classically used for stochastic differential equations, see for instance Theorem IX.1.7 of [29]. On the space  $\Omega \times \Omega_W \times \Omega_W$  with product measures, we consider the processes

$$Y_t'(\omega, \omega', \omega'') = Y_t(\omega, \omega'), \quad Y_t''(\omega, \omega', \omega'') = Y_t(\omega, \omega'').$$

Then  $Y_t'$  and  $Y_t''$  are two bounded martingales with the same final value  $g(X_1(\omega))$ , so they are identical. Thus  $Y_t$  can be defined on  $\Omega$ .  $\square$

Similarly to Corollary 5.2.9, we can deduce the following result.

**Corollary 5.3.2.** *The result of Theorem 5.3.1 holds under the probabilities  $\mathbb{P}^x$  if the symmetric diffusion satisfies the absolute continuity condition (5.2.10).*

The initial value of the  $\mathbb{P}^x$  martingale with final value  $g(X_t)$  is written as  $Q_t g(x)$ . We obtain as in Section 4 the nonlinear heat semigroup  $Q_t$ . It is well known that the energy is non increasing along the heat semigroup for smooth manifolds (in the case of maps with values in non positively curved manifolds, this is actually the classical method for proving the existence of a harmonic map in a prescribed homotopy class). Here, this property also holds true.

**Proposition 5.3.3.** *Assume the conditions of Theorem 5.3.1 and the absolute continuity condition (5.2.10). Let  $g$  be in  $\mathbb{D}^b(N)$ . Then  $Q_t g$  is also in  $\mathbb{D}^b(N)$ , and  $t \mapsto \mathcal{E}(Q_t g)$  is non increasing.*

*Proof.* We return to the proof of Theorem 5.3.1. We have a function  $h^\varepsilon$  such that  $h^\varepsilon(X_t, U_t^\varepsilon)$  is a martingale up to the first hitting time of 0 by  $U_t^\varepsilon$ , and

$$\int_0^\infty \mathcal{E}(h^\varepsilon(\cdot, u)) \mu_U^\varepsilon(du) \leq \mathcal{E}_*^\varepsilon(h^\varepsilon) \leq \mathcal{E}_*^\varepsilon(g) = \mathcal{E}(g). \quad (5.3.4)$$

Put

$$Q_u^\varepsilon g(x) = h^\varepsilon(u, x).$$

In particular,  $Q_u^\varepsilon g$  converges to  $Q_u g$ . Consider nonnegative  $u$  and  $v$ ; after a translation, the diffusion  $U_t^\varepsilon$  starting at  $u+v$  and stopped when it hits  $u$  has the same law than the same diffusion starting at  $v$  and stopped when it hits 0. We can deduce from this fact that  $Q_u^\varepsilon$  is a semigroup, and (5.3.4) is written as

$$\int \mathcal{E}(Q_u^\varepsilon g) \mu_U^\varepsilon(du) \leq \mathcal{E}(g).$$

If we consider  $(\mu_U^\varepsilon)^{(2)} = \mu_U^\varepsilon \star \mu_U^\varepsilon$  (the convolution product), then

$$\begin{aligned} \int \mathcal{E}(Q_u^\varepsilon g) (\mu_U^\varepsilon)^{(2)}(du) &= \iint \mathcal{E}(Q_{u+v}^\varepsilon g) \mu_U^\varepsilon(dv) \mu_U^\varepsilon(du) \\ &\leq \int \mathcal{E}(Q_u^\varepsilon g) \mu_U^\varepsilon(du) \leq \mathcal{E}(g) \end{aligned}$$

and more generally

$$\begin{aligned} \mathcal{E}\left(\int_0^\infty Q_u^\varepsilon g (\mu_U^\varepsilon)^{(n)}(du)\right) &\leq \int_0^\infty \mathcal{E}(Q_u^\varepsilon g) (\mu_U^\varepsilon)^{(n)}(du) \\ &\leq \mathcal{E}(g) \end{aligned} \tag{5.3.5}$$

where the left-hand side is computed in  $\mathbb{R}^\ell$ . By taking  $n \sim t/\varepsilon$ , we can let the measure  $(\mu_U^\varepsilon)^{(n)}$  (which is a convolution of exponential distributions with mean  $\varepsilon$ ) tend to a Dirac mass at  $t$ . Moreover, the function which maps  $u$  to the hitting time of 0 by  $U_t^\varepsilon$  when it starts from  $u$  is continuous in probability, uniformly in  $\varepsilon$  for  $0 < \varepsilon < 1$ ; we deduce that  $Q_u^\varepsilon g(x)$  is continuous with respect to  $u$ , uniformly in  $\varepsilon$ . Thus the integral in left-hand side of (5.3.5) converges to  $Q_t g$ , so we deduce from the lower semicontinuity of the energy that  $\mathcal{E}(Q_t g) \leq \mathcal{E}(g)$ . Since  $Q_t$  is a semigroup, the proof is complete.  $\square$

*Remark 5.3.6.* One can proceed as in Corollary 5.2.11 for the continuity of  $Q_t g$ . This means that a continuity property for  $P_t f(x)$  is transferred to  $Q_t g(x)$ .

Numerous examples of Dirichlet spaces are of course known. If we restrict to trees, we can consider the diffusion on the continuum tree constructed in [21]. Here is a simpler example.

*Example 5.3.7.* Consider a space  $M$  which is the union of a countable number of segments  $S_i = [O, A_i]$  with length  $\ell_i > 0$ , such that the intersection of two different segments is reduced to the point  $O$ . We suppose that  $\sum \ell_i$  is finite. Then we put

$$\mathcal{E}(f) = \frac{1}{2} \sum_i \int_{S_i} f'(x)^2 \lambda_i(dx)$$

for the Lebesgue measure  $\lambda_i$  on  $S_i$ , where  $f$  is a Lipschitz function which is smooth on each segment  $S_i$  and is constant on all but a finite number of segments. After taking the closure, we obtain a regular Dirichlet form. Notice that in this example (and also for more general trees), elements of  $\mathbb{D}$  are automatically Hölder continuous with rate  $1/2$ , so this implies a regularity on harmonic maps in  $\mathbb{D}^b(N)$ .

## 6 Generalisation to trees

Up to now, we have focussed on our baby tree  $\mathbb{Y}^\ell$ . Our aim is now to generalise the results of previous sections to other trees.

### 6.1 The geometry of trees

The simplest generalisation is the family of complete finite metric trees. They consist of a finite number  $\ell$  of edges which are isometric to closed intervals of  $\mathbb{R}$ , some endpoints of which are glued, and which yield a connected and simply connected space (there is no loop). The endpoints of the edges are the vertices of the tree; a particular class of vertices are those which belong to only one edge; they are the leaves of the tree. Then  $N$  can be embedded in  $\mathbb{R}^\ell$  by letting the  $i$ th edge being parallel to the  $i$ th axis. More generally, we can consider  $\mathbb{R}$ -trees, see for instance the definition in [13].

**Definition 6.1.1.** *A  $\mathbb{R}$ -tree is a metric space  $(N, \delta)$  satisfying the two following properties for any  $y_1$  and  $y_2$  in  $N$ .*

1. *There is one and only one isometry from the interval  $[0, \delta(y_1, y_2)]$  into  $N$  mapping 0 to  $y_1$  and  $\delta(y_1, y_2)$  to  $y_2$ .*
2. *If  $\phi : [0, 1] \rightarrow N$  is injective,  $\phi(0) = y_1$ ,  $\phi(1) = y_2$ , then  $\phi([0, 1])$  is isometric to  $[0, \delta(y_1, y_2)]$  (it is the arc joining  $y_1$  and  $y_2$ ).*

*Since only  $\mathbb{R}$ -trees will be considered, we will simply call them ‘trees’. We say that  $N$  is a measurable tree if moreover it is endowed with a  $\sigma$ -algebra containing the balls. When  $N$  is separable, it will be implicitly endowed with its Borel  $\sigma$ -algebra.*

The unique arc between two points of  $N$  is a geodesic, and as in Definition 2.1.1, a function is said to be convex when it is convex on geodesics parameterised by arc length. For instance, the distance functions  $\delta(O, \cdot)$  are convex.

The trees that we will consider will often be assumed to be complete. If  $y$  is a point of  $N$  and if the number of connected components of  $N \setminus \{y\}$  is different from two, then  $y$  is a vertex of the tree; in particular, if this number is one, then  $y$  is a leaf. A ray is a subset of  $N$  which is isometric to  $\mathbb{R}_+$ , and an end of  $N$  is an equivalence class of rays, where two rays are equivalent if their intersection is a ray. Then if one fixes an origin  $O$  and an end with some parametrisation  $(\xi_t; t \in \mathbb{R}_+)$ , one can define the associated Busemann function by

$$\gamma_\xi(y) = \lim_{t \rightarrow \infty} (\delta(\xi_t, y) - \delta(\xi_t, O)) \quad (6.1.2)$$

where the limit does not depend on the parametrisation and is actually stationary (changing  $O$  only adds a constant). This is a convex function. If  $\xi$  is an end and  $y$  a point, we can consider the ray  $[y, \xi)$  with origin  $y$  and equivalent to  $\xi$ ; its elements are called the ancestors of  $y$  (relatively to  $\xi$ ); if  $y$  and  $y'$  are two points, their first common ancestor  $y \wedge y'$  is the origin of the ray  $[y, \xi) \cap [y', \xi)$ ; it is in the arc  $[y, y']$ , and

$$\begin{aligned}\delta(y, y') &= \delta(y, y \wedge y') + \delta(y \wedge y', y') \\ &= \gamma_\xi(y) + \gamma_\xi(y') - 2\gamma_\xi(y \wedge y').\end{aligned}\tag{6.1.3}$$

Contrary to  $\mathbb{Y}^\ell$ , Busemann functions will not be sufficient for the characterisation of continuous martingales; for instance, a bounded tree has no end. We also have to consider leaves. If  $y_0$  is a point of  $N$ , we define

$$\gamma_{y_0}(x) = \delta(y_0, x) - \delta(y_0, O).\tag{6.1.4}$$

The set of ancestors of  $y$  are in this case defined to be the arc  $[y_0, y]$ , the endpoints of  $[y_0, y] \cap [y_0, y']$  are  $y_0$  and the first common ancestor  $y \wedge y'$ , and (6.1.3) is again satisfied. The functions  $\gamma_\xi$  when  $\xi$  is an end, and  $\gamma_{y_0}$  when  $y_0$  is a leaf, will be called the basic convex functions.

There are relations between these functions. If  $\xi$  and  $\xi'$  are two different ends, we can consider the line  $(\xi, \xi')$  joining them, which can be parameterised by  $(\xi_t; t \in \mathbb{R})$ ; we can apply

$$\delta(y, y') + \delta(y, y'') = \delta(y', y'') + 2\delta(y, [y', y''])$$

to  $y' = \xi_t$ ,  $y'' = \xi_{-t}$ , let  $t$  tend to  $+\infty$ , and deduce that

$$\gamma_\xi(y) + \gamma_{\xi'}(y) = 2\delta(y, (\xi, \xi')) - 2\delta(O, (\xi, \xi')).\tag{6.1.5}$$

There are similar relations with functions  $\gamma_{y_0}$ , involving rays  $[y_0, \xi)$ , or segments  $[y_0, y_0']$ .

In previous sections, we have used the embedding of  $\mathbb{Y}^\ell$  into  $\mathbb{R}^\ell$ . Separable complete trees have a similar property.

**Lemma 6.1.6.** *Consider the Banach space  $\ell^1$  of absolutely convergent real series and its subset  $\ell_+^1$  of convergent nonnegative series. Let  $e_k$  be the canonical series having 1 at the  $k$ th row and 0 elsewhere. On the other hand, let  $N$  be a separable complete tree. Then there is an isometric embedding which maps  $N$  onto a closed subset of  $\ell_+^1$  containing 0 and which is the closure of a set of the form  $\bigcup [z_k, z_k + \alpha_k e_k]$ , for  $z_k \in \ell_+^1$  and  $\alpha_k \geq 0$ . Then the distance  $\delta(y, z)$  in  $N$  is equal to the norm  $|z - y|$  in  $\ell^1$ .*

*Proof.* Let  $(y_k)_{k \geq 0}$  be a dense subset of  $N$ . Let  $N_n$  be the convex hull of  $\{y_k; 0 \leq k \leq n\}$ . Then  $N_n$  is a finite tree with at most  $n$  edges. We can consider the embedding of  $N_n$  in  $\mathbb{R}_+^n$  as follows; we let  $N_0 = \{0\}$  and we let  $N_j \setminus N_{j-1}$ , if not empty, be a segment of type  $[z_j, z_j + \alpha_j e_j]$  for  $z_j$  in  $N_{j-1}$  and  $\alpha_j \geq 0$ . These embeddings are compatible, so  $\bigcup N_n$  can be embedded in the subset of  $\ell_+^1$  consisting of nonnegative sequences such that only a finite number of terms is positive. The distance on  $N$  corresponds to the  $\ell^1$  distance on this space, so by taking the closure,  $N$  is embedded in the Banach space  $\ell^1$ .  $\square$

*Example 6.1.7.* The star  $\mathbb{Y}^{\mathbb{N}}$  with countably many rays is an example of separable tree. The star of Example 5.3.7 is a subtree of  $\mathbb{Y}^{\mathbb{N}}$ .

The situation is more complicated with non separable trees. In this case, the Borel  $\sigma$ -algebra is generally too large, so that the number of  $N$ -valued measurable variables is too small. Thus we have to endow  $N$  with another  $\sigma$ -algebra, and this is why we have introduced in Definition 6.1.1 the notion of measurable tree; the minimal admissible  $\sigma$ -algebra corresponds to the Baire topology (which is generated by open balls). Notice that continuous functions are not necessarily measurable for these  $\sigma$ -algebras, so, when dealing for instance with convex functions, we will have to restrict to measurable ones. Though stochastic calculus can be worked out on non separable trees (see the results below), the problem of finding a martingale with given limit has generally no solution. However, we mention these spaces because there are classical examples of them, they are a simple description of what can happen on a non separable space, and some classical diffusions are defined on the following examples.

*Example 6.1.8.* Consider the space  $\mathbb{R}^2$  as a star with origin 0; this is the metric space  $N = \mathbb{Y}^{\mathbb{S}^1}$  where the rays are indexed by the circle  $\mathbb{S}^1$ . Let  $Y$  be a Borel measurable  $N$ -valued variable; then the angular coordinate  $\Theta(Y) \in \mathbb{S}^1$  should be measurable for the  $\sigma$ -algebra of all subsets of  $\mathbb{S}^1$  (because any union of rays  $R_\theta^*$  is open); thus it should be supported by a countable subset, and  $Y$  should therefore be supported by a countable number of rays. This shows that if we want to consider more general variables, we have to choose another  $\sigma$ -algebra satisfying the conditions of Definition 6.1.1. We can for instance use the Borel  $\sigma$ -algebra associated to the classical topology of  $\mathbb{R}^2$ .

*Example 6.1.9.* Consider the space  $\mathbb{G}$  of paths with birth and death times described as follows; the elements of  $\mathbb{G}$  are triples  $y = (\omega_y, S_y, T_y)$ , with  $S_y \leq T_y$  and  $(\omega_y(u); u \in \mathbb{R})$  is a càdlàg path with values in the disconnected union of  $\mathbb{R}$  and a point  $\{\partial\}$ , such that

$$\omega_y(u) \in \mathbb{R} \iff S_y \leq u < T_y, \quad \omega_y(S_y) = 0 \quad \text{if} \quad S_y < T_y.$$

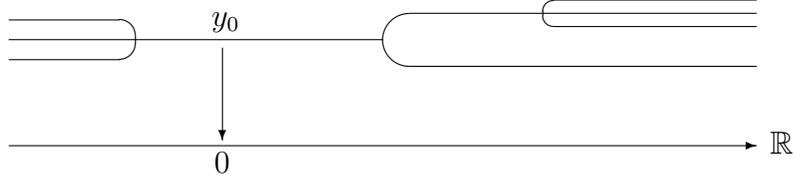


Figure 5: An example of oriented distance function  $\psi(y_0^\alpha, \cdot)$ .

The distance on  $\mathbb{G}$  is defined by

$$\delta(y, y') = T_y + T_{y'} - 2 \left( \inf \{u; \omega_y(u) \neq \omega_{y'}(u)\} \wedge T_y \wedge T_{y'} \right). \quad (6.1.10)$$

Then  $\mathbb{G}$  is a non separable tree. If we consider the end associated to the ray  $\{S_y = T_y \leq 0\}$ , the Busemann function is  $\gamma(y) = T_y$ ; if  $y \neq y'$ , their first common ancestor  $y \wedge y'$  consists of the paths  $y$  and  $y'$  killed when they begin to diverge, and the arcs joining  $y$  or  $y'$  to  $y \wedge y'$  are obtained by erasing  $y$  or  $y'$  progressively up to the divergence time. If we endow  $\mathbb{G}$  with the  $\sigma$ -algebra generated by the functions  $y \mapsto \omega_y(u)$ , it is a measurable tree. We can also consider the subtree  $\mathbb{G}^c$  consisting of the elements  $y$  such that the path  $\omega_y$  is continuous on  $[S_y, T_y)$ .

The basic geometric properties of  $\mathbb{Y}^\ell$  can be extended to more general trees. We first need to verify that the basic functions which were used in Section 2 are measurable, even if the tree is not separable.

**Lemma 6.1.11.** *Let  $N$  be a measurable tree. The distance functions  $\delta(y_0, \cdot)$  and the Busemann functions (6.1.2) are measurable. Moreover, if  $y_0^\alpha$  is a connected component of  $N \setminus \{y_0\}$ , then the oriented distance function (2.2.3) is measurable.*

*Proof.* The assertion concerning distance and Busemann functions is trivial since balls are measurable. For the oriented distance function, we have to check that  $y_0^\alpha$  is measurable. Choose an interval  $(y_0, y_1)$  in  $y_0^\alpha$  and a sequence  $y(n)$  in this interval converging to  $y_0$ ; then  $y$  is in  $y_0^\alpha$  if and only if  $y(n)$  belongs to the arc  $[y_0, y]$  for  $n$  large, so we can write

$$y_0^\alpha = \bigcup_n \left\{ y; \delta(y_0, y) = \delta(y_0, y(n)) + \delta(y(n), y) \right\},$$

and this set is measurable. □

The following result will also be useful later.

**Lemma 6.1.12.** *Let  $N$  be a complete tree, let  $\gamma$  be a distance or Busemann function, and let  $(y_t; t \geq 0)$  be a continuous path in  $N$  such that  $\gamma(y_t)$  converges in  $\mathbb{R}$  as  $t \uparrow \infty$ . Then  $y_t$  converges in  $N$ .*

*Proof.* Fix  $\varepsilon > 0$ . For  $t$  large enough, we have

$$\sup_{s \geq t} |\gamma(y_s) - \gamma(y_t)| \leq \varepsilon.$$

Moreover, we deduce from the continuity of the path that  $y_s \wedge y_t$  belongs to  $\{y_u; t \leq u \leq s\}$ , so

$$\gamma(y_t) - \gamma(y_s \wedge y_t) \leq \varepsilon$$

for  $s \geq t$ . Thus, from (6.1.3),

$$\begin{aligned} \delta(y_s, y_t) &= \gamma(y_s) + \gamma(y_t) - 2\gamma(y_s \wedge y_t) \\ &= \left( \gamma(y_s) - \gamma(y_t) \right) + 2\left( \gamma(y_t) - \gamma(y_s \wedge y_t) \right) \leq 3\varepsilon \end{aligned}$$

for  $s \geq t$ , and the convergence follows from the completeness of  $N$ .  $\square$

One can consider as in Definition 2.2.1 the barycentre for square integrable variables, then for integrable variables, and it is not difficult to extend the results of Section 2.

**Proposition 6.1.13.** *The results of Propositions 2.2.4, 2.3.5 and 2.3.7 hold true for measurable trees (for Proposition 2.3.5 one has to restrict to measurable functions  $f$ ).*

## 6.2 Stochastic calculus on a tree

As in Definition 3.1.1, an adapted càdlàg process  $Y_t$  in a measurable tree is said to be a semimartingale if  $f(Y_t)$  is a semimartingale for any measurable Lipschitz convex function  $f$ . In the general case, it does not seem easy to find a characterisation, but in the continuous case, a single function is again sufficient for testing the semimartingale property (as in Proposition 3.1.4).

**Proposition 6.2.1.** *Fix a root  $O$  in the measurable tree  $N$ , and let  $|y|$  be the distance  $\delta(O, y)$ . Then an adapted continuous process  $Y_t$  is a semimartingale if and only if  $|Y_t|$  is a semimartingale.*

*Proof.* Suppose that  $|Y_t|$  is a semimartingale. If  $f$  is a measurable non expanding convex function, we have to prove that  $f(Y_t)$  is a semimartingale. By a standard stopping argument, we can suppose that  $|Y_t|$  and the variation of its finite variation part are bounded by some constant  $C$ . Suppose also

that  $f$  is nonnegative and  $f(O) = 0$  (the general case immediately follows by adding to  $f$  a constant and a multiple of the distance to  $O$ ). Let us study  $f(Y_t)$  conditionally on  $\mathcal{F}_0$  for some fixed  $Y_0 = y_0$  (this is possible because conditional probabilities exist on  $\Omega$ , and we choose  $y_0$  so that  $|Y_t|$  is conditionally a semimartingale satisfying the above boundedness conditions). Fix some  $\eta > 0$ . The function  $f$  is convex on the arc  $[O, y_0]$ ; its right derivative  $f'$  is a right continuous non decreasing function, and we can choose a subdivision  $\zeta = (z_k; 0 \leq k \leq K)$  of the arc  $[O, y_0]$  such that the oscillation of  $f'$  on each arc  $[z_k, z_{k+1}]$  is at most  $\eta$ . The piecewise affine function defined on this arc by interpolation from  $f$  can be written as

$$\begin{aligned}\phi_\zeta(y) &= \sum_{k=0}^{K-1} (f(z_k) + \lambda_k \delta(z_k, y)) 1_{(z_k, z_{k+1}]}(y) \\ &= \sum_{k=0}^{K-1} \rho_k \delta(z_k, y) 1_{(z_k, y_0]}(y)\end{aligned}$$

where  $\rho_k = \lambda_k - \lambda_{k-1} \geq 0$  (put  $\lambda_{-1} = 0$ ), and

$$f'(z_k) \leq \lambda_k \leq f'(z_k) + \eta.$$

For  $0 \leq k \leq K-1$ , let  $z_k^0$  be the connected component of  $N \setminus \{z_k\}$  containing  $y_0$ , and put  $z_K^0 = \emptyset$ . Then  $\phi_\zeta$  coincides on  $[O, y_0]$  with the function

$$\begin{aligned}f_\zeta(y) &= \sum_{k=0}^{K-1} (f(z_k) + \lambda_k \delta(z_k, y)) 1_{z_k^0 \setminus z_{k+1}^0}(y) \\ &= \sum_{k=0}^{K-1} \rho_k \delta(z_k, y) 1_{z_k^0}(y)\end{aligned}\tag{6.2.2}$$

which is convex on  $N$ . Moreover, for  $y$  in  $z_k^0 \setminus z_{k+1}^0$  such that  $|y| \leq C$ , one has

$$\begin{aligned}f(y) &\geq f(z_k) + f'(z_k) \delta(z_k, y) \\ &\geq f(z_k) + (\lambda_k - \eta) \delta(z_k, y) \\ &= f_\zeta(y) - \eta \delta(z_k, y) \\ &\geq f_\zeta(y) - 2\eta C.\end{aligned}\tag{6.2.3}$$

On the other hand, from Lemma 3.1.5, the process

$$\delta(z_k, Y_t) 1_{z_k^0}(Y_t) = \delta(z_k, y_0) + \int_0^t 1_{z_k^0}(Y_s) d|Y_s| + \frac{1}{2} L_t^k$$

is a semimartingale, so, by applying (6.2.2),

$$f_\zeta(Y_t) = f_\zeta(y_0) + \int_0^t \left( \sum_{k=0}^{K-1} \rho_k 1_{z_k^0}(Y_s) \right) d|Y_s| + A_t$$

for a non decreasing process  $A_t$ . One has

$$\sum_{k=0}^{K-1} \rho_k = \lambda_{K-1} = \frac{f(z_K) - f(z_{K-1})}{\delta(z_{K-1}, z_K)} \leq 1$$

(recall that  $f$  is non expanding), so

$$f_\zeta(y_0) - \mathbb{E}[f_\zeta(Y_t) \mid \mathcal{F}_0] \leq \mathbb{E}\left[\int_0^t |dV_s| \mid \mathcal{F}_0\right],$$

where  $V_t$  is the finite variation part of  $|Y_t|$ , and where the conditional probability is taken for our fixed  $Y_0 = y_0$ . From (6.2.3) and since  $f_\zeta(y_0) = f(y_0)$ , we have

$$f(y_0) - f(Y_t) \leq f_\zeta(y_0) - f_\zeta(Y_t) + 2\eta C, \quad (6.2.4)$$

so

$$f(y_0) - \mathbb{E}[f(Y_t) \mid \mathcal{F}_0] \leq \mathbb{E}\left[\int_0^t |dV_s| \mid \mathcal{F}_0\right] + 2\eta C.$$

We can let  $\eta$  tend to 0, and allow  $y_0$  to vary, so

$$f(Y_0) - \mathbb{E}[f(Y_t) \mid \mathcal{F}_0] \leq \mathbb{E}\left[\int_0^t |dV_s| \mid \mathcal{F}_0\right].$$

If we write a similar inequality for any time interval, and by summing these inequalities for the intervals  $[t_j, t_{j+1}]$  of a subdivision of  $[0, t]$ , we obtain

$$\mathbb{E} \sum_j \left( f(Y_{t_j}) - \mathbb{E}[f(Y_{t_{j+1}}) \mid \mathcal{F}_{t_j}] \right)^+ \leq \mathbb{E} \int_0^t |dV_s| \leq C.$$

This implies as in the proof of Proposition 3.2.3 that  $f(Y_t)$  is a quasimartingale, and therefore a semimartingale.  $\square$

In particular, we see that the semimartingale property of  $|Y_t| = \delta(O, Y_t)$  does not depend on the choice of  $O$ .

**Corollary 6.2.5.** *In Proposition 6.2.1, we can replace  $|Y_t|$  by  $\gamma(Y_t)$ , where  $\gamma$  is any Busemann function (6.1.2).*

*Proof.* Suppose that  $\gamma(Y_t)$  is a semimartingale, where  $\gamma$  is the Busemann function associated to a ray  $R$ . We can again assume that  $Y_t$  lives in a bounded subset  $N_0$  of  $N$ . Let  $O$  be a point in the unbounded connected component of  $R \setminus N_0$ . Then  $|Y_t| = \delta(O, Y_t)$  is equal to  $\gamma(Y_t)$  modulo an additive constant, and is therefore a semimartingale. Thus we can apply Proposition 6.2.1.  $\square$

*Example 6.2.6.* Consider the measurable star  $\mathbb{Y}^{\mathbb{S}^1}$  of Example 6.1.8. Walsh processes  $X_t$  can be constructed similarly to Example 3.0.1 on this space by choosing a probability  $\mu$  on the circle, see [34]. As it was the case for  $\mathbb{Y}^\ell$ , they behave like a Brownian motion on each ray, and when they arrive at  $O$ , they choose a ray according to  $\mu$ . Then  $|X_t|$  is a reflected Brownian motion, so  $X_t$  is a semimartingale.

*Example 6.2.7.* Consider the measurable tree  $\mathbb{G}$  of Example 6.1.9, and the subtree  $\mathbb{G}_0 = \{y; S_y = 0\}$ . The Brownian snake ([22, 23]) and the Poisson snake (see for instance [1]) are two diffusions  $X_t$  on  $\mathbb{G}_0$  such that the lifetime  $T_{X_t}$  (the height of the snake) is a reflected Brownian motion. If we recall that  $\gamma(y) = T_y$  is a Busemann function, we deduce from Corollary 6.2.5 that these snakes are semimartingales.

*Example 6.2.8.* By a slight modification of the construction, we can also consider Brownian and Poisson snakes in  $\mathbb{G}$  with real Brownian height (the snake becomes a worm). More precisely, if  $X'_t$  is a Brownian or Poisson snake in  $\mathbb{G}_0$  with initial value  $X'_0 = O$  defined by  $S_O = T_O = 0$ , a snake  $X_t$  in  $\mathbb{G}$  with initial value  $X_0 = x$  can be constructed as follows. The height  $T_{X'_t}$  is a reflected Brownian motion that we write as

$$T_{X'_t} = B_t - \inf_{s \leq t} B_s$$

for a standard real Brownian motion  $B_t$ , and  $X_t$  is defined by

$$T_{X_t} = T_x + B_t, \quad S_{X_t} = S_x \wedge \inf\{T_{X'_s}; 0 \leq s \leq t\},$$

and

$$\omega_{X_t}(S_{X_t} + u) = \omega_{X'_t}(u).$$

These diffusions are again semimartingales and will be further studied later. Notice that the Brownian snake lives in the subtree  $\mathbb{G}^c$  of  $x$  such that  $\omega_x$  is continuous on  $[S_x, T_x]$ .

*Example 6.2.9.* Consider a complete separable tree without leaves, choose an end  $\xi$ , and let  $\gamma_\xi$  be the corresponding Busemann function (6.1.2). Evans has constructed in [13] a family of diffusions  $X_t$  on this space, so that  $\gamma_\xi(X_t)$  is a real Brownian motion. They are semimartingales.

In the above examples, the trees had no leaf. On the other hand, when the tree has many leaves, the semimartingale property is seldom satisfied by diffusions. Let us give an example.

*Example 6.2.10.* Consider the star and the diffusion  $X_t$  of Example 5.3.7. Fix a point  $A \neq O$  and let  $X_t$  start at  $O$  and stop at the first hitting time  $\tau_A$  of  $A$  (the process is recurrent and  $\tau_A$  is almost surely finite). We are going to prove that  $X_t$  is not a semimartingale by checking that the local time of  $|X_t|$  at point  $O$  and time  $\tau_A$  should be infinite. The local time at  $O$  on the segment containing  $A$  is some positive variable  $L_A$ . If  $N_0$  is a subtree consisting of  $n$  segments and containing  $A$ , the trace of  $X$  on  $N_0$  is the Walsh process with isotropic choice at  $O$  and reflection at the endpoints. In particular, the local times at  $O$  on the different segments are equal, so the total local time is equal to  $nL_A$ . By letting  $n \uparrow \infty$ , we see that the total local time should be infinite.

### 6.3 Quasimartingales

Quasimartingales can be defined as in Definition 3.2.1, and one can wonder whether the first two conditions of Proposition 3.2.3 are again equivalent. One can indeed prove similarly that quasimartingales are transformed into real quasimartingales by measurable Lipschitz convex functions (in particular they are semimartingales). However, the converse is true for finite trees, but is not evident on infinite trees; we can actually describe a counterexample for a non separable tree. This shows that quasimartingales are probably difficult to handle on general non positively curved spaces.

*Example 6.3.1.* Consider the Brownian snake  $X_t$  in  $\mathbb{G}$  of Example 6.2.8. Let  $X_0 = x_0$  and let us look for the barycentre  $\mathbb{B}[X_t]$ . Consider the end  $\xi$  associated to the ray  $\{S_y = T_y \leq 0\}$  and its Busemann function  $\gamma(y) = T_y$ ; the first common ancestor  $y \wedge y'$  for this end has been described in Example 6.1.9. Let  $x$  be some point of  $\mathbb{G}$ ; the path  $\omega_{X_t}$  has a part which is common with  $\omega_{x_0}$ , and the other part is governed by the Wiener measure, so has almost surely no interval common with  $\omega_x$ . This implies that the only intervals where  $\omega_x$  and  $\omega_{X_t}$  can coincide are those where  $\omega_x$  and  $\omega_{x_0}$  coincide, so  $x \wedge x_0$  is in the arc joining  $x$  and  $X_t$ , and

$$\delta(x, X_t) = \delta(x, x \wedge x_0) + \delta(x \wedge x_0, X_t).$$

We want to minimise the quadratic mean of the left hand side, and it is clear that we must have  $x = x \wedge x_0$ , so  $x = \mathbb{B}[X_t]$  should be an ancestor of  $x_0$ . We have to compute  $T_x$ , and we know that  $T_x \leq T_{x_0}$ . The fact that  $x$  is an

ancestor of  $x_0$  implies that  $x_0 \wedge X_t$  is in the arc  $[x, X_t]$ , so we have

$$\begin{aligned}\delta(x, X_t) &= \delta(x, x_0 \wedge X_t) + \delta(x_0 \wedge X_t, X_t) \\ &= \left| T_x - \inf_{s \leq t} T_{X_s} \right| + T_{X_t} - \inf_{s \leq t} T_{X_s}.\end{aligned}$$

Moreover, we have  $T_{X_s} = T_{x_0} + B_s$  for a Brownian motion  $B_s$ , so

$$\delta(x, X_t) = \left| T_x - T_{x_0} - \inf_{s \leq t} B_s \right| + B_t - \inf_{s \leq t} B_s.$$

The difference  $\delta(x_0, x) = T_{x_0} - T_x$  is obtained by minimising the quadratic mean of this variable, and the scaling property of the Brownian motion shows that  $\delta(x_0, x) = c\sqrt{t}$ . More generally,

$$\delta(X_s, \mathbb{B}[X_t | \mathcal{F}_s]) = c\sqrt{t-s},$$

so  $X_t$  is not a quasimartingale. However,  $\gamma(X_t) = T_{X_t}$  is a real Brownian motion, and

$$\gamma(x) + \delta(A, x) = 2\delta([A, \xi], x) + \gamma(A)$$

for any point  $A$ . The process  $\delta([A, \xi], X_t)$  is from Lemma 3.1.5 a submartingale, so  $\delta(A, X_t)$  is also a submartingale. The method used in Proposition 6.2.1 enables to deduce that  $f(X_t)$  is a quasimartingale for any Lipschitz convex function, so Proposition 3.2.3 does not hold true for this space.

## 6.4 Continuous martingales

Continuous martingales are defined as in Definition 3.3.3 (one has to restrict to measurable functions  $f$ ). Proposition 3.3.10 again holds true (a limit in probability of martingales is a martingale). The analogue of Proposition 3.3.4 is given as follows.

**Proposition 6.4.1.** *In a measurable complete tree  $N$ , a continuous adapted process  $Y_t$  is a martingale if and only if  $\gamma(Y_t)$  is a local submartingale for any basic convex function  $\gamma$  of type (6.1.2) for an end or (6.1.4) for a leaf.*

*Proof.* We have to prove that the condition is sufficient. Fix some point  $O$ , a connected component  $N_0$  of  $N \setminus \{O\}$ , and let us first prove that  $|Y_t|_{1_{N_0}}(Y_t)$  is a local submartingale. We suppose that  $O$  is not a leaf (otherwise this is included in the assumption), we choose an end or a leaf of  $N$  which is also an end or leaf of  $N \setminus N_0$ , and we let  $\gamma$  be the basic convex function associated to this end or leaf. By applying Lemma 3.1.5, we have

$$|Y_t|_{1_{N_0}}(Y_t) = |Y_0|_{1_{N_0}}(Y_0) + \int_0^t 1_{N_0}(Y_s) d\gamma(Y_s) + \frac{L_t}{2},$$

so the left hand side is a local submartingale. By adding expressions of this type, we can consider functions  $f_\zeta$  of type (6.2.2) and deduce that  $f_\zeta(Y_t)$  is a local submartingale. Now consider a measurable Lipschitz convex function  $f$  and let us study  $f(Y_t)$ . We can restrict ourselves to the case where  $Y_t$  lives in a bounded part of  $N$ , and by modifying  $f$  outside this part, we can suppose that  $f$  is minimal at some point  $O$ . We apply the method of Proposition 6.2.1 and in particular (6.2.4) with  $\eta \downarrow 0$  to deduce that

$$f(Y_0) \leq \mathbb{E}[f(Y_t) \mid \mathcal{F}_0].$$

A similar inequality for more general time intervals shows that  $f(Y_t)$  is a submartingale.  $\square$

*Example 6.4.2.* As in Proposition 3.3.4, on a finite tree, a continuous martingale is a process which is a local martingale on edges and which, when hitting a vertex, chooses no edge with probability greater than 1/2 (in the sense of (3.3.5)).

**Corollary 6.4.3.** *Let  $Y_t$  be a continuous adapted process. If  $\gamma(Y_t)$  is a local martingale for some basic convex function  $\gamma$ , then  $Y_t$  is a martingale.*

*Proof.* Suppose for instance that  $\gamma = \gamma_\xi$  for some end  $\xi$ , and let us check that  $\gamma_{\xi'}(Y_t)$  is a local submartingale for other ends  $\xi'$  (the case of leaves is dealt with similarly). Consider the line  $(\xi, \xi')$ . By applying Lemma 3.1.5, we check that  $\delta((\xi, \xi'), Y_t)$  is a local submartingale, and we conclude by means of (6.1.5).  $\square$

*Example 6.4.4.* The Evans processes (Example 6.2.9) are clearly martingales from Corollary 6.4.3.

*Example 6.4.5.* In  $\mathbb{G}$ , if we use the Busemann function  $\gamma(y) = T_y$  (see Example 6.1.9), Corollary 6.4.3 shows that Brownian and Poisson snakes (Example 6.2.8) are martingales. On the other hand, if we consider the snakes  $X_t$  in  $\mathbb{G}_0$  (Example 6.2.7), the process  $T_{X_t}$  is a reflected Brownian motion, so we have to study more carefully the behaviour of  $X_t$  when it is at the point  $O$  defined by  $T_O = 0$ . It appears that the Brownian snake is again a martingale, but not the Poisson snake.

**Proposition 6.4.6.** *If  $(Y_t; t > 0)$  is a continuous martingale of class (D) on a complete tree  $N$ , then  $Y_t$  has almost sure limits as  $t \downarrow 0$  and  $t \uparrow \infty$ , and  $(Y_t; 0 \leq t \leq \infty)$  is a martingale of class (D).*

*Proof.* If  $\gamma$  is a distance or Busemann function, then  $\gamma(Y_t)$  is a submartingale of class (D), so converges. Thus  $Y_t$  converges from Lemma 6.1.12. The extension of the martingale property to the compact interval  $[0, \infty]$  is not difficult.  $\square$

## 6.5 Martingales with prescribed limit

The uniqueness of a martingale of class (D) with prescribed limit is stated in the following result which is proved like the results it refers to.

**Proposition 6.5.1.** *If  $\delta$  is measurable on  $N \times N$ , the statements of Proposition 3.3.11 and Corollaries 3.3.12 and 3.3.13 hold true for measurable trees (for Corollary 3.3.13 we suppose that the tree is complete and separable).*

For the existence, we have to extend the results of Sections 4 and 5. However, this can be done only for separable trees.

**Theorem 6.5.2.** *The results of Sections 4.1 and 4.2 (Theorem 4.1.4, Corollaries 4.1.8 and 4.1.13, Propositions 4.2.1 and 4.2.2) hold true when  $N$  is a separable complete tree.*

*Proof.* We only consider Theorem 4.1.4. If  $N$  is compact, the proof is similar. Otherwise, one considers a dense sequence  $(y_k)$  and the compact subtrees  $N_K$  generated by  $(y_k; k \leq K)$ . There exist martingales converging to the projections  $Y_\infty^K$  of  $Y_\infty$  on  $N_K$ , and at the limit, we solve the problem on  $N$  by applying the analogue of Corollary 3.3.13.  $\square$

However, this result cannot be extended to non separable measurable trees.

*Example 6.5.3.* Consider the star  $N = \mathbb{Y}^{\mathbb{S}^1}$  of Example 6.1.8, and let us look for a martingale converging to a variable  $Y$ . If  $N$  is endowed with its Borel  $\sigma$ -algebra, we have seen that  $Y$  must be supported by a separable subtree, so we are reduced to the above result. On the other hand, if we use Borel measurability for the Euclidean topology, the existence of a continuous martingale does not always hold; suppose for instance that  $Y$  is the value at time 1 of a 2-dimensional Brownian motion; then

$$\mathbb{P}[\Theta(Y_t) = \Theta(Y)] = 0$$

for any adapted process  $Y_t$  and any  $t < 1$ ; thus  $Y_t$  cannot converge to  $Y$  as  $t \uparrow 1$  (for the tree metric) because this probability should converge to 1; this means that there does not exist any adapted continuous process with value  $Y$  at time 1.

One can also adapt the Dirichlet form technique to separable complete trees. The space  $\mathbb{D}^b(N)$  is defined as in Definition 5.1.9. and we have the following analogue of Lemma 5.1.10.

**Lemma 6.5.4.** *Let  $N$  be a separable complete tree which is embedded into  $\ell^1$  as in Lemma 6.1.6. Then a function  $f : M \rightarrow N$  is in  $\mathbb{D}^b(N)$  if and only if its components  $f_i$  are in  $\mathbb{D}^b$  and the series  $\sum \mathcal{E}(f_i)$  converges. In this case, (5.1.11) holds true.*

*Proof.* Use the notations of Lemma 6.1.6. Lemma 5.1.10 is easily extended to finite trees, so in particular to the finite subtrees  $N_n$  of  $N$ . Define

$$g_n(x) = \sum_{i=1}^n f_i(x)$$

which is the projection of  $f(x)$  on  $N_n$ . If  $f$  is in  $\mathbb{D}^b(N)$ , it is clear from the definition of this space that  $f_i$  should be in  $\mathbb{D}^b$ ,  $g_n$  should be in  $\mathbb{D}^b(N)$ , and  $\mathcal{E}(g_n) \leq \mathcal{E}(f)$ . Since  $g_n$  is  $N_n$ -valued, we have

$$\mathcal{E}(g_n) = \sum_{i=1}^n \mathcal{E}(f_i),$$

so  $\sum \mathcal{E}(f_i)$  converges and is dominated by  $\mathcal{E}(f)$ . Conversely, suppose that  $\sum \mathcal{E}(f_i)$  converges, and let  $\phi : N \rightarrow \mathbb{R}$  be a non expanding function. Then  $\phi \circ g_n$  is in  $\mathbb{D}^b$  and

$$\mathcal{E}(\phi \circ g_n) \leq \mathcal{E}(g_n) \leq \sum \mathcal{E}(f_i).$$

The sequence  $(\phi \circ g_n)$  converges to  $\phi \circ f$ , so this function is in  $\mathbb{D}^b$  and  $\mathcal{E}(\phi \circ f)$  is dominated by  $\sum \mathcal{E}(f_i)$ . Thus the definition of  $\mathbb{D}^b(N)$  shows that  $f$  is in this space and  $\mathcal{E}(f)$  is dominated by  $\sum \mathcal{E}(f_i)$ . To complete the proof, we have to study  $|f|$ . We have

$$\mathcal{E}(|f| - |g_n|) = \mathcal{E}\left(\sum_{i>n} f_i\right) \leq \liminf_m \mathcal{E}\left(\sum_{i=n+1}^m f_i\right) = \sum_{i>n} \mathcal{E}(f_i)$$

which converges to 0, and  $\mathcal{E}(|g_n|)$  is  $\sum_1^n \mathcal{E}(f_i)$ ; we easily deduce that  $\mathcal{E}(|f|)$  is  $\sum \mathcal{E}(f_i)$ .  $\square$

In particular, projection on the finite subtrees  $N_n$  defines an approximation procedure for elements of  $\mathbb{D}^b(N)$  (they can be approximated for the  $\mathcal{E}$  seminorm).

**Lemma 6.5.5.** *Lemma 5.2.1 holds true for separable trees.*

*Proof.* Write again the tree as in Lemma 6.1.6, and let us check formula (5.2.2) for the finite trees  $N_n$ . Consider an edge  $[z_i, z_i + \alpha_i e_i]$  with  $\alpha_i > 0$ , and the oriented function  $\psi_i$  of (2.2.3) associated to the connected component of  $N_n \setminus \{z_i\}$  containing  $z_i + \alpha_i e_i$ . Then, if  $w$  is the midpoint of  $u$  and  $v$ , we have

$$w_i = \min\left(\left(\psi_i \circ u + \psi_i \circ v\right)^+ / 2, \alpha_i\right).$$

This is the analogue of (5.2.3). Then one proceeds as in Lemma 5.2.1 with  $\psi_i$  instead of  $\gamma_i$ , and

$$A_i = \left\{x \in M; [u(x), v(x)] \cap [z_i, z_i + \alpha_i e_i] \neq \emptyset\right\}.$$

The measure  $\mu_{\langle w_i \rangle}$  is supported by the disjoint sets  $\{0 < w_i < \alpha_i\}$ , and each of these sets is included in  $A_i$ . We deduce (5.2.5) and (5.2.2) for  $N_n$ , and extend the result to  $N$  by approximating  $u$  and  $v$  in  $\mathcal{E}$ -norm with their projections on  $N_n$ .  $\square$

**Theorem 6.5.6.** *All the results stated in subsections 5.2 and 5.3 hold true for separable complete trees.*

*Proof.* Proposition 5.2.6 is proved similarly. For the martingale property of Theorem 5.2.8, we have to prove that  $(\gamma \circ h)(X_{t \wedge \tau})$  is a submartingale for any basic convex function  $\gamma$ . If  $\gamma$  is a Busemann function, we use a perturbation  $h^\varepsilon$  as in Theorem 5.2.8; by using the embedding of  $N$  into  $\ell^1$  (Lemma 6.1.6), we have

$$\gamma(y) = \sum \lambda_i y_i$$

for some  $\lambda_i = \pm 1$ , so  $\mathcal{E}(f) = \mathcal{E}(\gamma \circ f)$  for any  $f$  of  $\mathbb{D}^b(N)$ , and we conclude as in Theorem 5.2.8. If  $\gamma = \gamma_{y_0}$  for some leaf  $y_0$ , it is difficult to use the same perturbation  $h^\varepsilon$  (it will not be well defined if  $h$  is too close to the leaf). However, we can append a ray to  $N$  at the point  $y_0$ ; the function  $\gamma_{y_0}$  becomes a Busemann function for the new tree; the function  $h$  is also energy minimising with respect to the new tree, so we again deduce the submartingale property, and the extension of Theorem 5.2.8 is checked. For Corollary 5.2.9, the limit as  $t \downarrow 0$  is obtained from Lemma 6.1.12. The other results are proved as in Section 5.  $\square$

*Remark 6.5.7.* If now we consider connected spaces  $N$  which look locally like trees (namely graphs), the definition of continuous martingales can be localised similarly to manifolds. The uniqueness result (Corollary 3.3.12) does not hold any more since there can exist loops, and therefore closed geodesics; if  $(g(t); 0 \leq t \leq 1)$  is a loop and if  $B_t$  is a real Brownian motion stopped when it hits 1, then  $Y_t = g(B_t)$  and  $Y'_t = g(0)$  are two different bounded

martingales with limit  $g(0)$ . For the existence, we can use a method similar to Riemannian manifolds with nonpositive sectional curvature, namely use the universal cover. Here we associate a tree to the graph  $N$ ; more precisely, we fix a point  $O$  and the tree  $N_O$  is the set of geodesics  $(g(t); 0 \leq t \leq L)$  with unit speed and origin  $g(0) = O$ , endowed with a distance similar to (6.1.10); the map  $\pi : N_O \rightarrow N$  defined by  $\pi(g) = g(L)$  is a projection. If  $Y_\infty$  is an integrable  $N$ -valued variable, we can find an integrable  $N_O$ -valued variable  $Y'_\infty$  such that  $\pi(Y'_\infty) = Y_\infty$ , and construct the  $N_O$ -valued martingale  $Y'_t$  of class (D) with limit  $Y'_\infty$ ; then  $Y_t = \pi(Y'_t)$  is a  $N$ -valued martingale with limit  $Y_\infty$ .

## 7 Coupling of diffusions on trees

In Section 4, we have seen that when the probability space is generated by a diffusion on a separable metric space  $(M, d)$ , then the construction of  $N$ -valued continuous martingales and the properties of the semigroup  $Q_t$  can be deduced from coupling properties of the diffusion. We now explain how one can couple some diffusions when  $M$  is a tree (or more generally a graph) so that (4.1.6) is satisfied ((4.1.5) is generally easy to get); notice however that  $d$  will not always be the tree distance (we have to use another one when  $M$  is a non separable tree).

### 7.1 Coupling of spiders

As a first example, suppose that  $M$  is our baby tree  $\mathbb{Y}^\ell$ , and let us construct a coupling  $\mathbb{P}^{x, x'}$  for Walsh processes (Example 3.0.1); let  $\sigma$  be the first meeting time, and suppose that the two processes  $X_t$  and  $X'_t$  coincide after  $\sigma$ . If  $L_t$  and  $L'_t$  are the total local times at  $O$ , then, in the isotropic case  $p_i = 1/\ell$ , by applying the method used in Proposition 3.3.11, it is possible to prove (see Lemma 3.3 of [33]) the equality modulo martingales

$$d(X_t, X'_t) \sim \frac{\ell - 2}{2\ell} (L_{t \wedge \sigma} + L'_{t \wedge \sigma}).$$

In particular, if we want the expectation of this expression to be small, then the processes should meet before or shortly after the first one hits  $O$ ; one cannot use a non coalescent coupling. Thus we have to construct a coupling so that the probability of  $\{\sigma > t\}$  is small. There are several possibilities; we can use independent processes, or try to adapt Kendall's technique. We will describe a simple coupling for which the probabilities are easily estimated. To simplify the notation, we will subsequently omit the superscript in  $\mathbb{P}^{x, x'}$ .

**Proposition 7.1.1.** *If  $X_t$  is a Walsh process on  $M = \mathbb{Y}^\ell$ , then there exists an admissible coalescent coupling so that the meeting time  $\sigma$  satisfies*

$$\mathbb{P}^{x, x'}[\sigma > t] \leq 3 \frac{d(x, x')}{\sqrt{\pi t}}. \quad (7.1.2)$$

*In particular, we can apply Theorem 4.1.4 and other results of Section 4 when  $X_t$  is a Walsh process. The nonlinear semigroup  $Q_t$  is regularising ( $Q_t g$  is Lipschitz for any  $t > 0$  and any bounded map  $g$ ).*

*Proof.* We consider separately three cases according to the position of the initial points  $x$  and  $x'$ ,

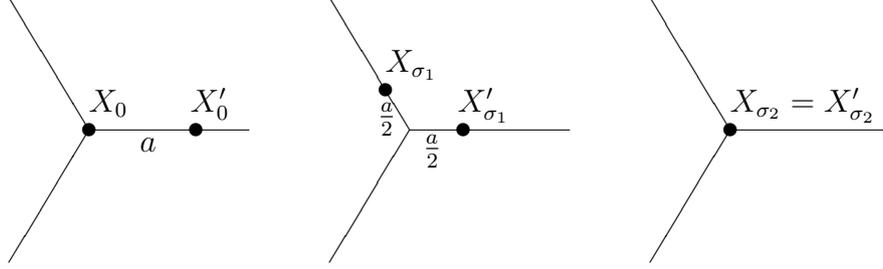


Figure 6: An example of coupling on  $\mathbb{Y}^3$ .

1. one of the initial points is  $O$ ,
2. the initial points are not in the same ray,
3. the initial points are in the same ray.

In all the proof, we will denote by  $T_a$  the law of the first hitting time of  $a$  by a standard Brownian motion. Notice that

$$\mathbb{P}[T_a > t] = \frac{1}{\sqrt{2\pi t}} \int_{-a}^a e^{-x^2/(2t)} dx \leq a\sqrt{\frac{2}{\pi t}}. \quad (7.1.3)$$

*First case.* Suppose for instance that  $x = O$  and  $x' = ae_1$  for  $a > 0$  (Figure 6). In a first step, we construct the processes up to the first hitting time  $\sigma_1$  of  $ae_1/2$  by  $X'_t$ ; notice that this time is distributed like  $T_{a/2}$ . Up to  $\sigma_1$ , we want  $d(x, X_t)$  and  $d(x', X'_t)$  to be identical reflected Brownian motions, so, for each excursion of this process, we have to choose the side of  $X'_t$  on  $R_1$  with respect to  $x'$ , and the ray in which  $X_t$  evolves. This can be done from a Walsh process  $\xi_t$  on a star  $\mathbb{Y}^{2\ell}$  with rays  $R_i^\pm$  by putting, up to  $\sigma_1$ ,

$$X_t = |\xi_t|e_i, \quad X'_t = (a \pm |\xi_t|)e_1 \quad \text{on } \{\xi_t \in R_i^\pm\}.$$

The Walsh process  $\xi_t$  is determined by the probabilities  $p_i^\pm$  of  $R_i^\pm$  which should satisfy

$$p_i^+ + p_i^- = p_i, \quad \sum_i p_i^+ = \sum_i p_i^- = 1/2.$$

This means that we are reduced to find a coupling between a variable in  $\{1, \dots, \ell\}$  with law  $(p_i)$ , and a variable in  $\{+, -\}$  with law  $(1/2, 1/2)$ . Then

$$\sigma_1 = \inf \left\{ t \geq 0; \xi_t \in \{ae_i^-/2; 1 \leq i \leq \ell\} \right\},$$

and the processes  $X_t$  and  $X'_t$  meet at this time if  $\xi_{\sigma_1} = ae_1^-/2$ ; this happens with probability

$$\mathbb{P}[X_{\sigma_1} = X'_{\sigma_1}] = p_1^- / \sum_i p_i^- = 2p_1^-.$$

If the processes  $X$  and  $X'$  do not meet at  $\sigma_1$ , then at this time both of them are at a distance  $a/2$  from the origin, but on two different rays  $R_1$  and, say,  $R_2$ . In this case, we extend the coupling after  $\sigma_1$  by using a standard coalescent Brownian coupling on the geodesic  $R_1 \cup R_2$  (Example 4.1.10), namely

$$X_t = \left(\frac{a}{2} + B_t - B_{\sigma_1}\right)e_1, \quad X'_t = \left(\frac{a}{2} + B_t - B_{\sigma_1}\right)e_2$$

for a Brownian motion  $B_t$ ; the processes meet at the first hitting time  $\sigma_2$  of  $O$  by  $X'_t$ , which is distributed like  $T_a$ . Thus  $\sigma$  is  $\sigma_1$  or  $\sigma_2$  with  $\sigma_1 < \sigma_2 \sim T_a$ , so

$$\mathbb{P}[\sigma > t] \leq \mathbb{P}[T_a > t] \leq a\sqrt{\frac{2}{\pi t}} \quad (7.1.4)$$

with  $a = d(x, x')$ , and (7.1.2) is proved.

*Second case.* Let  $x$  and  $x'$  be in two different rays, say  $R$  and  $R'$ , and suppose for instance that  $|x| \leq |x'|$ . Then we consider the usual coalescent Brownian coupling on  $R \cup R'$  up to the first time  $\sigma_0$  at which  $X_t$  hits  $O$ , and the coupling of the first case after this time. The meeting time is again dominated by the first hitting time of  $O$  by  $X'_t$ , so (7.1.4) holds with  $a$  replaced by  $|x'| \leq d(x, x')$ , and (7.1.2) is again proved.

*Third case.* Let  $x$  and  $x'$  be in the same ray, say  $R_1$ , and suppose for instance that  $|x| < |x'|$ . We consider the coalescent Brownian coupling on  $R_1$  up to the first time  $\sigma_0$  at which either the processes meet, or  $X_t$  hits  $O$ ; in the latter case, which happens with probability

$$\mathbb{P}[X_{\sigma_0} = O] = \frac{|x'| - |x|}{|x'| + |x|} = \frac{d(x, x')}{|x'| + |x|},$$

we use after  $\sigma_0$  the coupling of the first case with  $a = |x| + |x'|$ . The time  $\sigma_0$  is dominated in law by  $T_{d(x, x')/2}$ , so

$$\begin{aligned} \mathbb{P}[\sigma > t] &\leq \mathbb{P}[\sigma_0 > t/2] + \mathbb{P}[\sigma - \sigma_0 > t/2] \\ &\leq \mathbb{P}\left[T_{d(x, x')/2} > t/2\right] + \frac{d(x, x')}{|x'| + |x|} \mathbb{P}\left[T_{|x| + |x'|} > t/2\right] \\ &\leq \frac{d(x, x')}{\sqrt{\pi t}} + 2 \frac{d(x, x')}{|x'| + |x|} \frac{|x| + |x'|}{\sqrt{\pi t}} = 3 \frac{d(x, x')}{\sqrt{\pi t}}. \end{aligned}$$

□

*Remark 7.1.5.* In order to minimise  $\sigma$  in the first case, it seems reasonable to maximise  $p_1^-$ . For instance, in the isotropic case  $p_i = 1/3$  (for  $\ell = 3$ ), the probabilities  $p_i^\pm$  can be specified by

	1	2	3
+	0	1/4	1/4
-	1/3	1/12	1/12

The same procedure can be applied to the star  $\mathbb{Y}^{\mathbb{N}}$  with countably many rays. If now we consider Walsh processes on  $M = \mathbb{Y}^{\mathbb{S}^1}$  (Example 6.2.6), we cannot use the tree distance  $d_1$  (the space is not separable), so we identify  $M$  with  $\mathbb{R}^2$  and use the Euclidean distance  $d_2$ .

**Proposition 7.1.6.** *Consider  $M = \mathbb{Y}^{\mathbb{S}^1}$  and a Walsh process on it. If  $M$  is endowed with the Euclidean distance of  $\mathbb{R}^2$ , one can construct a coupling satisfying the requirements of Theorem 4.1.4; in particular, the semigroup  $Q_t$  exists. Moreover, if  $g$  is bounded and  $t > 0$ , then  $Q_t g$  is Lipschitz for the tree distance.*

*Proof.* We use the coupling of Proposition 7.1.1. Then (7.1.2) holds for the tree distance  $d_1$ , but not for  $d_2$ , because  $x$  and  $x'$  can be close to each other for  $d_2$  but not for  $d_1$ ; this happens when they are not in the same ray (second case of the above proof). So suppose that the initial points are  $x = \xi e$  and  $x' = \xi' e'$  with  $|e| = |e'| = 1$ ,  $e \neq e'$  and  $0 < \xi \leq \xi'$ ; then, up to the first hitting time  $\sigma_0$  of  $O$  by  $X_t$ , one has

$$\begin{aligned} d_2(X_t, X'_t) &= |(\xi' + B_t)e' - (\xi + B_t)e| \\ &\leq \xi' - \xi + (\xi + B_t)|e' - e|. \end{aligned}$$

The last term is a martingale so

$$\mathbb{E}[d_2(X_{t \wedge \sigma_0}, X'_{t \wedge \sigma_0})] \leq \xi' - \xi + \xi|e' - e| \leq C d_2(x, x').$$

so (4.1.6) holds for  $d_2$  if we restrict the expectation to  $\{\sigma_0 \geq t\}$ . On the other hand, on  $\{\sigma_0 < t\}$ , we can estimate the meeting time  $\sigma$  as in (7.1.4) to get

$$\mathbb{P}[\sigma > t \mid \mathcal{F}_{\sigma_0}'' ] \leq (\xi' - \xi) \sqrt{\frac{2}{\pi(t - \sigma_0)}},$$

and by proceeding as in Corollary 4.1.8, we obtain

$$\mathbb{E}[d_2(X_t, X'_t) \wedge 1 \mid \mathcal{F}_{\sigma_0}'' ] \leq \phi(\xi' - \xi) \leq \phi(d_2(x, x'))$$

with  $\lim_0 \phi = 0$ . We deduce (4.1.6) and the conclusion of Theorem 4.1.4. The regularising property of  $Q_t$  for the tree distance  $d_1$  is obtained from (7.1.2) which holds for  $d_1$ .  $\square$

Then we can consider finite graphs  $M$  where each point has a neighbourhood which is isometric to a neighbourhood of the origin in a star. The coupling method of Proposition 7.1.1 can be localised in order to study Walsh processes on  $M$ . It is sufficient to consider the case where  $x$  and  $x'$  are close to each other; in this case, there is at most one vertex between them, and there is a vertex  $O$  which minimises the distance to  $x$ . We consider the subset of  $M$  consisting of the edges starting at  $O$ , and apply the coupling of Proposition 7.1.1 on this subset; this can be done up to the first time at which one of the processes hits a vertex different from  $O$ ; the probability that this occurs before the meeting time  $\sigma$  is small, so the requirements of Theorem 4.1.4 are again fulfilled.

## 7.2 Coupling of snakes

Let us describe a coupling technique which can be used for some diffusions on trees, including the Brownian snake (Examples 6.2.7 and 6.2.8) and the Evans process (Example 6.2.9).

Fix a measurable tree  $M$  (recall Definition 6.1.1) and a height function  $\gamma$  which is either a Busemann function (6.1.2) or a distance function (6.1.4). If  $Z$  is a process, we will denote by  $Z_{st}$  the restricted process  $(Z_u; s \leq u \leq t)$ .

**Definition 7.2.1.** *A continuous diffusion  $X_t$  on  $M$  will be called a snake process with height  $\Gamma_t = \gamma(X_t)$  if, under  $\mathbb{P}^x$ , the relations between conditional laws*

$$\mathcal{L}(\Gamma_{t\infty} | \Gamma_{0t}) = \mathcal{L}(\Gamma_{t\infty} | \Gamma_t) \tag{7.2.2}$$

and

$$\mathcal{L}(X_{st} | X_{0s}, \Gamma_{0\infty}) = \mathcal{L}(X_{st} | X_s, \Gamma_{st}) \tag{7.2.3}$$

hold for  $s \leq t$ .

Condition (7.2.2) means that the process  $\Gamma_t$  is a Markov process for its natural filtration; condition (7.2.3) means that the conditional law of  $X$  given  $\Gamma$  is Markovian, and that estimating  $(X_s; s \leq t)$  from  $\Gamma$  only uses the values of  $\Gamma$  up to time  $t$  (write (7.2.3) for  $s = 0$ ).

**Lemma 7.2.4.** *If  $X_t$  is a snake with height  $\Gamma_t$ , then  $\Gamma_t$  and  $X_t$  are Markov processes for the filtration of  $X$ .*

*Proof.* We have, for  $s \leq t$  and bounded measurable  $f$  and  $\phi$ ,

$$\begin{aligned}\mathbb{E}[f(\Gamma_{t\infty})\phi(X_{0t})] &= \mathbb{E}\left[f(\Gamma_{t\infty})\mathbb{E}[\phi(X_{0t}) \mid \Gamma_{0\infty}]\right] \\ &= \mathbb{E}\left[f(\Gamma_{t\infty})\mathbb{E}[\phi(X_{0t}) \mid \Gamma_{0t}]\right] \\ &= \mathbb{E}\left[\mathbb{E}[f(\Gamma_{t\infty}) \mid \Gamma_t]\phi(X_{0t})\right]\end{aligned}$$

where we have used (7.2.3) for  $s = 0$  in the second line, and (7.2.2) in the third line. We deduce that  $\Gamma_t$  is Markovian for the filtration of  $X$ . On the other hand,

$$\begin{aligned}\mathbb{E}[f(X_{st}) \mid X_{0s}] &= \mathbb{E}\left[\mathbb{E}[f(X_{st}) \mid X_{0s}, \Gamma_{0\infty}] \mid X_{0s}\right] \\ &= \mathbb{E}\left[\mathbb{E}[f(X_{st}) \mid X_s, \Gamma_{st}] \mid X_{0s}\right] \\ &= \mathbb{E}[f(X_{st}) \mid X_s]\end{aligned}$$

where we have used (7.2.3) in the second line and the Markov property of  $\Gamma$  in the third line.  $\square$

Notice that when  $\Gamma_t$  decreases, then  $X_t$  is forced to follow the branch of its ancestors; on the other hand, when  $\Gamma_t$  increases, then  $X_t$  has a choice of branches at each vertex, and this is where the conditional law of  $X$  given  $\Gamma$  is involved.

*Example 7.2.5.* On the star  $\mathbb{Y}^\ell$ , Walsh processes  $X_t$  are snakes for the height function  $\delta(O, \cdot)$ , and  $\Gamma_t$  is a reflected Brownian motion. One can also choose for the height function any of the Busemann functions  $\gamma_i$ , and in this case  $\Gamma_t$  is a skew Brownian motion. Coupling for these diffusions have already been studied in Proposition 7.1.1.

*Example 7.2.6.* On a separable complete tree without leaves, Evans processes (Example 6.2.9) are snakes for the Busemann function  $\gamma_\xi$ , and  $\Gamma_t$  is a real Brownian motion.

*Example 7.2.7.* The Brownian and Poisson snakes on  $\mathbb{G}$  (Example 6.2.8) are snakes for the Busemann function associated to the end  $\{S_y = T_y \leq 0\}$ , and  $\Gamma_t = T_{X_t}$  is a real Brownian motion.

*Example 7.2.8.* The Brownian and Poisson snakes on  $\mathbb{G}_0$  (Example 6.2.7) are snakes for the height function  $\gamma = \delta(O, \cdot)$  where  $O$  is the point  $S_O = T_O = 0$ , and  $\Gamma_t = T_{X_t}$  is a reflected Brownian motion.

**Proposition 7.2.9.** *Consider an Evans process on a separable complete tree without leaves. Then there exists a coalescent coupling such that the meeting time  $\sigma$  satisfies*

$$\mathbb{P}[\sigma > t] \leq \frac{C}{\sqrt{t}} d(x, x') (1 + \log^+(\sqrt{t}/d(x, x'))). \quad (7.2.10)$$

*In particular, we obtain a regularising semigroup  $Q_t$  such that  $Q_t g$  is Hölder continuous for  $t > 0$ .*

*Proof.* The idea is to use a standard Brownian coalescent coupling of the height processes  $\Gamma_t$  and  $\Gamma'_t$  (notice that  $\Gamma_t$  and  $\Gamma'_t$  have the same filtration), and then to use a conditionally independent coupling for  $X$  and  $X'$  defined by

$$\mathbb{E}^{x, x'} [f(X)g(X') | \Gamma] = \mathbb{E}^x [f(X) | \Gamma] \mathbb{E}^{x'} [g(X') | \Gamma'].$$

Let us prove that this is an admissible coupling such that  $(X, X')$  is Markovian, and that the first meeting time  $\sigma$  satisfies (7.2.10); then we will modify  $(X, X')$  after  $\sigma$  so that they coincide. It is clear that  $X$  and  $X'$  have the correct law. From the conditional independence and (7.2.3),

$$\begin{aligned} \mathbb{E}[f(X_{st})g(X'_{st}) \mid X_{0s}, X'_{0s}, \Gamma_{0\infty}] & \\ &= \mathbb{E}[f(X_{st}) \mid X_{0s}, \Gamma_{0\infty}] \mathbb{E}[g(X'_{st}) \mid X'_{0s}, \Gamma_{0\infty}] \\ &= \mathbb{E}[f(X_{st}) \mid X_s, \Gamma_{st}] \mathbb{E}[g(X'_{st}) \mid X'_s, \Gamma'_{st}] \quad (7.2.11) \\ &= \mathbb{E}[f(X_{st})g(X'_{st}) \mid X_s, X'_s, \Gamma_{st}, \Gamma'_{st}], \end{aligned}$$

so  $(X, X')$  satisfies a relation similar to (7.2.3). By proceeding as in Lemma 7.2.4, we check that  $\Gamma$ ,  $\Gamma'$  and  $(X, X')$  are Markovian for the filtration of  $(X, X')$ . Moreover

$$\begin{aligned} \mathbb{E}[f(X_{st}) \mid X_{0s}, X'_{0s}] &= \mathbb{E}\left[\mathbb{E}[f(X_{st}) \mid X_{0s}, X'_{0s}, \Gamma_{0\infty}] \mid X_{0s}, X'_{0s}\right] \\ &= \mathbb{E}\left[\mathbb{E}[f(X_{st}) \mid X_s, \Gamma_{st}] \mid X_{0s}, X'_{0s}\right] \\ &= \mathbb{E}[f(X_{st}) \mid X_s] \end{aligned}$$

where we have used (7.2.11) in the second line and

$$\mathcal{L}(\Gamma_{st} | X_{0s}, X'_{0s}) = \mathcal{L}(\Gamma_{st} | \Gamma_s) = \mathcal{L}(\Gamma_{st} | X_s)$$

in the third line, so  $X$  is Markovian for the filtration of  $(X, X')$ . The same property of course holds for  $X'$ , so the coupling is admissible. If  $\gamma(x) \geq u$ , denote by  $\Pi_u(x)$  the ancestor of  $x$  at height  $u$ . Notice that  $X_t = \Pi_{\Gamma_t}(x)$  at

each time at which  $\Gamma_t = \inf_{s \leq t} \Gamma_s$ . Moreover, the first meeting time  $\sigma$  of  $X$  and  $X'$  is the first time at which

$$\Gamma_t = \Gamma'_t = \inf_{s \leq t} (\Gamma_s \wedge \Gamma'_s) \wedge \gamma(x \wedge x').$$

Let us estimate this time. By adding a constant to  $\gamma$ , we can suppose for instance that

$$\gamma(x) = a, \quad \gamma(x') = -a, \quad \gamma(x \wedge x') = -b,$$

for  $0 < a \leq b = d(x, x')/2$ . The process  $\Gamma_t$  is a Brownian motion starting at  $a$ ; the first hitting time  $\sigma_0$  of 0 is also the meeting time of  $\Gamma$  and  $\Gamma'$  and is distributed like  $T_a$  (the hitting time of  $a$  by a standard Brownian motion). Then let

$$\Gamma^* = \sup\{\Gamma_s; s \leq \sigma_0\} = -\inf\{\Gamma'_s; s \leq \sigma_0\}$$

which is  $\mathcal{F}_{\sigma_0}$ -measurable. The meeting time  $\sigma$  is

$$\sigma = \inf\{t \geq \sigma_0; -\Gamma_t \geq \Gamma^* \vee b\}. \quad (7.2.12)$$

One has

$$\mathbb{P}[\sigma > t] \leq \mathbb{P}[\sigma_0 > t/2] + \mathbb{P}[\sigma - \sigma_0 > t/2].$$

The first probability is estimated from (7.1.3) since  $\sigma_0 \sim T_a$ , and for the second one, notice that conditionally on  $\mathcal{F}_{\sigma_0}$ , the variable  $\sigma - \sigma_0$  is distributed like the first hitting time of  $\Gamma^* \vee b$  by an independent Brownian motion, so

$$\mathbb{P}[\sigma - \sigma_0 > t/2 \mid \mathcal{F}_{\sigma_0}] = \frac{2}{\sqrt{\pi t}} \int_0^{\Gamma^* \vee b} e^{-z^2/t} dz.$$

On the other hand, for  $z \geq a$ , saying that  $\Gamma^* \geq z$  means that the Brownian process  $\Gamma_t - \Gamma_0$  quits the interval  $[-a, z - a]$  at the point  $z - a$ , so

$$\mathbb{P}[\Gamma^* \geq z] = \frac{a}{z}.$$

Thus

$$\begin{aligned} \mathbb{P}[\sigma - \sigma_0 > t/2] &= \frac{2}{\sqrt{\pi t}} \int_0^{+\infty} \mathbb{P}[\Gamma^* \vee b \geq z] e^{-z^2/t} dz \\ &= \frac{2a}{\sqrt{\pi t}} \int_b^{+\infty} \frac{e^{-z^2/t}}{z} dz + \frac{2}{\sqrt{\pi t}} \int_0^b e^{-z^2/t} dz \\ &= \frac{2a}{\sqrt{\pi t}} \int_{b/\sqrt{t}}^{+\infty} \frac{e^{-z^2}}{z} dz + \frac{2}{\sqrt{\pi t}} \int_0^b e^{-z^2/t} dz \\ &\leq \frac{C}{\sqrt{t}} a \left(1 + \log^+(\sqrt{t}/b)\right) + \frac{C}{\sqrt{t}} b \\ &\leq \frac{C'}{\sqrt{t}} d(x, x') \left(1 + \log^+(\sqrt{t}/d(x, x'))\right). \end{aligned}$$

We can deduce (7.2.10).  $\square$

**Proposition 7.2.13.** *Consider the Brownian snake  $X_t$  in  $\mathbb{G}^c$  of Example 6.2.8, and denote the tree distance by  $d_1$ . There exists another distance  $d_2$  which makes  $\mathbb{G}^c$  separable, and for which  $X_t$  is continuous, measurable and satisfies the assumptions of Theorem 4.1.4. If  $Q_t$  is the resulting semigroup,  $Q_t g$  is Hölder continuous for the distance  $d_1$  for  $g$  bounded and  $t > 0$ .*

*Proof.* The distance  $d_2$  is defined by

$$d_2(x, x') = |T_{x'} - T_x| + |S_{x'} - S_x| + \sup_{S_x \vee S_{x'} \leq u \leq T_x \wedge T_{x'}} |\omega_{x'}(u) - \omega_x(u)|. \quad (7.2.14)$$

Then  $\mathbb{G}^c$  is separable, and it is not difficult to check that  $X_t$  is continuous and measurable. We have to construct a coupling satisfying (4.1.6) for  $d_2$ . Consider

$$\mathbb{G}_-^c = \{y \in \mathbb{G}^c; S_y \leq 0\}.$$

This set acts on  $\mathbb{G}^c$  as follows; if  $x \in \mathbb{G}^c$  and  $y \in \mathbb{G}_-^c$ , we define  $z = y \star x$  in  $\mathbb{G}^c$  by

$$\begin{aligned} T_z &= T_x + T_y, & S_z &= S_x \wedge (T_x + S_y), \\ \omega_z(u) &= \begin{cases} \omega_x(u) & \text{for } S_x \leq u < T_x + S_y, \\ \omega_x(T_x + S_y) + \omega_y(u - T_x) & \text{for } T_x + S_y \leq u < T_x + T_y \end{cases} \end{aligned}$$

if  $S_x < T_x + S_y$ , and

$$\omega_z(u) = \omega_y(u - T_x) \quad \text{for } T_x + S_y \leq u < T_x + T_y$$

otherwise. Intuitively, the action of  $y$  consists in erasing the path  $\omega_x$  between  $T_x + S_y$  and  $T_x$ , and then in completing this path between  $T_x + S_y$  and  $T_x + T_y$  by using  $\omega_y$  as increments. In particular, if  $O$  is the point  $S_O = T_O = 0$ , then  $O \star x = x$ . Notice also that

$$T_x = T_{x'} \implies d_2(y \star x, y \star x') \leq d_2(x, x'). \quad (7.2.15)$$

If  $Z_t$  is a Brownian snake with initial value  $Z_0 = O$ , then  $Z_t$  lives in  $\mathbb{G}_-$  and we can check that  $Z_t \star x$  is a Brownian snake with initial value  $x$  (the Brownian snake can be viewed as a Lévy process). The coupling  $\mathbb{P}^{x, x'}$  is then defined as follows; by viewing  $X_t$  as a snake, we first use the coupling of Proposition 7.2.9 up to the time  $\sigma$  of (7.2.12); after that time, we consider an independent Brownian snake  $Z_t$  with  $Z_0 = O$ , and put

$$X_t = Z_{t-\sigma} \star X_\sigma, \quad X'_t = Z_{t-\sigma} \star X'_\sigma.$$

This is an admissible coupling, and (7.2.15) enables to show that

$$d_2(X_t, X'_t) \leq d_2(X_\sigma, X'_\sigma) \leq d_2(x, x')$$

on  $\{t \geq \sigma\}$ . The probability of  $\{\sigma > t\}$  is estimated by (7.2.10), and an adaptation of Corollary 4.1.8 enables to obtain (4.1.6). The regularising property of  $Q_t$  is obtained with the method of Proposition 7.2.9.  $\square$

*Remark 7.2.16.* A modification of the method enables to study the Brownian snake in  $\mathbb{G}_0$ .

*Remark 7.2.17.* The Poisson snake is more delicate to study. It lives in the subset of  $\mathbb{G}$  with piecewise constant integer-valued paths, but (7.2.14) does not make this space separable. We can use the distance

$$d_2(x, x') = |T_{x'} - T_x| + |S_{x'} - S_x| + \int_{S_x \vee S_{x'}}^{T_x \wedge T_{x'}} \mathbf{1}_{\{x(u) \neq x'(u)\}} du.$$

However, the above proof cannot be directly extended because (7.2.15) is false for this distance.

## 8 Stochastic calculus with jumps

Let us now study càdlàg processes in trees. We have already considered càdlàg semimartingales, so let us focus on martingales. The case of smooth manifolds has been considered in [26]. For trees, we look for a definition generalising the notion of discrete martingale used in (4.0.1) or [32].

### 8.1 Martingales with jumps

We have seen in Section 2.3 that the barycentre of a variable is not characterised by the Jensen inequality, but by its semi-localised version. The idea is to define càdlàg martingales by means of this inequality and by using partially convex functions, similarly to the definition of continuous martingales using globally convex functions. We again consider a measurable tree  $N$ .

**Definition 8.1.1.** *A càdlàg adapted process  $Y_t$  is said to be a martingale if the following condition is satisfied; for any connected open subset  $G$  of  $N$ , for any  $\varepsilon > 0$ , for any Lipschitz measurable function  $f$ , and for any stopping times  $\tau_0 \leq \tau_1$ , if  $f$  is convex on all the geodesics intersecting  $G$  and if  $\delta(Y_t, G^c) > \varepsilon$  for  $\tau_0 \leq t < \tau_1$ , then  $f(Y_t)$  is a local submartingale on the interval  $[\tau_0, \tau_1]$ .*

*Example 8.1.2.* Continuous martingales of Definition 3.3.3 are also martingales in this sense because one can modify  $f$  outside  $G$  in order to obtain a function which is convex on  $N$ .

*Example 8.1.3.* If  $(Y_n; n \in \mathbb{N})$  is a discrete martingale in the sense that it is integrable and  $Y_n$  is  $\mathbb{B}[Y_{n+1}|\mathcal{F}_n]$ , then it follows from Proposition 2.3.5 that the piecewise constant process which is equal to  $Y_n$  on  $[n, n+1)$  is a martingale for the piecewise constant filtration.

*Remark 8.1.4.* In Definition 8.1.1, we can use

$$\tau_1^\varepsilon = \inf \left\{ t \geq \tau_0; \delta(Y_t, G^c) \leq \varepsilon \right\}$$

and we obtain the stopped process  $V_t^\varepsilon = f(Y_{t \wedge \tau_1^\varepsilon})$  which is a local submartingale for  $t \geq \tau_0$ . If we let  $\varepsilon \downarrow 0$ , we obtain a limiting local submartingale  $V_t$ , but  $V_t$  is not always  $f(Y_{t \wedge \tau_1^0})$ . It is equal to  $f(Y_t)$  on  $\{\tau_0 \leq t < \tau_1^0\}$ , and on  $\{t \geq \tau_1^0\}$ , it is  $f(Y_{\tau_1^0})$  if  $Y_{\tau_1^0-}$  is in  $G$ , and it is  $f(Y_{\tau_1^0-})$  otherwise.

Martingales are of course semimartingales, and by using  $f(y) = |y|$ , we see that they are locally of class (D). If  $f$  is a Lipschitz measurable function which is a difference of convex functions, then  $f(Y_t)$  is a semimartingale which is the sum of a local martingale and a predictable process  $V_t^f$  with finite variation,

and if  $f$  is convex on the geodesics intersecting  $G$ , then Definition 8.1.1 says that  $V^f$  should not decrease on  $\{Y_- \in G\}$ .

Similarly to Proposition 6.4.1, we are going to check that it is sufficient to consider some functions  $f$ .

**Proposition 8.1.5.** *In a complete measurable tree, let  $Y_t$  be a càdlàg adapted process which is locally of class (D), and suppose that the condition of Definition 8.1.1 holds true for oriented functions  $f = \psi(y_0^\alpha, \cdot)$  and  $G = y_0^\alpha$  (recall (2.2.3)). Then  $Y_t$  is a martingale.*

*Proof.* By a stopping argument, we can suppose that  $Y_t$  is of class (D). We first check as in Proposition 6.4.1 that if  $O$  is a point and if  $N_0$  is a connected component of  $N \setminus \{O\}$ , then

$$V_t = |Y_t|1_{N_0}(Y_t) \tag{8.1.6}$$

is a submartingale. Notice that this process is the positive part of  $U_t = \psi(N_0, Y_t)$ . We follow the method of Lemma 3.1.5 and introduce  $\tau'_0 = 0$ ,

$$\tau_k = \inf \left\{ t \geq \tau'_{k-1}; V_t \leq \varepsilon/2 \right\}, \quad \tau'_k = \inf \left\{ t \geq \tau_k; V_t \geq \varepsilon \right\}.$$

Our assumption implies that  $U_t$  is a submartingale on  $[\tau'_{k-1}, \tau_k]$ ; we deduce that  $V_t \vee \varepsilon = U_t \vee \varepsilon$  is also a submartingale on these intervals. On the intervals  $[\tau_k, \tau'_k]$ , the process  $V_t \vee \varepsilon$  is non decreasing (it is  $\varepsilon$  up to  $\tau'_k$  and can have a positive jump at this time), so this process is a submartingale on the whole time interval. By letting  $\varepsilon \downarrow 0$ , we deduce that  $V_t$  is a submartingale. Now consider a function  $f$ , a subset  $G$  as in Definition 8.1.1, and stop  $Y_t$  when it is at a distance less than  $\varepsilon$  from the complement of  $G$ . We suppose that  $G$  is bounded, that  $y_0$  is in  $G$ , and we want to prove that

$$\mathbb{E}[f(Y_t)|\mathcal{F}_0] \geq f(y_0) \tag{8.1.7}$$

for  $Y_0 = y_0$  (other time intervals are dealt with similarly). There are two cases.

*First case.* If the restriction of  $f$  to  $G$  is minimal at some point  $O$  of  $G$ , we can consider the arc  $[O, y_0]$  and introduce the function  $f_\zeta$  of (6.2.2) associated to a subdivision of the arc. The fact that  $f$  is convex on geodesics containing  $O$  implies that (6.2.3) and (6.2.4) again hold. The processes  $\delta(z_k, Y_t)1_{z_k^0}(Y_t)$  are of type (8.1.6), so are submartingales. Thus  $f_\zeta(Y_t)$  is a submartingale, and by letting  $\eta \downarrow 0$  in (6.2.4), we obtain (8.1.7).

*Second case.* Otherwise, the infimum of  $f$  on  $G$  is obtained at some point  $O$  on the boundary. We again consider the arc  $[O, y_0]$  and a function  $f_\zeta$

similar to (6.2.2), but the first term  $\rho_0 \delta(z_0, y) 1_{z_0^0}(y)$  (where  $z_0 = O$  and  $z_0^0$  is the connected component of  $N \setminus \{O\}$  containing  $y_0$ ) has to be replaced by  $\rho_0 \psi(z_0^0, y)$ . Then (6.2.3) holds with this change. The process  $\psi(z_0^0, Y_t)$  is a submartingale from our assumption, so  $f_\zeta(Y_t)$  is again a submartingale, and we conclude as in the first case.  $\square$

*Remark 8.1.8.* Not all the points  $y_0$  have to be considered in Proposition 8.1.5 because some functions  $\psi(y_0^\alpha, \cdot)$  are obtained by translation from each other; in the case of a finite tree with  $\ell$  edges, we only need  $2\ell$  functions; for the star  $\mathbb{Y}^\ell$ , we need the functions  $\pm \gamma_i$ .

Proposition 3.3.10 (the limit in probability of a sequence of martingales is a martingale) cannot be extended to càdlàg processes without additional integrability conditions, but this is not surprising since it is already false for real martingales (consider  $Y_t^n = Z_t^n - t$  where  $Z_t^n$  is a Poisson process with intensity  $1/n$  and jumps of size  $n$ , which converges to  $Y_t = -t$ ).

The analogue of Proposition 6.4.6 is the following one.

**Proposition 8.1.9.** *If  $(Y_t; 0 < t < \infty)$  is a càdlàg martingale of class (D) in a separable complete tree  $N$ , then  $Y_t$  has almost sure limits as  $t \downarrow 0$  and as  $t \uparrow \infty$ .*

*Proof.* Contrary to the continuous case (where it was sufficient to apply Lemma 6.1.12 for the two convergences), we need separate proofs.

*Convergence as  $t \downarrow 0$ .* We use the embedding of  $N$  into  $\ell^1$  of Lemma 6.1.6, and the subtrees  $N_n$ . The projection onto  $N_n$  is defined by  $\Pi_n(y) = \sum_1^n y_i$ . For  $\varepsilon > 0$ , we have

$$\mathbb{E} \delta(Y_1, N_n) = \mathbb{E} \delta(Y_1, \Pi_n(Y_1)) = \mathbb{E} \sum_{i=n+1}^{\infty} |Y_1^i| \leq \varepsilon.$$

if  $n$  large enough. On the other hand, the process  $\delta(Y_t, N_n)$  is a submartingale of class (D), so

$$\mathbb{P} \left[ \sup_{0 < t \leq 1} \delta(Y_t, \Pi_n(Y_t)) \geq \sqrt{\varepsilon} \right] \leq \sqrt{\varepsilon}.$$

We deduce that there exists a subsequence such that  $\Pi_n(Y_t)$  converges almost surely uniformly to  $Y_t$ . But for  $n$  fixed, the components of  $\Pi_n(Y_t)$  are submartingales of class (D), so they converge as  $t \downarrow 0$  and we can conclude.

*Convergence as  $t \uparrow \infty$ .* Let  $\gamma$  be a Busemann function (we can append a ray to  $N$  if it has no end). Fix  $\varepsilon > 0$ , let  $Z_t$  be the ancestor of  $Y_t$  at height

$$\gamma(Z_t) = \gamma(Y_t) - \varepsilon,$$

let  $Z_t^0$  be the connected component of  $N \setminus \{Z_t\}$  containing  $Y_t$ , and put

$$\tau = \tau(t) = \inf \left\{ s \geq t; Y_s \notin Z_t^0 \right\}.$$

Consider some  $\eta < \varepsilon$  which will be chosen later (it will be a function of  $\varepsilon$ ). We know that  $\gamma(Y_t)$  is a submartingale of class (D), so it converges almost surely, and

$$\mathbb{P}[A_t] \leq \eta \quad \text{with} \quad A_t = \left\{ \sup_{s \geq t} |\gamma(Y_s) - \gamma(Y_t)| \geq \eta \right\} \quad (8.1.10)$$

if  $t$  is large enough. We consider the oriented distance function  $\psi(Z_t^0, \cdot)$  which is convex on geodesics intersecting  $Z_t^0$ . By applying Remark 8.1.4, we obtain a process  $(V_s; s \geq t)$  which is a submartingale of class (D); we have

$$V_s = \begin{cases} \psi(Z_t^0, Y_s) & \text{for } t \leq s < \tau, \\ \psi(Z_t^0, Y_\tau) & \text{for } s \geq \tau \text{ and } Y_{\tau-} \neq Z_t, \\ 0 & \text{for } s \geq \tau \text{ and } Y_{\tau-} = Z_t. \end{cases} \quad (8.1.11)$$

We know that  $V_s$  converges to a  $V_\infty$ , and

$$\varepsilon = V_t = \mathbb{E}V_t \leq \mathbb{E}V_\tau. \quad (8.1.12)$$

Since  $Y$  is of class (D), the family of all possible variables  $V_\tau$  is uniformly integrable; by applying (8.1.10), we deduce that

$$\mathbb{E}[|V_\tau|1_{A_t}] \leq \phi(\eta) \quad (8.1.13)$$

with  $\lim_0 \phi = 0$ . On the other hand, on the event  $A_t^c$ , the third case of (8.1.11) cannot happen, so

$$\begin{aligned} V_\tau &= \psi(Z_t^0, Y_\tau) = -\delta(Z_t, Y_\tau) = \gamma(Z_t) - \gamma(Y_\tau) = \gamma(Y_t) - \gamma(Y_\tau) - \varepsilon \\ &\leq \eta - \varepsilon \end{aligned} \quad (8.1.14)$$

on  $A_t^c \cap \{\tau < \infty\}$ . On  $A_t^c \cap \{\tau = \infty\}$ , we have

$$V_s = \delta(Z_t, Y_s) = \gamma(Y_s) - \gamma(Z_t) = \gamma(Y_s) - \gamma(Y_t) + \varepsilon \leq \eta + \varepsilon$$

for  $s \geq t$ , so

$$V_\tau \leq \eta + \varepsilon. \quad (8.1.15)$$

By using (8.1.12) and (8.1.13) on  $A_t$ , (8.1.14) on  $A_t^c \cap \{\tau < \infty\}$ , (8.1.15) on  $A_t^c \cap \{\tau = \infty\}$ , we obtain

$$\begin{aligned} \varepsilon &\leq \mathbb{E}[V_\tau 1_{A_t}] + \mathbb{E}[V_\tau 1_{A_t^c \cap \{\tau < \infty\}}] + \mathbb{E}[V_\tau 1_{A_t^c \cap \{\tau = \infty\}}] \\ &\leq \phi(\eta) - (\varepsilon - \eta)\mathbb{P}[A_t^c \cap \{\tau < \infty\}] + \varepsilon + \eta \end{aligned}$$

so

$$\mathbb{P}[A_t \cup \{\tau < \infty\}] = \mathbb{P}[A_t^c \cap \{\tau < \infty\}] + \mathbb{P}[A_t] \leq \eta + \frac{\phi(\eta) + \eta}{\varepsilon - \eta}$$

for  $t$  large enough. We can choose  $\eta$  so that the right hand side is bounded by  $\varepsilon$ . Thus the event  $A_t^c \cap \{\tau = \infty\}$  has probability at least  $1 - \varepsilon$ , and on this event, one has

$$\delta(Y_t, Y_s) \leq \delta(Z_t, Y_t) + \delta(Z_t, Y_s) \leq 2\varepsilon + \eta \leq 3\varepsilon,$$

so

$$\mathbb{P}\left[\sup_{s \geq t} \delta(Y_t, Y_s) > 3\varepsilon\right] \leq \varepsilon$$

for  $t$  large enough. The completeness of  $N$  enables to conclude.  $\square$

*Remark 8.1.16.* For the convergence as  $t \downarrow 0$ , the process  $Y_t$  has not to fully satisfy the conditions in the definition of martingales; it is sufficient to assume that it is transformed into submartingales by convex Lipschitz functions.

## 8.2 Martingales with prescribed limit

We first extend the uniqueness result.

**Proposition 8.2.1.** *Proposition 3.3.11 and Corollaries 3.3.12 and 3.3.13 hold true for càdlàg martingales in measurable trees (the tree has to be complete for Corollary 3.3.13).*

*Proof.* Let us first prove the extension of Proposition 3.3.11, namely that  $D_t = \delta(Y_t, Z_t)$  is a local submartingale for martingales  $Y_t$  and  $Z_t$ ; we can suppose that they are of class (D). With the notations of Proposition 3.3.11, if  $D_{\tau_k} \leq \varepsilon/2$ , then  $D_t \vee \varepsilon$  is non decreasing on  $I_k$ . Otherwise, we again take the midpoint  $A$  and we notice that  $Y$  and  $Z$  do not cross  $A$  on  $I_k$  except perhaps at time  $\tau_{k+1}$ . Let  $A^\alpha$  and  $A^\beta$  be respectively the connected components of  $N \setminus \{A\}$  containing  $Y_{\tau_k}$  and  $Z_{\tau_k}$ , and consider the oriented distance function  $\psi_\beta = \psi(A^\beta, \cdot)$ . Then we have

$$D_t \geq \psi_\beta(Z_t) - \psi_\beta(Y_t) \tag{8.2.2}$$

with equality for  $\tau_k \leq t < \tau_{k+1}$ . The function  $\psi_\beta$  is convex on geodesics intersecting  $A^\beta$ , and concave on geodesics intersecting  $A^\alpha$ , so our definition of martingales implies that the right-hand side of (8.2.2) is a submartingale on  $I_k$ ; thus  $D_t$  and  $D_t \vee \varepsilon$  are submartingales on  $I_k$ . We can deduce by letting  $\varepsilon \downarrow 0$  that  $D_t$  is a submartingale on the whole time interval. The proof of

Corollary 3.3.12 is straightforward. For Corollary 3.3.13, we prove like previously that  $Y^n$  has a limit  $Y$ , and that  $Y_t$  is of class (D) and transformed into a submartingale by any measurable Lipschitz convex function  $f$ . However we have to localise this property and prove the condition of Definition 8.1.1. So we let  $f$  and  $G$  be as in this definition, and consider

$$\tau = \inf \left\{ t \geq 0; \delta(Y_t, G^c) \leq \varepsilon \right\}, \quad \tau^n = \inf \left\{ t \geq 0; \delta(Y_t^n, G^c) \leq \varepsilon/2 \right\}. \quad (8.2.3)$$

Then  $f(Y_0)$  is the limit of  $f(Y_0^n)$ ,

$$f(Y_0^n) \leq \mathbb{E}[f(Y_{t \wedge \tau \wedge \tau^n}^n) \mid \mathcal{F}_0] \quad (8.2.4)$$

and

$$\mathbb{E} \left| f(Y_{t \wedge \tau \wedge \tau^n}^n) - f(Y_{t \wedge \tau}) \right| \leq C \mathbb{E} \delta(Y_{t \wedge \tau \wedge \tau^n}^n, Y_{t \wedge \tau \wedge \tau^n}) + C \mathbb{E} \delta(Y_{t \wedge \tau \wedge \tau^n}, Y_{t \wedge \tau}).$$

The variables involved in the right hand side are uniformly integrable, the first term tends to 0 from the convergence of  $Y^n$  to  $Y$ , and the second term tends to 0 because  $\mathbb{P}[\tau^n < \tau]$  tends to 0. Thus this expression tends to 0, and by taking the limit in (8.2.4), we obtain

$$f(Y_0) \leq \mathbb{E}[f(Y_{t \wedge \tau}) \mid \mathcal{F}_0].$$

We can replace 0 by another stopping time and deduce that  $Y_t$  satisfies the condition of Definition 8.1.1.  $\square$

For the existence, like previously, we consider successively the coupling and energy methods.

**Theorem 8.2.5.** *The results of Section 4 (Theorem 4.1.4, Corollaries 4.1.8 and 4.1.13, Propositions 4.2.1 and 4.2.2) hold true for càdlàg Markov processes and provide càdlàg martingales in separable complete trees.*

*Proof.* Let us look at Theorem 4.1.4 (the proof of other results is straightforward). We use as in Section 4 the discrete martingales  $Y_k = h^\Delta(t_k, X_{t_k})$  and obtain at the limit a continuous function  $h$  and a process  $Y_t = h(t, X_t)$  which is transformed into submartingales by Lipschitz convex functions. We have to check the condition of Definition 8.1.1. This is done as in the last part of Proposition 8.2.1 but we have to take into account the fact that the approximations  $Y_t^n = h^{\Delta_n}(t, X_t)$  are not defined for all time but only on the subdivisions  $\Delta_n$ . We define  $\tau$  and  $\tau^n$  as in (8.2.3), let  $\sigma^n$  be the first time after  $\tau$  in  $\Delta_n$ , take a dyadic time  $t$ , and replace (8.2.4) by

$$f(Y_0^n) \leq \mathbb{E}[f(Y_{t \wedge \sigma^n \wedge \tau^n}^n) \mid \mathcal{F}_0].$$

Moreover,

$$\begin{aligned} \mathbb{E} \left| f(Y_{t \wedge \sigma^n \wedge \tau^n}^n) - f(Y_{t \wedge \tau}) \right| &\leq C \mathbb{E} \delta(Y_{t \wedge \sigma^n \wedge \tau^n}^n, Y_{t \wedge \sigma^n \wedge \tau^n}) + C \mathbb{E} \delta(Y_{t \wedge \sigma^n \wedge \tau^n}, Y_{t \wedge \sigma^n}) \\ &\quad + C \mathbb{E} \delta(Y_{t \wedge \sigma^n}, Y_{t \wedge \tau}). \end{aligned}$$

The first term tends to 0 from the convergence of  $h^{\Delta_n}$  to  $h$  uniformly in  $t$ , the second term tends to 0 because  $\mathbb{P}[\sigma^n > \tau^n]$  tends to 0 (this also follows from the convergence of  $h^{\Delta_n}$ ), and the third one tends to 0 from the right continuity of  $Y$ .  $\square$

*Example 8.2.6.* The process  $X_t$  can be the solution of a stochastic differential equations with jumps; we obtain the existence of martingales with prescribed limit on Wiener-Poisson spaces. However, we do not know coalescent couplings in this setting.

We now apply the energy method and extend the results of Section 5. We consider a Dirichlet form which is defined through its Beurling-Deny decomposition (see Section 3.2 of [14]). We have

$$\mathcal{E}(f) = \mathcal{E}^c(f) + \frac{1}{2} \int |f(x_1) - f(x_2)|^2 J(dx_1, dx_2) \quad (8.2.7)$$

where  $\mathcal{E}^c$  is a strongly local Dirichlet form and  $J$  is a symmetric jump measure (we suppose that there is no killing inside  $M$ ). We can consider the energy measure  $\mu_{\langle f \rangle}^c$  associated to the local part  $\mathcal{E}^c$  so that  $\mathcal{E}^c(f) = \mu_{\langle f \rangle}^c(M)$ . It can be extended to functions  $f$  of  $\mathbb{D}^{loc}$ , and the space  $\mathbb{D}^b$  is the set of bounded functions  $f$  of  $\mathbb{D}^{loc}$  such that the energy  $\mathcal{E}(f)$  defined by (8.2.7) is finite. If  $M_0$  is a relatively compact open subset of  $M$ , we define like previously the space  $\mathbb{D}_0^b$ . We suppose that for each compact subset  $K_1$ , there exists another compact subset  $K_2 \supset K_1$  such that  $J(K_1 \times K_2^c) = 0$ . With this condition, one can extend the proof of Lemma 5.1.8.

The form  $\mathcal{E}$  is associated to a Hunt process  $X_t$ . Its jumps are described by the measure  $J$ . In particular, the above condition says that if  $X_{t-}$  is in  $K_1$ , then  $X_t$  must be in  $K_2$ .

If we now consider our tree  $N$ , we can consider the space  $\mathbb{D}^{loc}(N)$  of functions  $f$  such that  $\phi \circ f$  is in  $\mathbb{D}^{loc}$  for any Lipschitz function  $f$ . The energy  $\mathcal{E}^c(f)$  can be defined on this space as in the local case, and  $\mathbb{D}^b(N)$  is defined as the space of bounded functions  $f$  of  $\mathbb{D}^{loc}(N)$  such that

$$\mathcal{E}(f) = \mathcal{E}^c(f) + \frac{1}{2} \int \delta^2(f(x_1), f(x_2)) J(dx_1, dx_2) < \infty. \quad (8.2.8)$$

**Theorem 8.2.9.** *Consider a non local Dirichlet form of type (8.2.7) satisfying the above conditions and a separable complete tree  $N$ . Assume the absolute continuity condition (5.2.10) and suppose that bounded real functions which are harmonic on an open subset of  $M$  are continuous on this subset. Then, if  $M_0$  is a relatively compact open subset of  $M$  and if  $g$  is in  $\mathbb{D}^b(N)$ , there exists a unique energy minimising map  $h$  in  $\mathbb{D}_g^b(N)$ ; this map has a continuous modification, and  $h(X_{t \wedge \tau})$  is a martingale. Moreover, càdlàg martingales with prescribed limit exist if the form is conservative (Theorem 5.3.1), and the semigroup  $Q_t$  does not increase the energy (Proposition 5.3.3).*

*Proof.* The proofs of the statements of Theorem 5.3.1 and Proposition 5.3.3 are similar to the continuous case, so let us prove the first part of the theorem. By considering separately the local and jump parts, the inequality (5.2.2) holds for energies of two functions and their middle function; it indeed holds for the local part  $\mathcal{E}^c$ , and for the jump part, we use

$$\begin{aligned} \delta^2(w(x_1), w(x_2)) &\leq \frac{1}{2} \delta^2(u(x_1), u(x_2)) + \frac{1}{2} \delta^2(v(x_1), v(x_2)) \\ &\quad - \frac{1}{4} \left( \delta(u(x_1), v(x_1)) - \delta(u(x_2), v(x_2)) \right)^2 \end{aligned}$$

which is a consequence of the non positivity of the curvature (see for instance Corollary 2.1.3 of [20]). We deduce as in Proposition 5.2.6 the existence of a minimising function  $h$  and of the corresponding process  $Y_t = h(X_t)$ . We can also prove with the method of Theorem 5.2.8 or 6.5.6 that  $f(Y_t)$  is a submartingale for any Lipschitz convex function  $f$ . This is not sufficient to prove that  $Y_t$  is a martingale, but this is sufficient to apply the method of Corollary 5.2.11 and prove the existence of a continuous modification of  $h$  (apply Remark 8.1.16 to obtain  $h(x)$  as the limit of  $h(X_t)$  under  $\mathbb{P}^x$ ). Now (see Proposition 8.1.5), we have to consider an oriented distance function  $\psi = \psi(y_0^\alpha, \cdot)$  which is convex on geodesics intersecting the open subset  $G = y_0^\alpha$ . For  $\varepsilon > 0$ , consider the open sets

$$G_\varepsilon = \{y; \delta(y, G^c) > \varepsilon\}, \quad M_\varepsilon = \{x \in M_0; h(x) \in G_\varepsilon\}.$$

Let  $\rho$  be a nonnegative function of  $\mathbb{D}^b$  which is 0 outside  $M_\varepsilon$ . Then we can define  $\mathcal{I}_x^\eta$  as the translation of step  $\eta\rho(x)$  in the direction of  $y_0$  (it is well defined if  $\eta$  is small enough), and the corresponding perturbation  $h^\eta(x) = \mathcal{I}_x^\eta(h(x))$ . We apply (8.2.8) to express  $\mathcal{E}(f)$ , (8.2.7) to express  $\mathcal{E}(\psi \circ h)$ , and

notice that  $\mathcal{E}^c(h) = \mathcal{E}^c(\psi \circ h)$ . Thus

$$\begin{aligned} \mathcal{E}(h) &= \mathcal{E}(\psi \circ h) + \frac{1}{2} \int \delta^2(\psi \circ h(x_1), \psi \circ h(x_2)) J(dx_1, dx_2) \\ &\quad - \frac{1}{2} \int (\psi \circ h(x_2) - \psi \circ h(x_1))^2 J(dx_1, dx_2) \\ &= \mathcal{E}(\psi \circ h) + \frac{1}{2} \int \Psi(h(x_1), h(x_2)) \Psi(h(x_2), h(x_1)) J(dx_1, dx_2) \quad (8.2.10) \end{aligned}$$

with

$$\Psi(y_1, y_2) = \delta(y_1, y_2) + \psi(y_2) - \psi(y_1).$$

Notice that  $\Psi$  is nonnegative, and the product  $\Psi(y_1, y_2)\Psi(y_2, y_1)$  is non zero only when  $y_0, y_1$  and  $y_2$  are not aligned. If  $y_0, h(x_1)$  and  $h(x_2)$  are aligned, then the perturbation keeps them aligned so does not modify the term in the integral of (8.2.10). If they are not aligned, then after the perturbation, either they become aligned, or the perturbation on  $\Psi(h(x_1), h(x_2))$  is  $-2\eta\rho(x_2)$  for  $\eta$  small. Thus it appears that the perturbation cannot increase the integral of (8.2.10). Since it cannot decrease  $\mathcal{E}(h)$ , it cannot decrease  $\mathcal{E}(\psi \circ h)$ . We deduce as in Theorem 5.2.8 that  $\mathcal{E}(\psi \circ h, \rho)$  is non positive, so  $(\psi \circ h)(X_t)$  is from the analogue of Lemma 5.1.8 a submartingale up to the first exit time of  $M_\varepsilon$ . Thus  $h(X_t)$  is a martingale.  $\square$

## References

- [1] R. Abraham and L. Serlet, *Poisson snake and fragmentation*, Electron. J. Probab. **7** (2002), 1–15 (electronic).
- [2] M. Arnaudon, *Barycentres convexes et approximations des martingales continues dans les variétés*, Séminaire de Probabilités, XXIX, Lecture Notes in Math., vol. 1613, Springer, Berlin, 1995, pp. 70–85.
- [3] ———, *Differentiable and analytic families of continuous martingales in manifolds with connection*, Probab. Theory Relat. Fields **108** (1997), 219–257.
- [4] M. Barlow, J. Pitman, and M. Yor, *On Walsh’s Brownian motions*, Séminaire de Probabilités, XXIII, Lecture Notes in Math., vol. 1372, Springer, Berlin, 1989, pp. 275–293.
- [5] N. Bouleau and F. Hirsch, *Dirichlet forms and analysis on Wiener space*, Studies in Mathematics, vol. 14, de Gruyter, 1991.
- [6] Z. Q. Chen, *On reflected Dirichlet spaces*, Probab. Theory Related Fields **94** (1992), no. 2, 135–162.
- [7] R.W.R. Darling, *Martingales in manifolds — Definition, examples and behaviour under maps*, Séminaire de Probabilités XVI, Supplément: Géométrie différentielle stochastique (J. Azéma and M. Yor, eds.), Lect. Notes in Mathematics, vol. 921, Springer, 1982, pp. 217–236.
- [8] C. Dellacherie and P.A. Meyer, *Probabilités et potentiel, Chapitres V à VIII*, Hermann, 1980.
- [9] J. Eells and B. Fuglede, *Harmonic maps between Riemannian polyhedra*, Cambridge Tracts in Mathematics, vol. 142, Cambridge University Press, Cambridge, 2001.
- [10] M. Émery, *Stochastic calculus in manifolds*, Universitext, Springer, 1989.
- [11] M. Émery and G. Mokobodzki, *Sur le barycentre d’une probabilité dans une variété*, Séminaire de Probabilités, XXV, Lecture Notes in Math., vol. 1485, Springer, Berlin, 1991, pp. 220–233.
- [12] M. Émery and W. A. Zheng, *Fonctions convexes et semimartingales dans une variété*, Seminar on probability, XVIII, Lecture Notes in Math., vol. 1059, Springer, Berlin, 1984, pp. 501–518.

- [13] S. N. Evans, *Snakes and spiders: Brownian motion on  $\mathbf{R}$ -trees*, Probab. Theory Related Fields **117** (2000), no. 3, 361–386.
- [14] M. Fukushima, Y. Ōshima, and M. Takeda, *Dirichlet forms and symmetric Markov processes*, de Gruyter Studies in Mathematics, vol. 19, Walter de Gruyter & Co., Berlin, 1994.
- [15] J. Jost, *Equilibrium maps between metric spaces*, Calc. Var. Partial Differential Equations **2** (1994), no. 2, 173–204.
- [16] ———, *Riemannian geometry and geometric analysis*, Universitext, Springer-Verlag, Berlin, 1995.
- [17] ———, *Generalized Dirichlet forms and harmonic maps*, Calc. Var. Partial Differential Equations **5** (1997), no. 1, 1–19.
- [18] W.S. Kendall, *Probability, convexity, and harmonic maps with small image I: Uniqueness and fine existence*, Proc. London Math. Soc. **61** (1990), no. 3, 371–406.
- [19] ———, *Probability, convexity and harmonic maps II: Smoothness via probabilistic gradient inequalities*, J. Functional Analysis **126** (1994), no. 1, 226–257.
- [20] N. J. Korevaar and R. M. Schoen, *Sobolev spaces and harmonic maps for metric space targets*, Comm. Anal. Geom. **1** (1993), no. 3-4, 561–659.
- [21] W. B. Krebs, *Brownian motion on the continuum tree*, Probab. Theory Related Fields **101** (1995), no. 3, 421–433.
- [22] J.-F. Le Gall, *A class of path-valued Markov processes and its applications to superprocesses*, Probab. Theory Related Fields **95** (1993), no. 1, 25–46.
- [23] ———, *Spatial branching processes, random snakes and partial differential equations*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1999.
- [24] P.A. Meyer, *Géométrie stochastique sans larmes*, Séminaire de Probabilités XV (J. Azéma and M. Yor, eds.), Lect. Notes in Mathematics, vol. 850, Springer, 1981, pp. 44–102.
- [25] J. Picard, *Martingales on Riemannian manifolds with prescribed limit*, J. Functional Analysis **99** (1991), no. 2, 223–261.

- [26] ———, *Barycentres et martingales sur une variété*, Ann. Inst. Henri Poincaré, Probab. Stat. **30** (1994), no. 4, 647–702.
- [27] ———, *Smoothness of harmonic maps for hypoelliptic diffusions*, The Annals of Probability **28** (2000), no. 2, 643–666.
- [28] ———, *The manifold valued Dirichlet problem for symmetric diffusions*, Potential Analysis **14** (2001), 53–72.
- [29] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, Grundlehren der Mathematischen Wissenschaften, vol. 293, Springer-Verlag, Berlin, 1991.
- [30] C. Stricker, *Semimartingales et valeur absolue*, Séminaire de Probabilités, XIII, Lecture Notes in Math., vol. 721, Springer, Berlin, 1979, pp. 472–477.
- [31] K.-T. Sturm, *A semigroup approach to harmonic maps*, Preprint.
- [32] ———, *Nonlinear martingale theory for processes with values in metric spaces of nonpositive curvature*, Ann. Probab. **30** (2002), no. 3, 1195–1222.
- [33] B. Tsirelson, *Triple points: from non-Brownian filtrations to harmonic measures*, Geom. Funct. Anal. **7** (1997), no. 6, 1096–1142.
- [34] J. Walsh, *A diffusion with a discontinuous local time*, Astérisque **52–53** (1978), 37–45.
- [35] M. Yor, *Some aspects of Brownian motion. Part II, Some recent martingale problems*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1997.