

NACEREDDINE BELILI and HENRI HEINICH (Rouen)

## MEDIAN FOR METRIC SPACES

*Abstract.* We consider a Köthe space  $(\mathbb{E}, \|\cdot\|_{\mathbb{E}})$  of random variables (r.v.) defined on the Lebesgue space  $([0, 1], \mathbf{B}, \lambda)$ . We show that for any sub- $\sigma$ -algebra  $\mathcal{F}$  of  $\mathbf{B}$  and for all r.v.'s  $X$  with values in a separable finitely compact metric space  $(M, d)$  such that  $d(X, x) \in \mathbb{E}$  for all  $x \in M$  (we then write  $X \in \mathbb{E}(M)$ ), there exists a median of  $X$  given  $\mathcal{F}$ , i.e., an  $\mathcal{F}$ -measurable r.v.  $Y \in \mathbb{E}(M)$  such that  $\|d(X, Y)\|_{\mathbb{E}} \leq \|d(X, Z)\|_{\mathbb{E}}$  for all  $\mathcal{F}$ -measurable  $Z$ . We develop the basic theory of these medians, we show the convergence of empirical medians and we give some applications.

**1. Introduction.** It has been a long time since M. Fréchet [14] and S. Doss [8] defined the notion of expectation of a random variable (r.v.)  $X$  with values in a metric space. The development of this theory proceeded through the work of [17], [18] where the metric space has negative curvature and for example [12] and [25] for manifolds. More recently, let us quote [27] for a study of the law of large numbers and [13] for the construction of a canonical barycentre.

In contrast, we had to wait for the 1980s to see the concept of median for vectorial r.v. appear out of the discrete case (see for example [4], [10], [15] and [30]). In dimension 1, the median is a point where the expected absolute deviation is minimized. In dimension  $\geq 2$ , the point from which the expected distance to a random point of the distribution is the smallest can be called the spatial median. The reader can find a survey of work up to 1990 in Small [29], and the paper of Kemperman [20] is used as a reference for statisticians. We can find some applications in Schuster [28], and certain algorithms in Drezner [9].

---

2000 *Mathematics Subject Classification*: 60B05, 60G48, 62F10.

*Key words and phrases*: median, metric spaces, conditional median, convergence of medians, Doss expectations, Köthe spaces.

The definitions of expectation and of median for an integrable r.v.  $X$  exhibit certain similarities. Both are solutions of inequalities: for the expectation

$$(1) \quad |x - m| \leq E[|X - x|], \quad \forall x \in \mathbb{R},$$

and for the median

$$(2) \quad E[|X - m|] \leq E[|X - x|], \quad \forall x \in \mathbb{R}.$$

Just as for the expectation, the medians are a fundamental tool for estimation.

The passage to r.v.'s with values in a metric space  $(M, d)$  leads from (1) to the definition of *Doss expectation*: a point  $b \in M$  is an *expectation* (in the sense of Doss) of a r.v.  $X$  if

$$d(x, b) \leq E[d(X, x)], \quad \forall x \in M.$$

The condition (2), which was considered only in the case of Banach spaces, allows us to define, formally, a median: a point  $m \in M$  is a *median* of  $X$  if

$$(3) \quad E[d(X, m)] \leq E[d(X, x)], \quad \forall x \in M.$$

Note that the above two definitions are related to the  $\mathbb{L}^1$ -norm of the r.v.'s  $d(X, x)$ . We will assume that  $d(X, x)$  belongs to a Banach space  $(\mathbb{E}, \|\cdot\|_{\mathbb{E}})$  of r.v.'s, more precisely, to a Köthe space.

The condition (3) becomes our starting point in the following form: a point  $m \in M$  is a *median* of  $X$  if

$$\|d(X, m)\|_{\mathbb{E}} \leq \|d(X, x)\|_{\mathbb{E}}, \quad \forall x \in M.$$

Finally we note that if  $m$  is a median of  $X$ , then (3) can be interpreted to say that  $m$  is a *best approximant* of  $X$  with respect to the trivial  $\sigma$ -algebra [5], [19].

More generally, our problem is formulated as follows: Let  $(\mathbb{E}, \|\cdot\|_{\mathbb{E}})$  be a Banach space of real r.v. (on  $([0, 1], \mathbf{B}, \lambda)$ ), let  $X$  be a r.v. with values in a metric space  $(M, d)$  such that  $d(X, x) \in \mathbb{E}$  for all  $x \in M$ , and let  $\mathcal{F}$  be a sub- $\sigma$ -algebra of  $\mathbf{B}$ . An  $\mathcal{F}$ -measurable random variable  $Y$  with values in  $M$  is called a *conditional median* (or  $\mathcal{F}$ -*median*) of  $X$  given  $\mathcal{F}$  if

$$\|d(X, Y)\|_{\mathbb{E}} \leq \|d(X, Z)\|_{\mathbb{E}} \quad \text{for all } \mathcal{F}\text{-measurable } Z.$$

Let  $\mathcal{M}_{\mathbb{E}}(\mathcal{F}, X)$  denote the set of all  $\mathcal{F}$ -conditional medians of  $X$  for  $\mathbb{E}$ . When  $\mathcal{F} = \{\emptyset, [0, 1]\}$ , one simply says *median* for  $\mathcal{F}$ -median and writes  $\mathcal{M}_{\mathbb{E}}(X)$  for the set of medians.

In [22], C. A. León and J.-C. Massé give a counterexample to the existence of an  $\mathbb{L}^1$ -median (also called spatial median) for some Banach spaces. Therefore we must specify the metric space  $(M, d)$ .

Koltchinskii [21] develops a class of extensions of the univariate case (M-quantiles), related in a certain way to M-parameters of probability distributions and their M-estimators.

When  $M$  is a Banach lattice and  $|x| = x^+ + x^-$ , Heinich [16] adopts an alternative definition: an  $\mathcal{F}$ -measurable r.v.  $Y$  is an  $\mathcal{F}$ -median of  $X$  for the order if for any other  $\mathcal{F}$ -measurable r.v.  $Z$ ,

$$\| |X - Y| \|_{\mathbb{E}} \leq \| |X - Z| \|_{\mathbb{E}}.$$

**1.1. Summary of main results.** For simplicity, we assume all r.v.'s are defined on the Lebesgue probability space  $([0, 1], \mathbf{B}, \lambda)$ .

Let  $(M, d)$  be a separable metric space which is *finitely compact* (i.e., the closed balls are compact). Moreover,  $\mathbb{E}$  is a rearrangement invariant (r.i.) and weakly sequentially complete (w.s.c.) Köthe space (see Appendix). A random variable  $X$  defined on the Lebesgue space  $([0, 1], \mathbf{B}, \lambda)$  with values in  $M$  is  $\mathbb{E}$ -integrable (written  $X \in \mathbb{E}(M)$ ) if for all  $x$  in  $M$ ,  $d(X, x)$  is in  $\mathbb{E}$ .

**THEOREM 1.** *For every  $X \in \mathbb{E}(M)$  there exists a median of  $X$ . If  $\mathbb{E} = \mathbb{L}^1$  then for every finite sub- $\sigma$ -algebra  $\mathcal{F}$  and every  $X \in \mathbb{L}^1(M)$  there exists an  $\mathcal{F}$ -median for  $X$ .*

This theorem is extended and proved in Theorem 5 below.

To obtain a generalization to Köthe spaces and any sub- $\sigma$ -algebras, we prove the following measure-theoretic theorem:

**THEOREM 2.** *If  $(M, d)$  is Doss convex, then the space  $\mathbb{E}(M)$  of  $\mathbb{E}$ -integrable r.v.'s is sub-weakly sequentially complete.*

For the definition of Doss convexity and the proof of this result, see Appendix (Theorem 10). As a consequence of Theorem 6 below, we obtain

**THEOREM 3.** *If  $(M, d)$  is Doss convex, then for every sub- $\sigma$ -algebra  $\mathcal{F}$  and for every  $X \in \mathbb{E}(M)$ , there exists an  $\mathcal{F}$ -median for  $X$ .*

With a view to statistical applications, we also have

**THEOREM 4.** *The empirical medians of  $X$  form a relatively compact set in  $M$  whose limit values are medians of  $X$ .*

The paper concludes with examples and applications of medians to parametric estimation. In the Appendix we find the main facts on Doss expectation, Köthe spaces and the proof of Theorem 10.

## 2. Medians and conditional medians

**2.1. Existence of medians.** In this subsection we show the following theorem which generalizes Theorem 1. The finite compactness hypothesis can be weakened [6].

**THEOREM 5.** *Let  $(M, d)$  be a separable metric space with the property that if  $(x_n) \subset M$  is a sequence such that  $d(x_n, y)$  converges for all  $y \in M$  then there exists  $x \in M$  with  $d(x, y) \leq \lim d(x_n, y)$ . Let  $\mathbb{E}$  be a Banach space of real r.v.'s with the Fatou property (see Appendix). Then*

- (a)  $\mathcal{M}_{\mathbb{E}}(X) := \{m \in M : \|d(X, m)\|_{\mathbb{E}} \leq \|d(X, y)\|_{\mathbb{E}}, \forall y \in M\} \neq \emptyset$ .
- (b) *If  $\mathcal{F}$  is a finite sub- $\sigma$ -algebra and  $\mathbb{E} = \mathbb{L}^1$ , then  $\mathcal{M}_{\mathbb{E}}(\mathcal{F}, X) \neq \emptyset$ .*

*Proof.* For an  $\mathbb{E}$ -integrable r.v.  $X$ , the function  $x \mapsto \|d(X, x)\|_{\mathbb{E}}$  is continuous. Let  $(x_n) \subset M$  be a sequence with  $\|d(X, x_n)\|_{\mathbb{E}} \rightarrow \inf_{y \in M} \|d(X, y)\|_{\mathbb{E}}$ . The sequence  $(x_n)$  is bounded and passing to a subsequence we can suppose  $d(x_n, y)$  converges for all  $y \in M$  (see Appendix, proof of Theorem 10). Then there exists  $x \in M$  such that  $d(X, x) \leq \underline{\lim} d(X, x_n)$ , and Fatou's property gives (a).

The second part results from Proposition 1 below.

Let  $\mathcal{F}$  be a finite  $\sigma$ -algebra generated by a partition  $\{B_1, \dots, B_n\}$ . For fixed  $B_i$  denote by  $\lambda_i$  the trace of  $\lambda$  on  $B_i$ , and by  $\mathcal{M}_{\lambda_i}(X)$  the set of medians of  $X$  for  $\lambda_i$ . We remark that  $\mathcal{M}_{\lambda_i}(X) \neq \emptyset$  (consider the function  $x \mapsto E_{\lambda_i}[d(X, x)]$ ).

**PROPOSITION 1.** *For  $\mathcal{F}$  finite, we have  $Y \in \mathcal{M}(\mathcal{F}, X)$  if and only if*

$$Y = \sum_{i=1}^n x_i \mathbb{1}_{B_i} \quad \text{where } x_i \in \mathcal{M}_{\lambda_i}(X).$$

*Proof.* The r.v.  $Y = \sum_{i=1}^n x_i \mathbb{1}_{B_i}$  is  $\mathcal{F}$ -measurable and

$$E[d(X, Y)] = \sum_{i=1}^n E[d(X, x_i) \mathbb{1}_{B_i}] = \sum_{i=1}^n E[d(X, x_i) \mathbb{1}_{B_i}].$$

But

$$\frac{1}{\lambda(B_i)} E[d(X, x_i) \mathbb{1}_{B_i}] = E_{\lambda_i}[d(X, x_i)] \leq \frac{1}{\lambda(B_i)} E[d(X, a) \mathbb{1}_{B_i}]$$

for all  $a \in M$ . Therefore  $E[d(X, x_i) \mathbb{1}_{B_i}] \leq E[d(X, a) \mathbb{1}_{B_i}]$  for all  $a \in M$ , and consequently,  $E[d(X, Y)] \leq \sum_{i=1}^n E[d(X, a) \mathbb{1}_{B_i}]$  for all  $a \in M$ . This shows that

$$E[d(X, Y)] \leq E[d(X, Z)], \quad \forall Z \text{ } \mathcal{F}\text{-measurable} \quad (\text{i.e. } Y \in \mathcal{M}(\mathcal{F}, X)).$$

Conversely, for  $Y \in \mathcal{M}(\mathcal{F}, X)$ , let  $x_i$  be the value of  $Y$  on the set  $B_i$ ; we show that  $x_i$  is a median of  $X$  for  $\lambda_i$ .

Set  $J := \{j : x_j \notin \mathcal{M}_{\lambda_j}(X)\}$ . For each  $j \in J$ , there exists  $a_j \in M$  such that  $E_{\lambda_j}[d(X, a_j)] < E_{\lambda_j}[d(X, x_j)]$ . The r.v.  $Z = \sum_{i \notin J} x_i \mathbb{1}_{B_i} + \sum_{j \in J} a_j \mathbb{1}_{B_j}$  is  $\mathcal{F}$ -measurable and satisfies  $E[d(X, Z)] < E[d(X, Y)]$ , which is a contradiction. Thus  $J = \emptyset$  and so Proposition 1 and Theorem 5 are proved. ■

REMARK. This case is classical, both for spatial medians and  $\mathbb{L}^1$ -medians. It enables one to connect the conditional medians with medians. We have a similar situation to that of the conditional expectation which, in the case of a finite  $\sigma$ -algebra, is calculated starting from the expectation with respect to the trace probabilities.

**2.2. Existence of conditional medians.** We will generalize Theorem 7 of Bru and Heinich [5], which ensures the existence of conditional medians:

“Suppose a Köthe space  $\mathbb{E}$  has one of the following two properties:

- $\mathbb{E}$  is weakly sequentially complete;
- $\mathbb{E}$  has Fatou’s property and its bounded parts are  $\mathbb{L}^1$ -uniformly integrable.

Then, for every sub- $\sigma$ -algebra  $\mathcal{F}$  and every real random variable  $f$  in  $\mathbb{E}$ , there exists an  $\mathcal{F}$ -measurable real r.v.  $g_0$  in  $\mathbb{E}$  such that  $\|f - g_0\|_{\mathbb{E}} \leq \|f - g\|_{\mathbb{E}}$  for every  $\mathcal{F}$ -measurable real r.v.  $g$ .”

Furthermore the following theorem extends Theorem 3 above.

**THEOREM 6.** *Let  $M$  be a Doss convex finitely compact metric space, and  $\mathbb{E}$  a Köthe space with one of the two following properties:*

- $\mathbb{E}$  is weakly sequentially complete;
- $\mathbb{E}$  has Fatou’s property and its bounded parts are  $\mathbb{L}^1$ -uniformly integrable.

Then, for every sub- $\sigma$ -algebra  $\mathcal{F}$  and every r.v.  $X$  in  $\mathbb{E}(M)$ , there exists an  $\mathcal{F}$ -conditional median of  $X$ .

*Proof.* Let  $X$  be in  $\mathbb{E}(M)$  and  $\mathcal{F}$  be a sub- $\sigma$ -algebra. A sequence  $(X_n)$  of  $\mathcal{F}$ -measurable r.v.’s is an *approximant sequence* of  $X$  given  $\mathcal{F}$  if

$$\|d(X, X_n)\|_{\mathbb{E}} \rightarrow \inf\{\|d(X, Y)\|_{\mathbb{E}} : Y \text{ } \mathcal{F}\text{-measurable}\}.$$

For  $a$  in  $M$ , let  $Y_n := X_n \mathbb{1}_{B_n^c} + a \mathbb{1}_{B_n}$  where  $B_n$  is a sequence of events of  $\mathcal{F}$  such that  $\lambda(B_n) \rightarrow 0$ . Then

$$d(X_n, X) \mathbb{1}_{B_n} + d(X, Y_n) = d(a, X) \mathbb{1}_{B_n} + d(X_n, X).$$

This relation allows one to show, as in the real case [5], that the sequence  $(d(x, Y_n))$  is  $\mathbb{E}$ -uniformly integrable for all  $x$  in  $M$ , hence it is  $\sigma(\mathbb{E}, \mathbb{E}')$  relatively compact. Consequently, as in Theorem 10, we obtain a subsequence (denoted as the initial sequence) such that  $d(x, Y_n)$  weakly converges in  $\mathbb{E}$ .

The inequality  $d(X, Y_n) \leq d(x, X) + d(x, Y_n)$  shows that (a subsequence of)  $(d(X, Y_n))$  converges for  $\sigma(\mathbb{E}, \mathbb{E}')$ ; denote by  $U$  its weak limit.

By using Theorem 10, denote by  $Y_\infty$  the a.s. “sub-limit” of the sequence  $(Y_n)$ , i.e.,  $d(x, Y_\infty) \leq f^x$ .

Let us show that  $U \geq d(X, Y_\infty)$  a.s. Suppose first  $X$  is a simple r.v., i.e.,  $X = \sum x_i \mathbb{1}_{A_i}$  where the sets  $A_i$  form a Borel partition of  $[0, 1]$ . We saw that the sequence  $d(x, Y_n) = \sum d(x_i, Y_n) \mathbb{1}_{A_i}$  converges weakly in  $\mathbb{E}$  to

$$U = \sum f^{x_i} \mathbb{1}_{A_i} \geq \sum d(x_i, Y_\infty) \mathbb{1}_{A_i}.$$

This proves the property for  $X$  simple.

In the general case, consider the sequence  $(f_k)$  (see [13]) such that  $f_k(X)$  is a r.v. with finite values in  $M$  and  $\|d(f_k(X), X)\|_{\mathbb{E}} \rightarrow 0$ . One has

$$|d(f_k(X), Y_n) - d(X, Y_n)| \leq d(f_k(X), X) \xrightarrow{\mathbb{E}} 0 \quad \text{as } n \rightarrow \infty.$$

And, since  $f_k(X)$  is simple, without loss of generality, for all  $k$  we have

$$d(f_k(X), Y_n) \xrightarrow{\sigma(\mathbb{E}, \mathbb{E}')} U_k \geq d(f_k(X), Y_\infty).$$

Let us show the convergence of  $U_k$  to  $U$  in  $\mathbb{E}$ . Let  $\phi$  be a r.v. such that  $\|\phi\|_{\mathbb{E}'} \leq 1$ . Then

$$\left| \int \phi(U_k - U) \right| = \left| \lim_n \int \phi(d(f_k(X), Y_n) - d(X, Y_n)) \right| \leq \|d(f_k(X), X)\|_{\mathbb{E}}.$$

Hence  $\|U_k - U\|_{\mathbb{E}} \rightarrow 0$ . Taking the limit in the inequality  $U_k \geq d(f_k(X), Y_\infty)$ , we get the required relation  $U \geq d(X, Y_\infty)$ .

To conclude, notice that  $Y_\infty$  is  $\mathcal{F}$ -measurable because all the  $X_n$  are. Finally,

$$\|d(X, Y_\infty)\| \leq \|U\| \leq \lim \|d(X, Y_n)\| = \inf\{\|d(X, Z)\| : Z \mathcal{F}\text{-measurable}\}.$$

The r.v.  $Y_\infty$  is thus an  $\mathcal{F}$ -conditional median of  $X$ . ■

**2.3. Properties of conditional medians.** In order to simplify the statements, throughout this subsection we suppose  $(M, d)$  is a finitely compact and Doss convex separable metric space; and  $\mathbb{E}$  is a r.i. and w.s.c. Köthe space (see Appendix). We also write  $\mathcal{M}(\cdot)$  for  $\mathcal{M}_{\mathbb{E}}(\cdot)$ .

PROPOSITION 2. (i)  $\mathcal{M}_{\mathbb{E}}(X)$  depends only on the law of  $X$ .

(ii) The set  $\mathcal{M}(\mathcal{F}, X)$  is closed and convex (in Doss's sense).

(iii) If  $\mathbb{E}$  is strictly convex then for all  $Y_1$  and  $Y_2$  in  $\mathcal{M}(\mathcal{F}, X)$ , we have  $d(Y_1, X) = d(Y_2, X)$  a.s.

(iv) If  $X_n$  converges to  $X$  in  $\mathbb{E}(M)$  and  $Y_n \in \mathcal{M}(\mathcal{F}, X_n)$  then  $(Y_n)$  is an approximant sequence of  $X$ .

(v) If  $Y$  is an  $\mathcal{F}$ -conditional median of  $X$ , then

$$\|d(x, Y)\|_{\mathbb{E}} \leq 2\|E^{\mathcal{F}}[d(x, X)]\|_{\mathbb{E}} \quad \text{a.s.}$$

*Proof.* The assertion (i) is obvious.

(ii) If  $Y_1$  and  $Y_2$  are two medians of  $X$  given  $\mathcal{F}$ , then the r.v.  $Y = \frac{1}{2}Y_1 + \frac{1}{2}Y_2$  (Doss expectation) is a median of  $X$ . Indeed,  $Y$  is  $\mathcal{F}$ -measurable

and for all  $x$  in  $M$ ,

$$\begin{aligned} \|d(x, Y)\|_{\mathbb{E}} &\leq \frac{1}{2}(\|d(x, Y_1)\|_{\mathbb{E}} + \|d(x, Y_2)\|_{\mathbb{E}}), \\ \|d(X, Y)\|_{\mathbb{E}} &\leq \frac{1}{2}(\|d(X, Y_1)\|_{\mathbb{E}} + \|d(X, Y_2)\|_{\mathbb{E}}), \end{aligned}$$

which proves the assertion.

(iii) When  $\mathbb{E}$  is strictly convex, assume  $Y_1 \neq Y_2$ ; the last inequality above is then a strict inequality, which contradicts the fact that  $Y_1$  and  $Y_2$  are two distinct medians.

(iv) For  $Y \in \mathbb{E}(\mathcal{F}, M)$ , we have  $\|d(Y_n, X_n)\|_{\mathbb{E}} \leq \|d(Y, X_n)\|_{\mathbb{E}}$ . Therefore  $\|d(Y_n, X)\|_{\mathbb{E}} \leq \|d(Y_n, X_n)\|_{\mathbb{E}} + \|d(X, X_n)\|_{\mathbb{E}} \leq \|d(Y, X_n)\|_{\mathbb{E}} + \|d(X, X_n)\|_{\mathbb{E}}$ .

(v) The inequalities  $d(x, Y) \leq d(x, X) + d(X, Y)$  and  $\|d(X, Y)\|_{\mathbb{E}} \leq \|d(x, X)\|_{\mathbb{E}}$  give the required result. ■

REMARK. When in (iii) we replace “ $\mathbb{E}$  is strictly convex” by “ $\mathbb{E}$  is strictly monotone” (i.e.,  $0 \leq f \leq g, \|f\|_{\mathbb{E}} = \|g\|_{\mathbb{E}}$  implies  $f = g$  a.s. [5]), we obtain:

(iii)' If  $\mathbb{E}$  is strictly monotone, then for all  $Y_1, Y_2 \in \mathcal{M}(\mathcal{F}, X)$  we have

$$d(Y_1, X) = d(Y_2, X) \quad \text{a.s.}$$

The adaptation of Proposition 8 of [5] gives:

PROPOSITION 3. *If  $\mathbb{E}$  is a uniformly monotone Köthe space, then  $\mathcal{M}(\mathcal{F}, X)$  is weakly compact in  $\mathbb{E}(M)$ .*

*Proof.* Let  $(Y_n) \subset \mathcal{M}(\mathcal{F}, X)$ . This sequence is an approximant sequence of  $X$  given  $\mathcal{F}$ . By Theorem 10, for all  $x$  in  $M$ ,  $d(x, Y_n)$  is  $\mathbb{L}^1$ -uniformly integrable and, as  $\mathbb{E}$  is uniformly monotone,  $(d(x, Y_n))$  is  $\mathbb{E}$ -uniformly integrable. The sequence  $(d(X, Y_n))$  admits a subsequence which has a finite weak limit  $U$ . Moreover, by Theorem 11 and the Fatou property, there exists an  $\mathcal{F}$ -measurable r.v.  $Y$  such that  $d(X, Y) \leq U$ , i.e.,

$$\|d(X, Y)\|_{\mathbb{E}} \leq \|d(X, U)\|_{\mathbb{E}} \leq \underline{\lim} \|d(X, X_n)\|_{\mathbb{E}},$$

which completes the proof. ■

We can complete the preceding proposition by:

PROPOSITION 4. *Suppose that  $\mathbb{E}$  is a uniformly monotone Köthe space with the Kadec property: “if  $f_n \rightarrow f$  weakly in  $\mathbb{E}$  and  $\|f_n\|_{\mathbb{E}} \rightarrow \|f\|_{\mathbb{E}}$ , then  $\|f_n - f\|_{\mathbb{E}} \rightarrow 0$ ”. Then from any approximant sequence we can extract a subsequence converging to a median and  $\mathcal{M}(\mathcal{F}, X)$  is compact in  $\mathbb{E}(M)$ .*

Kemperman’s paper [20] concerns the case  $\mathbb{E} = \mathbb{L}^1$  and convexity of a norm on  $\mathbb{R}^d$ . We obtain a version of the preceding assertions (iii) and (iii)'.

PROPOSITION 5. *If  $M$  is  $d$ -strictly convex and  $\mathbb{E}$  is strictly monotone, then the set  $\mathcal{M}(\mathcal{F}, X)$  is a singleton.*

*Proof.* For two different r.v.'s  $Y_1$  and  $Y_2$  belonging to  $\mathcal{M}(\mathcal{F}, X)$ , the r.v.  $Y = \frac{1}{2}(Y_1 + Y_2)$  is also in  $\mathcal{M}(\mathcal{F}, X)$  and  $d(Y, X) < \frac{1}{2}(d(Y_1, X) + d(Y_2, X))$  a.s., which contradicts the equalities  $\|d(Y, X)\|_{\mathbb{E}} = \|d(Y_1, X)\|_{\mathbb{E}} = \|d(Y_2, X)\|_{\mathbb{E}}$ . ■

For  $(M, d) = (\mathbb{R}^d, \|\cdot\|)$ ,  $d > 1$ , the Euclidean norm is strictly convex if the points are not on the same line. So, one recovers Theorem 2.7 of Kemperman [20], when the support of  $X$  is not a.s. on the same line.

**3. Convergence of empirical medians.** In this section we study the convergence of empirical medians. The next theorem gives the convergence of the medians when  $X_n \rightarrow X$ .

**THEOREM 7.** *Suppose that  $M$  and  $\mathbb{E}$  satisfy the conditions of Theorem 5. Let  $(X_n) \subset \mathbb{E}(M)$  be such that  $\|d(X_n, X)\|_{\mathbb{E}} \rightarrow 0$ , and  $m_n \in \mathcal{M}(X_n)$ . Then there exists  $m \in \mathcal{M}(X)$  such that  $d(m, a) \leq \underline{\lim}_n d(m_n, a)$  for all  $a \in M$ . In particular if  $M$  is finitely compact, then there exists a subsequence such that  $m_{n_i} \rightarrow m \in \mathcal{M}(X)$ .*

*Proof.* By Proposition 2(v), the sequence  $(m_n)$  is bounded. The proof of Theorem 10 shows there exists  $m$  such that

$$d(m, a) \leq \underline{\lim}_n d(m_n, a), \quad \forall a \in M.$$

The Fatou property implies  $\|d(m, X)\|_{\mathbb{E}} \leq \underline{\lim} \|d(m_n, X)\|_{\mathbb{E}}$ . Hence  $m \in \mathcal{M}(X)$ . Indeed, the inequality  $d(m_n, X) \leq d(m_n, X_n) + d(X_n, X)$  gives  $\|d(m_n, X)\|_{\mathbb{E}} \leq \|d(a, X_n)\|_{\mathbb{E}} + \|d(X_n, X)\|_{\mathbb{E}}$ , so  $\|d(m, X)\|_{\mathbb{E}} \leq \|d(a, X)\|_{\mathbb{E}}$ . The rest of Theorem 7 is easy. ■

The convergence of empirical measures is a key point to statistics. We show the Glivenko–Cantelli Theorem, which is a tool for convergence of empirical medians.

For two probability measures  $P$  and  $Q$  on  $M$ , we define the Kantorovich–Köthe functional [1] by  $\mathcal{K}(P, Q) = \inf\{\|d(X, Y)\|_{\mathbb{E}} : X \sim P, Y \sim Q\}$ .

**THEOREM 8.** *Suppose that  $M$  and  $\mathbb{E}$  satisfy the conditions of Theorem 5 and  $\mathbb{E}$  is a r.i. Köthe space such that  $\mathbb{L}^p \subset \mathbb{E} \subset \mathbb{L}^1$  for  $1 \leq p < \infty$ . Let  $(X_n)$  be a sequence of i.i.d. random variables having the same law as  $X \in \mathbb{E}(M)$ ,  $X \sim P$ . Set  $\mu_n^t = n^{-1} \sum_{i=1}^n \delta_{X_i(t)}$ . Then, a.s.,  $\mathcal{K}(\mu_n^t, P) \rightarrow 0$  and for  $a \in M$ ,  $n^{-1} \sum_{i=1}^n d(X_i, a) \xrightarrow{\mathbb{E}} E[d(X, a)]$ .*

*Proof.* We have the convergence in law of the empirical laws to  $P$ . The Skorokhod Theorem on stochastic representation [11] ensures the existence of a sequence  $Y_n^t(\cdot)$  of r.v.'s defined on the canonical space, with law  $\mu_n^t$  and converging a.s. to a r.v.  $Y^t$  with law  $P$ . The sequence  $(d(Y_n^t, a))_n$  is  $\mathbb{E}$ -uniformly integrable and converges a.s. Furthermore  $d(Y_n^t, a) \xrightarrow{\mathbb{E}} d(Y^t, a)$ .



By Proposition 9 (Appendix), we obtain the second assertion. The first part is a consequence of Theorem 6 (see [1]). ■

We can state the following result:

**THEOREM 9.** *Suppose that  $M$  and  $\mathbb{E}$  satisfy the conditions of Theorem 8. For an i.i.d. sequence  $(X_n)$  having the same law as  $X \in \mathbb{E}(M)$ , denote by  $Y_n^t(\cdot)$  a r.v. with law  $\mu_n^t = n^{-1} \sum_{i=1}^n \delta_{X_i(t)}$  and  $m_n^t \in \mathcal{M}(Y_n^t)$ . Then there exists  $m \in \mathcal{M}(X)$  such that a.s.  $d(m, a) \leq \underline{\lim} d(m_n^t, a)$  for all  $a \in M$ . In particular, if  $M$  is finitely compact, then the sequence  $(m_n^t)_n$  is relatively compact in  $M$  and its limit points are medians of  $X$ . Moreover, if  $\mathbb{E}$  is strictly convex, then the sequence  $(m_n^t)_n$  converges to the median of  $X$ .*

*Proof.* To simplify the notations, we write  $m_n$  for  $m_n^t$  and  $Y_n$  for  $Y_n^t$ . By Proposition 2(i) and the Skorokhod Theorem, we can replace  $X$  by a r.v.  $Y \sim X$  and choose a sequence  $(Y_n)$   $\mathbb{E}$ -uniformly integrable which converges a.s. and in  $\mathbb{E}$ , i.e.,  $\|d(Y_n, Y)\|_{\mathbb{E}} \rightarrow 0$ . Then the assertion is a consequence of Theorem 7. ■

### 4. Examples and applications

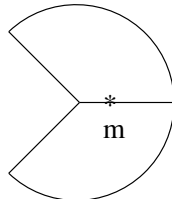
#### 4.1. Medians on $\mathbb{R}^2$

1) Let  $M_1, \dots, M_n$  be  $n$  non-collinear points of  $\mathbb{R}^2$ . The median  $M_0$  of these points is determined by  $\sum_{i=1}^n \vec{U}_i = 0$  or  $\sum_{i=1}^n \cos(\theta_i) = 0$  where  $\vec{U}_i$ , respectively  $\theta_i$ , is the unit vector joining  $M_0$  to  $M_i$ , resp. the angle formed by the vector  $M_0M_i$  and a fixed direction (cf. [15]).

2) Let  $X$  have a uniform law on the triangle with vertices  $\{(0, 0); (1, 1); (1, -1)\}$ . With the usual length, we have  $(0, 0.79313 \dots) = \mathcal{M}_{\mathbb{L}^1}(X)$  and  $(0, 3/4) = \mathcal{M}_{\mathbb{L}^2}(X)$ .

3) *Pac-man median.* Let  $X$  be a r.v. with uniform law on the Pac-man (see below). It is easy to see that the median  $m$  of  $X$  is the number  $m$  which minimizes

$$\int_0^{3\pi/4} \int_0^1 \sqrt{\varrho^2 - 2\varrho m \cos \theta} \varrho d\varrho d\theta.$$



We find  $m = 0.25$ .

4) *Poincaré’s half plane.* This framework can naturally be adapted to the hyperbolic space  $\mathbb{H} = M$  formed by Poincaré’s half-plane which satisfies

the conditions of our study. Let  $X$  be a r.v. with law  $\frac{1}{5}(\delta_{(0,y)} + \delta_{(1,1)} + \delta_{(-1,1)})$ , where  $y > 2.10487$ , defined on  $\mathbb{H}$ . Then  $(0, 2.10487) \in \mathcal{M}_{\mathbb{L}^1}(M)$ .

**4.2. Medians on  $\mathbb{N}$ .** Let  $X$  be a r.v. with law  $P(X = n) = \exp(-\lambda)\lambda^n/n!$ .  $\mathbb{N}$  is naturally a metric space and we will find  $\mathcal{M}(X)$ . For  $k \in \mathbb{N}$  let  $f(k) = \sum(\lambda^n/n!)|n - k|$ . The sign of  $f(k + 1) - f(k)$  shows that the median is  $\lfloor \lambda \rfloor$  or  $\lfloor \lambda \rfloor + 1$ , where  $\lfloor x \rfloor \in \mathbb{N}$  and  $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$ .

**4.3. Estimate of a translation parameter.** We suppose that a distance on  $\mathbb{R}$  is translation invariant. Let  $X$  be a real r.v. with density  $f$  differentiable and symmetric with respect of 0, i.e.,  $f(-x) = f(x)$ . Then  $E[d(X, x)] < \infty \Rightarrow 0 \in \mathcal{M}_{\mathbb{L}^1}(X)$ . Indeed,

$$\frac{d}{dx} E[d(X, x)]|_{x=0} = E \left[ \frac{d}{dx} d(X, x) \Big|_{x=0} \right] = 0.$$

By Theorem 7, we obtain:

**PROPOSITION 6.** *Suppose  $d(x, y) = d(x - y, 0)$  and  $X$  is a r.v. such that  $X \sim -X$  and  $E[d(X, 0)] < \infty$ . Then  $a \in \mathcal{M}_{\mathbb{L}^1}(X + a)$ .*

We can use this fact to estimate a translation parameter from an  $n$ -sample. We consider the Cauchy law. As we have supposed that the r.v.'s are integrable, we must take some other equivalent distance on  $\mathbb{R}$ , for example, the distance  $d_1(x, y) = \sqrt{|x - y|}$  which makes the Cauchy law integrable:

$$\int d_1(x, a) \frac{1}{1 + x^2} dx < \infty.$$

One can also take the bounded distance  $d_2(x, y) = |x - y|/(1 + |x - y|)$  (in this case the space  $(\mathbb{R}, d_2)$  is not finitely compact).

An application is, for example, the determination of the parameter of translation of the Cauchy law. The following values are the realization of a 20-sample of the standard Cauchy law:

- $x_1 := 8.04984124;$        $x_2 := 5.02404057;$        $x_3 := -7.17135224;$
- $x_4 := 1.09634317;$        $x_5 := 0.75504330;$        $x_6 := 0.01115475;$
- $x_7 := 4.38616372;$        $x_8 := 1.82269607;$        $x_9 := 0.14906384;$
- $x_{10} := -0.03833015;$        $x_{11} := -0.32844966;$        $x_{12} := -0.91846761;$
- $x_{13} := 4.21925516;$        $x_{14} := -0.22843143;$        $x_{15} := 7.92810925;$
- $x_{16} := -0.73455212;$        $x_{17} := 0.24745305;$        $x_{18} := -0.03568694;$
- $x_{19} := -1.82256075;$        $x_{20} := -1.54843756.$

For  $d_1$ , the median estimator is  $\hat{\theta} = x_6 = 0.01115475$ .

**REMARK.** For the distances considered below,  $\mathbb{R}$  is not Doss convex. Nevertheless  $\mathcal{M}(x_1 \dots, x_n) \subset \{x_1 \dots, x_n\}$  and the determination of medians is given by rearrangement.

**4.4. Traffic.** We are interested in the following type of problem: where should we put a bus-stop in order to serve an area whose population distribution follows a given law? In a similar way, where can a car-park be put in order to optimize the travels? The variations are numerous.

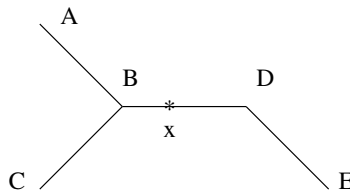
Only a very simplified version of this problem is studied here: a binary tree  $A, B, C, D, E$  whose edges are of length one. Various hypotheses are considered:

- ( $H_1$ ) The r.v.  $X$  has law  $\frac{1}{5}(\delta_A + \delta_B + \delta_C + \delta_D + \delta_E)$ .
- ( $H_2$ ) The law of  $X$  is uniform on the edges  $(A, B); (C, B); (B, D); (D, E)$ .

We can also suppose for example that the law is uniform on the edges  $(A, B); (C, B); (B, D); (B, E)$ .

Let  $x$  be a point of  $(B, D)$  situated at a distance  $x$  from  $B$ . Then we can show that

- ( $H_1$ )  $x = 0 \in \mathcal{M}_{L^1}(X); \quad x = 1/5 \in \mathcal{M}_{L^2}(X),$
- ( $H_2$ )  $x = 0 \in \mathcal{M}_{L^1}(X); \quad x = 1/4 \in \mathcal{M}_{L^2}(X).$



**4.5. Medians of probability measures.** Let  $\mathcal{P}_1$  be the set of probability measures,  $P$ , such that  $\int |x| dP < \infty$ . On this space we consider the  $l_1$ -distance (Wasserstein metric) defined by

$$l_1(P, Q) = \int_{\mathbb{R}} |F(t) - G(t)| dt = \int_0^1 |F^{-1}(t) - G^{-1}(t)| dt,$$

where  $F$  and  $G$  are the distribution functions (d.f.) of  $P$  and  $Q$  and  $F^{-1}(t) = \inf\{F(x) \geq t\}$  ([26], [24], [28]). For  $\alpha \in [0, 1]$  we can define the Doss expectation of  $\mu = \alpha\delta_P + (1 - \alpha)\delta_Q$  to be  $\alpha F + (1 - \alpha)G$  which is the d.f. of  $\alpha P + (1 - \alpha)Q$ . We can also choose  $\alpha F^{-1} + (1 - \alpha)G^{-1} = H^{-1}$ ; then  $H$  is the d.f. of the probability measure  $\alpha P \oplus (1 - \alpha)Q$ . So  $(\mathcal{P}_1, l_1)$  is Doss convex but not strictly convex.

Let  $N$  be a r.v. with values in  $(\mathcal{P}_1, l_1)$ . Then  $N$  defines a kernel by  $N(t, A) = N(t)(A)$ . We want to find  $\mathcal{M}(N)$ .

For the discrete case, we suppose

$$N(t) = \sum_{n=1}^{2n+1} P_i \mathbb{1}_{[(i-1)/(2n+1), i/(2n+1)]}(t),$$

and set  $F(t) = \text{median of } \{F_i(t) : i = 1, \dots, 2n + 1\}$ ,  $F_i$  d.f. of  $P_i$ . It is easy to see that  $F$  is the d.f. of a probability measure  $P$  belonging to  $\mathcal{M}_{\mathbb{L}^1}(N)$ . The odd case gives the same result for a suitable choice of medians. Theorem 9 gives the general case. Let us remark we can also achieve  $F^{-1}(t) \in \mathcal{M}(F_1^{-1}, \dots, F_n^{-1})$ . We obtain

PROPOSITION 7. *Let  $N$  be a r.v. with values in  $(\mathcal{P}_1, l_1)$  with  $l_1(N(\cdot), Q) \in \mathbb{L}^1$  for some  $Q \in \mathcal{P}_1$ . There exists  $P \in \mathcal{P}_1$  such that*

$$E[l_1(N(\cdot), P)] \leq E[l_1(N(\cdot), Q)], \quad \forall Q \in \mathcal{P}_1.$$

The d.f. of  $P$  is given by  $F(x)$  where  $F(x)$  is a median of the r.v.  $t \mapsto N(t, ]-\infty, x]) = F_t(x)$ , i.e.

$$\int_0^1 |F_t(x) - F(x)| dt \leq \int_0^1 |F_t(x) - y| dt, \quad \forall y.$$

We can also take  $F$  such that  $F^{-1}(x)$  is a median of  $t \mapsto F_t^{-1}(x)$ .

**4.6.  $\mathbb{L}^1$  conditional medians.** Here, we study conditional medians of  $(X, Y)$  given  $X$ , when  $\mathbb{E} = \mathbb{L}^1$ . The result is the same as the classical one for conditional expectation. Let  $\mathcal{F}$  be the  $\sigma$ -algebra generated by  $X$ ; for example, in the case of  $\mathbb{R}^2$ ,  $\mathcal{F} = \mathbf{B}(\mathbb{R}) \otimes \mathbb{R}$ .

PROPOSITION 8. *There exists a measurable function  $m$ ,  $x \mapsto m_x$ , such that  $m_x \in \mathcal{M}_{\mathbb{L}^1}(Y \setminus X = x)$  and  $m \in \mathcal{M}_{\mathbb{L}^1}(\mathbb{F}, (X, Y))$ .*

*Proof.* Let us start with a r.v.  $X$  with finite values. We find an  $\mathcal{F}$ -measurable map  $f$  such that

$$\sum d\left(\begin{pmatrix} x_i \\ y_i \end{pmatrix}, \begin{pmatrix} x_i \\ f(x_i) \end{pmatrix}\right) \leq \sum d\left(\begin{pmatrix} x_i \\ y_i \end{pmatrix}, \begin{pmatrix} x_i \\ g(x_i) \end{pmatrix}\right), \quad \forall g \mathcal{F}\text{-measurable.}$$

Considering the conditional law of  $Y$  given  $X = x_i$  and  $m_{x_i} \in \mathcal{M}_{\mathbb{L}^1}(Y | X = x_i)$ , we can see it suffices to take  $f(x_i) = m_{x_i}$ .

The general case is given by Theorem 7 or by a selection theorem [7]. ■

### Appendix

**5.1. Doss expectations.** In what follows,  $(M, d)$  is a separable complete metric space endowed with its Borel  $\sigma$ -algebra. A r.v.  $X$  defined on a probability space  $(\Omega, \mathcal{A}, P)$  with values in  $M$  is *integrable* if for some (and then for any) point  $a$  in  $M$  the r.v.  $d(X, a)$  is integrable. For an integrable r.v.  $X$ , the set

$$\{E[X]\} := \{m \in M : d(a, m) \leq E[d(a, X)], \forall a \in M\}$$

is the *expectation* of  $X$  in the sense of Doss. The set  $\{E[X]\}$  is closed (possibly empty) and depends only on the law of  $X$ ; we also write  $\{E[\mu]\}$  if  $\mu$  is the law of  $X$ .

The space  $M$  is called *Doss convex* (resp. *Doss strictly convex*) if for every probability measure  $\mu := \frac{1}{2}(\delta_x + \delta_y)$ , the set  $\{E[\mu]\}$  is non-empty (resp. is a singleton);  $\delta_a$  denotes the Dirac mass at the point  $a$ .

If  $M$  is Doss convex, then for  $a \in M$ , the function  $x \mapsto d(x, a)$  is *convex*, i.e., if  $z$  is an expectation of  $\alpha\delta_x + (1 - \alpha)\delta_y$ ,  $\alpha \in ]0, 1[$ , then  $d(z, a) \leq \alpha d(x, a) + (1 - \alpha)d(y, a)$ .

If  $M$  is *Doss strictly convex*, then for all  $(x, a) \in M^2$ ,  $x \mapsto d(x, a)$  is *strictly convex*, i.e., for all  $y \in M$ ,  $y \neq x$ , if  $z$  is an expectation of  $\alpha\delta_x + (1 - \alpha)\delta_y$ , then  $d(\alpha x + (1 - \alpha)y, a) < \alpha d(x, a) + (1 - \alpha)d(y, a)$ ,  $\alpha \in ]0, 1[$ .

Let  $\mathcal{F}$  be a sub- $\sigma$ -algebra of  $\mathbf{B}$ . An  $M$ -valued r.v.  $Y$  is a *conditional expectation* of  $X$  given  $\mathcal{F}$  if

- (i)  $Y$  is  $\mathcal{F}$ -measurable;
- (ii) for all  $a$  in  $M$ ,  $d(a, Y) \leq E^{\mathcal{F}}[d(a, X)]$  a.s.

We denote by  $\{E^{\mathcal{F}}[X]\}$  the set of all r.v.'s  $Y$  satisfying (i) and (ii).

Let us note the following property: if  $\mathcal{F}_1$  and  $\mathcal{F}_2$  are two sub- $\sigma$ -algebras such that  $\mathcal{F}_1 \subset \mathcal{F}_2$  then  $E^{\mathcal{F}_1}(E^{\mathcal{F}_2}) \subset E^{\mathcal{F}_1}$ .

A sequence  $(Y_n)_n$  adapted to a monotone filtration  $(\mathcal{F}_n)_n$  is a *martingale* if  $Y_n \in \{E^{\mathcal{F}_n}[Y_{n+1}]\}$ . Let us recall the theorem on a.s. convergence of martingales according to [8] or [3]:

*If  $M$  is finitely compact, then any martingale  $(Y_n)$  satisfying Doob's condition (i.e.,  $\sup_n E[d(x, Y_n)] < \infty$  for all  $x \in M$ ) converges a.s.*

**5.2. Köthe spaces.** A Banach space  $(\mathbb{E}, \|\cdot\|_{\mathbb{E}})$  of real r.v.'s defined on the Lebesgue space  $([0, 1], \mathbf{B}, \lambda)$  is called a *Köthe space* if

- (i)  $\mathbb{L}^{\infty} \subset \mathbb{E} \subset \mathbb{L}^1$ , the injections are continuous;
- (ii) if  $X$  and  $Y$  are two r.v.'s defined on  $([0, 1], \mathbf{B}, \lambda)$  such that  $|X| \leq |Y|$  with  $Y \in \mathbb{E}$ , then  $X \in \mathbb{E}$  and  $\|X\|_{\mathbb{E}} \leq \|Y\|_{\mathbb{E}}$ .

We adopt the terminology of [23]; in particular  $\mathbb{E}^*$  denotes the topological dual of  $\mathbb{E}$  and  $\mathbb{E}' := \{Y \in \mathbb{L}^1 : XY \in \mathbb{L}^1 \text{ for all } X \in \mathbb{E}\}$ . Let us recall that:

- (a)  $\mathbb{E}' = \mathbb{E}^*$  if and only if  $\mathbb{E}$  has a *continuous norm* with respect to the order, i.e., if  $X_{\alpha} \downarrow 0$  then  $\|X_{\alpha}\|_{\mathbb{E}} \downarrow 0$ ;
- (b)  $\mathbb{E} = \mathbb{E}''$  if and only if  $\mathbb{E}$  has the *Fatou property*, i.e., for every sequence  $(X_n) \subset \mathbb{E}$  such that  $X_n \uparrow X$  and  $\sup \|X_n\|_{\mathbb{E}} < \infty$  we have  $X \in \mathbb{E}$  and  $\lim \|X_n\|_{\mathbb{E}} = \|X\|_{\mathbb{E}}$ .

Moreover, as we want to study functional relations between the laws of r.v.'s it is natural to require the norm of  $X \in \mathbb{E}$  to depend only upon the law of  $X$ . This is the justification of the following definitions.

An *automorphism*  $\tau$  of  $([0, 1], \mathbf{B}, \lambda)$  is a map of  $[0, 1]$  onto  $[0, 1]$  such that  $B \in \mathbf{B}$  if and only if  $\tau(B) \in \mathbf{B}$  and  $\lambda(\tau(B)) = \lambda(B)$ . In particular, any automorphism  $\tau$  of  $[0, 1]$  induces a non-negative isometry on  $\mathbb{L}^1$ .

A Köthe space  $\mathbb{E}$  is *rearrangement invariant* (r.i.) if for all  $X \in \mathbb{E}$ , we have  $X \circ \tau \in \mathbb{E}$  and  $\|X\|_{\mathbb{E}} = \|X \circ \tau\|_{\mathbb{E}}$  for every automorphism  $\tau$ . In fact, we can always renorm  $\mathbb{E}$  to become r.i.

We have the following lemmas.

LEMMA 1 ([1], [23]). *A Köthe space  $\mathbb{E}$  is r.i. if and only if for any r.v.'s  $X$  and  $Y$  on  $([0, 1], \mathbf{B}, \lambda)$  such that  $X \sim Y$  and  $X \in \mathbb{E}$ , we have  $Y \in \mathbb{E}$  and  $\|X\|_{\mathbb{E}} = \|Y\|_{\mathbb{E}}$ .*

LEMMA 2 ([1]). *Let  $\mathbb{E}$  be a r.i. Köthe space having the Fatou property and let  $(X_n) \subset \mathbb{E}$  be a bounded sequence with respect to  $\|\cdot\|_{\mathbb{E}}$  such that  $X_n$  converges in law to  $X_0$ . Then  $X_0 \in \mathbb{E}$  and  $\|X_0\|_{\mathbb{E}} \leq \liminf \|X_n\|_{\mathbb{E}}$ .*

A Köthe space  $\mathbb{E}$  is said to be *uniformly monotone* if for each  $\varepsilon > 0$ , there exists  $\eta$  such that if  $X, Y \in \mathbb{E}_+$  (where  $\mathbb{E}_+$  denotes  $\{X \in \mathbb{E} : X \geq 0\}$ ),  $\|X\|_{\mathbb{E}} = 1$  and  $\|X + Y\|_{\mathbb{E}} \leq 1 + \eta$  then  $\|Y\|_{\mathbb{E}} \leq \varepsilon$ .

The concept of uniformly integrable sequences, which is equivalent in the  $\mathbb{L}^1$  space to being relatively compact for  $\sigma(\mathbb{L}^1, \mathbb{L}^\infty)$ , can be generalized to Köthe spaces as follows ([5], [6]).

A sequence  $(X_n) \subset \mathbb{E}$  is called  *$\mathbb{E}$ -uniformly integrable* if for every  $A_n \in \mathbf{B}$  such that  $\lambda(A_n) \rightarrow 0$ , we have  $\|X_n 1_{A_n}\|_{\mathbb{E}} \rightarrow 0$ .

Let us recall (see [1]):

“If  $\mathbb{E}$  is a r.i. Köthe space,  $(X_n) \subset \mathbb{E}$  is  $\mathbb{E}$ -uniformly integrable and  $Y_n \sim X_n$  for all  $n \in \mathbb{N}$ , then the sequence  $(Y_n)$  is  $\mathbb{E}$ -uniformly integrable.”

PROPOSITION 9 ([1]). *Let  $\mathbb{E}$  be a r.i. Köthe space. Then the following statements are equivalent:*

- (i) *The sequence  $(X_n)$  is  $\mathbb{E}$ -uniformly integrable and converges in probability.*
- (ii) *The sequence  $(X_n)$  converges in  $\mathbb{E}$ .*

A r.v.  $X$  with values in a metric space  $(M, d)$ , is called  *$\mathbb{E}$ -integrable* if  $d(x, X) \in \mathbb{E}$  for all  $x \in M$ ; we then write  $X \in \mathbb{E}(M)$ . We denote by  $\mathbb{E}(\mathcal{F}, M)$  the subset of  $\mathbb{E}(M)$  formed by the  $\mathcal{F}$ -measurable r.v.'s. When  $M = \mathbb{R}$  one writes  $\mathbb{E}(\mathcal{F})$  instead of  $\mathbb{E}(\mathcal{F}, \mathbb{R})$ .

**5.3.** *The space  $\mathbb{E}(\mathcal{F}, M)$  is sub-weakly sequentially complete.* To study the existence of conditional medians on a separable finitely compact metric space, we have to adapt the classical theorem on weak convergence. It is possible to replace the “finitely compact” assumption by the condition of Theorem 5.

The following theorem shows that if the Köthe space  $\mathbb{E}$  is weakly sequentially complete (w.s.c.), then, in a sense, so is  $\mathbb{E}(M)$ .

THEOREM 10. *Let  $\mathbb{E}$  be a w.s.c. Köthe space of r.v.'s defined on  $([0, 1], \mathbf{B}, \lambda)$  and let  $\mathcal{F}$  be a sub- $\sigma$ -algebra of  $\mathbf{B}$ . If a sequence  $(X_n) \subset \mathbb{E}(M)$  is*

such that for all  $x$  in  $M$ , the sequence  $(d(X_n, x))$  is relatively compact for  $\sigma(\mathbb{E}, \mathbb{E}')$ , then there exists a sequence  $(n_i)$  and a r.v.  $Y \in \mathbb{E}(\mathcal{F}, M)$  such that: for each  $x \in M$ , the sequence  $(d(X_{n_i}, x))_i$  converges weakly to a real r.v.  $f^x$  and  $d(Y, x) \leq E^{\mathcal{F}}(f^x)$  a.s. We say that the space  $\mathbb{E}(M)$  is  $\mathcal{F}$ -sub-weakly sequentially complete, or  $\mathbb{E}(\mathcal{F}, M)$  is sub-w.s.c.

*Proof.* First, by the separability of  $M$ , let us show that there exists a subsequence  $(n_i)$  such that  $(d(X_{n_i}, x))_i$  converges weakly to a real r.v.  $f^x$ .

Let  $(a_k)$  be a dense sequence in  $M$ . By the diagonal process, there is a subsequence, still denoted by  $(n_i)$ , such that for all  $k$ ,  $d(a_k, X_{n_i}) \rightarrow f^{a_k}$  for the  $\sigma(\mathbb{E}, \mathbb{E}')$  topology.

Let us show that this convergence takes place for any  $x$  in  $M$ .

To simplify, let  $(n_i)$  be the initial sequence. For points  $x, x_n, x_m$  in  $M$ , the inequalities  $d(x, x_n) \leq d(x, a_q) + d(a_q, x_n)$  and  $d(a_p, x_m) \leq d(a_p, x) + d(x, x_m)$  imply

$$\begin{aligned} d(x, x_n) + d(a_p, x_m) &\leq d(x, a_q) + d(a_q, x_n) + d(a_p, x) + d(x, x_m) \\ &\leq d(x, a_q) + d(x, a_p) + d(a_q, a_p) + d(a_p, x_n) + d(x, x_m), \end{aligned}$$

therefore

$$d(x, x_n) - d(x, x_m) \leq d(x, a_q) + d(x, a_p) + d(a_q, a_p) + d(a_p, x_n) - d(a_p, x_m).$$

If we choose the indices  $p$  and  $q$  so that  $a_p$  and  $a_q$  converge to  $x$ , we obtain

$$(*) \quad d(x, x_n) - d(x, x_m) \leq \varepsilon + d(a_k, x_n) - d(a_k, x_m),$$

as soon as these indices are large enough.

Take  $\phi$  in  $\mathbb{E}'_+$ , multiply both sides of the inequality  $(*)$  by  $\phi$  and integrate to obtain

$$\int \phi(d(x, X_n) - d(x, X_m)) d\lambda \leq \varepsilon \int \phi + \int \phi(d(a_k, X_n) - d(a_k, X_m)).$$

By interchanging the roles of  $n$  and  $m$ ,

$$d(x, x_m) - d(x, x_n) \leq \varepsilon + d(a_k, x_m) - d(a_k, x_n).$$

Thus, as soon as  $n$  and  $m$  are large enough, we have

$$\left| \int \phi(d(x, X_n) - d(x, X_m)) d\lambda \right| \leq \varepsilon \int \phi + \left| \int \phi(d(a_k, X_n) - d(a_k, X_m)) \right| \leq 2\varepsilon'.$$

We conclude that for all  $x$  in  $M$ ,  $d(x, X_{n_i})$  converges for  $\sigma(\mathbb{E}, \mathbb{E}')$ .

*Case when  $\mathcal{F}$  is a finite sub- $\sigma$ -algebra of  $\mathbf{B}$ .* One can define a representative r.v. of  $E^{\mathcal{F}}(X)$ , either using the canonical barycentre [13] when the space has negative curvature, or with the axiom of the choice in the general case. Let us present briefly this last possibility. Let  $\mathcal{F} = \sigma(B_1, \dots, B_k)$  and let  $E_k$  be the Doss expectation with respect to the probability trace  $\lambda_k$  on  $B_k$ . Then  $E^{\mathcal{F}}(X) = \{\sum y^k(X)\mathbb{1}_{B_k}\}$ , where  $y^k(X)$  is a point of  $E_k(X)$ . By

choosing a representative  $y^k(X)$  in  $E_k(X)$ , we determine a representative r.v. of  $E^{\mathcal{F}}(X)$ , denoted by  $Y$ .

On the other hand, the map  $E^{\mathcal{F}} : \mathbb{E} \rightarrow \mathbb{E}(\mathcal{F})$  is continuous and remains continuous for the weak topologies. It is a contraction if the space is r.i. [5]. Moreover, since the  $\sigma$ -algebra is finite, the weak and strong topologies of  $\mathbb{E}(\mathcal{F})$  are identical.

*For  $M$  finitely compact and  $\mathcal{F}$  a finite sub- $\sigma$ -algebra, let us show that the space  $\mathbb{E}(\mathcal{F}, M)$  is weakly sequentially complete.*

Indeed, let  $(X_n)$  in  $\mathbb{E}(\mathcal{F}, M)$  be such that  $(d(x, X_n))$  is relatively compact in  $\sigma(\mathbb{E}, \mathbb{E}')$ , for all  $x$  in  $M$ . We have seen that for a suitable subsequence,  $d(x, X_{n_i})$  converges weakly to a limit denoted by  $f^x$ .

For  $B_k \in \mathcal{F}$ , we have  $\lim d(x, X_{n_i})P(B_k) = \int_{B_k} f^x d\lambda$ . Therefore, the sequence  $(d(x, X_{n_i}))$  is bounded and, since  $M$  is finitely compact, one can extract a new subsequence  $(n_i^*)$  such that  $X_{n_i^*}$  converges a.s. to  $X$  in  $E(\mathcal{F}, M)$ .

*Let us show that for  $M$  finitely compact and  $\mathcal{F}$  a finite sub- $\sigma$ -algebra, the space  $\mathbb{E}(M)$  is  $\mathcal{F}$ -sub-weakly sequentially complete.*

Now, we suppose that  $(X_n) \subset \mathbb{E}(M)$  and the sequence  $(d(x, X_n))$  is weakly relatively compact in  $\mathbb{E}$  for all  $x \in M$ . Denote by  $Y_n$  a representative of  $E^{\mathcal{F}}(X_n)$ ,  $Y = \sum_{n=1}^K y_n^k \mathbb{1}_{B_k}$ ; let us establish that the sequence  $(y_n^k)_n$  is bounded, for all  $k$ . It suffices to show that any subsequence admits a new one  $(n_i^*)$  such that  $(y_{n_i^*}^k)$  is bounded in  $M$ . For simplicity, assume that the initial sequence  $(d(x, X_n))$  converges weakly.  $Y_n$  being in  $E^{\mathcal{F}}(X_n)$  implies that

$$d(x, y_n^k) \mathbb{1}_{B_k} \leq E^{\mathcal{F}}(d(x, X_n)) \mathbb{1}_{B_k} \quad \text{a.s.}$$

As  $E^{\mathcal{F}}(d(x, X_n))$  converges strongly to  $E^{\mathcal{F}}(f^x)$ , one deduces that

$$\overline{\lim}_n d(x, y_n^k) \mathbb{1}_{B_k} \leq E^{\mathcal{F}}(f^x) \mathbb{1}_{B_k} \quad \text{a.s.}$$

By taking, as before,  $x$  in a dense countable set and by passing to the limit, one has

$$\overline{\lim}_i d(x, y_{n_i}^k) \mathbb{1}_{B_k} \leq E^{\mathcal{F}}(f^x) \mathbb{1}_{B_k}, \quad \forall x \in M,$$

outside a null set and for a subsequence. The space  $M$  being finitely compact, we can extract a subsequence  $(n_j^*)$  such that  $y_{n_j^*}^k$  converges for all  $k$  to a point  $y^k$  in  $M$ . The r.v.  $Y = \sum y^k \mathbb{1}_{B_k}$  is  $\mathcal{F}$ -measurable and it is clear that a.s.  $Y_{n_i^*} \rightarrow Y$ ,  $d(x, Y_{n_i^*}) \rightarrow d(x, Y) \leq f^x$  in  $\mathbb{E}$ . This proves the required assertion.

*General case.* Consider a finite sequence of sub- $\sigma$ -algebras  $\mathcal{F}_p$  increasing to  $\mathcal{F}$  and denote by  $Y^p$  the r.v. corresponding to  $Y$  in the  $p$ th part, for each  $p$ . There exists a subsequence,  $(n_i(p))_i$ , such that, for representatives



$Y_{n_i(p)}^p$  in  $E^{\mathcal{F}_p}(X_{n_i(p)})$ , we have

$$Y_{n_i(p)}^p \xrightarrow{i} Y^p \quad \text{a.s.}, \quad d(x, Y^p) \leq \lim E^{\mathcal{F}_p}(d(x, X_{n_i(p)})) \quad \text{a.s.}$$

Notice that we can choose the same subsequence for all  $p$  and, for this reason, we simplify the notations by denoting this subsequence like the initial sequence: for a fixed  $p$ , we have

$$(**) \quad \begin{aligned} Y_n^p &\text{ in } E^{\mathcal{F}_p}(X_n), & Y_n^p &\xrightarrow{n} Y^p \text{ a.s.}, \\ d(x, Y_n^p) &\xrightarrow{n} d(x, Y^p) &\text{ in } \mathbb{E}, & d(x, Y^p) \leq f^x \text{ a.s.} \end{aligned}$$

The proof is easy when we can choose  $Y^p$  in  $E^{\mathcal{F}_p}(Y^{p+1})$ . Indeed, the sequence  $(Y^p)$  is a martingale in the sense of Doss, and Doob's condition  $\sup_p E[d(x, Y^p)] < \infty$  is satisfied. Then the Doss Theorem ensures the a.s. convergence of  $Y^p$  to a r.v.  $Y_\infty$  which satisfies the conclusion of the theorem.

We will circumvent this difficulty. Without loss of generality, assume that  $(**)$  holds. For any fixed  $p$ , let  $\tilde{Y}_p^p := Y_p^p$  and for all  $k < p$ , choose a r.v.  $Y_k^p$  in  $E^{\mathcal{F}_k}(Y_{k+1}^p)$ . The finite sequence  $(\tilde{Y}_1^p, \dots, \tilde{Y}_{p-1}^p, \tilde{Y}_p^p)$  is a martingale. Since the spaces  $\mathbb{E}(\mathcal{F}_k, M)$  are weakly sequentially complete, we can extract from  $(\tilde{Y}_p^p)_p$  a subsequence which converges. Indeed, the inequalities  $d(x, \tilde{Y}_k^p) \leq E^{\mathcal{F}_k}(Y_{k+1}^p) \leq E^{\mathcal{F}_k}(Y^p) \leq E^{\mathcal{F}_k}(f^x)$  a.s. show that the sequence  $(d(x, \tilde{Y}_k^p))_p$  is weakly relatively compact in  $\mathbb{E}$ .

Varying  $k$  and by the diagonal procedure, there exists a subsequence  $(p_i)_i$  such that  $(\tilde{Y}_k^{p_i})_i$  converges a.s. for all  $k$ . For simplicity, let the subsequence be the initial sequence and let  $Y_k$  be the limit. For all  $k$  we have  $d(x, \tilde{Y}_k^p) \rightarrow d(x, Y_k)$  in  $\mathbb{E}$  and  $Y_k$  is  $\mathcal{F}_k$ -measurable. Since, for  $p \geq k + 1$ ,  $E^{\mathcal{F}_k}(d(x, \tilde{Y}_{k+1}^p)) \geq d(x, \tilde{Y}_k^p)$  a.s., taking the limit we deduce that

$$E^{\mathcal{F}_k}(d(x, Y_{k+1})) \geq d(x, Y_k) \quad \text{a.s.}$$

The sequence  $(Y_k)$  is a martingale in the sense of Doss.

Now, we prove that  $(Y_k)$  converges a.s. to a r.v.  $Y_\infty$ . To use the Doss Theorem, it suffices to verify that  $\sup_k E[d(x, Y_k)] < \infty$ . This results from

$$\begin{aligned} E[d(x, Y_k)] &= \lim_p E[d(x, \tilde{Y}_k^p)] \leq \overline{\lim}_p E[d(x, Y^p)] \leq \lim_n E[d(x, Y_k)] \\ &= E[f^x] < \infty. \end{aligned}$$

To complete the proof of the Theorem, let us show that  $d(x, Y_\infty) \leq f^x$  a.s. Let  $\phi$  be a real  $\mathcal{F}_k$ -measurable r.v. belonging to  $\mathbb{E}'_+$  or  $\mathbb{L}^\infty_+$ . For  $x \in M$ , we have

$$\int \phi d(x, X_n) = \int \phi E^{\mathcal{F}_k}[d(x, X_n)] \rightarrow \int \phi E^{\mathcal{F}_k}[f^x] = \int \phi f^x.$$

However if  $p \geq k$ , then

$$\int \phi d(x, X_n) = \int \phi E^{\mathcal{F}_p} [d(x, X_n)] \geq \int \phi d(x, Y_n^p) \xrightarrow{n} \int \phi x Y_k.$$

Consequently,  $d(x, Y_k) \leq E^{\mathcal{F}_k}(f^x)$  a.s. and taking the limit, we obtain the desired property  $d(x, Y_\infty) \leq f^x$  a.s. ■

Here is another version of the theorem.

**THEOREM 11.** *If a space  $\mathbb{E}$  has the Fatou property and its bounded parts are uniformly integrable in  $\mathbb{L}^1$ , then  $\mathbb{E}(M)$  is  $\mathcal{F}$ -sub-w.s.c. for all sub- $\sigma$ -algebras  $\mathcal{F}$ .*

*Proof.* Indeed, Theorem 10 shows that  $\mathbb{L}^1(M)$  is  $\mathcal{F}$ -sub-w.s.c. and since the sequence  $(d(x, X_n))$  is bounded in  $\mathbb{E}$ , its limit is in  $\mathbb{E}$  by Lemma 2. ■

### References

- [1] N. Belili and H. Heinich, *Kantorovich functional revisited*, preprint INSA, 2000.
- [2] —, —, *Transport problem and derivation*, Appl. Math. (Warsaw) 26 (1999), 299–314.
- [3] V. E. Benes, *Martingales on metric spaces*, Theor. Probab. Appl., 7 (1962), 82–83.
- [4] B. M. Brown, *Statistical uses of the spatial median*, J. Roy. Statist. Soc. Ser. B 45 (1983), 25–30.
- [5] B. Bru et H. Heinich, *Meilleures approximations et médianes conditionnelles*, Ann. Inst. H. Poincaré 21 (1985), 197–224.
- [6] B. Bru, H. Heinich et J.-C. Lootgieter, *Distances de Lévy et extensions des théorèmes de la limite centrale et de Glivenko–Cantelli*, Publ. Inst. Statist. Univ. Paris 37 (1993), no. 3–4, 29–42.
- [7] C. Dellacherie et P. A. Meyer, *Probabilités et potentiel*, Hermann, Paris, 1983.
- [8] S. Doss, *Moyennes conditionnelles et martingales dans un espace métrique*, C. R. Acad. Sci. Paris Sér. I 254 (1962), 3630–3632.
- [9] Z. Drezner, *On the conditional p-median problem*, Comput. Oper. Res. 22 (1995), 525–530.
- [10] G. R. Ducharme and P. Milasevic, *Spatial median and directional data*, Biometrika 74 (1987), 212–215.
- [11] R. M. Dudley, *Probability and Metrics*, Aarhus Univ., Aarhus, 1976.
- [12] M. Émery et G. Mokobodzki, *Sur le barycentre d’une probabilité dans une variété*, in: Séminaire de probabilités XXV (Strasbourg), Lecture Notes in Math. 1485, Springer, 1991, 220–233.
- [13] A. Es-Sahib et H. Heinich, *Barycentre canonique pour un espace métrique? Courbure négative*, in: Séminaire de probabilités XXXIII (Strasbourg), Lecture Notes in Math. 1709, Springer, 1999, 355–370.
- [14] M. Fréchet, *Les éléments aléatoires de nature quelconque*, Ann. Inst. H. Poincaré 14 (1948), 215–310.
- [15] J. B. S. Haldane, *Note on the median of a multivariate distribution*, Biometrika 35 (1948), 414–415.
- [16] H. Heinich, *Médianes vectorielles*, Ann. Inst. H. Poincaré Probab. Statist. 26 (1990), 375–385.

- [17] W. Herer, *Mathematical expectation and martingales of random subsets of a metric space*, Probab. Math. Statist. 11 (1991), 291–304.
- [18] —, *Mathematical expectation and strong law of large numbers for random variables with values in a metric space of negative curvature*, *ibid.* 13 (1992), 59–70.
- [19] N. Herrndorf, *Best  $\phi$  and  $N_\phi$ -approximations in Orlicz spaces of vector valued functions*, Z. Wahrsch. Verw. Gebiete 58 (1981), 309–329.
- [20] J. H. B. Kemperman, *The median of a finite measure on a Banach space*, in: Y. Dodge (ed.), *Statistical Data Analysis Based on the  $L_1$ -Norm*, Kluwer, Dordrecht, 1987, 217–230.
- [21] V. I. Koltchinskii, *M-estimation, convexity and quantiles*, Ann. Statist. 25 (1997), 435–477.
- [22] C. A. León and J.-C. Massé, *A counterexample on the existence of the  $L_1$ -median*, Statist. Probab. Lett. 92 (1992), 117–120.
- [23] J. Lindenstrauss and L. Tzafriri, *Classical Banach Spaces*, Springer, Berlin, 1996.
- [24] P. Major, *On the invariance principle for sums of independent identically distributed random variables*, J. Multivariate Anal. 8 (1978), 487–517.
- [25] J. Picard, *Barycentres et martingales sur une variété*, Ann. Inst. H. Poincaré 30 (1994), 647–702.
- [26] S. T. Rachev and L. Rüschendorf, *Mass Transportation Problems*, Springer, New York, 1998.
- [27] P. Raynaud de Fitte, *Théorème ergodique ponctuel et lois fortes des grands nombres pour des points aléatoires d'un espace métrique à courbure négative*, Ann. Probab. 25 (1997), 738–766.
- [28] E. F. Schuster, *Minimizing  $l_1$ -distance between distribution functions*, Probab. Math. Statist. 12 (1991), 265–270.
- [29] C. G. Small, *A survey of multidimensional medians*, Internat. Statist. Rev. 58 (1990), 263–277.
- [30] M. Valadier, *La multi-application médianes conditionnelles*, Z. Wahrsch. Verw. Gebiete 67 (1984), 279–282.

N. Belili  
UPRES-A 6085  
Analyse et Modèles Stochastiques  
Université de Rouen  
76821 Mont-Saint-Aignan Cedex, France  
E-mail: Nacereddine.Belili@univ-rouen.fr

H. Heinich  
UPRES-A 6085, INSA de Rouen  
Département de Génie Mathématiques  
Place E. Blondel  
76131 Mont-Saint-Aignan, France  
E-mail: heinich@insa-rouen.fr