

# Proximality and uniformly approximable sets in $L^p$

Guillaume Grelier\* & Jaime San Martín†

September 7, 2022

## Abstract

For any  $p \in [1, \infty]$ , we prove that the set of simple functions taking at most  $k$  different values is proximal in  $L^p$  for all  $k \geq 1$ . We introduce the class of uniformly approximable subsets of  $L^p$ , which is larger than the class of uniformly integrable sets. This new class is characterized in terms of the  $p$ -variation if  $p \in [1, \infty)$  and in terms of covering numbers if  $p = \infty$ . We study properties of uniformly approximable sets. In particular, we prove that the convex hull of a uniformly approximable bounded set is also uniformly approximable and that this class is stable under Hölder transformations. We also prove that, for  $p \in [1, \infty)$ , the unit ball of  $L^p$  is uniformly approximable if and only if  $L^p$  is finite-dimensional, while for  $p = \infty$  the unit ball is always uniformly approximable.

**Key words:** Proximal, Chebyshev, Variation, Covering numbers.

**MSC2010:** Primary 41A50, 26A42; Secondary 41A30, 28Axx.

## 1 Introduction

In this paper we study the approximation of measurable functions by simple functions taking at most  $k$  values, for  $k \in \mathbb{N}$ . This problem has important consequences in multiple applications, where for example, one seeks for reduction of dimensionality, among many others. For example, the embedding of metric spaces into finite-dimensional normed spaces with small dimension is one of the main issues in non-linear analysis (see [5, 10, 11]). These results have deep consequences in order to design approximation algorithms, for instance for the Sparsest Cut problem (see [3]). When we approximate a given function  $f \in L^p(\Omega, \mathcal{F}, \mu)$  by simple functions, the number of terms in those approximations grows to infinity in general. Here, a main concern is what we can say if we restrict the number of terms in the approximations. In particular, what we can say about subsets of  $L^p(\Omega, \mathcal{F}, \mu)$  that can be uniformly approximated by simple functions taking  $k$  values, as  $k$  grows to  $\infty$ . As we shall see, this new concept is more general than uniform integrability or compactness, and we fully characterize it in terms of a new measure of variation defined for functions in  $L^p(\Omega, \mathcal{F}, \mu)$  for  $p \in [1, \infty)$ , and in terms of covering numbers in the case of  $p = \infty$ .

Let us fix some notations we need to explain the main results of this paper. Consider  $(\Omega, \mathcal{F}, \mu)$  a measure space. For any  $k \geq 1$ , we denote by  $\mathcal{G}_{p,k}(\Omega, \mathcal{F}, \mu)$ , or simply  $\mathcal{G}_{p,k}$  when the measure space  $(\Omega, \mathcal{F}, \mu)$  is clear from the context, the set of simple functions given by

$$\mathcal{G}_{p,k} = \left\{ \sum_{i=1}^l a_i \mathbb{1}_{A_i} \in L^p(\Omega, \mathcal{F}, \mu) : \{A_i\}_{1 \leq i \leq l} \text{ measurable partition of } \Omega, a_i \in \mathbb{R} \text{ for all } i, l \leq k \right\}.$$

\*Departamento de Matemáticas, Universidad de Murcia, Campus de Espinardo, 30100 Espinardo, Murcia, Spain  
email: g.grelier@um.es

†CMM-DIM; Universidad de Chile; UMI-CNRS 2807; Casilla 170-3 Correo 3 Santiago; Chile.  
email: jsanmart@dim.uchile.cl

**Remark.** Assume  $\mu$  is a finite measure. Then  $\mathcal{G}_{p,k} = \mathcal{G}_{1,k}$  for all  $p \in [1, \infty]$  is just the set of simple measurable functions that takes at most  $k$  values. If  $\mu$  is an infinite measure, then  $h = \sum_{i=1}^k a_i \mathbb{1}_{A_i}$ , where  $\{A_i\}_{1 \leq i \leq k}$  is a measurable partition, belongs to  $\mathcal{G}_{p,k}$ , for  $p \in [1, \infty)$ , if and only if  $\mu(A_i) = \infty$  implies  $a_i = 0$ . So, again for all  $p \in [1, \infty)$  it holds that  $\mathcal{G}_{p,k} = \mathcal{G}_{1,k} \subset \mathcal{G}_{\infty,k}$ , and the latter is the set of all simple measurable functions that takes at most  $k$  values.

We recall some notions from approximation theory. Let  $X$  be a Banach space and let  $K$  be a closed subset of  $X$ . The *metric projection* on  $K$  is the multi-valued mapping  $P_K : X \rightrightarrows K$  defined by  $P_K(x) = \{y \in K : \|x - y\| = d(x, K)\}$  (where  $d(A, B)$  is the distance between two subsets  $A$  and  $B$  of  $X$ ). If  $P_K(x)$  is not empty for all  $x \in X$ , we say that  $K$  is *proximal*. If  $P_K(x)$  is a singleton for all  $x \in X$ , we say that  $K$  is *Chebyshev*. Section 2 is devoted to show the following result.

**Theorem 1.1.** Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space and  $p \in [1, +\infty]$ . Then  $\mathcal{G}_{p,k}$  is proximal in  $L^p(\Omega, \mathcal{F}, \mu)$  for all  $k \geq 1$ .

In other words, the distance of a function  $f$  to  $\mathcal{G}_{p,k}$  is attained at some  $g \in \mathcal{G}_{p,k}$ . Note that most of the classical results on the existence of a solution cannot be used in this case since  $\mathcal{G}_{p,k}$  is obviously not compact in the strong topology, nor convex, and as we will see, it is not closed in the weak topology, in general. The proof of this result is divided into several steps. We first deal with the case  $p \in [1, \infty)$  and we also prove that a minimum can be chosen to have a particular form (see Theorem 2.3 when  $\mu$  is finite and Theorem 2.7 if not). The proof is rather technical since we deal with any kind of measure (not only finite or  $\sigma$ -finite). In case  $\mu$  is finite we also give conditions to ensure that there is a unique minimizer (see Theorem 2.6). In general, the set of minimizers is not a singleton, it can even exist a continuum of minimizers. Then, it makes sense to study if the metric projection  $P_{\mathcal{G}_{p,k}}$  has a continuous selection. In general, there is no continuous selection, unless  $L^p(\Omega, \mathcal{F}, \mu)$  is finite dimensional (see the Remark before Section 3).

When  $p = \infty$ , in Proposition 2.9 we prove that  $\mathcal{G}_{\infty,k}$  is proximal. The proofs we provide are somehow constructive in nature, but still there is a long way to go for obtaining useful algorithms, which in itself, we think, will be important in many applications.

An important role in this section is played by  $\mathcal{M}_p(f, A)$ , the  $p$ -th mean of  $f$  on a set  $A$  (see Definition 2.1). In particular, for  $p = 2$  we have  $\mathcal{M}_2(f, A) = \frac{1}{\mu(A)} \int_A f(x) d\mu(x)$ . A well-known approximation associated to a finite measurable partition  $\mathcal{P} = \{A_i\}_{1 \leq i \leq k}$  is given by

$$\mathbb{E}^{\mathcal{P}}(f) = \sum_{i=1}^k \mathcal{M}_2(f, A_i) \mathbb{1}_{A_i},$$

which corresponds to the conditional expectation of  $f$  over the  $\sigma$ -field generated by  $\mathcal{P}$ .

In Section 3, we introduce the  $p$ -variation  $\text{Var}_{p,k}(f)$  of a function  $f \in L^p(\Omega, \mathcal{F}, \mu)$ , for  $p \in [1, \infty)$  and we studied some of its properties. The  $p$ -variation of a function allow us to control the distance of  $f$  to the sets  $\mathcal{G}_{p,k}$ , up to a factor of 2 (see Proposition 3.3). This notion will be a useful tool to characterize the uniform approximability of sets in the following section and whose definition is the following:

**Definition 1.1.** Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space and  $p \in [1, +\infty]$ . Let  $\mathcal{A} \subset L^p(\Omega, \mathcal{F}, \mu)$ . For  $\varepsilon > 0$ , we define

$$N_{p,\varepsilon}(\mathcal{A}) = \inf\{k \geq 1 : \forall f \in \mathcal{A}, \exists h \in \mathcal{G}_{p,k} \ \|f - h\|_p \leq \varepsilon\}.$$

As usual if the set where the infimum is taken is empty we set  $N_{p,\varepsilon}(\mathcal{A}) = \infty$ . We say that  $\mathcal{A}$  is *uniformly approximable* (in short UA) in  $L^p(\Omega, \mathcal{F}, \mu)$  if  $N_{p,\varepsilon}(\mathcal{A}) < \infty$  for any  $\varepsilon > 0$ .

Concretely a set  $\mathcal{A}$  is UA in  $L^p(\Omega, \mathcal{F}, \mu)$  if for any  $\varepsilon > 0$  there exists  $k \geq 1$  such that any function in  $\mathcal{A}$  can be  $\varepsilon$ -approximated in  $L^p(\Omega, \mathcal{F}, \mu)$  by simple functions taking less than  $k$  different values. Notice that  $\mathcal{A}$  is UA if and only if

$$\lim_{k \rightarrow \infty} \sup_{f \in \mathcal{A}} \inf \{ \|f - g\|_p : g \in \mathcal{G}_{p,k} \} = 0.$$

We point out that a similar quantity leads to relatively compactness of  $\mathcal{A}$ . Indeed, if  $1 \leq p < \infty$ , a result inspired by M. Riesz (see Theorem 4.7.28 in [4]) says that  $K \subset L^p(\Omega, \mathcal{F}, \mu)$  is relatively compact if and only if  $K$  is bounded in  $L^p(\Omega, \mathcal{F}, \mu)$  and

$$\inf_{\mathcal{P}} \sup_{f \in K} \|f - \mathbb{E}^{\mathcal{P}}(f)\|_p = 0.$$

We point out that for every finite measurable partition  $\mathcal{P}$ , with at most  $k$  atoms it holds

$$\inf \{ \|f - g\|_p : g \in \mathcal{G}_{p,k} \} \leq \|f - \mathbb{E}^{\mathcal{P}}(f)\|_p,$$

so relatively compactness implies UA, a fact that can be easily proved directly.

The last part of the paper, Section 4, is dedicated to the study of uniformly approximable sets. We will give some examples of UA sets and prove that it is a larger class than the class of uniformly integrable sets. We also characterize this property in terms of covering numbers if  $p = \infty$  and in terms of the  $p$ -variation if  $p < \infty$ . The covering numbers  $\mathcal{N}(f, \varepsilon)$  of a function  $f$  is simply defined as the covering number of its range, up to measure 0. We will prove the following two results:

**Theorem 1.2.** *Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space and let  $\mathcal{A} \subset L^\infty(\Omega, \mathcal{F}, \mu)$ . The following assertions are equivalent:*

- (i)  $\mathcal{A}$  is UA;
- (ii)  $\sup_{f \in \mathcal{A}} \mathcal{N}(f, \varepsilon) < \infty$  for all  $\varepsilon > 0$ .

**Theorem 1.3.** *Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space,  $p \in [1, \infty)$  and let  $\mathcal{A} \subset L^p(\Omega, \mathcal{F}, \mu)$ . Then, the following are equivalent*

- (i)  $\mathcal{A}$  is UA in  $L^p(\Omega, \mathcal{F}, \mu)$ ;
- (ii)  $\lim_{k \rightarrow \infty} \sup_{f \in \mathcal{A}} \text{Var}_{p,k}(f) = 0$ .

Then we investigate when the unit ball of  $L^p(\Omega, \mathcal{F}, \mu)$  is UA. If  $1 \leq p < \infty$ , this happens, as one can expect, if and only if  $L^p(\Omega, \mathcal{F}, \mu)$  is finite dimensional (see Theorem 4.7). We conclude this section by establishing some stability properties of the class of UA sets. In particular, a nice use of the Rademacher type allows us to prove that if  $\mathcal{A}$  is a bounded UA set in  $L^p(\Omega, \mathcal{F}, \mu)$  for  $p \in (1, \infty)$  then its closed convex hull also is UA (see Theorem 4.12). For more information about Rademacher type and cotype, we refer the reader to [1] (chapter 6).

In what follows all the measures considered are assumed non trivial, that is, different from the 0 measure, unless it is explicitly stated. We believe that our notation is quite standard. For example, the closure of a set  $A$  is denoted by  $\bar{A}$  and the distance between two subsets  $A$  and  $B$  in a metric space is denoted by  $d(A, B)$ . The complement of a set  $A$  is denoted by  $A^c$ . In some of the results we will need to consider diffuse and atomic measures. For that reason we fix some notations at this respect. We recall that an *atom* in a measure space  $(\Omega, \mathcal{F}, \mu)$  is a measurable set  $A$  that satisfies:  $\mu(A) > 0$  and if  $B \subset A$  is a measurable set such that  $\mu(B) < \mu(A)$  then  $\mu(B) = 0$ . Notice that if  $A_1, A_2$  are two atoms with finite measure, then either  $\mu(A_1 \cap A_2) = 0$  or they differ on a set of measure 0, that is,  $\mu(A_1 \Delta A_2) = 0$  (where  $\Delta$  is the symmetric difference). A measurable

space is said to be *atomic* if every measurable set of positive measure contains an atom. An atomic space is said to have a finite number of atoms of finite measure, up to measure 0, if there exists a finite collection (eventually empty)  $\mathcal{A}$  of atoms of finite measure such that for any atom  $B$  either  $\mu(B) = \infty$  or there exists  $A \in \mathcal{A}$  such that  $\mu(A\Delta B) = 0$ . A *diffuse* measure, is a measure that has no atoms. Notice that the measure  $\mu \equiv 0$  is by definition diffuse, and we refer to this case as the trivial one. More information about measure theory can be found in [4].

## 2 Minimizing the distance to the sets $\mathcal{G}_{p,k}$

The main objective of this section is to prove that  $\mathcal{G}_{p,k}$  proximal, i.e. given some  $f \in L^p(\Omega, \mathcal{F}, \mu)$ , the distance from  $f$  to  $\mathcal{G}_{p,k}$  is reached at some function  $g \in \mathcal{G}_{p,k}$  (see Theorem 1.1). We denote by

$$\mathcal{D}_{p,k}(f) = \inf\{\|f - h\|_p : h \in \mathcal{G}_{p,k}\}.$$

for all  $p \in [1, \infty]$ , that is the distance between  $f$  and  $\mathcal{G}_{p,k}$ . A function  $g \in P_{\mathcal{G}_{p,k}}$  will be called a *minimizer*. As we mentioned in the introduction, the classical results of optimization do not apply in this case since  $\mathcal{G}_{p,k}$  is not convex nor compact. Even in the reflexive case (that is  $1 < p < \infty$ ), it is not clear if the problem admits a solution. However, if  $1 < p < \infty$  and  $\mathcal{G}_{p,k}$  is weakly closed, it is easy to see that there exists a minimizer. In fact, let  $(g_n)_n \subset \mathcal{G}_{p,k}$  such that  $\|g_n - f\| \rightarrow \mathcal{D}_{p,k}(f)$ . In particular,  $(g_n)_n$  is bounded and then admits a subsequence  $(g_{n'})_{n'}$  that weakly converges to some  $g \in \mathcal{G}_{p,k}$ . Since the norm is weakly lower semicontinuous, we obtain that

$$\mathcal{D}_{p,k}(f) \leq \|f - g\|_p \leq \liminf_{n'} \|f - g_{n'}\|_p = \mathcal{D}_{p,k}(f),$$

implying that  $\mathcal{D}_{p,k}(f) = \|f - g\|_p$ . Unfortunately, as the following discussion will show,  $\mathcal{G}_{p,k}$  is not weakly closed in general, a fact that depends strongly on the measure space. On the one hand, in the case of the  $\ell_p$  spaces for  $1 \leq p < \infty$ , every  $\mathcal{G}_{p,k}$  is closed under the weak topology. This follows directly from the fact that if  $(f_n)_n \subset \mathcal{G}_{p,k}$  converges to  $f$  weakly in  $\ell_p$ , then  $(f_n)_n$  converges pointwise to  $f$ . From this fact it follows that  $f(\mathbb{N})$  is a finite set with cardinality at most  $k$  and therefore  $f \in \mathcal{G}_{p,k}$ . On the other extreme we have the following result:

**Proposition 2.1.** *Consider  $([0, 1], \mathcal{L}, dx)$  the Lebesgue measure and let  $p \in [1, \infty)$ . Then  $\mathcal{G}_{p,k}$  is weakly dense in  $L^p([0, 1], \mathcal{L}, dx)$  for all  $k \geq 2$ .*

*Proof.* It is enough to prove the case  $k = 2$ . Consider an integer  $r \geq 2$ . Every  $x \in [0, 1)$  has a unique expansion

$$x = \sum_{n=1}^{\infty} \zeta_n^r(x) r^{-n}.$$

where  $\zeta_n^r(x) \in \{0, \dots, r-1\}$  and  $(\zeta_n^r(x))_n$  is not eventually constant  $r-1$ . For  $x = 1$ , we define  $\zeta_n^r(x) = r-1$  for all  $n \geq 1$ .

Let us prove that for every  $A \in \mathcal{L}$ , the sequence  $(\mathbb{1}_{A \cap \{\zeta_n^2=1\}})_n$  converges weakly to the function  $f = \frac{1}{2} \mathbb{1}_A$ . Indeed, assume first that  $A = [0, 1]$ . For  $n \geq 1$ , let  $\Psi_n : [0, 1] \rightarrow [0, 1]$  be the bi-measurable and measure preserving transformation which flips the  $n$ -th binary digit. Then for all continuous functions  $g : [0, 1] \rightarrow \mathbb{R}$  it holds

$$\int_{\{\zeta_n^2=1\}} g(x) dx = \int_{\{\zeta_n^2=0\}} g(x) dx + R_n,$$

where  $R_n = \int_{\{\zeta_n^2=1\}} g(x) - g(\Psi_n(x)) dx$ . The continuity of  $g$ , allow us to prove that  $R_n$  converges to zero. This shows  $\mathbb{1}_{\{\zeta_n^2=1\}}$  converges weakly to  $\frac{1}{2} \mathbb{1}_{[0,1]}$ . Thus, for all  $h \in L^q$ , where  $q$  is the conjugated index of  $p$ , and all  $A \in \mathcal{L}$  we have

$$\lim_{n \rightarrow \infty} \int \mathbb{1}_{\{\zeta_n^2=1\}}(x) \mathbb{1}_A(x) h(x) dx = \frac{1}{2} \int \mathbb{1}_A(x) h(x) dx,$$

showing that  $(\mathbb{1}_{A \cap \{\zeta_n^2=1\}})_n$  converges weakly to  $\frac{1}{2}\mathbb{1}_A$ .

In a similar way, it is shown that for all  $A \in \mathcal{L}$ , any integer number  $r \geq 2$ , any  $m \in \{1, \dots, r\}$  and all  $0 \leq t_1 < t_2 \dots < t_m \leq r-1$ , the sequence

$$f_n = \mathbb{1}_{A \cap \cup_{j=1}^m \{\zeta_n^r = t_j\}} = \sum_{j=1}^m \mathbb{1}_{A \cap \{\zeta_n^r = t_j\}} \in \mathcal{G}_{p,2},$$

converges weakly to  $\frac{m}{r}\mathbb{1}_A$ .

Now, for any  $\ell \geq 1$ , any partition  $\{A_j\}_{1 \leq j \leq \ell}$  of measurable sets, any collection  $\{r_j\}_{1 \leq j \leq \ell}$  of integer numbers greater or equal than 2, any collection  $\{m_j\}_{1 \leq j \leq \ell}$  such that  $m_j \in \{1, \dots, r_j\}$  and any collection of integer numbers  $\{t_{j,i} : 1 \leq i \leq m_j, 1 \leq j \leq \ell\}$  such that  $0 \leq t_{j,1} < \dots < t_{j,m_j} \leq r_j - 1$ , we obtain that the sequence

$$f_n = \sum_{j=1}^{\ell} \mathbb{1}_{A_j \cap \cup_{i=1}^{m_j} \{\zeta_n^{r_j} = t_{j,i}\}} = \sum_{j=1}^{\ell} \sum_{i=1}^{m_j} \mathbb{1}_{A_j \cap \{\zeta_n^{r_j} = t_{j,i}\}},$$

converges weakly to  $\sum_{j=1}^{\ell} \frac{m_j}{r_j} \mathbb{1}_{A_j}$ . We notice that  $f_n = \mathbb{1}_{B_n}$ , where

$$B_n = \bigcup_{j=1}^{\ell} \bigcup_{i=1}^{m_j} A_j \cap \{\zeta_n^{r_j} = t_{j,i}\},$$

so  $f_n \in \mathcal{G}_{p,2}$ . This shows that the weak closure of  $\mathcal{G}_{p,2}$  contains all the simple functions of the form

$f = \sum_{j=1}^{\ell} \alpha_j \mathbb{1}_{A_j}$ , where  $\ell \geq 1$ ,  $\{A_j\}_{1 \leq j \leq \ell}$  is any finite measurable partition and  $\alpha_j \in [0, 1]$  for all  $j \in \{1, \dots, \ell\}$ . Moreover, any such simple function is the weak limit of a sequence  $(\mathbb{1}_{F_n})_n$  for some sequence  $(F_n)_n$  of measurable sets. From here it follows that the weak closure of  $\mathcal{G}_{p,2}$  contains all

the simple functions. Indeed, consider a simple function  $f = \sum_{j=1}^{\ell} a_j \mathbb{1}_{A_j}$ , with  $\ell \geq 1$  and  $a_j \in \mathbb{R}$  for

all  $j \in \{1, \dots, \ell\}$ . By adding a large constant  $C$ , we have  $f + C = \sum_{j=1}^{\ell} b_j \mathbb{1}_{A_j}$ , where  $b_j = a_j + C > 0$

for all  $j \in \{1, \dots, \ell\}$ . Letting  $D = \max_{1 \leq j \leq \ell} b_j$ , we deduce that  $\frac{1}{D}(f + C)$  is the weak limit of a sequence  $(\mathbb{1}_{F_n})_n$  for some sequence of measurable sets  $(F_n)_n$ . Then

$$f_n := D\mathbb{1}_{F_n} - C = (D - C)\mathbb{1}_{F_n} - C\mathbb{1}_{F_n^c} \in \mathcal{G}_{p,2},$$

converges weakly to  $f$ . The density of the simple functions in  $L^p$ , in the strong topology, shows the result.  $\square$

The previous result implies obviously that  $\mathcal{G}_{p,k}$  is not weakly closed in general, and the usual optimization methods do not work in this context, we have to find a minimizer by a more constructive way.

**Definition 2.1.** In what follows, for a measurable set  $A$  of positive and finite measure, we consider  $\mathcal{M}_p(f, A)$  as one of the  $p$ -th means of  $f$  on  $A$  where  $p \in [1, \infty)$ . The function

$$a \mapsto \int_A |f(x) - a|^p d\mu(x)$$

is convex, nonnegative and finite on  $\mathbb{R}$ , which converges to  $\infty$  as  $a \rightarrow \pm\infty$ . Therefore, this function has at least one global minimum. For  $p = 1$ , the set of minima is a bounded interval with extremes  $a^*$  and  $b^*$  and it is customary to take, the median, as

$$\mathcal{M}_1(f, A) = \frac{a^* + b^*}{2}.$$

For  $p > 1$  the minimum is unique due to strict convexity and we denote it by  $\mathcal{M}_p(f, A)$ . For example, for  $p = 2$

$$\mathcal{M}_2(f, A) = \frac{1}{\mu(A)} \int_A f(x) d\mu(x),$$

is the mean of  $f$  over the set  $A$ . If a set has measure 0, we simply put  $\mathcal{M}_p(f, A) = 0$ .

The next concept will play an important role in what follows.

**Definition 2.2.** Assume  $f \in L^p(\Omega, \mathcal{F}, \mu)$ ,  $p \in [1, \infty)$ . A function  $g \in \mathcal{G}_{p,k}$

$$g = \sum_{i=1}^q a_i \mathbb{1}_{C_i},$$

with  $1 \leq q \leq k$ , is said in *f-special form* if there exist  $-\infty \leq r_1 < \dots < r_k < r_{k+1} \leq \infty$  such that

- $C_i = f^{-1}([r_i, r_{i+1}))$  for all  $i \in \{1, \dots, q-1\}$ ,  $C_q = f^{-1}([r_q, r_{q+1}))$  and  $\{C_i\}_{1 \leq i \leq q}$  is a partition of  $\Omega$ ;
- $-\infty < a_1 < \dots < a_q < \infty$ ;
- for all  $i \in \{1, \dots, q\}$  such that  $\mu(C_i) < \infty$ , it holds  $a_i$  is a  $p$ -th mean of  $f$  on  $C_i$ .

Suppose that  $g = \sum_{i=1}^q a_i \mathbb{1}_{C_i}$  is in *f-special form*. Note that if  $\mu$  is an infinite measure there exists a unique  $1 \leq s \leq q$  such that  $a_s = 0$  and  $\mu(C_i) < \infty$  for all  $i \neq s$ . We also have that  $a_i = \mathcal{M}_p(f, C_i)$  for all  $i \in \{1, \dots, q\}$  if  $p > 1$ . Moreover notice that  $g = h \circ f$ , where  $h = \sum_{i=1}^{q-1} a_i \mathbb{1}_{[r_i, r_{i+1})} + a_q \mathbb{1}_{[r_q, r_{q+1})}$  is a Borel function and  $g$  is  $f$ -measurable, that is,  $g$  is measurable with respect  $\sigma(f) = f^{-1}(\mathcal{B})$ , where  $\mathcal{B}$  is the Borel  $\sigma$ -field in  $\mathbb{R}$ .

## 2.1 The case of a finite measure, $p \in [1, \infty)$

If the measure is finite, we start by proving that there exists an approximation sequence which is uniformly bounded:

**Lemma 2.2.** Let  $(\Omega, \mathcal{F}, \mu)$  be a finite measure space and  $p \in [1, \infty)$ . Let  $f \in L^p(\Omega, \mathcal{F}, \mu)$  and  $k \geq 1$ . Then there exists a uniformly bounded sequence  $(g_n)_n \subset \mathcal{G}_{p,k}$  such that

$$\|f - g_n\|_p \rightarrow \mathcal{D}_{p,k}(f).$$

*Proof.* Let  $(h_n)_n \in \mathcal{G}_{p,k}$  be a sequence such that  $\|f - h_n\|_p \rightarrow \mathcal{D}_{p,k}(f)$ . Assume that  $h_n = \sum_{i=1}^{m(n)} c_{i,n} \mathbb{1}_{A_{i,n}}$ , where  $(c_{i,n})_{1 \leq i \leq m(n)}$  are all different,  $\{A_{i,n}\}_{1 \leq i \leq m(n)}$  is a measurable partition with sets of positive measure and  $m(n) \leq k$ . We assume that  $m(n) = m$  is constant by passing to a subsequence if necessary. We modify this approximating sequence by considering  $a_{i,n} = \mathcal{M}_p(f, A_{i,n})$  any of the  $p$ -th means of  $f$  in  $A_{i,n}$ . By definition of the  $p$ -th means we have, for all  $i \in \{1, \dots, m\}$

$$\int_{A_{i,n}} |f(x) - a_{i,n}| d\mu(x) \leq \int_{A_{i,n}} |f(x) - c_{i,n}| d\mu(x),$$

showing that  $\tilde{h}_n = \sum_{i=1}^m a_{i,n} \mathbb{1}_{A_{i,n}} \in \mathcal{G}_k$  is a minimizing sequence since

$$\mathcal{D}_{p,k}(f) \leq \|f - \tilde{h}_n\|_p \leq \|f - h_n\| \rightarrow \mathcal{D}_{p,k}(f)$$

If  $m < k$ , we define  $a_{i,n} = 0$  and  $A_{i,n} = \emptyset$  for  $i \in \{m+1, \dots, k\}$ . We assume that  $\{A_{i,n}\}_{1 \leq i \leq k}$  are ordered in decreasing order according to their measure

$$\mu(A_{1,n}) \geq \mu(A_{2,n}) \geq \dots \geq \mu(A_{k,n}) \geq 0.$$

In this way, the vector  $v_n = (\mu(A_{1,n}), \mu(A_{2,n}), \dots, \mu(A_{k,n}))$  belongs to the compact set in  $\mathbb{R}^k$

$$\Delta = \left\{ x \in \mathbb{R}^k : x_1 \geq x_2 \geq \dots \geq x_k \geq 0, \sum_i x_i = \mu(\Omega) \right\}$$

By passing to a subsequence if necessary, we can assume that  $(v_n)_n$  converges to some vector  $v = (v_1, v_2, \dots, v_k) \in \Delta$ . If  $q$  is the largest index such that  $v_q > 0$  ( $q$  could be exactly  $k$ ) then, we have  $q \geq 1$  and  $v_1 \geq \dots \geq v_q > 0 = v_{q+1} = \dots = v_k$ . We notice that  $q \leq m$ . Now, define  $B_n = \bigcup_{i=q+1}^k A_{i,n}$  for all  $n \in \mathbb{N}$ , that we take as the empty set if  $q = k$ , so

$$\lim_{n \rightarrow \infty} \mu(B_n) = \lim_{n \rightarrow \infty} \sum_{i=q+1}^k \mu(A_{i,n}) = 0.$$

On the other hand, for all  $i \in \{1, \dots, q\}$  we have

$$\lim_{n \rightarrow \infty} \mu(A_{i,n}) = v_i > 0,$$

and so, passing to a further subsequence we can assume there exists a finite constant  $\Gamma$  such that for all  $n$  and all  $i \in \{1, \dots, q\}$  it holds

$$\frac{1}{\mu(A_{i,n})} \leq \Gamma \tag{2.1}$$

The finite measure  $\nu$  defined by

$$\nu(A) = \int_A |f(x)|^p d\mu(x),$$

is absolutely continuous with respect to  $\mu$ , which means that, for all  $\rho > 0$  there exists a  $\delta > 0$  such that, for any measurable set  $A$  if  $\mu(A) \leq \delta$  then  $\nu(A) = \int_A |f(x)|^p d\mu(x) \leq \rho$ . This property shows that

$$\lim_{n \rightarrow \infty} \int_{B_n} |f(x)|^p d\mu(x) = 0.$$

Now, we modify further the approximation sequence by defining

$$b_{i,n} = \begin{cases} a_{i,n} & \text{for } i \in \{1, \dots, q\} \\ 0 & \text{for } i \in \{q+1, \dots, k\} \end{cases}, \tag{2.2}$$

and define

$$g_n = \sum_{i=1}^k b_{i,n} \mathbb{1}_{A_{i,n}} = \sum_{i=1}^q \mathcal{M}_p(f, A_{i,n}) \mathbb{1}_{A_{i,n}} + 0 \mathbb{1}_{B_n} \in \mathcal{G}_{p,k}. \tag{2.3}$$

We need to show that  $(g_n)_n$  is a good approximation sequence and it is uniformly bounded. For the first claim notice that for  $i \in \{q+1, \dots, k\}$ , we have

$$\int_{A_{i,n}} |f(x) - \tilde{h}_n(x)|^p d\mu(x) = \int_{A_{i,n}} |f(x) - \mathcal{M}_p(f, A_{i,n})|^p d\mu(x) \leq \int_{A_{i,n}} |f(x)|^p d\mu(x),$$

where we have used the optimality of  $\mathcal{M}_p(f, A_{i,n})$  in the last inequality. This shows that

$$\begin{aligned} \mathcal{D}_{p,k}(f)^p &\leq \|f - \tilde{h}_n\|_p^p = \sum_{i=1}^m \int_{A_{i,n}} |f(x) - \mathcal{M}_p(f, A_{i,n})|^p d\mu(x) \\ &\leq \sum_{i=1}^q \int_{A_{i,n}} |f(x) - \mathcal{M}_p(f, A_{i,n})|^p d\mu(x) + \int_{B_n} |f(x)|^p d\mu(x) = \|f - g_n\|_p^p \\ &\leq \sum_{i=1}^m \int_{A_{i,n}} |f(x) - \mathcal{M}_p(f, A_{i,n})|^p d\mu(x) + \int_{B_n} |f(x)|^p d\mu(x) \\ &\leq \|f - \tilde{h}_n\|_p^p + \int_{B_n} |f(x)|^p d\mu(x) \rightarrow \mathcal{D}_{p,k}(f)^p \end{aligned}$$

Now, we prove that  $(g_n)_n$  is uniformly bounded. We notice that  $g_n = 0$  on  $B_n$ , so we must study  $g_n$  on  $B_n^c$ . For  $i \in \{1, \dots, q\}$  and  $x \in A_{i,n}$  we have  $g_n(x) = \mathcal{M}_p(f, A_{i,n})$  and so

$$\|\mathcal{M}_p(f, A_{i,n}) \mathbb{1}_{A_{i,n}}\|_p \leq \|(f - \mathcal{M}_p(f, A_{i,n})) \mathbb{1}_{A_{i,n}}\|_p + \|f \mathbb{1}_{A_{i,n}}\|_p \leq 2\|f \mathbb{1}_{A_{i,n}}\|_p \leq 2\|f\|_p,$$

where we have used again the optimality of  $\mathcal{M}_p(f, A_{i,n})$ . This shows that

$$|\mathcal{M}_p(f, A_{i,n})| \leq 2 \frac{\|f\|_p}{\mu(A_{i,n})^{\frac{1}{p}}} \leq 2\|f\|_p \Gamma^{\frac{1}{p}},$$

where  $\Gamma$  is the constant obtained in (2.1). □

The next result proves that  $\mathcal{G}_{p,k}$  is proximal in case of finite measure spaces. Remember that  $P_K$  is the metric projection over  $K$ .

**Theorem 2.3.** *Let  $(\Omega, \mathcal{F}, \mu)$  be a finite measure space,  $p \in [1, \infty)$  and  $k \geq 1$ . Then  $\mathcal{G}_{p,k}$  is proximal.*

Moreover, if  $f \in L^p(\Omega, \mathcal{F}, \mu)$  and  $g = \sum_{i=1}^q b_i \mathbb{1}_{A_i} \in P_{\mathcal{G}_{p,k}}(f)$  is a minimizer with  $q \leq k$ ,  $-\infty < b_1 < \dots < b_q < \infty$  and  $\{A_i\}_{1 \leq i \leq q}$  a partition of  $\Omega$  with sets of positive measure, there exists a minimizer  $\tilde{g} \in P_{\mathcal{G}_{p,q}}(f)$  in  $f$ -special form:

$$\tilde{g} = \sum_{i=1}^q \mathcal{M}_p(f, f^{-1}(C_i)) \mathbb{1}_{f^{-1}(C_i)}$$

where

- $r_1 = -\infty, r_{q+1} = \infty$  and  $r_i = \frac{b_{i-1} + b_i}{2}$  for all  $i \in \{2, \dots, q\}$ ;
- $C_i = f^{-1}([r_i, r_{i+1}))$  for all  $i \in \{1, \dots, q-1\}$  and  $C_q = f^{-1}([r_q, r_{q+1}])$ ;
- $b_i$  is a  $p$ -th mean of  $f$  on  $f^{-1}(C_i)$  for all  $i \in \{1, \dots, q\}$  such that  $\mu(f^{-1}(C_i)) > 0$ .

If  $q$  is the smallest among all minimizers, then  $\mu(C_i) > 0$  for all  $i \in \{1, \dots, q\}$ .

*Proof.* By Lemma 2.2, let  $(g_n)_n \subset \mathcal{G}_{p,k}$  be a uniformly bounded sequence such that  $\|f - g_n\|_p \rightarrow \mathcal{D}_{p,k}(f)$ . Let  $C > 0$  such that  $|g_n| < C$  for all  $n \in \mathbb{N}$ . We write  $g_n = \sum_{i=1}^k b_{i,n} \mathbb{1}_{A_{i,n}}$  where  $\{A_{i,n}\}_{1 \leq i \leq k}$  is a partition of  $\Omega$  and  $-C \leq b_{1,n} \leq \dots \leq b_{k,n} \leq C$ . The vector  $u_n = (b_{1,n}, \dots, b_{k,n})$  belongs to the compact set  $[-C, C]^k$  and therefore, by taking a subsequence if necessary, we can assume that  $(u_n)_n$  converges to some  $u = (b_1, \dots, b_k) \in [-C, C]^k$  with  $b_1 \leq \dots \leq b_k$ . Some of the entries in  $u$  can be equal, for that we consider  $z_1 < \dots < z_l$  the distinct entries in  $u$  where  $1 \leq l \leq k$ . We define  $r_1 = -\infty, r_{l+1} = \infty$  and  $r_j = \frac{z_{j-1} + z_j}{2}$  for  $j \in \{2, \dots, l\}$ . Consider the intervals  $I_j = [r_j, r_{j+1})$  for



$j \in \{1, \dots, l-1\}$  and  $I_l = [r_l, r_{l+1}]$ . For  $j \in \{1, \dots, l\}$ , we also define  $L_j = \{i \in \{1, \dots, k\} : b_i = z_j\}$ , which is a partition of  $\{1, \dots, k\}$ . For all  $n \in \mathbb{N}$ , consider the function

$$\tilde{g}_n = \sum_{i=1}^k b_i \mathbb{1}_{A_{i,n}}.$$

Then, we have

$$\|f - \tilde{g}_n\|_p \leq \|f - g_n\|_p + \|g_n - \tilde{g}_n\|_p \leq \|f - g_n\|_p + \max_{1 \leq i \leq k} |b_{i,n} - b_i| \mu(\Omega)^{\frac{1}{p}} \rightarrow \mathcal{D}_{p,k}(f),$$

proving that  $(\tilde{g}_n)_n$  is also a minimizing sequence. Finally, our candidate for minimizer is the function  $g = \sum_{j=1}^l z_j \mathbb{1}_{f^{-1}(I_j)} \in \mathcal{G}_{p,k}$ . For all  $i \in \{1, \dots, k\}$ , all  $j \in \{1, \dots, l\}$  and all  $n$ , we have

$$\int_{f^{-1}(I_j) \cap A_{i,n}} |f(x) - z_j|^p d\mu(x) \leq \int_{f^{-1}(I_j) \cap A_{i,n}} |f(x) - b_i|^p d\mu(x).$$

This is clear if  $i \in L_j$  because in that case  $z_j = b_i$ . Now, if  $i \in L_{j'}$  with  $j' \neq j$ , we have  $b_i = z_{j'}$  and for all  $x \in f^{-1}(I_j)$  it holds  $|f(x) - z_j| \leq |f(x) - z_{j'}| = |f(x) - b_i|$ . Now, summing over  $i, j$  we get for all  $n$  that

$$\begin{aligned} \mathcal{D}_{p,k}(f)^p &\leq \|f - g\|_p^p = \sum_{i,j} \int_{f^{-1}(I_j) \cap A_{i,n}} |f(x) - z_j|^p d\mu(x) \leq \sum_{i,j} \int_{f^{-1}(I_j) \cap A_{i,n}} |f(x) - b_i|^p d\mu(x) \\ &\leq \|f - \tilde{g}_n\|_p^p \rightarrow \mathcal{D}_{p,k}(f)^p, \end{aligned}$$

proving that  $g \in P_{\mathcal{G}_{p,k}}(f)$ .

Now, we prove the last part of the Theorem. Assume that  $g = \sum_{i=1}^q b_i \mathbb{1}_{A_i} \in P_{\mathcal{G}_{p,k}}(f)$  is a minimizer, with  $b_1 < \dots < b_q$ ,  $\{A_i\}_{1 \leq i \leq q}$  a partition of  $\Omega$  where all the sets  $A_i$  have positive measure and  $q \leq k$ . Let  $r_1 = -\infty, r_{q+1} = \infty, r_i = \frac{b_{i-1} + b_i}{2}$  for  $i = \{2, \dots, q\}$  and

$$C_i = f^{-1}([r_i, r_{i+1})) \text{ for } i \in \{1, \dots, q-1\}, C_q = f^{-1}([r_q, r_{q+1})).$$

For all  $i \in \{1, \dots, q\}$ , we modify the sets  $A_i$  as

$$\tilde{A}_i = (A_i \cup f^{-1}(\{r_i\})) \setminus f^{-1}(\{r_{i+1}\}). \quad (2.4)$$

Let us prove that  $\mu(\tilde{A}_1 \Delta C_1) = \mu(A_j \cap E_2) = 0$  for all  $j > 2$  where  $E_2 = f^{-1}(\{r_2\})$ . Define

$$g' = b_1 \mathbb{1}_{\tilde{A}_1} + \sum_{i=3}^q b_i \mathbb{1}_{A_i \setminus E_2} + b_2 \mathbb{1}_{A_2 \cup E_2} \in \mathcal{G}_{p,k}$$

and note that  $\{\tilde{A}_1, \{A_j \setminus E_2\}_{j>2}, A_2 \cup E_2\}$  is a partition of  $\Omega$ . Consider the following decomposition

$$\begin{aligned} \|f - g\|_p^p &= \sum_j \int_{A_j} |f(x) - b_j|^p d\mu(x) = \int_{\tilde{A}_1} |f(x) - b_1|^p d\mu(x) + \sum_{j>2} \int_{A_j \setminus E_2} |f(x) - b_j|^p d\mu(x) \\ &+ \sum_{j>2} \int_{A_j \cap E_2} |f(x) - b_j|^p d\mu(x) + \int_{A_1 \cap E_2} |f(x) - b_1|^p d\mu(x) + \int_{A_2} |f(x) - b_2|^p d\mu(x) \\ &\geq \int_{\tilde{A}_1} |f(x) - b_1|^p d\mu(x) + \sum_{j>2} \int_{A_j \setminus E_2} |f(x) - b_j|^p d\mu(x) + \int_{A_2 \cup E_2} |f(x) - b_2|^p d\mu(x) \\ &+ (b_3 - b_2)^p \sum_{j>2} \mu(A_j \cap E_2) \\ &= \|f - g'\|_p^p + (b_3 - b_2)^p \sum_{j>2} \mu(A_j \cap E_2), \end{aligned}$$

where the second equality follows from the fact that  $\tilde{A}_1 = A_1 \setminus E_2$  (up to a set of measure zero). The inequality is proved noting that, for  $x \in A_j \cap E_2$  with  $j > 2$  it holds  $|f(x) - b_2| = b_2 - r_2 <$

$b_3 - r_2 \leq b_j - r_2 = |f(x) - b_j|$ , which implies  $|f(x) - b_j| \geq |f(x) - b_2| + b_3 - b_2$ , and for  $x \in A_1 \cap E_2$  it holds  $|f(x) - b_1| = |f(x) - b_2|$ . So, since  $g$  is a minimizer we deduce that  $\mu(A_j \cap E_2) = 0$  for all  $j > 2$ . Thus, we get

$$\|f - g\|_p^p = \int_{\tilde{A}_1} |f(x) - b_1|^p d\mu(x) + \sum_{j>2} \int_{A_j} |f(x) - b_j|^p d\mu(x) + \int_{A_2 \cup f^{-1}(\{r_2\})} |f(x) - b_2|^p d\mu(x),$$

showing that

$$b_1 \mathbb{1}_{\tilde{A}_1} + b_2 \mathbb{1}_{A_2 \cup f^{-1}(\{r_2\})} + \sum_{j>2} b_j \mathbb{1}_{A_j} \in P_{\mathcal{G}_{p,k}}(f).$$

A similar argument shows that  $\mu(A_j \cap C_1) = 0$  and  $\mu(\tilde{A}_1 \cap C_j) = 0$  for all  $j \geq 2$ . Since  $\{A_i\}_{1 \leq i \leq q}$  is a partition we conclude that  $\mu(C_1) = \sum_i \mu(A_i \cap C_1) = \mu(C_1 \cap A_1) = \mu(C_1 \cap \tilde{A}_1)$ , proving that  $C_1 \subset \tilde{A}_1$  except for a set of measure 0. On the other hand, using again that  $\{\tilde{A}_1, A_2 \cup f^{-1}(\{r_2\}), \{A_j \setminus f^{-1}(\{r_2\})\}_{j>2}\}$  is also a partition, we conclude that  $\tilde{A}_1 \subset C_1$  except for a set of measure 0. In a similar way, we prove  $\mu(\tilde{A}_i \cap C_j) = \mu(\tilde{A}_i \Delta C_i) = 0$ , for all  $i \neq j$ .

At this point we should mention that some of the  $\tilde{A}_i$  could have measure 0. For example this occurs if  $A_1 = f^{-1}(\{r_2\})$ . In any case, we have

$$\|f - g\|_p^p = \sum_{i=1}^q \int_{\tilde{A}_i} |f(x) - b_i|^p d\mu(x) = \sum_{i=1}^q \int_{C_i} |f(x) - b_i|^p d\mu(x),$$

showing that

$$\hat{g} = \sum_{i=1}^q b_i \mathbb{1}_{C_i} \in P_{\mathcal{G}_{p,k}}(f),$$

is a minimizer. On the other hand, if  $\mu(C_i) > 0$  we have  $\int_{C_i} |f(x) - \mathcal{M}_p(f, C_i)|^p d\mu(x) \leq \int_{C_i} |f(x) - b_i|^p d\mu(x)$ . The inequality cannot be strict, otherwise we contradict the minimality of  $\hat{g}$ , showing that  $b_i$  is a  $p$ -th means of  $f$  on  $C_i$ , and therefore,

$$\tilde{g} = \sum_{i=1}^q \mathcal{M}_p(f, C_i) \mathbb{1}_{C_i} \in P_{\mathcal{G}_{p,k}}(f),$$

is a minimizer in  $f$ -special form, as we wanted to prove. In case that  $q$  is the minimal among all minimizers, we conclude that  $\mu(C_i) > 0$  for all  $i$ .  $\square$

**Remark.** *In the last part of the Theorem, for any minimizer  $g$ , we have constructed a minimizer  $\tilde{g}$  in  $f$ -special form, but it may happens that some of the sets  $(C_i)_i$  have measure 0, which can be discarded to get a minimizer with fewer terms. An interesting question is if this procedure applied to any minimizer gives always a minimizer with the smallest possible number of terms (see Proposition 2.12).*

Recall that given  $f \in L^p(\Omega, \mathcal{F}, \mu)$ , the distribution of  $f$  is the measure  $\mu_f$  defined on  $(\mathbb{R}, \mathcal{B})$  given by, for all  $B \in \mathcal{B}$

$$\mu_f(B) = \mu(f^{-1}(B)).$$

Let  $g$  be a minimizer of  $f$  in  $\mathcal{G}_{p,k}$  in  $f$ -special form provided by Theorem 2.3

$$g = \sum_{i=1}^q a_i \mathbb{1}_{f^{-1}([r_i, r_{i+1}))},$$

So,  $g = \ell \circ f$  with

$$\ell = \sum_{i=1}^q a_i \mathbb{1}_{[r_i, r_{i+1})},$$

and

$$\begin{aligned} \|f - g\|_p^p &= \int_{\Omega} |f(x) - g(x)|^p d\mu(x) = \int_{\Omega} |f(x) - \ell(f(x))|^p d\mu(x) = \int_{\mathbb{R}} |y - \ell(y)|^p d\mu_f(y) \\ &= \|\text{id} - \ell\|_{L^p(\mathbb{R}, \mathcal{B}, \mu_f)}^p. \end{aligned}$$

Thus, the problem of finding a minimizer for  $f$  is equivalent to find a minimizer for the identity function  $\text{id}$  in  $\mathcal{G}_{p,k}(\mathbb{R}, \mathcal{B}, \mu_f)$ . The following result shows that when  $\mu_f$  is continuous, this search can be done over the subclass of simple functions in  $f$ -special form. Before stating the result, let us fix some notations. The cumulative distribution associated to  $\mu_f$  is the function  $F_f(x) = \mu_f((-\infty, x])$ . Notice that  $F_f(-\infty) = 0$  and  $F_f(\infty) = \mu(\Omega)$ . The convex support of  $\mu_f$  is the interval  $[\mathbf{a}_f, \mathbf{b}_f]$ , where

$$\mathbf{a}_f = \sup\{z : F_f(z) = 0\}, \mathbf{b}_f = \inf\{z : F_f(z) = F_f(\infty)\}.$$

The following lemma is needed to study the uniqueness of minimizers, where  $p$ -th means are characterized as roots of certain equations, suitable for our purposes. We include a proof, inspired by exercise 1.4.23 in [13], for the sake of completeness.

**Lemma 2.4.** *Let  $(\Omega, \mathcal{F}, \mu)$  be a finite measure space and  $f \in L^p(\Omega, \mathcal{F}, \mu)$ , for  $p \in [1, \infty)$ . For  $p = 1$ , we also assume that  $F_f$  is continuous and strictly increasing on  $[\mathbf{a}_f, \mathbf{b}_f]$ . Let  $I \subset \mathbb{R}$  be an interval with extremities  $c, d \in \overline{\mathbb{R}}$  such that  $\mu_f(I) > 0$ . Then, the  $p$ -th mean  $m = \mathcal{M}_p(f, f^{-1}(I), \mu) = \mathcal{M}_p(\text{id}, I, \mu_f)$  is characterized as the unique solution of the equation*

$$\int_{I \cap (-\infty, m]} (m - x)^{p-1} d\mu_f(x) = \int_{I \cap (m, \infty)} (x - m)^{p-1} d\mu_f(x), \quad (2.5)$$

which for  $p = 1$  is equivalent to

$$F_f(m) - F_f(c) = \frac{1}{2}(F_f(d) - F_f(c)). \quad (2.6)$$

*Proof.* The proof is based on the following equality for all  $x, b \in \mathbb{R}$

$$|x - b|^p = p \int_{-\infty}^b (t - x)^{p-1} \mathbb{1}_{\{x \leq t\}} dt + p \int_b^{\infty} (x - t)^{p-1} \mathbb{1}_{\{t < x\}} dt,$$

which implies that

$$|x - b|^p - |x - a|^p = p \int_a^b ((t - x)^{p-1} \mathbb{1}_{\{x \leq t\}} - (x - t)^{p-1} \mathbb{1}_{\{t < x\}}) dt.$$

Fix  $m \in \mathbb{R}$ . Define a function  $L : \mathbb{R} \rightarrow \mathbb{R}$  by

$$L(b) = \int_I |x - b|^p d\mu_f(x) - \int_I |x - m|^p d\mu_f(x).$$

It is clear that  $m$  is a minimum of  $L$  if and only if  $m$  is a  $p$ -th mean. Using Fubini's Theorem and the previous equality, we obtain that for all  $b \in \mathbb{R}$

$$L(b) = p \int_m^b \int_{I \cap (-\infty, t]} (t - x)^{p-1} d\mu_f(x) dt - p \int_m^b \int_{I \cap (t, \infty)} (x - t)^{p-1} d\mu_f(x) dt$$

Note that  $L$  is convex, coercive and continuous and then reaches a minimum.

Suppose  $p > 1$ . The functions  $t \mapsto \int_{I \cap [-\infty, t]} (t-x)^{p-1} d\mu_f(x)$  and  $t \mapsto \int_{I \cap (t, \infty)} (x-t)^{p-1} d\mu_f(x)$  are continuous, and therefore  $L$  is strictly convex and continuously differentiable, which proves that  $L'(b) = 0$  is the equation for the unique minima, that is,

$$L'(b) = p \int_{I \cap (-\infty, b]} (b-x)^{p-1} d\mu_f(x) - p \int_{I \cap (b, \infty)} (x-b)^{p-1} d\mu_f(x) = 0.$$

It follows that  $m$  is the  $p$ -th mean if and only if  $m$  fulfills (2.5).

For  $p = 1$ , using that  $F_f$  is continuous, we have

$$L(b) = \int_m^b F_f(t) - F_f(c) - (F_f(d) - F_f(t)) dt.$$

Again, since  $F_f$  is continuous we obtain that  $L$  is continuously differentiable. Then, if  $b$  is any minima for  $L$ , it holds that  $L'(b) = 0$ , that is,  $F_f(b) - F_f(c) = \frac{1}{2}(F_f(d) - F_f(c))$ . Since  $F_f$  is assumed to be strictly increasing, this equation has a unique solution, and then  $L$  has exactly one minimum. Then,  $m$  is a 1-th if and only if  $F_f(m) - F_f(c) = \frac{1}{2}(F_f(d) - F_f(c))$ .  $\square$

Notice that in the previous Lemma we can replace  $(-\infty, m]$  by  $(-\infty, m)$  and  $(m, \infty)$  by  $[m, \infty)$  in 2.5, because  $x = m$  does not add to the integrals. In the case  $F_f$  is just increasing,  $I = (c, d]$  and  $p = 1$ , all the 1-th means satisfy the equations  $L'^+(m) \geq 0$  and  $L'^-(m) \leq 0$ , which are equivalent to

$$F_f(m) - F_f(c) \geq \frac{1}{2}(F_f(d) - F_f(c)), \quad F_f(m-) - F_f(c) \leq \frac{1}{2}(F_f(d) - F_f(c)),$$

and the solution set is, in general, an interval.

The next result shows that when  $F_f$  is continuous all minimizers are in  $f$ -special form.

**Corollary 2.5.** *Let  $(\Omega, \mathcal{F}, \mu)$  be a finite measure space,  $p \in [1, \infty)$  and  $k \geq 1$ . Let  $f \in L^p(\Omega, \mathcal{F}, \mu)$  and assume that  $F_f$  is continuous. Then any minimizer  $g \in \mathcal{P}_{\mathcal{G}_{p,k}}(f)$  is of the  $f$ -special form*

$$g = \sum_{i=1}^k a_i \mathbb{1}_{f^{-1}([r_i, r_{i+1}))}, \quad (2.7)$$

where

- $[r_i, r_{i+1})$  has positive  $\mu_f$ -measure for all  $i \in \{1, \dots, k\}$ ;
- $\mathbf{a}_f = r_1 < \dots < r_k < r_{k+1} = \mathbf{b}_f$  and  $r_i = \frac{a_i + a_{i+1}}{2}$  for all  $i \in \{2, \dots, k\}$ ;
- $a_i$  is a  $p$ -th mean of  $id$  on  $[r_i, r_{i+1})$  under  $\mu_f$  for all  $i \in \{1, \dots, k\}$ . Moreover, if  $F_f$  is strictly increasing on  $[\mathbf{a}_f, \mathbf{b}_f]$ , then  $a_i = \mathcal{M}_p(id, [r_i, r_{i+1}), \mu_f)$  for all  $i \in \{1, \dots, k\}$ .

*Proof.* Notice first that  $f \notin \mathcal{G}_{p,k}$ , because the image of  $f$  cannot be a finite set a.e., since  $\mu_f$  is not atomic. This implies that there is no minimizer in  $\mathcal{G}_{p,q}$ , with  $q < k$  (see Proposition 2.12 below). So, any minimizer has the structure

$$g = \sum_{i=1}^k a_i \mathbb{1}_{A_i}$$

where  $a_1 < \dots < a_k$ ,  $\{A_i\}_{1 \leq i \leq k}$  is a partition with sets of positive measure and  $a_i$  is a  $p$ -th mean of  $f$  in  $A_i$ , for all  $i \in \{1, \dots, k\}$ . In the previous proof, we then modify this minimizer to get one in  $f$ -special form. If one goes over that proof and using the fact that  $F_f$  is continuous, one realizes

that in equation (2.4), we get  $\tilde{A}_i = A_i$  a.e. and then  $A_i = C_i$  a.e., proving that  $A_i = f^{-1}([r_i, r_{i+1}))$  a.e.

The fact that  $a_i$  is a  $p$ -th mean is just the fact that  $g$  is a minimizer. For  $p > 1$ , the uniqueness of the  $p$ -th mean shows that  $a_i = \mathcal{M}_p(\text{id}, [r_i, r_{i+1}), \mu_f)$ . This is also true for  $p = 1$ , when  $F_f$  is continuous and strictly increasing in  $[\mathbf{a}_f, \mathbf{b}_f]$  (see Lemma 2.4).  $\square$

**Remark.** *The previous result could be used as the basis of an algorithm to approximate a minimizer. Assume that  $\mu_f$  is a continuous distribution. For any  $s \in \mathbb{R}$ , were  $s$  plays the role of  $r_2$  in the representation (2.7), we define  $r_1(s) = -\infty, r_2(s) = s$  and  $a_1 = a_1(s) = \mathcal{M}_p(\text{id}, (-\infty, s), \mu_f)$ . Then, we define  $a_2(s) = 2r_2(s) - a_1(s)$ , which is a relation that should satisfy any minimizer. Then, compute  $r_3(s)$  so that*

$$a_2(s) = \mathcal{M}_p(\text{id}, [r_2(s), r_3(s)), \mu_f).$$

and continue in this way defining  $a_3(s), r_4(s), \dots, a_k(s), r_{k+1}(s)$ . It may happens that at some iteration  $r_{i+1}(s)$  is not well defined for some  $i \leq k-1$  because, for all  $t \in [r_i(s), \infty]$

$$a_i(s) > \mathcal{M}_p(\text{id}, [r_i(s), t), \mu_f),$$

which shows that there is no minimizer starting with  $r_2 = s$ . So, we say  $s$  is admissible if  $r_{k+1}(s)$  is well defined. For every admissible  $s$  we have a candidate

$$\ell_s = \sum_{i=1}^{k-1} a_i(s) \mathbb{1}_{[r_i(s), r_{i+1}(s))} + \mathcal{M}_p(\text{id}, [r_k(s), \infty), \mu_f) \mathbb{1}_{[r_k(s), \infty)}$$

and we can compute  $R(s) = \int |x - \ell_s(x)|^p d\mu_f(x)$ . For  $s$  which is not admissible put  $R(s) = \infty$ . Then a minimizer of  $R$  gives a minimizer for  $f$ . One expects that the set of admissible values of  $s$  is an interval. We shall work on this algorithm in a forcoming paper.

For example, if  $\mu_f$  is a normal  $\mathcal{N}(0, 1)$ ,  $p = 2$  and  $k = 3$ , this algorithm gives the following approximation

$$h \approx -1.2 \mathbb{1}_{(-\infty, -0.6)} + 0 \mathbb{1}_{[-0.6, 0.6)} + 1.2 \mathbb{1}_{[0.6, \infty)}.$$

Notice that  $\|id\|_2 = 1$  and  $(\mathcal{D}_{2,3}(id))^2 \approx 0.18$ , which means that, in the language of statistics, 82% of the variance of  $f$  is explained by a simple function taking 3 values.

Uniqueness of minimizers is a much harder problem. Here, we present a partial result in case  $\mu_f$  satisfies a certain monotone likelihood ratio property.

**Theorem 2.6.** *Let  $(\Omega, \mathcal{F}, \mu)$  be a finite measure space,  $p \in [1, \infty)$  and  $f \in L^p(\Omega, \mathcal{F}, \mu)$ . Assume  $\mu_f$  has a density with respect to the Lebesgue measure  $\Psi : (\mathbf{a}_f, \mathbf{b}_f) \rightarrow (0, \infty)$ , which we extend by 0 outside this interval. Consider for  $s \in (0, \mathbf{b}_f - \mathbf{a}_f)$  the function  $G_s : (\mathbf{a}_f, \mathbf{b}_f) \rightarrow [0, \infty)$  given by  $G_s(y) = \frac{\Psi(y+s)}{\Psi(y)}$  and assume that  $G_s$  is decreasing. Moreover, we suppose that one of the following hypotheses hold:*

- (H1) either  $\mathbf{a}_f$  or  $\mathbf{b}_f$  is finite;
- (H2)  $G_s((\mathbf{a}_f, \mathbf{b}_f))$  is an infinite set;
- (H3)  $\Psi$  is continuous.

Then there exists a unique minimizer for  $f$  in  $\mathcal{G}_{p,k}$  for all  $k \geq 1$ .

*Proof.* The case  $k = 1$  is direct from the fact that  $F_f$  is strictly increasing (because  $\Psi$  is strictly positive on  $(\mathbf{a}_f, \mathbf{b}_f)$ ) and therefore the  $p$ -th means are unique. So, we assume that  $k \geq 2$ .

According to Corollary 2.5, and since  $F_f$  is strictly increasing all minimizers for  $f$  have the  $f$ -special form given in (2.7). Fix one of them  $g = \ell \circ f$ , where

$$\ell = \sum_{i=1}^k a_i \mathbb{1}_{I_i}$$

with  $r_1 = \mathbf{a}_f, r_{k+1} = \mathbf{b}_f, r_i = \frac{a_{i-1} + a_i}{2}$  for  $i \in \{2, \dots, k\}$ ,  $I_i = [r_i, r_{i+1})$  for  $i \in \{1, \dots, k-1\}$ ,  $I_k = [r_k, r_{k+1}]$ ,  $\mu_f(I_i) > 0$  for  $i \in \{1, \dots, k\}$  and  $a_i = \mathcal{M}_p(\text{id}, I_i, \mu_f)$  for  $i \in \{1, \dots, k\}$ . Assume there exists another minimizer  $\tilde{g} = \tilde{\ell} \circ f$  with

$$\tilde{\ell} = \sum_{i=1}^k \tilde{a}_i \mathbb{1}_{\tilde{I}_i}$$

where  $\tilde{r}_1 = \mathbf{a}_f, \tilde{r}_{k+1} = \mathbf{b}_f, \tilde{r}_i = \frac{\tilde{a}_{i-1} + \tilde{a}_i}{2}$  for  $i \in \{2, \dots, k\}$ ,  $\tilde{I}_i = [\tilde{r}_i, \tilde{r}_{i+1})$  for  $i \in \{1, \dots, k-1\}$ ,  $\tilde{I}_k = [\tilde{r}_k, \tilde{r}_{k+1}]$ ,  $\mu_f(\tilde{I}_i) > 0$  for  $i \in \{1, \dots, k\}$  and  $\tilde{a}_i = \mathcal{M}_p(\text{id}, \tilde{I}_i, \mu_f)$  for  $i \in \{1, \dots, k\}$ . We need to prove that  $\ell = \tilde{\ell}$ . Consider  $s = \tilde{r}_2 - r_2$ . Switching  $g$  and  $\tilde{g}$  if necessary, we can suppose without loss of generality that  $s \geq 0$ , and since  $r_2, \tilde{r}_2 \in (\mathbf{a}_f, \mathbf{b}_f)$  then  $s < \mathbf{b}_f - \mathbf{a}_f$ . Define  $\delta_i = \tilde{a}_i - a_i$  for all  $i \in \{1, \dots, k\}$  and  $\eta_i = \tilde{r}_i - r_i$  for all  $i \in \{2, \dots, k\}$ .

Case 1: Suppose that  $s = 0$ . We shall prove that  $\ell = \tilde{\ell}$ . Notice that  $I_1 = \tilde{I}_1$  and  $\tilde{a}_1 = a_1 = \mathcal{M}_p(\text{id}, I_1, \mu_f)$  and by uniqueness of the  $p$ -th mean we deduce that  $a_1 = \tilde{a}_1$ . But  $\tilde{a}_1, \tilde{a}_2$  and  $\tilde{r}_2$  are related by  $\tilde{a}_2 = 2\tilde{r}_2 - \tilde{a}_1 = 2r_2 - a_1 = a_2$ , showing that  $\tilde{a}_2 = a_2$ . Using the fact that  $a_2 = \mathcal{M}_p(\text{id}, I_2, \mu_f) = \mathcal{M}_p(\text{id}, \tilde{I}_2, \mu_f)$  and Lemma 2.4, we have that

$$\begin{aligned} \int_{a_2}^{\tilde{r}_3} (x - a_2)^{p-1} \Psi(x) dx &= \int_{\tilde{a}_2}^{\tilde{r}_3} (x - a_2)^{p-1} \Psi(x) dx = \int_{\tilde{r}_2}^{\tilde{a}_2} (a_2 - x)^{p-1} \Psi(x) dx \\ &= \int_{r_2}^{a_2} (a_2 - x)^{p-1} \Psi(x) dx = \int_{a_2}^{r_3} (x - a_2)^{p-1} \Psi(x) dx. \end{aligned}$$

Since  $\tilde{r}_3 \leq \mathbf{b}_f$  and since  $\Psi$  is strictly positive, we conclude that  $\tilde{r}_3 = r_3$ . Repeating this argument we conclude that  $\tilde{r}_i = r_i$  for all  $i \in \{2, \dots, k\}$  and  $\tilde{a}_i = a_i$  for all  $i \in \{1, \dots, k\}$ . Thus  $\tilde{\ell} = \ell$ .

Case 2: Suppose that  $s > 0$  and let us arrive to a contradiction if we suppose that (H1), (H2) or (H3) holds. This part will be divided in several steps.

Step 1: We are going to show that the following properties hold:

- (a)  $\delta_k \geq \eta_k \geq \delta_{k-1} \geq \eta_{k-1} \geq \dots \geq \eta_2 \geq \delta_1$ ;
- (b) if one of these inequalities is strict then all the inequalities on the left are also strict;
- (c) all of these inequalities are in fact equalities if and only if  $\mathbf{a}_f = -\infty$  and for all  $i \in \{1, \dots, k-1\}$  it holds

$$\forall y \in (r_i, r_{i+1}) \quad \frac{\Psi(y+s)}{\Psi(y)} = \frac{\Psi(a_i+s)}{\Psi(a_i)}.$$

Define a function  $\phi : z \mapsto \int_{\mathbf{a}_f}^{a_1+z} (z + a_1 - x)^{p-1} \Psi(x) dx - \int_{a_1+z}^{\tilde{r}_2} (x - a_1 - z)^{p-1} \Psi(x) dx$ . It is clear that  $\phi$  is strictly increasing. Recall that  $\tilde{a}_1 = \mathcal{M}_p(\text{id}, [\mathbf{a}_f, \tilde{r}_2), \mu_f)$ , which is characterized by

$$\int_{\mathbf{a}_f}^{\tilde{a}_1} (\tilde{a}_1 - x)^{p-1} \Psi(x) dx = \int_{\tilde{a}_1}^{\tilde{r}_2} (x - \tilde{a}_1)^{p-1} \Psi(x) dx,$$

so  $\phi(\delta_1) = 0$ . Note also that

$$\begin{aligned}\phi(0) &= \int_{\mathfrak{a}_f}^{a_1} (a_1 - x)^{p-1} \Psi(x) dx - \int_{a_1}^{\tilde{r}_2} (x - a_1)^{p-1} \Psi(x) dx \\ &= \int_{a_1}^{r_2} (x - a_1)^{p-1} \Psi(x) dx - \int_{a_1}^{\tilde{r}_2} (x - a_1)^{p-1} \Psi(x) dx < 0,\end{aligned}$$

since  $\tilde{r}_2 - r_2 = s > 0$ . Moreover, we have

$$\begin{aligned}\int_{\mathfrak{a}_f}^{a_1+s} (a_1 + s - x)^{p-1} \Psi(x) dx &\geq \int_{\mathfrak{a}_f+s}^{a_1+s} (a_1 + s - x)^{p-1} \Psi(x) dx = \int_{\mathfrak{a}_f}^{a_1} (a_1 - x)^{p-1} \frac{\Psi(x+s)}{\Psi(x)} \Psi(x) dx \\ &\geq \frac{\Psi(a_1+s)}{\Psi(a_1)} \int_{\mathfrak{a}_f}^{a_1} (a_1 - x)^{p-1} \Psi(x) dx = \frac{\Psi(a_1+s)}{\Psi(a_1)} \int_{a_1}^{r_2} (x - a_1)^{p-1} \Psi(x) dx \\ &\geq \int_{a_1}^{r_2} (x - a_1)^{p-1} \frac{\Psi(x+s)}{\Psi(x)} \Psi(x) dx = \int_{a_1+s}^{\tilde{r}_2} (x - a_1 - s)^{p-1} \Psi(x) dx\end{aligned}$$

proving that  $\phi(s) \geq 0$ . It follows that  $0 < \delta_1 \leq s = \eta_2$ . The only way that  $\delta_1 = \eta_2 = s$  is that the previous inequalities are only equalities, which means that  $\mathfrak{a}_f = -\infty$  and  $\frac{\Psi(y+s)}{\Psi(y)} = \frac{\Psi(a_1+s)}{\Psi(a_1)}$  holds for all  $y \in (\mathfrak{a}_f, r_2)$   $dy$ -a.e., but since  $G_s$  is decreasing this property holds for all  $y \in (\mathfrak{a}_f, r_2)$ . We summarize this condition for future reference

$$\mathfrak{a}_f = -\infty \text{ and } \forall y \in (\mathfrak{a}_f, r_2) \quad \frac{\Psi(y+s)}{\Psi(y)} = \frac{\Psi(a_1+s)}{\Psi(a_1)}. \quad (2.8)$$

On the other hand, since  $\tilde{r}_2 = \frac{\tilde{a}_1 + \tilde{a}_2}{2}$ , we deduce

$$\tilde{a}_2 = 2\tilde{r}_2 - \tilde{a}_1 = 2r_2 - a_1 + 2\eta_2 - \delta_1 = a_2 + 2\eta_2 - \delta_1,$$

from where we deduce that  $\delta_2 = 2\eta_2 - \delta_1 \geq \eta_2$ , with equality  $\delta_2 = \eta_2$  if and only if  $\delta_2 = \eta_2 = \delta_1 = s$ . Now, if there exists  $\tilde{r}_3 \leq \mathfrak{b}_f$  such that

$$\mathcal{M}_p(\text{id}, [\tilde{r}_2, \tilde{r}_3], \mu_f) = \tilde{a}_2,$$

we deduce that  $\eta_3 \geq \delta_2$ . Indeed, this follows from the inequalities

$$\begin{aligned}\int_{\tilde{a}_2}^{\tilde{r}_3} (x - \tilde{a}_2)^{p-1} \Psi(x) dx &= \int_{\tilde{r}_2}^{\tilde{a}_2} (\tilde{a}_2 - x)^{p-1} \Psi(x) dx = \int_{r_2+\eta_2}^{a_2+\delta_2} (\tilde{a}_2 - x)^{p-1} \Psi(x) dx \\ &\geq \int_{r_2+\delta_2}^{a_2+\delta_2} (\tilde{a}_2 - x)^{p-1} \Psi(x) dx = \int_{r_2}^{a_2} (a_2 - x)^{p-1} \Psi(x + \delta_2) dx \\ &= \int_{r_2}^{a_2} (a_2 - x)^{p-1} \frac{\Psi(x + \delta_2)}{\Psi(x)} \Psi(x) dx \geq \frac{\Psi(a_2 + \delta_2)}{\Psi(a_2)} \int_{r_2}^{a_2} (a_2 - x)^{p-1} \Psi(x) dx \\ &= \frac{\Psi(a_2 + \delta_2)}{\Psi(a_2)} \int_{a_2}^{r_3} (x - a_2)^{p-1} \Psi(x) dx \geq \int_{a_2}^{r_3} (x - a_2)^{p-1} \frac{\Psi(x + \delta_2)}{\Psi(x)} \Psi(x) dx \\ &= \int_{a_2}^{r_3} (x - a_2)^{p-1} \Psi(x + \delta_2) dx = \int_{a_2+\delta_2}^{r_3+\delta_2} (x - \tilde{a}_2)^{p-1} \Psi(x) dx \\ &= \int_{\tilde{a}_2}^{r_3+\delta_2} (x - \tilde{a}_2)^{p-1} \Psi(x) dx\end{aligned}$$

proving that  $\tilde{r}_3 \geq r_3 + \delta_2$ , i.e.  $\eta_3 \geq \delta_2$ . Also, we notice that  $\eta_3 = \delta_2$  if and only if

$$\eta_2 = \delta_2 \text{ and } \forall y \in (r_2, r_3) \quad \frac{\Psi(y + \delta_2)}{\Psi(y)} = \frac{\Psi(a_2 + \delta_2)}{\Psi(a_2)}, \quad (2.9)$$

which in particular implies that  $\eta_3 = \eta_2 = \delta_2 = \delta_1 = s$  and (2.8) holds. Iterating this idea, we complete Step 1.

Step 2: Since  $\tilde{g}$  has  $f$ -special form, we have that  $\tilde{a}_k = \mathcal{M}_1(\text{id}, [\tilde{r}_k, \mathbf{b}_f], \mu_f)$ . Since  $\delta_k \geq \eta_k$  by (a) in Step 1, we obtain that

$$\begin{aligned}
\int_{\tilde{a}_k}^{\mathbf{b}_f} (x - \tilde{a}_k)^{p-1} \Psi(x) dx &= \int_{\tilde{r}_k}^{\tilde{a}_k} (\tilde{a}_k - x)^{p-1} \Psi(x) dx = \int_{r_k + \eta_k}^{a_k + \delta_k} (\tilde{a}_k - x)^{p-1} \Psi(x) dx \\
&\geq \int_{r_k + \delta_k}^{a_k + \delta_k} (\tilde{a}_k - x)^{p-1} \Psi(x) dx = \int_{r_k}^{a_k} (a_k - x)^{p-1} \Psi(x + \delta_k) dx \\
&= \int_{r_k}^{a_k} (a_k - x)^{p-1} \frac{\Psi(x + \delta_k)}{\Psi(x)} \Psi(x) dx \geq \frac{\Psi(a_k + \delta_k)}{\Psi(a_k)} \int_{r_k}^{a_k} (a_k - x)^{p-1} \Psi(x) dx \\
&= \frac{\Psi(a_k + \delta_k)}{\Psi(a_k)} \int_{a_k}^{\mathbf{b}_f} (x - a_k)^{p-1} \Psi(x) dx \geq \frac{\Psi(a_k + \delta_k)}{\Psi(a_k)} \int_{a_k}^{\mathbf{b}_f - \delta_k} (x - a_k)^{p-1} \Psi(x) dx \\
&\geq \int_{a_k}^{\mathbf{b}_f - \delta_k} (x - a_k)^{p-1} \frac{\Psi(x + \delta_k)}{\Psi(x)} \Psi(x) dx = \int_{a_k + \delta_k}^{\mathbf{b}_f} (x - \tilde{a}_k)^{p-1} \Psi(x) dx \\
&= \int_{\tilde{a}_k}^{\mathbf{b}_f} (x - \tilde{a}_k)^{p-1} \Psi(x) dx
\end{aligned}$$

It follows that all the inequalities are in fact equalities and then the following properties hold:

- (d)  $\mathbf{b}_f = \infty$ ;
- (e)  $\eta_k = \delta_k$ ;
- (f)  $\forall y \in (r_k, \infty) \quad \frac{\Psi(y + \delta_k)}{\Psi(y)} = \frac{\Psi(a_k + \delta_k)}{\Psi(a_k)}$ .

We notice that (e) implies that all inequalities in (a) are equalities and then (c) holds. This together with (d) and (f) implies that if  $\tilde{g}$  is a minimizer then

- (g)  $\delta_k = \eta_k = \delta_{k-1} = \eta_{k-1} = \dots = \eta_2 = \delta_1 = s$ ;
- (h)  $\mathbf{a}_f = -\infty, \mathbf{b}_f = \infty$  and for all  $i \in \{1, \dots, k\}$  it holds

$$\forall y \in (r_i, r_{i+1}) \quad \frac{\Psi(y + s)}{\Psi(y)} = \frac{\Psi(a_i + s)}{\Psi(a_i)}.$$

Step 3: Conclusion. Clearly under (H1) or (H2) the function  $\tilde{g}$  cannot be a minimizer. It remains to consider that (H3) holds. From (h) and the continuity of  $\Psi$  it holds that

$$\forall y \in (-\infty, \infty) \quad \frac{\Psi(y + s)}{\Psi(y)} = \frac{\Psi(a_1 + s)}{\Psi(a_1)} = C \in (0, \infty)$$

Iterating this equality we have  $\Psi(y + 2s) = \frac{\Psi(y + 2s)}{\Psi(y + s)} \frac{\Psi(y + s)}{\Psi(y)} \Psi(y) = C^2 \Psi(y)$ , and then for all  $n \in \mathbb{Z}$  and all  $y$

$$\Psi(y + ns) = C^n \Psi(y).$$

Then, if  $C \geq 1$ , we have

$$C^n \int_0^1 \Psi(y) dy = \int_0^1 \Psi(y + ns) dy = \int_{ns}^{ns+1} \Psi(y) dy \xrightarrow[n \rightarrow \infty]{} 0,$$



which is a contradiction. A similar contradiction is obtained if  $C < 1$ , because  $\Psi(y - ns) = C^{-n}\Psi(y)$  and then

$$C^{-n} \int_{-1}^0 \Psi(y) dy = \int_{-1}^0 \Psi(y - ns) dy = \int_{-(1+ns)}^{-ns} \Psi(y) dy \xrightarrow{n \rightarrow \infty} 0,$$

proving that  $\tilde{g}$  cannot be a minimizer, and the result is shown.  $\square$

**Remark.** *Examples of distributions that satisfies the hypothesis of the previous Proposition are the exponential distribution  $\mu_f(dx) = e^{-x} dx$  for  $x \geq 0$ , the normal distribution  $\mathcal{N}(0, 1)$  and the uniform distribution  $\mu_f(dx) = dx$  for  $x \in [0, 1]$ . In the uniform case, we obtain an explicit solution for the minimizer of  $f \in L^p(\Omega, \mathcal{F}, \mu)$ . For all  $k \geq 1$  this unique minimizer is  $g = \ell \circ f$ , where*

$$\ell = \sum_{i=1}^k \frac{2i-1}{2k} \mathbb{1}_{[\frac{i-1}{k}, \frac{i}{k}]},$$

independently of  $p \in [1, \infty)$ .

## 2.2 The case of an infinite measure, $p \in [1, \infty)$

The case of infinite measure needs an extra work and use some ideas already developed in the finite measure case.

**Theorem 2.7.** *Let  $(\Omega, \mathcal{F}, \mu)$  be an infinite measure space,  $p \in [1, \infty)$  and  $k \geq 1$ . Then  $\mathcal{G}_{p,k}$  is proximal.*

Moreover, if  $f \in L^p(\Omega, \mathcal{F}, \mu)$  and  $g = \sum_{i=1}^q b_i \mathbb{1}_{A_i} \in P_{\mathcal{G}_{p,k}}(f)$  is a minimizer, with  $q \leq k$ ,  $-\infty < b_1 < \dots < b_q < \infty$ ,  $\{A_i\}_{1 \leq i \leq q}$  a partition of  $\Omega$  such that  $\mu(A_i) > 0$  for all  $i \in \{1, \dots, q\}$  and a unique  $1 \leq s \leq q$  such that  $b_s = 0$ . Then, there exists a minimizer  $\tilde{g} \in P_{\mathcal{G}_{p,q}}(f)$  in  $f$ -special form

$$\tilde{g} = \sum_{i=1, i \neq s}^q \mathcal{M}_p(f, f^{-1}(C_i)) \mathbb{1}_{f^{-1}(C_i)} + 0 \mathbb{1}_{f^{-1}(C_s)},$$

where

- $r_1 = -\infty, r_{q+1} = \infty$  and  $r_i = \frac{b_{i-1} + b_i}{2}$  for all  $i \in \{2, \dots, q\}$  (notice that  $r_s < 0 < r_{s+1}$ );
- $C_i = f^{-1}([r_i, r_{i+1}))$  for  $i \in \{1, \dots, q-1\}$  and  $C_q = f^{-1}([r_q, r_{q+1}])$ ;
- if  $\mu(f^{-1}(C_i)) > 0$  and  $i \neq s$ , then  $b_i$  is a  $p$ -th mean of  $f$  on  $f^{-1}(C_i)$ ;

If  $q$  is the smallest among all minimizers, then  $\mu(C_i) > 0$  for all  $i$ .

*Proof.* For  $k = 1$  the result is obvious since  $\mathcal{G}_{p,1} = \{0\}$ . So for the rest of the proof we assume that  $k \geq 2$ .

Let  $f \in L^p(\Omega, \mathcal{F}, \mu)$  and consider a sequence  $(g_n)_n \in \mathcal{G}_{p,k}$  such that  $g_n = \sum_{i=1}^{q(n)} a_{i,n} \mathbb{1}_{A_{i,n}}$ , where  $a_{1,n} < \dots < a_{q(n),n} \in \mathbb{R}$ ,  $\{A_{i,n}\}_{1 \leq i \leq q(n)}$  is a measurable partition with sets of positive measure and  $q(n) \leq k$  for all  $n \in \mathbb{N}$ , and such that

$$\|f - g_n\|_p \rightarrow \mathcal{D}_{p,k}(f).$$

Since  $g_n \in L^p(\Omega, \mathcal{F}, \mu)$  there exists a unique  $1 \leq s(n) \leq q(n)$  such that  $a_{s(n),n} = 0$  and we have that  $\mu(A_{i,n}) < \infty$  for all  $i \neq s(n)$ . Passing to a subsequence, we can assume that  $1 \leq s(n) = s \leq q(n) = q \leq k$ . Define  $r_1 = r_{1,n} = -\infty$ ,  $r_{q+1,n} = \infty$  and  $r_{i,n} = \frac{a_{i-1,n} + a_{i,n}}{2}$  for  $i \in \{2, \dots, q\}$ . We

point out that if  $q = 1$ , then  $g_n = 0$ , for all  $n$  and so  $h = 0$  is a minimizer. Then for the rest of the proof, we assume  $q \geq 2$ .

Now, consider  $I_{i,n} = [r_{i,n}, r_{i+1,n})$  and the corresponding  $C_{i,n} = f^{-1}(I_{i,n})$  for all  $i \in \{1, \dots, q\}$ . For all  $n \in \mathbb{N}$ , define

$$\tilde{g}_n = \sum_{i=1}^q a_{i,n} \mathbb{1}_{f^{-1}(I_{i,n})}.$$

If  $i, j \in \{1, \dots, q\}$ , we have that  $|f(x) - a_{i,n}| \geq |f(x) - a_{j,n}|$  for all  $x \in C_j$ . It follows that for all  $n \in \mathbb{N}$

$$\begin{aligned} \|f - g_n\|_p^p &= \sum_{i=1}^q \int_{A_i} |f(x) - a_{i,n}|^p d\mu(x) = \sum_{j=1}^q \sum_{i=1}^q \int_{A_i \cap C_j} |f(x) - a_{i,n}|^p d\mu(x) \\ &\geq \sum_{j=1}^q \sum_{i=1}^q \int_{A_i \cap C_j} |f(x) - a_{j,n}|^p d\mu(x) \\ &= \sum_{j=1}^q \int_{C_j} |f(x) - a_{j,n}|^p d\mu(x) = \|f - \tilde{g}_n\|_p^p \end{aligned}$$

proving that  $(\tilde{g}_n)_n$  is also a minimizing sequence.

For all  $i \in \{1, \dots, q\}$ , the sequence  $(a_{i,n})_n$  has a convergent subsequence in  $\overline{\mathbb{R}}$ . Then we can also assume that  $a_{i,n} \rightarrow a_i \in \overline{\mathbb{R}}$  for all  $i \in \{1, \dots, q\}$ . We denote by  $z_1 < \dots < z_\ell$  the different values in  $\{a_1, \dots, a_q\}$ , where  $\ell \leq q$ . We point out that  $z_t = 0$  for some  $1 \leq t \leq \ell$ . For each  $1 \leq m \leq \ell$ , we denote  $L_m = \{i : 1 \leq i \leq q \text{ and } a_i = z_m\}$ . Each  $L_m$  is an interval in  $\mathbb{N}$ , because we have assumed  $a_{1,n} < \dots < a_{i,n} < \dots < a_{q,n}$ , for each  $n$ . We define  $i_m^- = \min\{L_m\}$  and  $i_m^+ = \max\{L_m\}$  for all  $m \in \{1, \dots, \ell\}$  and also  $i_{\ell+1}^- = \ell + 1$ . Note that  $L_m = \{i_m^-, \dots, i_m^+\}$  for all  $m \in \{1, \dots, \ell\}$ .

Assume that  $z_\ell = \infty$  or  $z_1 = -\infty$ . In this situation  $\ell \geq 2$ , because  $z_t = 0$ . As in the case of finite measure we can modify  $(\tilde{g}_n)_n$  to get a uniformly bounded minimizing sequence. Consider first the case  $z_\ell = \infty$  and recall that  $i_\ell^- = \min\{L_\ell\} \in \{s+1, \dots, q\}$ . Then, we have

$$r_{i_\ell^-, n} = \frac{a_{i_\ell^- - 1, n} + a_{i_\ell^-, n}}{2} \rightarrow \infty,$$

because  $a_{i_\ell^- - 1, n} \geq a_{s, n} = 0$  and then  $a_{i_\ell^- - 1, n} \rightarrow a_{i_\ell^- - 1} = z_{\ell-1} \in [0, \infty)$ . Consider

$$\hat{g}_n = \sum_{i < i_\ell^-} a_{i,n} \mathbb{1}_{f^{-1}(I_{i,n})} + a_{i_\ell^- - 1, n} \mathbb{1}_{f^{-1}([r_{i_\ell^- - 1, n}, \infty))} = \sum_{i < i_\ell^- - 1} a_{i,n} \mathbb{1}_{f^{-1}(I_{i,n})} + a_{i_\ell^- - 1, n} \mathbb{1}_{f^{-1}([r_{i_\ell^- - 1, n}, \infty))}$$

An important fact is that  $\int_{\{f \geq r_{i_\ell^- - 1, n}\}} |f(x) - a_{i_\ell^- - 1, n}|^p d\mu(x) \rightarrow 0$ , because  $(a_{i_\ell^- - 1, n})_n$  is a bounded sequence and  $r_{i_\ell^- - 1, n} \rightarrow \infty$ . Then

$$\|f - \hat{g}_n\|_p^p \leq \|f - \tilde{g}_n\|_p^p + \int_{\{f > r_{i_\ell^- - 1, n}\}} |f(x) - a_{i_\ell^- - 1, n}|^p d\mu(x) \rightarrow (\mathcal{D}_{p,k}(f))^p.$$

Then, the sequence  $(\hat{g}_n)_n$  is a minimizing sequence, which is uniformly upper bounded. Similarly, we can modify this sequence to get a minimizing sequence, which is uniformly bounded. Then, in what follows we assume  $(\tilde{g}_n)_n$  is uniformly bounded and  $-\infty < z_1, z_\ell < \infty$ .

Now we consider 2 different cases.

**Case 1:**  $\ell = 1$ . In this situation  $a_1 = \dots = a_q = 0$ . Notice that  $r_{q,n} = \frac{a_{q-1,n} + a_{q,n}}{2} \rightarrow 0$ , so if  $0 < f(x) < \infty$ , then  $\tilde{g}_n(x) = a_{q,n}$  for all large  $n$  and  $\tilde{g}_n(x) \rightarrow 0$ . In the same way, if  $f(x) < 0$ , then

$\tilde{g}_n(x) = a_{1,n}$  for all large  $n$  and  $\tilde{g}_n(x) \rightarrow 0$ . On the other hand, if  $f(x) = 0$ , then  $\tilde{g}_n(x) = 0$ . Then by Fatou's Lemma we conclude

$$\liminf_n \|f - \tilde{g}_n\|_p^p \geq \int \liminf_n |f(x) - \tilde{g}_n|^p d\mu(x) = \|f\|_p^p,$$

and we obtain  $\mathcal{D}_{p,k}(f) \geq \|f\|_p$ , showing that  $h = 0$  is a minimizer.

**Case 2:**  $\ell \geq 2$ . For all  $m \in \{1, \dots, \ell\}$  recall that  $i_m^- = \min\{L_m\}$  and  $i_{\ell+1}^- = \ell + 1$ . Then, for all  $2 \leq m \leq \ell$

$$r_{i_m^-, n} \rightarrow r_m := \frac{z_{m-1} + z_m}{2}.$$

and  $r_1 = -\infty < r_2 < \dots < r_\ell < r_{\ell+1} := \infty$ . Now, we choose a particular subsequence  $(n')_{n'}$ . We start with  $(r_{i_2^-, n})_n$ . If there exist an increasing subsequence of  $(r_{i_2^-, n})_n$ , we fix one of these subsequences as  $(n^{(2)})$  and we put  $T(2) = \text{in}$ , for increasing. Otherwise we take  $(n^{(2)})$  so that  $(r_{i_2^-, n^{(2)}})_{n^{(2)}}$  is strictly decreasing, and we put  $T(2) = \text{sd}$ , for strictly decreasing. We repeat this procedure for  $(r_{i_3^-, n^{(2)}})_{n^{(2)}}$ , to obtain, if possible,  $(n^{(3)})$  a subsequence of  $(n^{(2)})$  so  $(r_{i_3^-, n^{(3)}})_{n^{(3)}}$  is increasing, and put  $T(3) = \text{in}$ . Otherwise we take  $(n^{(3)})$  a subsequence of  $(n^{(2)})$  so that  $(r_{i_3^-, n^{(3)}})_{n^{(3)}}$  is strictly decreasing, and we put  $T(3) = \text{sd}$ . We continue until  $m = \ell$ . We also put  $T(1) = \text{in}$  and  $T(\ell + 1) = \text{in}$ . Denote by  $(n') = (n^{(\ell)})$ .

Now, we define the intervals that give a minimizer. For all  $m \in \{1, \dots, \ell\}$  let

$$I_m = \begin{cases} [r_m, r_{m+1}) & \text{if } T(m) = \text{in}, T(m+1) = \text{in} \\ [r_m, r_{m+1}) & \text{if } T(m) = \text{in}, T(m+1) = \text{sd} \\ (r_m, r_{m+1}) & \text{if } T(m) = \text{sd}, T(m+1) = \text{in} \\ (r_m, r_{m+1}) & \text{if } T(m) = \text{sd}, T(m+1) = \text{sd} \end{cases} \quad (2.10)$$

We notice that  $\cup_{m=1}^\ell I_m = [-\infty, \infty)$ , and for all  $m \in \{1, \dots, \ell\}$  and all  $n'$ , we define  $J_{m,n'} = \cup_{i \in L_m} [r_{i,n'}, r_{i+1,n'}) = [r_{i_m^-, n'}, r_{i_{m+1}^-, n'})$ . Then, it holds

$$\mathbb{1}_{f^{-1}(J_{m,n'})} \rightarrow \mathbb{1}_{f^{-1}(I_m)} \text{ a.e.}$$

The last piece of information we need is that the set  $\cup_{m \neq t} f^{-1}(J_{m,n'})$  is contained in a fixed set of finite measure  $\tilde{A}$  for large  $n'$ . If  $t = \ell$ , then  $\cup_{m \neq t} f^{-1}(J_{m,n'}) \subset f^{-1}((-\infty, r_{t,n'}]) \subset \tilde{A} = f^{-1}((-\infty, \frac{r_t}{2}])$ , for large  $n'$ , because  $r_{t,n'} \rightarrow r_t = \frac{z_{t-1}}{2} < z_t = 0$ , and then  $\tilde{A}$  has finite measure. Similarly, if  $t = 1$ , then  $\cup_{m \neq t} f^{-1}(J_{m,n'}) \subset f^{-1}([r_{2,n'}, \infty]) \subset \tilde{A} = f^{-1}([\frac{r_2}{2}, \infty))$ , for large  $n'$ . This set has finite measure because  $r_2 > 0$ . In the general case,  $1 < t < q$ , we have for large  $n'$

$$\cup_{m \neq t} f^{-1}(J_{m,n'}) \subset \tilde{A} = f^{-1}((-\infty, r_t/2]) \cup f^{-1}([r_{t+1}/2, \infty)),$$

which has finite measure because  $r_t < 0 < r_{t+1}$ .

Now, consider the decomposition

$$\|f - \tilde{g}_{n'}\|_p^p = \int |f(x) - \tilde{g}_{n'}(x)|^p \mathbb{1}_{f^{-1}(J_{t,n'})} d\mu(x) + \sum_{m \neq t} \int_{f^{-1}(J_{m,n'})} |f(x) - \tilde{g}_{n'}(x)|^p d\mu(x).$$

We use now Fatou's Lemma for the first term and the Dominated Convergence Theorem for the second term. In the first term, we have the a.e. convergence

$$|f - \tilde{g}_{n'}| \mathbb{1}_{f^{-1}(J_{t,n'})} \rightarrow |f| \mathbb{1}_{f^{-1}(I_t)}$$

With respect to the second term, for large  $n'$ , we have  $\max_{1 \leq i \leq q} |a_{i,n'}| \leq \max\{|z_1|, |z_\ell|\} + 1 := C$ , also

$$|f(x) - \tilde{g}_{n'}(x)|^p \mathbb{1}_{\cup_{m \neq t} f^{-1}(J_{m,n'})} \leq 2^{p-1} (|f(x)|^p + C^p) \mathbb{1}_{\tilde{A}} \in L^1(\Omega, \mathcal{F}, \mu),$$

$\tilde{g}_{n'} \mathbb{1}_{\cup_{m \neq t} f^{-1}(J_{m,n'})} \rightarrow \sum_{m \neq t} z_m \mathbb{1}_{f^{-1}(I_m)}$  a.e. and  $f \mathbb{1}_{\cup_{m \neq t} f^{-1}(J_{m,n'})} \rightarrow f \mathbb{1}_{f^{-1}(\cup_{m \neq t} I_m)}$  a.e. So, we get

$$\liminf_{n'} \|f - \tilde{g}_n\|_p^p \geq \int_{f^{-1}(I_t)} |f(x)|^p d\mu(x) + \sum_{m \neq t} \int_{f^{-1}(I_m)} |f(x) - z_m|^p d\mu(x)$$

and then  $h = \sum_{m=1}^{\ell} z_m \mathbb{1}_{f^{-1}(I_m)}$  is a minimizer, where the intervals  $\{I_m\}_{1 \leq m \leq \ell}$  are either open, closed or semi-closed, they are disjoint and  $\cup_{m=1}^{\ell} I_m = \mathbb{R}$  (see (2.10)).

From here it is clear that a minimizer exists in  $f$ -special form as we have done in the finite measure case. Also notice that if  $m \neq t$  and  $0 < \mu(f^{-1}(I_m))$ , we must have  $z_m$  is a  $p$ -th mean for  $f$  in  $f^{-1}(I_m)$ , since  $h$  is a minimizer.  $\square$

### 2.3 The case $p = \infty$

In this section we shall prove that  $\mathcal{G}_{\infty,k}$  is proximal. We start with a lemma.

**Lemma 2.8.** *Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space and  $f \in L^\infty(\Omega, \mathcal{F}, \mu)$ . Then, for all  $k \geq 1$ , we have that  $\mathcal{D}_{\infty,k}(f) = \eta_k(f)$  where*

$$\eta_k(f) = \inf_{h=f \text{ a.e.}} \inf\{\alpha > 0 \mid h(\Omega) \text{ can be covered by at most } k \text{ closed balls of radius } \alpha\}.$$

*Proof.* Let  $\varepsilon > 0$  and let  $g \in \mathcal{G}_{\infty,k}$  such that  $\|f - g\|_\infty \leq \mathcal{D}_{\infty,k}(f) + \varepsilon$ . Write  $g = \sum_{i=1}^k a_i \mathbb{1}_{A_i}$  where  $\{A_i\}_{1 \leq i \leq k}$  is a partition of  $\Omega$ . For every  $i$ , the set  $C_i = \{x \in A_i : |f(x) - a_i| > \|(f - g)\mathbb{1}_{A_i}\|_\infty\}$  has measure 0 and therefore  $h = f \mathbb{1}_{\Omega \setminus \cup_j C_j} + \sum_{j=1}^k a_j \mathbb{1}_{C_j}$  satisfies  $h = f$  a.e. and

$$h(\Omega) \subset \bigcup_{i=1}^k [a_i - \mathcal{D}_{\infty,k}(f) - \varepsilon, a_i + \mathcal{D}_{\infty,k}(f) + \varepsilon].$$

It follows that  $\eta_k(f) \leq \mathcal{D}_{\infty,k}(f) + \varepsilon$  and since  $\varepsilon$  is arbitrary, we obtain that  $\eta_k(f) \leq \mathcal{D}_{\infty,k}(f)$ . To prove the other inequality, let again  $\varepsilon > 0$  and pick  $l \leq k$ ,  $a_1, \dots, a_l \in \mathbb{R}$  and  $h = f$  a.e. such that

$$h(\Omega) \subset \bigcup_{i=1}^l [a_i - \eta_k(f) - \varepsilon, a_i + \eta_k(f) + \varepsilon].$$

For  $1 \leq i \leq l$ , define  $A_i = h^{-1}([a_i - \eta_k(f) - \varepsilon, a_i + \eta_k(f) + \varepsilon]) \in \mathcal{F}$ . Now define  $B_1 = A_1$  and  $B_i = A_i \setminus \bigcup_{j=1}^{i-1} A_j$  for  $i \in \{2, \dots, l\}$ . Then  $\{B_i\}_{1 \leq i \leq l}$  is a partition of  $\Omega$ . Defining  $g = \sum_{i=1}^l a_i \mathbb{1}_{B_i} \in \mathcal{G}_{\infty,k}$ , it is clear that  $\|f - g\|_\infty \leq \eta_k(f) + \varepsilon$ . It follows that  $\mathcal{D}_{\infty,k}(f) \leq \eta_k(f) + \varepsilon$  and then  $\mathcal{D}_{\infty,k}(f) \leq \eta_k(f)$ .  $\square$

**Proposition 2.9.** *Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space. Then  $\mathcal{G}_{\infty,k}$  is proximal for all  $k \geq 1$ .*

*Proof.* For all  $n \in \mathbb{N}$ , let  $\alpha_n = \eta_k(f) + \frac{1}{n}$ . So, for all  $n \in \mathbb{N}$ , there exist  $a_1^n, \dots, a_{l_n}^n \in \mathbb{R}$  with  $1 \leq l_n \leq k$  and  $h_n = f$  a.e. such that  $h_n(\Omega) \subset \bigcup_{i=1}^{l_n} [a_i^n - \alpha_n, a_i^n + \alpha_n]$ . Of course there exists  $i_0$  such that  $\mu(h_n^{-1}([a_{i_0}^n - \alpha_n, a_{i_0}^n + \alpha_n])) > 0$ . If for some  $i$  it holds  $\mu(h_n^{-1}([a_i^n - \alpha_n, a_i^n + \alpha_n])) = 0$ , we can redefine  $h_n$  on a set of measure 0, to have  $h_n(w) = a_{i_0}$  for all  $w \in h_n^{-1}([a_i^n - \alpha_n, a_i^n + \alpha_n])$ . So, we can assume for all  $i$  it holds  $\mu(h_n^{-1}([a_i^n - \alpha_n, a_i^n + \alpha_n])) > 0$ . Consider

$$t_{i,n} = \frac{1}{\mu(h_n^{-1}([a_i^n - \alpha_n, a_i^n + \alpha_n]))} \int_{h_n^{-1}([a_i^n - \alpha_n, a_i^n + \alpha_n])} h_n(x) d\mu(x) \in [a_i^n - \alpha_n, a_i^n + \alpha_n],$$

which obviously satisfies  $|t_{i,n}| \leq \|h_n\|_\infty = \|f\|_\infty$ . Then, for all  $i, n$ , it holds

$$|a_i^n| \leq |t_{i,n}| + |a_i^n - t_{i,n}| \leq \|f\|_\infty + \alpha_n \leq \|f\|_\infty + \eta_k(f) + 1,$$

which implies that the set  $\{a_i^n : 1 \leq i \leq l_n, n \in \mathbb{N}\}$  is bounded.

Considering a subsequence if necessary, we can suppose that  $l_n = l \in \{1, \dots, k\}$  for all  $n \in \mathbb{N}$ . By compactness and taking a further subsequence, we can also assume that  $a_i^n \rightarrow a_i$  for all  $i \in \{1, \dots, l\}$ . Define

$$C = \{\omega \in \Omega \mid \forall n \in \mathbb{N} f(\omega) = h_n(\omega)\} \in \mathcal{F}$$

and note that  $\mu(C^c) = 0$ . Let us show that  $f(C) \subset \bigcup_{i=1}^l [a_i - \eta_k(f), a_i + \eta_k(f)]$ . In fact, if  $\omega \in C$  then for all  $n \in \mathbb{N}$  there exists  $i(\omega, n) \in \{1, \dots, l\}$  such that  $f(\omega) = h_n(\omega) \in [a_{i(\omega, n)}^n - \alpha_n, a_{i(\omega, n)}^n + \alpha_n]$ . There exists a subsequence  $\phi(n) = \phi(n)(\omega)$  such that the sequence  $(i(\omega, \phi(n)))_n$  is constant and equal to some  $i_0(\omega) \in \{1, \dots, l\}$ . It follows that  $f(\omega) \in [a_{i_0(\omega)} - \eta_k(f), a_{i_0(\omega)} + \eta_k(f)] \subset \bigcup_{i=1}^l [a_i - \eta_k(f), a_i + \eta_k(f)]$ . Define  $h = f \mathbb{1}_C + t \mathbb{1}_{C^c}$  where  $t$  is any real belonging to  $\bigcup_{i=1}^l [a_i - \eta_k(f), a_i + \eta_k(f)]$ . We have that  $f = h$  a.e. and  $h(\Omega) \subset \bigcup_{i=1}^l [a_i - \eta_k(f), a_i + \eta_k(f)]$ . For  $1 \leq i \leq l$ , define  $A_i = h^{-1}([a_i - \eta_k(f), a_i + \eta_k(f)]) \in \mathcal{F}$ . Now define  $B_1 = A_1$  and  $B_i = A_i \setminus \bigcup_{j=1}^{i-1} A_j$  for  $i \in \{2, \dots, l\}$ . Then  $\{B_i\}_{1 \leq i \leq l}$  is a partition of  $\Omega$ . Defining  $g = \sum_{i=1}^l a_i \mathbb{1}_{B_i} \in \mathcal{G}_{\infty, k}$ , it is clear that  $\|f - g\|_{\infty} = \|h - g\|_{\infty} \leq \eta_k(f)$ . Moreover, we have that  $\mathcal{D}_{\infty, k}(f) = \eta_k(f)$  by the previous Lemma and so we conclude that  $\|f - g\|_{\infty} = \mathcal{D}_{\infty, k}(f)$ .  $\square$

## 2.4 Extra properties of minimizers and the sets $(\mathcal{G}_{p, k})_{p, k}$

In this section we include some extra properties of the sets  $(\mathcal{G}_{p, k})_{p, k}$  as well as some natural questions like uniqueness of minimizers and the existence of a continuous selection for  $P_{\mathcal{G}_{p, k}}$ .

Let us start by proving that  $\mathcal{G}_{p, k}$  is a closed set, for all  $p \geq 1, k \geq 1$ , something that it is not straightforward to do. Nevertheless, this is a direct consequence of the previous results.

**Corollary 2.10.** *Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space,  $p \in [1, \infty]$  and  $k \geq 1$ . Then  $\mathcal{G}_{p, k}$  is closed.*

*Proof.* Assume  $(g_n)_n \subset \mathcal{G}_{p, k}$  converges in  $L^p(\Omega, \mathcal{F}, \mu)$  to  $g$ . Then

$$\inf\{\|g - h\|_p : h \in \mathcal{G}_{p, k}\} = 0.$$

From the previous results, there exists a minimizer  $\bar{h} \in \mathcal{G}_{p, k}$ , that is  $g = \bar{h}$  a.e. and the result is shown.  $\square$

A question that appears when proving the existence of minimizers is the following. Assume there exists a best approximation of  $f$  by an element of  $\mathcal{G}_{p, k}$  which is in fact an element of  $\mathcal{G}_{p, m}$  for some  $m < k$ , then it is natural to think that  $f$  should belong to  $\mathcal{G}_{p, m}$ . This is true when  $p \in [1, \infty)$  and it is not true for  $p = \infty$ . Before doing that we require the following lemma.

**Lemma 2.11.** *Assume that  $f \in L^p(\Omega, \mathcal{F}, \mu)$ , for  $1 \leq p < \infty$ , and  $A = f^{-1}(I)$  is a set of positive and finite measure, where  $I$  is an interval. Assume  $b$  is a  $p$ -th mean of  $f$  on  $A$ , then  $b \in \bar{I}$ .*

*Proof.* Assume the interval  $\bar{I} = [c, d]$ , where  $c, d \in \overline{\mathbb{R}}$  and let us prove that  $b \geq c$ . If  $c = -\infty$  it is clear that  $c < b$ . So assume  $c$  is finite. By contradiction, if  $b < c$  we have  $|f(x) - b| = f(x) - c + (c - b) > f(x) - c = |f(x) - c|$ , for all  $x \in f^{-1}(I)$  and then, since  $\mu(f^{-1}(I)) > 0$ , we get

$$\int_{f^{-1}(I)} |f(x) - b|^p d\mu(x) > \int_{f^{-1}(I)} |f(x) - c|^p d\mu(x),$$

which is contradiction. Similarly, it is shown that  $b \leq d$ .  $\square$

**Proposition 2.12.** *Assume  $f \in L^p(\Omega, \mathcal{F}, \mu)$  with  $p \in [1, \infty)$ . Let  $m, k \in \mathbb{N}$  such that  $1 \leq m < k$ . Suppose that there exists  $g \in \mathcal{G}_{p, m} \cap P_{\mathcal{G}_{p, k}}(f)$ . Then  $f \in \mathcal{G}_{p, m}$ .*

*Proof.* Suppose that the measure is finite. We can assume that

$$g = \sum_{i=1}^r b_i \mathbb{1}_{f^{-1}(I_i)},$$

where  $r \leq m$ ,  $\{I_i\}_{1 \leq i \leq r}$  is a family of disjoint intervals such that  $\{f^{-1}(I_i)\}_{1 \leq i \leq r}$  is a partition of  $\Omega$  and  $b_1 < \dots < b_r$ . Suppose by contradiction that  $f \notin \mathcal{G}_{p,m}$ . Then in particular it holds that  $\mu(f^{-1}(\{b_1, \dots, b_r\}^c)) > 0$ . Since  $\Theta = \{b_1, \dots, b_r\}^c$  is open, it is a countable union of open intervals  $(J_n)_n$  and therefore for some  $n_0$  we should have  $\mu(f^{-1}(J_{n_0})) > 0$ . By the continuity of the measure, there exists a closed bounded interval  $J \subset J_{n_0}$  such that  $\mu(f^{-1}(J)) > 0$ , and therefore  $\mu(f^{-1}(J \cap I_{i_0})) > 0$ , for some  $i_0$ . Hence, we obtain

$$\int_{f^{-1}(I_{i_0} \cap J)} |f - \mathcal{M}_p(f, f^{-1}(I_{i_0} \cap J))|^p d\mu(x) < \int_{f^{-1}(I_{i_0} \cap J)} |f - b_{i_0}|^p d\mu(x),$$

since an equality in the previous formula would imply that  $b_{i_0} \in \overline{I_{i_0} \cap J} \subset J \subset \{b_1, \dots, b_r\}^c$ , by the previous Lemma. If we define

$$h = \sum_{i=1, i \neq i_0}^r b_i \mathbb{1}_{f^{-1}(I_i)} + b_{i_0} \mathbb{1}_{f^{-1}(I_{i_0} \cap J^c)} + \mathcal{M}_p(f, f^{-1}(I_{i_0} \cap J)) \mathbb{1}_{f^{-1}(I_{i_0} \cap J)} \in \mathcal{G}_{p,r+1} \subset \mathcal{G}_{p,k},$$

we have that  $\|f - h\|_p < \|f - g\|_p$  which contradicts the minimality of  $g$ . We conclude that  $f \in \mathcal{G}_{p,m}$ .

In case the measure is infinite, with the same notation as above, we know that  $b_{i_1} = 0$  for some  $i_1$ . As above there exists a closed and bounded interval  $J \subset \{b_1, \dots, b_r\}^c \subset \{0\}^c$ , such that  $\mu(f^{-1}(J)) > 0$ . Without loss of generality we can assume that  $J \subset [a, \infty)$ , for some  $a > 0$ . Then

$$\mu(f^{-1}(J))a^p \leq \|f\|_p^p,$$

proving that  $f^{-1}(J)$  has finite and positive measure. The argument now goes as in the case of finite measure.  $\square$

The following result shows that, for  $p \in [1, \infty)$ , the error in the approximation by functions in  $\mathcal{G}_{p,k}$  decreases strictly with  $k$  until eventually reaching zero.

**Corollary 2.13.** *Assume that  $f \in L^p(\Omega, \mathcal{F}, \mu)$  with  $p \in [1, \infty)$  and consider  $\mathcal{D}_{p,\infty}(f) = 0$ . Define  $k^* = \min\{k : \mathcal{D}_{p,k}(f) = 0\} \in [1, \infty]$ . Then,  $(\mathcal{D}_{p,k}(f))_{k \leq k^*}$  is strictly decreasing and  $\mathcal{D}_{p,k}(f) = 0$  for all  $k \geq k^*$ , that is*

$$k^* = \min\{k : \mathcal{D}_{p,k+1}(f) = \mathcal{D}_{p,k}(f)\} = \min\{k : \mathcal{D}_{p,k}(f) = 0\}.$$

The previous results are not true for  $p = \infty$ . In fact, we have the following example:

**Example.** *Consider the Lebesgue measure in  $[0, 1]$ , the function*

$$f(x) = \begin{cases} x & \text{for } x \notin (\frac{1}{3}, \frac{2}{3}) \\ \frac{1}{3} & \text{for } x \in (\frac{1}{3}, \frac{2}{3}) \end{cases}$$

and  $k = 3$ . It is not difficult to show that  $\mathcal{D}_{\infty,3}(f) = \frac{1}{6}$ , where there are multiple minimizers, for example

$$h = \frac{1}{6} \mathbb{1}_{[0, \frac{1}{3}]} + \frac{1}{2} \mathbb{1}_{(\frac{1}{3}, \frac{2}{3})} + \frac{5}{6} \mathbb{1}_{[\frac{2}{3}, 1]}$$

is a minimizer, but also

$$g = \frac{1}{6} \mathbb{1}_{[0, \frac{2}{3}]} + \frac{5}{6} \mathbb{1}_{[\frac{2}{3}, 1]} \in \mathcal{G}_{\infty,2}$$

is a minimizer, in particular  $\mathcal{D}_{\infty,3}(f) = \mathcal{D}_{\infty,2}(f) > 0$ . Nevertheless,  $f \notin \mathcal{G}_{\infty,k}$  for all  $k$ . This also shows that  $\mathcal{G}_{\infty,3}$  is not Chebyshev.

In Proposition 2.6, we have shown that under certain conditions on  $f$ , there exists a unique minimizer. An important question then is if  $\mathcal{G}_{p,k}$  is Chebyshev, that is, if there is a unique minimizer for all  $f$ . As we have seen in the previous example this is not true for  $p = \infty$ , and we complement this for all  $p$ .

**Example.** Consider again the Lebesgue measure in  $[0, 1]$ . Then  $\mathcal{G}_{p,2}$  is not Chebyshev for any  $p \in [1, \infty]$ . To see that, let  $f = -\mathbb{1}_{[0, \frac{1}{3}]} + 0\mathbb{1}_{[\frac{1}{3}, \frac{2}{3}]} + \mathbb{1}_{[\frac{2}{3}, 1]}$ . A possible minimizer in  $\mathcal{G}_{2,2}$  has the form  $g = a\mathbb{1}_{f^{-1}((-\infty, r_2))} + b\mathbb{1}_{f^{-1}([r_2, \infty])}$ , for suitable  $a, b, r_2$  (see Theorem 2.3). If  $r_2 \leq -1$  or  $r_2 > 1$ , a candidate to be a minimizer is  $g_1 = 0$ . For  $-1 < r_2 \leq 0$  the candidate is  $g_2 = -\mathbb{1}_{[0, \frac{1}{3}]} + \frac{1}{2}\mathbb{1}_{[\frac{1}{3}, 1]}$ . Finally, for  $0 < r_2 < 1$  the candidate is  $g_3 = -\frac{1}{2}\mathbb{1}_{[0, \frac{2}{3}]} + \mathbb{1}_{[\frac{2}{3}, 1]}$ . The corresponding errors are

$$\|f - g_1\|_2^2 = \frac{2}{3}, \quad \|f - g_2\|_2^2 = \|f - g_3\|_2^2 = \frac{1}{6},$$

showing that  $g_2$  and  $g_3$  are two minimizers and then  $\mathcal{G}_{2,2}$  is not Chebyshev. Finally, for every  $p \in [1, \infty]$  both  $g_2$  and  $g_3$  are minimizers in  $\mathcal{G}_{p,2}$ , showing that this set is not Chebyshev for any  $p$ . Moreover, for  $1 < p < \infty$ , it can be proved that  $g_2, g_3$  are the only minimizers. For  $p = 1$ , there is a continuum of minimizers since

$$g_a = -\mathbb{1}_{[0, \frac{1}{3}]} + a\mathbb{1}_{[\frac{1}{3}, 1]}$$

is a minimizer for all  $a \in [0, 1]$ . For  $p = \infty$ , there is also a continuum of minimizers since

$$h_b = b\mathbb{1}_{[0, \frac{1}{3}]} + \frac{1}{2}\mathbb{1}_{[\frac{1}{3}, 1]}$$

is a minimizer for all  $b \in [-\frac{3}{2}, \frac{-1}{2}]$ .

**Remark.** We have proved that  $\mathcal{G}_{p,k}$  is proximal and closed for all  $k \geq 1$  and  $p \in [1, \infty]$ . However,  $\mathcal{G}_{p,k}$  is not Chebyshev in general as we have shown in the previous examples. Then, it is natural to ask if  $P_{\mathcal{G}_{p,k}}$  admits a continuous selection. If such continuous selection exists, then  $\mathcal{G}_{p,k}$  has to be almost-convex (see Lemma 5 in [14]). Remember that a subset  $K$  of a Banach space is said to be almost-convex (see [14]) if for every closed ball  $B$  such that  $K \cap B = \emptyset$ , there exists a closed ball  $B'$  of arbitrary large radius such that  $K \cap B' = \emptyset$  and  $B \subset B'$ . If  $p \in (1, \infty)$ , a subset  $K$  is almost-convex if and only if  $K$  is convex (see Lemma 2 in [14]). So, the question is if  $\mathcal{G}_{p,k}$  can be convex. For  $k \geq 2$  and  $p < \infty$ ,  $\mathcal{G}_{p,k}$  is convex if and only if  $L^p(\Omega, \mathcal{F}, \mu)$  is finite dimensional and  $L^p(\Omega, \mathcal{F}, \mu) = \mathcal{G}_{p,k}$ . Indeed, assume  $k \geq 2$  and that  $\mathcal{G}_{p,k}$  is convex. Then it is direct to show that  $\mathcal{G}_{p,k}$  is a vector space, because it is homogeneous. Then  $\mathcal{G}_{p,\ell} = \mathcal{G}_{p,k}$ , for all  $\ell \geq k$ . This is done by induction, so the only interesting case is  $\ell = k + 1$ . Take  $g = \sum_{i=1}^{k+1} a_i \mathbb{1}_{A_i}$ , which can be seen as the sum of three elements  $g_1, g_2, g_3 \in \mathcal{G}_{p,k}$

$$g_1 = \sum_{i=1}^{k-1} a_i \mathbb{1}_{A_i} + 0\mathbb{1}_{A_k \cup A_{k+1}}, \quad g_2 = a_k \mathbb{1}_{A_k} + 0\mathbb{1}_{\cup_{j \neq k} A_j}, \quad g_3 = a_{k+1} \mathbb{1}_{A_{k+1}} + 0\mathbb{1}_{\cup_{j \neq k+1} A_j}.$$

Therefore,  $\mathcal{G}_{p,k} = \cup_{\ell} \mathcal{G}_{p,\ell}$  is dense and closed in  $L^p(\Omega, \mathcal{F}, \mu)$ , which implies  $\mathcal{G}_{p,k} = L^p(\Omega, \mathcal{F}, \mu)$ . The conclusion is that the unit ball of  $L^p(\Omega, \mathcal{F}, \mu)$  is UA and then  $L^p(\Omega, \mathcal{F}, \mu)$  is finite dimensional (see Theorem 4.7 in Section 4.3).

### 3 The $p$ -variation

In this part we introduce a new notion of variation for functions in  $L^p(\Omega, \mathcal{F}, \mu)$ . There are several notions of variation or oscillation for functions. Our notion could be contrasted with the definition of oscillation given in [4] (p.296), which helps to characterize compact sets in  $L^1$ . However, both concepts are not comparable, in general.

**Definition 3.1.** Let  $p \in [1, \infty)$ . For  $f \in L^p(\Omega, \mathcal{F}, \mu)$  and  $A$  a measurable set of finite measure, we define  $\text{var}_p(f, A)$ , the  $p$ -variation of  $f$  in  $A$ , as

$$\text{var}_p(f, A)^p = \begin{cases} \frac{1}{\mu(A)} \int_{A \times A} |f(x) - f(y)|^p d\mu(x)d\mu(y) & \text{if } \mu(A) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Given  $\mathcal{P} = (A_i)_i$ , a finite collection of disjoint measurable sets each one of finite measure, which we also assume it contains at least one set of positive measure, we define the total  $p$ -variation of  $f$  in  $\mathcal{P}$  as

$$\text{var}_p(f, \mathcal{P}) = \left( \sum_i \text{var}_p(f, A_i)^p \right)^{1/p} = \left( \sum_{i: \mu(A_i) > 0} \frac{1}{\mu(A_i)} \int_{A_i \times A_i} |f(x) - f(y)|^p d\mu(x)d\mu(y) \right)^{1/p}.$$

For a measurable set  $A$  of finite measure, we define the  $k$ -th total  $p$ -variation of  $f$  as

$$\text{Var}_{p,k}(f, A) = \inf \left\{ \text{var}_p(f, \mathcal{P}) : \mathcal{P} \text{ is a partition of } A, |\mathcal{P}| \leq k \right\}$$

where the infimum is taken over the set of finite measurable partitions of  $A$  consisting of at most  $k$  measurable sets. Finally, we define the total  $p$ -th variation of  $f$  as

$$\text{Var}_{p,k}(f) = \sup_{\substack{A \in \mathcal{F} \\ \mu(A) < \infty}} \text{Var}_{p,k}(f, A)$$

Note that if  $\mu$  is finite then  $\text{Var}_{p,k}(f, \Omega) \leq \text{Var}_{p,k}(f)$ , and it is not clear if both measures of total variation are equivalent, something that we study below (see Proposition 3.3).

**Remark.** Notice that the sets in  $\mathcal{P}$  that have measure 0 can be removed by gluing them to an element of  $\mathcal{P}$  with positive measure. We redefine a new collection  $\tilde{\mathcal{P}}$ , which has fewer elements and  $\text{var}_p(f, \mathcal{P}) = \text{var}_p(f, \tilde{\mathcal{P}})$ . So, in what follows, we can always assume that  $\mathcal{P}$  is a collection with sets of positive and finite measure.

We compile some basic properties of  $\text{Var}_{p,k}(\bullet, \Omega)$  and  $\text{Var}_{p,k}(\bullet)$  in the next result:

**Proposition 3.1.** Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space and  $p \in [1, \infty)$ . Then  $(\text{Var}_{p,k})_{k \geq 1}$  is a decreasing family of continuous semi-norms on  $L^p(\Omega, \mathcal{F}, \mu)$  such that  $\text{Var}_{p,k}(\bullet) \leq 2\|\bullet\|_p$  for all  $k \geq 1$ . The same properties hold for  $(\text{Var}_{p,k}(\bullet, \Omega))_{k \geq 1}$ , in the case  $\mu$  is a finite measure.

*Proof.* The fact that  $\text{Var}_{p,k}$  is a semi-norm is easy and is left to the reader. The monotony of  $(\text{Var}_{p,k})_{k \geq 1}$  follows directly from the definition. Let  $f \in L^p(\Omega, \mathcal{F}, \mu)$  and  $A$  be a measurable set of finite and positive measure. First note that  $\text{var}_p(f, A) \leq 2\|f\mathbb{1}_A\|_p$ . In fact, using that  $(a+b)^p \leq 2^{p-1}(a^p + b^p)$  holds for all nonnegative numbers  $a, b$  and Fubini's theorem, we have that

$$\begin{aligned} \text{var}_p(f, A)^p &= \frac{1}{\mu(A)} \int_{A \times A} |f(x) - f(y)|^p d\mu(x)d\mu(y) \\ &\leq \frac{2^{p-1}}{\mu(A)} \int_{A \times A} |f(x)|^p + |f(y)|^p d\mu(x)d\mu(y) \\ &= 2^p \|f\mathbb{1}_A\|_p^p. \end{aligned} \tag{3.1}$$

It follows that if  $\mathcal{P}$  is a finite measurable partition of  $A$  then  $\text{Var}_p(f, \mathcal{P}) \leq 2\|f\mathbb{1}_A\|_p$  and therefore, we deduce that  $\text{Var}_{p,k}(f) \leq 2\|f\|_p$ . In particular,  $\text{Var}_{p,k}$  is continuous. In case the measure is finite we have

$$\text{Var}_{p,k}(f, \Omega) \leq \text{Var}_{p,k}(f) \leq 2\|f\|_p.$$

□



**Remark.** Assume that  $\mu$  is a finite measure. We notice that for a fixed function  $f \in L^p(\Omega, \mathcal{F}, \mu)$ , we have  $\lim_{k \rightarrow \infty} \text{Var}_{p,k}(f, \Omega) = 0$ . Indeed, let  $k \in \mathbb{N}$  and define the sets

$$A_i = \left\{ x : \frac{i}{k} \leq f(x) < \frac{i+1}{k} \right\}, \text{ for } i \in \{-k^2, \dots, k^2 - 2\},$$

$A_{k^2-1} = \{x : k - \frac{1}{k} \leq f(x) \leq k\}$  and  $A_{k^2} = \{x : |f(x)| > k\}$ . Then, we have, for all  $i \in \{-k^2, \dots, k^2 - 1\}$

$$\text{var}_p(f, A_i) \leq \frac{\mu(A_i)^{\frac{1}{p}}}{k}$$

and for  $i = k^2$

$$\text{var}_p(f, A_{k^2}) \leq 2\|f \mathbb{1}_{A_{k^2}}\|_p.$$

Thus,

$$\text{Var}_{p,k}(f, \Omega)^p \leq \frac{1}{k^p} \mu(\Omega) + 2^p \int_{|f|>k} |f(x)|^p d\mu(x),$$

and then  $\lim_k \text{Var}_{p,k}(f, \Omega) = 0$ .

We also notice that the same property holds for  $(\text{Var}_{p,k}(f))_k$ , in general measure spaces, but its proof is more involved and we postponed to Corollary 3.4.

The following lemma proves that the variation of a function can always be computed on a  $\sigma$ -finite set if the measure has no atoms of infinite mass.

**Lemma 3.2.** Assume  $(\Omega, \mathcal{F}, \mu)$  is a measurable space such that  $\mu$  has no atoms of infinite mass and  $p \in [1, \infty)$ . Let  $f \in L^p(\Omega, \mathcal{F}, \mu)$  and fix  $k \geq 1$ . Then there exists an increasing sequence of finite measure sets  $(\Omega_n^*)_n \subset \mathcal{F}$  such that

$$\text{Var}_{p,k}(f) = \lim_n \text{Var}_{p,k,n}(f) = \text{Var}_{p,k,*}(f),$$

where

- $\text{Var}_{p,k,*}(f)$  is the total variation of  $f|_{\Omega^*}$  computed in  $(\Omega^*, \mathcal{F}|_{\Omega^*}, \mu|_{\Omega^*})$  with  $\Omega^* = \bigcup_n \Omega_n^*$ ;
- $\text{Var}_{p,k,n}(f)$  is the total variation of  $f|_{\Omega_n^*}$  computed in  $(\Omega_n^*, \mathcal{F}|_{\Omega_n^*}, \mu|_{\Omega_n^*})$ .

*Proof.* We can obviously suppose that  $\mu$  is infinite. Define  $F = \{x : f(x) = 0\}$ . In case  $F$  has finite measure, we define  $\tilde{D} = F$ . If  $F$  has infinite measure, we consider a subset  $\tilde{D} \subset F$  which is  $\sigma$ -finite and of infinite measure. Note that such a set exists. Indeed, take

$$a = \sup_{D \in \mathcal{F}, D \subset F, \mu(D) < \infty} \mu(D).$$

Let us prove that  $a = \infty$ . Consider a sequence  $(D_l)_l$  of subsets of  $F$ , each one of finite measure such that  $\lim_l \mu(D_l) = a$ . It is clear that  $\tilde{D}_l = \cup_{i \leq l} D_i$  is an increasing sequence of sets of finite measure, included in  $F$  which satisfies  $\mu(D_l) \leq \mu(\tilde{D}_l)$ , proving that  $\mu(\tilde{D}_l) \uparrow a$  and  $\tilde{D} = \cup_l \tilde{D}_l$  satisfies  $\mu(\tilde{D}) = a$ . If  $a$  is finite then,  $F \setminus \tilde{D}$  has infinite measure. By hypothesis this set contains a set  $H$  of finite and positive measure. Then  $\mu(\tilde{D} \cup H) = \mu(\tilde{D}) + \mu(H) > a$ , which is a contradiction.

Now consider a sequence of sets of finite measure  $(A_n)_n$  such that

$$\text{Var}_{p,k}(f, A_n) \geq \text{Var}_{p,k}(f) - \frac{1}{n}.$$

For every  $m \geq 1$  the set  $C_m = \{x : |f(x)| > \frac{1}{m}\}$  has finite measure. The set  $\Omega^* = \cup_n A_n \cup \cup_m C_m \cup \tilde{D}$  is  $\sigma$ -finite and it has infinite measure, because  $\mu(\bigcup_m C_m \cup F) = \mu(\Omega) = \infty$ . We consider

$$\Omega_n^* = \begin{cases} \cup_{i \leq n} A_i \cup C_i \cup \tilde{D} & \text{if } \mu(\tilde{D}) < \infty \\ \cup_{i \leq n} A_i \cup C_i \cup \tilde{D}_i & \text{if } \mu(\tilde{D}) = \infty \end{cases},$$

which is an increasing sequence of sets of finite and positive measure, such that  $\Omega_n^* \uparrow \Omega^*$ . Define  $\text{Var}_{p,k,n}(f)$  the total variation of  $f|_{\Omega_n^*}$  computed in  $(\Omega_n^*, \mathcal{F}|_{\Omega_n^*}, \mu|_{\Omega_n^*})$ , that is

$$\text{Var}_{p,k,n}(f) = \sup_{A \in \mathcal{F}, A \subset \Omega_n^*} \text{Var}_{p,k}(f|_{\Omega_n^*}, A) = \sup_{A \in \mathcal{F}, A \subset \Omega_n^*} \text{Var}_{p,k}(f, A) \leq \text{Var}_{p,k}(f).$$

Similarly, we define  $\text{Var}_{p,k,*}(f)$ , which is the total variation of  $f|_{\Omega^*}$  computed in  $(\Omega^*, \mathcal{F}|_{\Omega^*}, \mu|_{\Omega^*})$ . It is clear that for every  $n$ , by construction,

$$\text{Var}_{p,k}(f) - \frac{1}{n} \leq \text{Var}_{p,k}(f, A_n) \leq \text{Var}_{p,k,n}(f) \leq \text{Var}_{p,k,*}(f) \leq \text{Var}_{p,k}(f),$$

and also that  $(\text{Var}_{p,k,n}(f))_n$  is increasing, showing that

$$\text{Var}_{p,k,n}(f) \uparrow \text{Var}_{p,k}(f)$$

and  $\text{Var}_{p,k,*}(f) = \text{Var}_{p,k}(f)$ . □

The next proposition shows that the variation and  $\mathcal{D}_{p,k}$  have the same behaviour. This will be a fundamental tool to characterize the uniform approximability of sets.

**Proposition 3.3.** *Assume  $(\Omega, \mathcal{F}, \mu)$  is a measurable space and  $p \in [1, \infty)$ . For any  $k \geq 1$  and any  $f \in L^p(\Omega, \mathcal{F}, \mu)$ , we have*

(i)

$$\mathcal{D}_{p,k+1}(f) \leq \text{Var}_{p,k}(f) \leq 2\mathcal{D}_{p,k}(f).$$

(ii) *If the measure  $\mu$  is finite, it holds*

$$\mathcal{D}_{p,k}(f) \leq \text{Var}_{p,k}(f, \Omega) \leq \text{Var}_{p,k}(f) \leq 2\mathcal{D}_{p,k}(f) \leq 2\text{Var}_{p,k}(f, \Omega).$$

(iii) *If  $\mu$  has no atoms of infinite mass, we have that*

$$\mathcal{D}_{p,k}(f) \leq \text{Var}_{p,k}(f) \leq 2\mathcal{D}_{p,k}(f).$$

*Proof.* (i) For the upper bound, consider  $g \in \mathcal{G}_k \cap L^p(\Omega, \mathcal{F}, \mu)$  a function such that

$$\mathcal{D}_{p,k}(f)^p = \|f - g\|_p^p.$$

Assume that  $g = \sum_{i=1}^k c_i \mathbb{1}_{A_i}$ , where  $\{A_i\}_{1 \leq i \leq k}$  is a finite partition of  $\Omega$ . Clearly if  $\mu(A_i) = \infty$ , then  $c_i = 0$ . For  $A$  a set of finite measure define a partition of  $A$  by  $\mathcal{P} = \{A \cap A_i\}_{1 \leq i \leq k}$ . Using that  $(a+b)^p \leq 2^{p-1}(a^p + b^p)$  for all positive numbers  $a$  and  $b$ , we get for all  $i$  such that  $\mu(A \cap A_i) > 0$

$$\text{var}_p(f, A \cap A_i)^p = \frac{1}{\mu(A \cap A_i)} \int_{(A \cap A_i) \times (A \cap A_i)} |f(x) - f(y)|^p d\mu(x) d\mu(y) \leq 2^p \int_{A \cap A_i} |f(z) - c_i|^p d\mu(z).$$

Then,

$$\text{var}_p(f, \mathcal{P})^p \leq 2^p \sum_i \int_{A \cap A_i} |f(z) - c_i|^p d\mu(z) \leq 2^p \|f - g\|_p^p.$$

Therefore, we get

$$\text{Var}_{p,k}(f, A) \leq 2\|f - g\|_p = 2\mathcal{D}_{p,k}(f).$$

For the lower bound let  $\varepsilon > 0$  and take a set  $A$  of finite measure such that  $\|f\mathbb{1}_{A^c}\|_p^p < \varepsilon$ . By definition of  $\text{Var}_{p,k}(f, A)$ , there exists a finite partition  $\mathcal{P} = \{A_i\}_{1 \leq i \leq n}$  of  $A$ , with  $n \leq k$ , such that (we assume all the sets in  $\mathcal{P}$  has positive measure)

$$\begin{aligned} \text{var}_p(f, \mathcal{P})^p &= \sum_i \frac{1}{\mu(A_i)} \int_{A_i \times A_i} |f(x) - f(y)|^p d\mu(x) d\mu(y) \\ &\leq (\text{Var}_{p,k}(f, A))^p + \varepsilon \leq (\text{Var}_{p,k}(f))^p + \varepsilon. \end{aligned}$$

For every  $i \leq n$  by the definition of  $\mathcal{M}_p(f, A_i)$ , we have

$$\mathbb{1}_{A_i}(y) \int_{A_i} |f(x) - \mathcal{M}_p(f, A_i)|^p d\mu(x) \leq \mathbb{1}_{A_i}(y) \int_{A_i} |f(x) - f(y)|^p d\mu(x),$$

and therefore, integrating over  $y$  we get

$$\int_{A_i} |f(x) - \mathcal{M}_p(f, A_i)|^p d\mu(x) \leq \frac{1}{\mu(A_i)} \int_{A_i \times A_i} |f(x) - f(y)|^p d\mu(x),$$

and then

$$\sum_i \int_{A_i} |f(x) - \mathcal{M}_p(f, A_i)|^p d\mu(x) \leq \text{var}_p(f, \mathcal{P})^p.$$

Finally, define  $g = \sum_i \mathcal{M}_p(f, A_i) \mathbb{1}_{A_i} + 0 \mathbb{1}_{A^c} \in \mathcal{G}_{p,k+1} \cap L^p(\Omega, \mathcal{F}, \mu)$  to obtain that

$$\|f\mathbb{1}_A - g\|_p^p \leq \text{var}_p(f, \mathcal{P})^p \leq (\text{Var}_{p,k}(f))^p + \varepsilon$$

To finish this part, notice that

$$\begin{aligned} \|f - g\|_p^p &= \|f\mathbb{1}_A - g\mathbb{1}_A\|_p^p + \|f\mathbb{1}_{A^c} - g\mathbb{1}_{A^c}\|_p^p \\ &= \|f\mathbb{1}_A - g\|_p^p + \|f\mathbb{1}_{A^c}\|_p^p \leq (\text{Var}_{p,k}(f))^p + 2\varepsilon \end{aligned}$$

which implies that  $\mathcal{D}_{p,k+1}(f) \leq \text{Var}_{p,k}(f)$ .

(ii) The proof is similar to (i). The upper bound follows immediately from the lower bound to be proved. For the lower estimate, in the above proof we can take  $A = \Omega$ .

(iii) Let  $(\Omega_n^*)_n \subset \mathcal{F}$  and  $\Omega^* = \bigcup_n \Omega_n^*$  given by Lemma 3.2, such that

$$\text{Var}_{p,k}(f) = \lim_n \text{Var}_{p,k,n}(f) = \text{Var}_{p,k,*}(f).$$

We first assume that  $f$  is bounded by some constant  $C > 0$ . Then, using the result we have shown for the finite measure case, we have on  $\Omega_n^*$

$$\inf\{\|f|_{\Omega_n^*} - g\|_p : g \in \mathcal{G}_{p,k}(\Omega_n^*)\} \leq \text{Var}_{p,k,n}(f).$$

By Theorem 2.3, the left hand side is attained at some function  $g_n$  defined in  $\Omega_n^*$ , which is also bounded by  $C$ . We can assume this minimizer has the following form

$$g_n = \sum_{i=1}^{q(n)} b_{i,n} \mathbb{1}_{B_{i,n}},$$

where  $\{B_{i,n}\}_{1 \leq i \leq q(n)}$  is a partition of sets of positive measure of  $\Omega_n^*$  and

$$-C \leq b_{1,n} < \dots < b_{q(n),n} \leq C,$$

$$r_{1,n} = -C - 1, \quad r_{q(n)+1,n} = C + 1, \quad r_{i,n} = \frac{b_{i-1,n} + b_{i,n}}{2} \quad \text{for } i \in \{2, \dots, q(n)\},$$

$$B_{i,n} = f^{-1}([r_{i,n}, r_{i+1,n}]) \cap \Omega_n^* \quad \text{for } i \in \{1, \dots, q(n) - 1\} \quad \text{and} \quad B_{q(n),n} = f^{-1}([r_{q(n),n}, r_{q(n)+1,n}]) \cap \Omega_n^*,$$

$$b_{i,n} = \mathcal{M}_p(f, B_{i,n}) \quad \text{for } i \in \{1, \dots, q(n)\}$$

and  $q(n) \leq k$ . As before, we can assume by passing to a subsequence that  $q(n) = q$  is constant and the vector  $v_n = (r_{1,n}, b_{1,n}, r_{2,n}, \dots, r_{q,n}, b_{q,n}, r_{q+1,n})$  converges to a vector in  $[-C-1, C+1]^{3q}$ , which we denote by  $v = (r_1, b_1, r_2, \dots, r_q, b_q, r_{q+1})$ . Also we denote by  $L = (b_1, \dots, b_q)$  and  $w_1 < \dots < w_m$  the different values in  $L$ , where  $m \leq q$ .

Let us show that  $w_{t^*} = 0$  for some  $t^*$ . For that, remark that  $\Omega_n^* = \cup_{i=1}^q B_{i,n}$  and therefore, there exists an index  $i(n)$ , such that

$$\mu(B_{i(n),n}) \geq \frac{1}{q} \mu(\Omega_n^*),$$

showing that  $\lim_n \mu(B_{i(n),n}) = \infty$ . We can assume that  $i(n) = i$  is constant, by passing to a subsequence if necessary. Using the optimality of  $b_{i,n} = \mathcal{M}_p(f, B_{i,n})$ , we get

$$|b_{i,n}|^p \mu(B_{i,n}) \leq 2^{p-1} \left( \int_{B_{i,n}} |f(x) - \mathcal{M}_p(f, B_{i,n})|^p d\mu(x) + \int_{B_{i,n}} |f(x)|^p d\mu(x) \right) \leq 2^p \|f\|_p^p.$$

This shows that  $b_{i,n} \rightarrow b_i = 0$ , and the claim holds by taking  $t^*$  such that  $w_{t^*} = b_i = 0$ .

Consider  $I_t = \{j \in \{1, \dots, q\} : b_j = w_t\}$  for  $t \in \{1, \dots, m\}$ . Notice that each  $I_t$  is a nonempty interval of  $I = \{1, \dots, q\}$ . Assume that  $I_t = \{l(t), \dots, u(t)\}$ , then we have  $r_{l(t)+1,n} \rightarrow w_t, \dots, r_{u(t),n} \rightarrow w_t, b_{l(t),n} \rightarrow w_t, \dots, b_{u(t),n} \rightarrow w_t$  and

$$\lim_n r_{l(t),n} = r_{l(t)} = \frac{b_{l(t)-1} + w_t}{2} < w_t < \frac{b_{u(t)+1} + w_t}{2} = r_{u(t)+1} = \lim_n r_{u(t)+1,n},$$

with the obvious modifications in the case  $l(t) = 1$  or  $u(t) = q$ . By construction we have for all  $i < l(t^*)$  it holds  $r_{i+1,n} \leq r_{l(t^*),n} < \frac{r_{l(t^*)}}{2} = r_-$ , for all large  $n$ , because  $r_{l(t^*)} < 0$ . Similarly, for all  $i \geq u(t^*)$  we have  $r_{i+1,n} \geq r_{u(t^*)+1,n} > \frac{r_{u(t^*)+1}}{2} = r_+ > 0$ , for all large  $n$ . This implies that, for large  $n$

$$\bigcup_{i < l(t^*)} B_{i,n} \subset f^{-1}([-C-1, r_-]),$$

which is a set of finite measure:  $\mu(f^{-1}([-C-1, r_-])) < \infty$ . Consider a modification of  $g_n$  given by

$$\ell_n = \sum_{i \notin [l(t^*), u(t^*)]} b_i \mathbb{1}_{B_{i,n}} + \sum_{l(t^*) \leq i \leq u(t^*)} b_{i,n} \mathbb{1}_{B_{i,n}}.$$

We have  $\|g_n - \ell_n\|_p$  converges to zero. Indeed, this follows from the inequality

$$\begin{aligned} \|g_n - \ell_n\|_p^p &= \sum_{i \notin \{l(t^*), \dots, u(t^*)\}} |b_i - b_{i,n}|^p \mu(B_{i,n}) \\ &\leq \max_j |b_j - b_{j,n}|^p \mu(f^{-1}([-C-1, r_-]) \cup f^{-1}([r_+, C+1])) \rightarrow 0. \end{aligned}$$

Using the triangular inequality and the optimality of  $g_n$ , we get

$$\|f|_{\Omega_n^*} - g_n\|_p \leq \|f|_{\Omega_n^*} - \ell_n\|_p \leq \|f|_{\Omega_n^*} - g_n\|_p + \|g_n - \ell_n\|_p,$$

and we plan to use Fatou's Lemma. Before doing that, we will fix a subsequence with certain monotonic properties. Since  $r_{1,n}, r_{q+1,n}$  are constant, there is no restriction here. For  $i \in \{2, \dots, q\}$  we choose a subsequence in the following order. If  $(r_{2,n})_n$  has an strictly decreasing subsequence, we consider this as  $n^{(2)}$  and define  $T(2) = \text{sd}$  (for strictly decreasing) otherwise, we consider  $n^{(2)}$  so that  $(r_{2,n})_n$  is increasing along this subsequence and  $T(2) = \text{in}$  (for increasing). Now, we construct  $n^{(3)}$ . If  $(r_{3,n^{(2)}})_{n^{(2)}}$  has an strictly decreasing subsequence we take this as  $n^{(3)}$  and  $T(3) = \text{sd}$ , otherwise we take  $n^{(3)}$  so that  $(r_{3,n^{(3)}})_{n^{(3)}}$  is increasing, and  $T(3) = \text{in}$ . We continue in this way until we define  $n^{(q)}$ . We put  $T(1) = \text{in}$  and  $T(q+1) = \text{sd}$ . We call  $n' = n^{(q)}$ . In this way we have the a.e. convergence

$$\mathbb{1}_{f^{-1}([r_{i,n'}, r_{i+1,n'}]) \cap \Omega_{n'}^*} \rightarrow \begin{cases} \mathbb{1}_{f^{-1}([r_i, r_{i+1}]) \cap \Omega^*} & \text{if } T(i) = \text{in}, T(i+1) = \text{sd} \\ \mathbb{1}_{f^{-1}([r_i, r_{i+1}]) \cap \Omega^*} & \text{if } T(i) = \text{in}, T(i+1) = \text{in} \\ \mathbb{1}_{f^{-1}([r_i, r_{i+1}]) \cap \Omega^*} & \text{if } T(i) = \text{sd}, T(i+1) = \text{sd} \\ \mathbb{1}_{f^{-1}([r_i, r_{i+1}]) \cap \Omega^*} & \text{if } T(i) = \text{sd}, T(i+1) = \text{in} \end{cases}$$

We call  $\mathcal{J}_i$  the interval, with extremes  $r_i, r_{i+1}$ , according to the above classification. An important remark is that  $\cup_i \mathcal{J}_i = [-C-1, C+1]$ .

Using the Dominated Convergence Theorem we conclude that

$$\sum_{i \notin \{l(t^*), \dots, u(t^*)\}} \int_{B_{i,n'}} |f(x) - b_i|^p d\mu(x) \rightarrow \sum_{i \notin \{l(t^*), \dots, u(t^*)\}} \int_{f^{-1}(\mathcal{J}_i)} |f(x) - b_i|^p d\mu(x).$$

On the other hand, using Fatou's Lemma we conclude

$$\begin{aligned} \liminf_{n'} \sum_{i \in \{l(t^*), \dots, u(t^*)\}} \int_{B_{i,n'}} |f(x) - b_{i,n'}|^p d\mu(x) &\geq \int \liminf_{n'} |f(x) - g_n(x)|^p \mathbb{1}_{B_{n'}} d\mu(x) \\ &\geq \int |f(x)|^p \mathbb{1}_{f^{-1}(\tilde{\mathcal{J}}) \cap \Omega^*} d\mu(x). \end{aligned}$$

where  $B_{n'} = \cup_{i \in \{l(t^*), \dots, u(t^*)\}} B_{i,n'}$  and  $\tilde{\mathcal{J}} = \cup_{i \in \{l(t^*), \dots, u(t^*)\}} \mathcal{J}_i$ . Here, we have used that for all  $x \in B_{n'}$  we have

$$|g_{n'}(x)| \leq \max_{i \in \{l(t^*), \dots, u(t^*)\}} |b_{i,n'}| \rightarrow 0.$$

Hence,

$$\begin{aligned} |f(x)| \mathbb{1}_{B_{i,n'}} &\leq |f(x) - g_{n'}(x)| \mathbb{1}_{B_{i,n'}} + \max_{i \in \{l(t^*), \dots, u(t^*)\}} |b_{i,n'}| \mathbb{1}_{B_{i,n'}} \\ |f(x) - g_{n'}(x)| \mathbb{1}_{B_{i,n'}} &\leq |f(x)| \mathbb{1}_{B_{i,n'}} + \max_{i \in \{l(t^*), \dots, u(t^*)\}} |b_{i,n'}| \mathbb{1}_{B_{i,n'}}, \end{aligned}$$

showing that

$$\liminf_{n'} |f(x) - g_{n'}(x)| \mathbb{1}_{B_{i,n'}} = |f(x)| \liminf_{n'} \mathbb{1}_{B_{i,n'}}.$$

Putting all together, we conclude that

$$\text{Var}_{p,k}(f) \geq \liminf_{n'} \|f|_{\Omega_{n'}^*} - g_{n'}\|_p \geq \|f|_{\Omega^*} - \ell\|_p,$$

where the function  $\ell \in \mathcal{G}_{p,q}(\Omega^*)$  is defined on  $\Omega^*$  as

$$\ell = \sum_{i \notin \{l(t^*), \dots, u(t^*)\}} b_i \mathbb{1}_{f^{-1}(\mathcal{J}_i) \cap \Omega^*} + 0 \mathbb{1}_{f^{-1}(\tilde{\mathcal{J}}) \cap \Omega^*}.$$

Notice that for  $i \notin \{l(t^*), \dots, u(t^*)\}$ , we have  $\mu(f^{-1}(\mathcal{J}_i) \cap \Omega^*) \leq \mu(f^{-1}([-C-1, r_-] \cup f^{-1}([r_+, C+1]))) < \infty$  and so  $\mu(f^{-1}(\tilde{\mathcal{J}}) \cap \Omega^*) = \infty$ .

Since  $b_i = 0$  for some  $i$  and  $f = 0$  outside  $\Omega^*$ , we can extend  $\ell$  by 0 outside  $\Omega^*$  and still this extension  $\bar{\ell}$  belongs to  $\mathcal{G}_{p,q} \subset \mathcal{G}_{p,k}$ . So, we get that

$$\mathcal{D}_{p,k}(f) \leq \|f - \bar{\ell}\|_p \leq \text{Var}_{p,k}(f).$$

and the result is shown in the case  $f$  is bounded.

Now, for the general case, consider  $\varepsilon > 0$  and a large  $C > 0$ , such that  $\|f \mathbb{1}_{|f|>C}\|_p \leq \varepsilon$ . From the domination  $\text{Var}_{p,k}(\bullet) \leq 2\|f\|_p$ , and the seminorm property of  $\text{Var}_{p,k}$  we conclude

$$\text{Var}_{p,k}(f \mathbb{1}_{|f|\leq C}) \leq \text{Var}_{p,k}(f) + \text{Var}_{p,k}(f - f \mathbb{1}_{|f|\leq C}) \leq \text{Var}_{p,k}(f) + 2\varepsilon.$$

Using what we have shown, we get there exists and  $\ell \in \mathcal{G}_{p,k}$  such that

$$\|f \mathbb{1}_{|f|\leq C} - \ell\|_p \leq \text{Var}_{p,k}(f \mathbb{1}_{|f|\leq C}) \leq \text{Var}_{p,k}(f) + 2\varepsilon.$$

On the other hand, we have

$$\|f - \ell\|_p \leq \|f \mathbb{1}_{|f|\leq C} - \ell\|_p + \|f \mathbb{1}_{|f|\leq C} - f\|_p \leq \|f \mathbb{1}_{|f|\leq C} - \ell\|_p + \varepsilon,$$

which shows that

$$\mathcal{D}_{p,k}(f) \leq \text{Var}_{p,k}(f) + 3\varepsilon,$$

and the result is shown.  $\square$

**Remark.** *Examples that satisfies (iii) in the previous Proposition are the  $\sigma$ -finite measures. In particular, it can be applied to  $\ell^p = L^p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \delta)$ , where  $\delta$  is the counting measure. But, there are non  $\sigma$ -finite measures that satisfies that hypothesis as well, the counting measures on any uncountable space.*

**Corollary 3.4.** *Assume  $(\Omega, \mathcal{F}, \mu)$  is a measurable space and  $p \in [1, \infty)$ . For all  $f \in L^p(\Omega, \mathcal{F}, \mu)$  it holds that  $\lim_{k \rightarrow \infty} \text{Var}_{p,k}(f) = 0$ .*

*Proof.* This follows directly from the previous proposition since  $\lim_k \mathcal{D}_{p,k}(f) = 0$  by density of the simple functions.  $\square$

A question of some interest is when  $\text{Var}_{p,k}(f) = 0$ , for a function  $f \in L^p(\Omega, \mathcal{F}, \mu)$ . Clearly, if  $f \in \mathcal{G}_{p,k}$  then  $\text{Var}_{p,k}(f) = 0$ . The next result answers the converse.

**Proposition 3.5.** *Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space and  $p \in [1, \infty)$ . Let  $k \geq 1$ . We have*

(i) *if  $\mu$  is a general measure, then  $\mathcal{G}_{p,k} \subset \text{Var}_{p,k}^{-1}(\{0\}) \subset \mathcal{G}_{p,k+1}$ .*

(ii) *if  $\mu$  has no atoms of infinite mass, then  $\text{Var}_{p,k}^{-1}(\{0\}) = \mathcal{G}_{p,k}$ .*

*Proof.* (i). Let  $f \in L^p(\Omega, \mathcal{F}, \mu)$  satisfying  $\text{Var}_{p,k}(f) = 0$ . From (i) of Proposition 3.3, we have

$$\mathcal{D}_{p,k+1}(f) \leq \text{Var}_{p,k}(f) = 0,$$

which implies that  $f = g$  a.e. for some  $g \in \mathcal{G}_{p,k+1}$  (see Theorems 2.3 and 2.7). The other inclusion is obvious.

(ii) The proof is similar to (i) and uses (iii) in Proposition 3.3.  $\square$

**Remark.** *Notice that if  $\mu$  has an atom of infinite mass it may happens that  $\text{Var}_{p,k}(f) = 0$ , but  $f \in \mathcal{G}_{p,k+1} \setminus \mathcal{G}_{p,k}$ . Indeed, assume  $\Omega = \{1, \dots, k+1\}$ , where the mass of each atom in  $\{1, \dots, k\}$  is one and the mass at atom  $\{k+1\}$  is infinite. Every function in  $L^p(\Omega, \mathcal{F}, \mu)$  for  $p \in [1, \infty)$  satisfies  $f(k+1) = 0$ . The function  $f$  given by  $f(i) = i$  for  $i \in \{1, \dots, k\}$  and  $f(k+1) = 0$  belongs to  $\mathcal{G}_{p,k+1} \setminus \mathcal{G}_{p,k}$ . Nevertheless,  $\text{Var}_{p,k}(f) = 0$ , which is exactly the case (i) in Proposition 3.5. Also, this example explains why the lower bound in Proposition 3.3 (i) is computed over  $\mathcal{G}_{p,k+1}$  and not over  $\mathcal{G}_{p,k}$ , in general.*

## 4 Uniform approximability

In this section, we investigate some properties of uniformly approximable sets (see Definition 1.1).

### 4.1 Uniform integrability

In this subsection, we prove that the class of uniform approximable sets is strictly larger than the class of uniform integrable sets. Assume that  $(\Omega, \mathcal{F}, \mu)$  is a measure space and let  $p \in [1, \infty)$ . Remember that a subset  $\mathcal{A} \subset L^p(\Omega, \mathcal{F}, \mu)$  is *uniformly integrable* (in short, UI) if

$$\inf_{g \in L^p_+(\Omega, \mathcal{F}, \mu)} \sup_{f \in \mathcal{A}} \int_{|f| > g} |f(x)|^p d\mu(x) = 0,$$

where  $L^p_+(\Omega, \mathcal{F}, \mu)$  is the set of nonnegative functions in  $L^p(\Omega, \mathcal{F}, \mu)$ . Note that if  $\mu$  is a finite measure, then this definition coincides with the usual one, that is  $\mathcal{A}$  is UI in  $L^p(\Omega, \mathcal{F}, \mu)$  if and only if (see [7], page 254)

$$\lim_{a \rightarrow \infty} \sup_{f \in \mathcal{A}} \int_{|f(x)| \geq a} |f^p(x)| d\mu(x) = 0.$$

**Proposition 4.1.** *Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space and  $p \in [1, +\infty)$ . Assume  $\mathcal{A} \subset L^p(\Omega, \mathcal{F}, \mu)$  is UI. Then,  $\mathcal{A}$  is UA.*

*Proof.* Consider  $\varepsilon > 0$ , and take  $g \in L^p_+(\Omega, \mathcal{F}, \mu)$ , such that

$$\sup_{f \in \mathcal{A}} \int_{|f| > g} |f(x)|^p d\mu(x) \leq \frac{\varepsilon^p}{3}.$$

Fix  $f \in \mathcal{A}$ . Consider  $n \in \mathbb{N}$ , large enough such that  $\int_{g > n} g^p(x) d\mu(x) + \int_{g < \frac{1}{n}} g^p(x) d\mu(x) \leq \frac{\varepsilon^p}{3}$ . The set  $B_n = \{x : \frac{1}{n} \leq g(x) \leq n\} \cap \{x : |f(x)| \leq g(x)\} \subset C_n = \{x : \frac{1}{n} \leq g(x) \leq n\}$  has finite measure. Notice that over  $B_n$  we have  $|f| \leq n$ . Take now  $k \geq 2$  such that  $(\frac{n}{k})^p \mu(C_n) \leq \frac{\varepsilon^p}{3}$  and define

$$A_i = B_n \cap \left\{ x : \frac{ni}{k} \leq f(x) < \frac{n(i+1)}{k} \right\}$$

for  $i \in \{-k, \dots, k-2\}$ ,  $A_{k-1} = B_n \cap \left\{ x : \frac{n(k-1)}{k} \leq f(x) \leq n \right\}$  and

$$h = \sum_{i=-k}^{k-1} \frac{in}{k} \mathbb{1}_{A_i} + 0 \mathbb{1}_{B_n^c} \in \mathcal{G}_{p, 2k+1}.$$

Then, we have

$$\begin{aligned} \int_{B_n} |f(x) - h(x)|^p d\mu(x) &= \sum_{i=-k}^{k-1} \int_{A_i} |f(x) - h(x)|^p d\mu(x) \leq \left(\frac{n}{k}\right)^p \sum_{i=-k}^{k-1} \mu(A_i) \\ &\leq \left(\frac{n}{k}\right)^p \mu(B_n) \leq \left(\frac{n}{k}\right)^p \mu(C_n) \leq \frac{\varepsilon^p}{3}. \end{aligned}$$

On the other hand,  $B_n^c = (C_n^c \cap \{x : |f(x)| \leq g(x)\}) \cup \{x : |f(x)| > g(x)\}$  and so

$$\begin{aligned} \int_{B_n^c} |f(x) - h(x)|^p d\mu(x) &= \int_{B_n^c} |f(x)|^p d\mu(x) = \int_{C_n^c \cap |f| \leq g} |f(x)|^p d\mu(x) + \int_{|f| > g} |f(x)|^p d\mu(x) \\ &\leq \int_{C_n^c} g^p(x) d\mu(x) + \int_{|f| > g} |f(x)|^p d\mu(x) \leq \frac{2\varepsilon^p}{3} \end{aligned}$$

Finally, we have  $\|f - h\|_p \leq \varepsilon$ , and the result is shown.  $\square$

**Remark.** Note that the converse of Proposition 4.1 is not true in general. In fact  $\mathcal{G}_{p,2}$  is UA, but this set is not UI in  $L^p(\Omega, \mathcal{F}, \mu)$  in general. Indeed, assume the space has finite measure and there exists a sequence  $(B_n)_n$  of measurable sets with positive measure such that  $\mu(B_n) \rightarrow 0$ . Then  $f_n = \mu(B_n)^{-1/p} \mathbb{1}_{B_n}$  belongs to  $\mathcal{G}_{p,2}$ , each one has norm 1 and the subfamily  $(f_n)_n$  is not UI, since for all  $a \geq 0$ , we have

$$\sup_n \int_{f_n > a} f_n^p d\mu = 1.$$

The conclusion is that UA is weaker than UI.

**Remark.** If  $(\Omega, \mathcal{F}, \mu)$  is a finite measure space, the following examples are UI in  $L^p(\Omega, \mathcal{F}, \mu)$ , for  $p \in [1, \infty)$ ,

- $\mathcal{A}$  is bounded in  $L^q(\Omega, \mathcal{F}, \mu)$  for some  $q > p$ ;
- $\mathcal{A}$  is bounded by a fixed function  $g \in L^p(\Omega, \mathcal{F}, \mu)$ .

The following result can be prove using the fact that totally boundedness implies UI in  $L^p(\Omega, \mathcal{F}, \mu)$  for  $p \in [1, \infty)$ . However, since the case  $p = \infty$  needs a proof, we give a more direct argument:

**Proposition 4.2.** Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space and  $p \in [1, +\infty]$ . If  $\mathcal{A}$  is totally bounded in  $L^p(\Omega, \mathcal{F}, \mu)$  then  $\mathcal{A}$  is UA.

*Proof.* Let  $\varepsilon > 0$ . There exist finitely many functions  $f_1, \dots, f_n$  such that  $\mathcal{A} \subset \bigcup_{j=1}^n B(f_j, \varepsilon)$ . By density of the simple functions, there exist  $k \in \mathbb{N}$  and  $g_j \in \mathcal{G}_{p,k}$  such that  $\|f_j - g_j\|_p \leq \varepsilon$  for all  $j \in \{1, \dots, n\}$ . Now if  $f \in \mathcal{A}$  then, there exists  $j_0 = j_0(f) \in \{1, \dots, n\}$  such that  $\|f - f_{j_0}\|_p \leq \varepsilon$ . It follows that

$$\|f - g_{j_0}\|_p \leq \|f - f_{j_0}\|_p + \|f_{j_0} - g_{j_0}\|_p \leq 2\varepsilon$$

and the proof is complete.  $\square$

## 4.2 Characterization of the uniform approximability

If  $M$  is a metric space, we recall that the covering numbers of  $M$  are defined for every  $\varepsilon > 0$  by

$$\mathcal{N}(M, \varepsilon) = \inf \{N \geq 1 : M \text{ can be covered by } N \text{ closed balls of radius } \varepsilon\}.$$

For more informations about covering numbers and its applications to Machine Learning, we refer the reader to [15] and [2].

If  $(\Omega, \mathcal{F}, \mu)$  is a measure space, we define the covering numbers of a measurable function  $f$  by

$$\mathcal{N}(f, \varepsilon) = \inf \{\mathcal{N}(g(\Omega), \varepsilon) : g \text{ measurable function such that } f = g \text{ a.e.}\}.$$

This notion allows us to characterize the uniform approximability in  $L^\infty(\Omega, \mathcal{F}, \mu)$  in terms of uniformly bounded covering numbers. Before doing that, we notice that if  $f \in L^\infty(\Omega, \mathcal{F}, \mu)$ , then  $\mathcal{N}(f, \varepsilon) < \infty$ . Indeed, we know that  $|f| \leq \|f\|_\infty$  holds a.e., so by considering  $g = f \mathbb{1}_{\{|f| \leq \|f\|_\infty\}}$ , we have  $g = f$  a.e. and

$$\mathcal{N}(f, \varepsilon) \leq \frac{2}{\varepsilon} \|f\|_\infty + 1.$$

If  $f$  is a measurable function and  $\mathcal{N}(f, \varepsilon) < \infty$  then  $f \in L^\infty(\Omega, \mathcal{F}, \mu)$ . On the other hand, by definition of infimum, there exists a measurable function  $g$  such that  $f = g$  a.e., and

$$\mathcal{N}(f, \varepsilon) \leq \mathcal{N}(g(\Omega), \varepsilon) \leq \mathcal{N}(f, \varepsilon) + \frac{1}{2},$$

showing that  $\mathcal{N}(f, \varepsilon) = \mathcal{N}(g(\Omega), \varepsilon)$ .



**Theorem 4.3.** Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space and let  $\mathcal{A} \subset L^\infty(\Omega, \mathcal{F}, \mu)$ . The following assertions are equivalent:

- (i)  $\mathcal{A}$  is UA;
- (ii)  $\sup_{f \in \mathcal{A}} \mathcal{N}(f, \varepsilon) < \infty$  for all  $\varepsilon > 0$ .

In this case, we have that  $N_{\infty, \varepsilon}(\mathcal{A}) = \sup_{f \in \mathcal{A}} \mathcal{N}(f, \varepsilon)$  for all  $\varepsilon > 0$ .

*Proof.* Let  $\varepsilon > 0$  and suppose  $\sup_{f \in \mathcal{A}} \mathcal{N}(f, \varepsilon) = \infty$ . Fix  $k \geq 1$  and choose  $f \in \mathcal{A}$  such that  $\mathcal{N}(f, \varepsilon) \geq 10(k+1)$ . Changing the representant of  $f$  is necessary, we can suppose that  $m_\varepsilon := \mathcal{N}(f, \varepsilon) = \mathcal{N}(f(\Omega), \varepsilon)$ . So there exists a collection  $\mathcal{J}$  of closed balls  $\{I_i = [a_i, b_i]\}_{1 \leq i \leq m_\varepsilon}$ , of radius  $\varepsilon$ , such that

$$f(\Omega) \subset \bigcup_{i=1}^{m_\varepsilon} I_i$$

Using the minimality of this covering each interval cannot be covered by the other intervals, so for  $i \neq j$  we have  $I_i \setminus I_j \neq \emptyset$ . Consider the measurable sets  $A_i = f^{-1}(I_i)$ . We shall prove that the minimality of  $\mathcal{J}$  implies that  $\mu(A_i) > 0$ . Indeed, assume that for some  $i$  we have  $\mu(A_i) = 0$ . Take any  $j \neq i$  (notice that we have assumed that  $m_\varepsilon$  is at least  $10(k+1) > 2$ ) and  $a \in I_j \setminus I_i$ . The measurable function

$$h = f \mathbb{1}_{A_j^c} + a \mathbb{1}_{A_i}$$

coincides with  $f$  up to measure 0 and  $h(\Omega) \subset \bigcup_{r \neq i} I_r$ , so  $\mathcal{N}(h(\Omega), \varepsilon) \leq m_\varepsilon - 1$ , which is a contradiction.

We say that a subcollection  $\mathcal{C} \subset \mathcal{J}$  is  $\varepsilon$ -separated if for two different intervals  $I, J \in \mathcal{C}$ , we have the distance between them  $d(I, J)$  is greater than  $\varepsilon$ . Notice that a collection with only one interval from  $\mathcal{J}$  is  $\varepsilon$ -separated. Take  $\mathcal{C}^*$  a maximal  $\varepsilon$ -separated subcollection with respect to inclusion. Now, if  $I \in \mathcal{J} \setminus \mathcal{C}^*$  there exists an interval  $L = [a, b] \in \mathcal{C}^*$  such that  $d(I, L) \leq \varepsilon$ , otherwise the maximality of  $\mathcal{C}^*$  is contradicted. Then,

$$I \subset [a - 4\varepsilon, a - 2\varepsilon] \cup [a - 2\varepsilon, a] \cup [a, b] \cup [b, b + 2\varepsilon] \cup [b + 2\varepsilon, b + 4\varepsilon],$$

showing that the collection

$$\mathcal{D} = \{[a_i - 4\varepsilon, a_i - 2\varepsilon], [a_i - 2\varepsilon, a_i], [a_i, b_i], [b_i, b_i + 2\varepsilon], [b_i + 2\varepsilon, b_i + 4\varepsilon] : [a_i, b_i] \in \mathcal{C}^*\},$$

is a covering of  $f(\Omega)$  with closed balls of radius  $\varepsilon$ . Therefore

$$m_\varepsilon \leq |\mathcal{D}| \leq 5|\mathcal{C}^*|,$$

showing that  $n = |\mathcal{C}^*| \geq \frac{1}{5}m_\varepsilon$  (here  $|\mathcal{C}^*|$  is the cardinal of  $\mathcal{C}^*$ ).

Consider now  $g \in \mathcal{G}_{\infty, k}$ . We say that an interval  $I \in \mathcal{C}^*$  is unmarked if  $d(g(\Omega), I) > \varepsilon$ . There are at least  $n - 2k \geq \frac{1}{5}m_\varepsilon - 2k = 2(\frac{1}{10}m_\varepsilon - k) > 1$  unmarked intervals in  $\mathcal{C}^*$ . Consider  $I_i \in \mathcal{C}^*$  any unmarked interval, then for all  $x \in A_i = f^{-1}(I_i)$ , we have

$$|f(x) - g(x)| > \varepsilon.$$

Since  $\mu(A_i) > 0$ , we conclude that  $\|f - g\|_\infty > \varepsilon$  and therefore  $N_{\infty, \varepsilon}(\mathcal{A}) > \frac{m_\varepsilon}{10} - 1$ , showing that  $N_{\infty, \varepsilon}(\mathcal{A}) = \infty$ . So we have proved that (i) implies (ii).

Now, let us show that (ii) implies (i). So, we are assuming that  $M_\varepsilon = \sup_{f \in \mathcal{A}} \mathcal{N}(f, \varepsilon) < \infty$  for all  $\varepsilon > 0$ . Fix  $\varepsilon > 0$  and let  $f \in \mathcal{A}$ . Suppose that  $m_\varepsilon := \mathcal{N}(f, \varepsilon) = \mathcal{N}(f(\Omega), \varepsilon) \leq M_\varepsilon$ . Again we can write

$$f(\Omega) \subset \bigcup_{i=1}^{m_\varepsilon} I_i$$

where  $I_i = [a_i, b_i]$  are closed balls of radius  $\varepsilon$ . We assume that the left extremes are ordered increasingly:  $a_1 < a_2 < \dots < a_{m_\varepsilon}$ . We define recursively  $\tilde{a}_1 = a_1, \tilde{b}_1 = b_1$  and for  $i \geq 2$

$$\tilde{a}_i = \max\{b_{i-1}, a_i\}, \tilde{b}_i = b_i.$$

Define  $\tilde{I}_i = [\tilde{a}_i, \tilde{b}_i]$  for  $i \in \{1, \dots, m_\varepsilon\}$ . The fact that every interval  $I_i$  cannot be covered by the intervals  $\{I_j\}_{j \neq i}$  allows us to show the following facts about the new intervals  $\{\tilde{I}_i\}_{1 \leq i \leq m_\varepsilon}$

$$\forall i \quad \tilde{I}_i = I_i \setminus \bigcup_{j=1}^{i-1} [a_j, b_j], \quad \forall i \quad \bigcup_{j=1}^i I_j = \bigcup_{j=1}^i \tilde{I}_j,$$

$$\text{int}(\tilde{I}_i) = (\tilde{a}_i, \tilde{b}_i) \neq \emptyset,$$

$$\forall i < j : \tilde{I}_i \cap \tilde{I}_j \subset \begin{cases} \emptyset & \text{if } j - i \geq 2 \\ \{\tilde{b}_i\} & \text{if } j = i + 1 \end{cases}.$$

Thus,  $\{\tilde{I}_i\}_{1 \leq i \leq m_\varepsilon}$  is a collection of closed balls of radii at most  $\varepsilon$ , that covers  $f(\Omega)$ , which are disjoint except for consecutive intervals that can intersect at one extreme.

With this new intervals we can produce a partition of  $f(\Omega)$ , by choosing  $\hat{I}_1 = \tilde{I}_1$  and for  $i \geq 2$

$$\hat{I}_i = \begin{cases} (\tilde{a}_i, \tilde{b}_i] & \text{if } \tilde{I}_i \cap \tilde{I}_{i-1} \neq \emptyset \\ \tilde{I}_i & \text{otherwise} \end{cases}$$

We now define  $\hat{A}_i = f^{-1}(\hat{I}_i)$ , which is a partition of  $\Omega$  (maybe some of them are empty). If  $x \in \hat{A}_i$  then  $f(x) \in \hat{I}_i \subset I_i = [a_i, b_i]$  and therefore  $|f(x) - \frac{a_i + b_i}{2}| \leq \varepsilon$ . Define the simple function

$$g(x) = \sum_{i=1}^{m_\varepsilon} \frac{a_i + b_i}{2} \mathbb{1}_{\hat{A}_i},$$

that belongs to  $\mathcal{G}_{\infty, m_\varepsilon}$  and satisfies for all  $x \in \Omega$

$$|f(x) - g(x)| \leq \varepsilon,$$

showing that  $\|f - g\|_\infty \leq \varepsilon$ . We conclude that  $N_{\infty, \varepsilon}(\mathcal{A}) \leq \sup_{f \in \mathcal{A}} \mathcal{N}(f, \varepsilon)$ .

To finish, we prove that  $N_{\infty, \varepsilon}(\mathcal{A}) = \sup_{f \in \mathcal{A}} \mathcal{N}(f, \varepsilon)$ . For that purpose consider  $k = N_{\infty, \varepsilon}(\mathcal{A})$ , which means that for all  $f \in \mathcal{A}$ , there exists  $g \in \mathcal{G}_{\infty, k}$ , such that  $\|f - g\|_\infty \leq \varepsilon$ . We assume that  $g = \sum_{i=1}^k c_i \mathbb{1}_{B_i}$ , where  $(B_i)_{i=1}^k$  is a partition of  $\Omega$ . For any  $i \in \{1, \dots, k\}$  we have

$$\|(f - c_i) \mathbb{1}_{B_i}\|_\infty \leq \|f - g\|_\infty \leq \varepsilon,$$

which means that  $A_i = \{x \in B_i : |f(x) - c_i| > \varepsilon\}$  is a measurable set of measure 0. Since  $\mu(\Omega) > 0$ , not all the sets  $B_j$  can have measure 0, so we assume without loss of generality that  $\mu(B_1) > 0$ .

Consider  $B = \Omega \setminus \bigcup_{i=1}^k A_i$ ,  $h = f \mathbb{1}_B + c_1 \mathbb{1}_{B^c}$  and  $\tilde{g} = c_1 \mathbb{1}_{B_1 \cup B^c} + \sum_{i=2}^k c_i \mathbb{1}_{B_i \setminus A_i}$ . We notice that  $f = h$  a.e. and  $\tilde{g} = g$  a.e. On the other hand,  $B_1 \cup B^c, B_2 \setminus A_2, \dots, B_k \setminus A_k$  is a partition and  $\tilde{g} \in \mathcal{G}_{\infty, k}$ . Also, it is clear that  $B_1 \setminus A_1, \dots, B_k \setminus A_k, B^c$  is a partition and

$$\tilde{g} = g \mathbb{1}_B + c_1 \mathbb{1}_{B^c}.$$

With these modifications, we have for all  $x \in \Omega$

$$|h(x) - \tilde{g}(x)| \leq \varepsilon.$$

This is clear for  $x \in B$ . For  $x \in B^c$ , we have  $h(x) = c_1 = \tilde{g}(x)$  and the claim is shown. Finally, the collection of closed ball of radius  $\varepsilon$  given by:  $\{[c_i - \varepsilon, c_i + \varepsilon]\}_{1 \leq i \leq k}$  is an  $\varepsilon$ -cover of  $h(\Omega)$ , showing that  $\mathcal{N}(f, \varepsilon) \leq k$ . The conclusion is that

$$\sup_{f \in \mathcal{A}} \mathcal{N}(f, \varepsilon) \leq k = N_{\infty, \varepsilon}(\mathcal{A}),$$

and the result is shown.  $\square$

**Corollary 4.4.** *Let  $(\Omega, \mathcal{F}, \mu)$  be a finite measure space and  $\mathcal{A}$  be a set of measurable functions. Assume that  $\mathcal{A}$  is UA in  $L^q(\Omega, \mathcal{F}, \mu)$  for some  $q \in [1, \infty]$ , then  $\mathcal{A}$  is UA in  $L^p(\Omega, \mathcal{F}, \mu)$  for all  $p \in [1, q]$  and for all  $\varepsilon > 0$  it holds*

$$N_{p, \varepsilon}(\mathcal{A}) \leq N_{q, \varepsilon \mu(\Omega)^{-r}}(\mathcal{A}),$$

where  $r = \frac{1}{p} - \frac{1}{q}$ .

*In particular if  $\sup_{f \in \mathcal{A}} \mathcal{N}(f, \varepsilon) < \infty$  for all  $\varepsilon > 0$ , then  $\mathcal{A}$  is UA in  $L^p(\Omega, \mathcal{F}, \mu)$  for all  $p \in [1, \infty]$  and for all  $\varepsilon > 0$  it holds  $N_{p, \varepsilon}(\mathcal{A}) \leq \sup_{f \in \mathcal{A}} \mathcal{N}(f, \varepsilon \mu(\Omega)^{\frac{1}{p}})$ .*

*Proof.* This is a direct consequence of Hölder's inequality. In fact, assume that  $p \leq q$  and consider  $g \in \mathcal{G}_{p, k}$ ,  $f \in \mathcal{A}$  then, we have

$$\|f - g\|_p \leq \|f - g\|_q (\mu(\Omega))^r$$

where  $r = \frac{1}{p} - \frac{1}{q}$ . From this it follows that  $\mathcal{A}$  is UA in  $L^p(\Omega, \mathcal{F}, \mu)$  and

$$N_{p, \varepsilon}(\mathcal{A}) \leq N_{q, \varepsilon \mu(\Omega)^{-r}}(\mathcal{A}).$$

The second assertion follows from Theorem 4.3  $\square$

The previous result gives a large class of UA sets when the measure is finite. For example suppose that  $\Omega$  is a bounded metric space,  $\mathcal{F}$  is the Borel  $\sigma$ -algebra and  $\mu$  is a finite measure on  $\mathcal{F}$ . Then the set of 1-Lipschitz functions is UA in  $L^p(\Omega, \mathcal{F}, \mu)$  for any  $p \in [1, +\infty]$ .

The following result is a characterization of UA in  $L^p$  for  $p \in [1, \infty)$ , where we shall prove that a class is UA if and only  $\text{Var}_{p, k}(f)$  converges toward 0, when  $k \rightarrow \infty$ , uniformly in the class.

**Theorem 4.5.** *Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space,  $p \in [1, \infty)$  and let  $\mathcal{A} \subset L^p(\Omega, \mathcal{F}, \mu)$ . Then, the following are equivalent*

- (i)  $\mathcal{A}$  is UA in  $L^p(\Omega, \mathcal{F}, \mu)$ ;
- (ii)  $\lim_{k \rightarrow \infty} \sup_{f \in \mathcal{A}} \text{Var}_{p, k}(f) = 0$ .

*In this case if we define  $r_\varepsilon(\mathcal{A}) = \min\{k \in \mathbb{N} : \sup_{f \in \mathcal{A}} \text{Var}_{p, k}(f) \leq \varepsilon\}$ , we have that for all  $\varepsilon > 0$*

$$r_{2\varepsilon}(\mathcal{A}) \leq N_{p, \varepsilon}(\mathcal{A}) \leq r_\varepsilon(\mathcal{A}) + 1.$$

*Moreover, if the measure  $\mu$  is finite both properties (i), (ii) are equivalent to*

- (iii)  $\lim_{k \rightarrow \infty} \sup_{f \in \mathcal{A}} \text{Var}_{p, k}(f, \Omega) = 0$ .

In this case if we define  $m_\varepsilon(\mathcal{A}) = \min\{k \in \mathbb{N} : \sup_{f \in \mathcal{A}} \text{Var}_{p,k}(f, \Omega) \leq \varepsilon\}$ , we have that for all  $\varepsilon > 0$

$$m_{2\varepsilon}(\mathcal{A}) \leq r_{2\varepsilon}(\mathcal{A}) \leq N_{p,\varepsilon}(\mathcal{A}) \leq m_\varepsilon(\mathcal{A}) \leq r_\varepsilon(\mathcal{A})$$

*Proof.* Suppose that  $\mathcal{A}$  is UA and fix  $\varepsilon > 0$ . Then we have that  $\mathcal{D}_{p,k}(f) \leq \varepsilon$  for all  $f \in \mathcal{A}$ , where  $k = N_{p,\varepsilon}(\mathcal{A})$ . By Proposition 3.3, we deduce that  $\text{Var}_{p,k}(f) \leq 2\varepsilon$  for all  $f \in \mathcal{A}$ . It follows that  $r_{2\varepsilon}(\mathcal{A}) \leq N_{p,\varepsilon}(\mathcal{A})$ , implying that (ii) holds. Now suppose that (ii) holds. Using Proposition 3.3 again, it is easy to see that  $N_{p,\varepsilon}(\mathcal{A}) \leq r_\varepsilon(\mathcal{A}) + 1$ , from what we deduce that (i) is true. In the case that  $\mu$  is finite, the equivalence between (ii) and (iii) and the last assertion of the theorem follow directly from Proposition 3.3.  $\square$

### 4.3 The unit ball of $L^p$

Now we investigate when the unit ball of  $L^p(\Omega, \mathcal{F}, \mu)$ , denoted by  $B_{L^p(\Omega, \mathcal{F}, \mu)} = \{f \in L^p : \|f\|_p \leq 1\}$ , is UA. The case  $p = \infty$  is simple:

**Proposition 4.6.** *Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space. Then  $B_{L^\infty(\Omega, \mathcal{F}, \mu)}$  is UA. More precisely we have that  $N_{\infty,\varepsilon}(B_{L^\infty(\Omega, \mathcal{F}, \mu)}) \leq \lceil \frac{2}{\varepsilon} \rceil + 1$  (where  $\lceil \cdot \rceil$  is the integer part) for all  $\varepsilon > 0$ .*

*Proof.* It is a direct consequence of theorem 4.3.  $\square$

The main objective of this section is to prove the following result:

**Theorem 4.7.** *Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space and  $p \in [1, +\infty)$ . The following assertions are equivalent:*

- (i)  $B_{L^p(\Omega, \mathcal{F}, \mu)}$  is UA;
- (ii)  $L^p(\Omega, \mathcal{F}, \mu)$  is finite dimensional;
- (iii)  $\mu$  is atomic and has only a finite number of atoms with finite measure, up to measure 0.

More precisely, if the previous assertions are false then  $N_{p,\varepsilon}(B_{L^p(\Omega, \mathcal{F}, \mu)}) = \infty$  for all  $\varepsilon \in (0, 1)$ .

This theorem will be proved thanks to several intermediary results. We start with the following result:

**Proposition 4.8.** *Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space and  $p \in [1, +\infty)$ . Suppose that there exists a sequence of disjoint measurable sets  $(A_n)_n$  of positive measure such that  $\mu(A_n) \rightarrow 0$ . Then  $B_{L^p(\Omega, \mathcal{F}, \mu)}$  is not UA. More precisely, we have that  $N_{p,\varepsilon}(B_{L^p(\Omega, \mathcal{F}, \mu)}) = \infty$  for all  $\varepsilon \in (0, 1)$ .*

*Proof.* We are going to prove that for all  $k \geq 1$

$$\sup_{f \in L^p(\Omega, \mathcal{F}, \mu), f \neq 0} \inf_{h \in \mathcal{G}_{p,k}} \frac{\|f - h\|_p^p}{\|f\|_p^p} = 1. \quad (4.1)$$

Note that this equality implies easily that  $N_{p,\varepsilon}(B_{L^p(\Omega, \mathcal{F}, \mu)}) = \infty$  for all  $\varepsilon \in (0, 1)$ . Consider  $r \geq 2$  and consider a subsequence  $(n_k)_k$  such that  $\mu(A_{n_k}) \in (r^{-n_k-1}, r^{-n_k}]$ . We further assume that  $n_{k+1} - n_k \geq 2$ . With this sequence we consider

$$f_N(x) = \sum_{j=1}^N \mu(A_{n_j})^{-\frac{1}{p}} \mathbb{1}_{A_{n_j}}(x).$$

for all  $N \geq 1$  and note that  $\|f_N\|_p^p = N$ . Let  $h \in \mathcal{G}_{p,k}$  and  $N > 2 + 2k$ . We say that an index  $1 < j < N$  is unmarked if  $\text{Im}(h) \cap \left(r^{\frac{n_j-1}{p}}, r^{\frac{n_j+1}{p}}\right) = \emptyset$ . Note that there are at least  $N - 2 - 2k$  unmarked indexes. For such unmarked index  $j$ , we have for  $x \in A_{n_j}$

$$f_N(x) = \mu(A_{n_j})^{\frac{-1}{p}} \geq r^{\frac{n_j}{p}} > r^{\frac{n_j-1}{p}} > r^{\frac{n_j-1}{p}} \geq h(x), \text{ or}$$

$$f_N(x) = \mu(A_{n_j})^{\frac{-1}{p}} < r^{\frac{n_j+1}{p}} < r^{\frac{n_j+1}{p}} \leq h(x).$$

In the first case we have

$$f_N(x) - h(x) \geq f_N(x) - r^{\frac{n_j-1}{p}} \geq f_N(x) - r^{\frac{-1}{p}} f_N(x) = \left(1 - r^{\frac{-1}{p}}\right) f_N(x).$$

In the second case we get

$$h(x) - f_N(x) \geq r^{\frac{n_j+1}{p}} - f_N(x) \geq r^{\frac{n_j+2}{p}} - f_N(x) = r^{\frac{1}{p}} r^{\frac{n_j+1}{p}} - f_N(x) > \left(r^{\frac{1}{p}} - 1\right) f_N(x).$$

Notice that

$$\theta = 1 - r^{\frac{-1}{p}} = \frac{r^{\frac{1}{p}} - 1}{r^{\frac{1}{p}}} < r^{\frac{1}{p}} - 1.$$

So, we have  $|f_N(x) - h(x)| \geq \theta f_N(x)$  on  $A_{n_j}$ . Then, we conclude that

$$\|f_N - h\|_p^p \geq \sum_{j:\text{unmarked}} \int_{A_{n_j}} |f_N(x) - h(x)|^p dx \geq \theta^p \sum_{j:\text{unmarked}} \int_{A_{n_j}} |f_N(x)|^p dx \geq \theta^p (N - 2 - 2k).$$

Hence, we have

$$\frac{\|f_N - h\|_p^p}{\|f_N\|_p^p} \geq \theta^p \frac{N - 2 - 2K}{N},$$

and we get

$$\sup_{f \in L^p(\Omega, \mathcal{F}, \mu), f \neq 0} \inf_{h \in \mathcal{G}_{p,k}} \frac{\|f - h\|_p^p}{\|f\|_p^p} \geq \theta^p = (1 - r^{\frac{-1}{p}})^p.$$

Now, it is enough to make  $r \uparrow \infty$ . □

An immediate corollary is obtained for diffuse measures.

**Corollary 4.9.** *Assume that  $\mu$  is a non trivial diffuse measure, then for all  $p \in [1, \infty)$  the unit ball  $B_{L^p(\Omega, \mathcal{F}, \mu)}$  is not UA and  $N_{p,\varepsilon}(B_{L^p(\Omega, \mathcal{F}, \mu)}) = \infty$ , for all  $\varepsilon \in (0, 1)$ .*

*Proof.* This follows directly from Sierpiński's theorem (see [12]). In fact, consider a measurable set  $B_0$  such that  $0 < \mu(B_0) = a < \infty$  (if such set does not exist then  $\Omega$  is an atom of  $\mu$ ). Then, there exists  $B_1 \subset B_0$  such that  $\mu(B_1) = \frac{a}{2}$ . Applying the same idea to  $B_0 \setminus B_1$ , there exists  $B_2 \subset (B_0 \setminus B_1)$  such that  $\mu(B_2) = \frac{\mu(B_0 \setminus B_1)}{2} = \frac{a}{4}$ . Inductively, we construct a sequence of disjoint subsets  $(B_k)_k$  such that

$$B_{k+1} \subset B_0 \setminus \bigcup_{i=1}^k B_i,$$

and

$$\mu(B_{k+1}) = \frac{\mu\left(B_0 \setminus \bigcup_{i=1}^k B_i\right)}{2} = \frac{a}{2^{k+1}}$$

for all  $k \in \mathbb{N}$ . The result follows from the previous Proposition. □

**Proposition 4.10.** *Assume that  $(\Omega, \mathcal{F}, \mu)$  is an atomic measure space and  $p \in [1, \infty)$ . Then the following are equivalent:*

- (i)  $\mu$  has a finite number of atoms of finite measure, up to measure 0;
- (ii) The space  $L^p(\Omega, \mathcal{F}, \mu)$  is finite dimensional;
- (iii) The unit ball  $B_{L^p(\Omega, \mathcal{F}, \mu)}$  is UA.

Moreover, if the previous assertions are false then  $N_{p, \varepsilon}(B_{L^p(\Omega, \mathcal{F}, \mu)}) = \infty$  for all  $\varepsilon \in (0, 1)$ .

*Proof.* Assume (i) holds. Denote by  $\{A_k\}_{1 \leq k \leq n}$  a finite collection of atoms of finite measure, such that all other atom  $C$  of finite measure coincides with some of them up to measure 0. Take  $B = \Omega \setminus \cup_{k=1}^n A_k$ . If  $\mu(B) > 0$  there exists an atom  $C \subset B$ . This atom  $C$  satisfies that  $\mu(C \setminus A_k) = \mu(C) > 0$  and it cannot coincide with  $A_k$  up to measure 0. Then  $C$  has infinite measure. Then either  $\mu(B) = 0$  or  $\mu(B) = \infty$  and contains no measurable subset of positive finite measure. Then,  $L^p(\Omega, \mathcal{F}, \mu)$  is generated by the finite collection  $\{\mathbb{1}_{A_k}\}_{1 \leq k \leq n}$ , so (ii) holds. Clearly (ii)  $\Rightarrow$  (iii).

So, for the rest of the proof we assume that there exists a countable collection of disjoint atoms  $(A_n)_n$  each one of finite positive measure. Here there are two different situations. The first one is the existence of an infinite subsequence of atoms  $(A_{n_k})_k$  such that  $\mu(A_{n_k}) \rightarrow 0$ . Then, we can apply Theorem 4.8, to conclude that the unit ball  $B_{L^p(\Omega, \mathcal{F}, \mu)}$  is not UA.

The second possibility is the existence of a constant  $a > 0$  such that  $\mu(A_n) \geq a$ , for all  $n$ . We now proceed to prove that  $B_{L^p(\Omega, \mathcal{F}, \mu)}$  is not UA. We do it for  $p = 1$ , the other cases are treated similarly.

In what follows we fix  $k \geq 2$  and  $R > 1$ , and we consider the partial sums

$$S_i = \sum_{2^{i-1} \leq j < 2^i} \mu(A_j) \geq a2^{i-1},$$

for  $i \geq 1$ , and we construct a strictly increasing sequence of integers  $(t_q)_q$  such that the interval  $[R^{t_q}, R^{t_q+1})$  contains at least one of these partial sums. We call  $S_{i_q}$  any such partial sums, for example the smallest one, that is, for  $q$  such that  $[R^{t_q}, R^{t_q+1}) \cap \{S_i\}_{i \geq 1} \neq \emptyset$ , we take

$$i_q = \min\{r \in \mathbb{N} : R^{t_q} \leq S_r < R^{t_q+1}\}.$$

We also define

$$B_q = \bigcup_{j=2^{i_q-1}}^{2^{i_q}-1} A_j,$$

the union of the atoms that has mass  $S_{i_q}$ . We consider the function

$$f = \sum_{q=3}^{M+2} R^{-t_q} \mathbb{1}_{B_q},$$

where  $M$  is a large integer. For the moment we choose  $M > 2kR$ . Take  $h \in \mathcal{G}_{1, k}$  and as before we say that  $3 \leq q \leq M+2$  is an unmarked index if

$$\text{Im}(h) \cap (R^{-t_q-1}, R^{-t_q+1}) = \emptyset.$$

There are at least  $M - 2k$  unmarked indexes. For an unmarked index  $q$  and  $x \in B_q$ , we either have

$$f(x) - h(x) \geq R^{-t_q} - R^{-t_q-1} \geq f(x) \left(1 - \frac{1}{R}\right) = f(x) \frac{R-1}{R}, \text{ or}$$

$$h(x) - f(x) \geq R^{-t_q+1} - R^{-t_q} \geq f(x)(R-1).$$

In any case, we have for  $x \in B_q$

$$|f(x) - h(x)| \geq f(x) \frac{R-1}{R},$$

and then

$$\begin{aligned} \|f - h\|_1 &\geq \frac{R-1}{R} \sum_{q: \text{ unmarked}} R^{-t_q} \mu(B_q) = \frac{R-1}{R} \left( \sum_q R^{-t_q} \mu(B_q) - \sum_{q: \text{ marked}} R^{-t_q} \mu(B_q) \right) \\ &\geq \frac{R-1}{R} (\|f\|_1 - 2kR) = \|f\|_1 \frac{R-1}{R} \left( 1 - \frac{2kR}{\|f\|_1} \right) \end{aligned}$$

Now, we estimate the norm of  $f$ . Clearly, we have  $\|f\|_1 = \sum_q R^{-t_q} \mu(B_q)$ , which gives the lower estimate

$$M \leq \|f\|_1, \quad (4.2)$$

and then the lower bound

$$\|f - h\|_1 \geq \|f\|_1 \frac{R-1}{R} \left( 1 - \frac{2kR}{M} \right). \quad (4.3)$$

So, we conclude that for  $\tilde{f} = f/\|f\|_1$

$$\inf\{\|\tilde{f} - g\|_1 : g \in \mathcal{G}_{1,k}\} \geq \frac{R-1}{R} \left( 1 - \frac{2kR}{M} \right),$$

and therefore

$$\sup_{f \in B_{L^1(\Omega, \mathcal{F}, \mu)}} \inf\{\|f - g\|_1 : g \in \mathcal{G}_{1,k}\} \geq \frac{R-1}{R} \left( 1 - \frac{2kR}{M} \right)$$

Taking  $M \uparrow \infty$ , we conclude that

$$\sup_{f \in B_{L^1(\Omega, \mathcal{F}, \mu)}} \inf\{\|f - g\|_1 : g \in \mathcal{G}_{1,k}\} \geq \frac{R-1}{R}.$$

Now we take  $R \uparrow \infty$ , to get finally that

$$\sup_{f \in B_{L^1(\Omega, \mathcal{F}, \mu)}} \inf\{\|f - g\|_1 : g \in \mathcal{G}_{1,k}\} \geq 1$$

independently of  $k \geq 2$ . For  $k = 1$ , we point out that  $\mathcal{G}_{1,1} = \{0\}$  and so

$$\sup_{f \in B_{L^1(\Omega, \mathcal{F}, \mu)}} \inf\{\|f - g\|_1 : g \in \mathcal{G}_{1,1}\} = 1.$$

Hence,  $N_{1,\varepsilon}(B_{L^1(\Omega, \mathcal{F}, \mu)}) = \infty$ , for all  $\varepsilon < 1$ .  $\square$

In order to prove Theorem 4.7, we shall use a result in [9], where the notion of atomic and nonatomic are different from the (standard) notions we are using. In this discussion we add an  $*$  to distinguish the notions we are using and the corresponding in [9]. According to [9] a measurable set  $A$  is an  $*$ -atom if  $\mu(A) > 0$  and for all  $E \in \mathcal{F}$  either  $\mu(A \cap E) = 0$  or  $\mu(A \setminus E) = 0$ . It is direct to show that if  $A$  is an  $*$ -atom for  $\mu$ , then it is an atom for  $\mu$ . Indeed, assume that  $B \subset A$  satisfies  $\mu(B) < \mu(A)$ , then  $\mu(A \setminus B) = \mu(A) - \mu(B) > 0$  and we conclude that  $0 = \mu(A \cap B) = \mu(B)$ , proving that  $A$  is an atom for  $\mu$ . The converse is not always true (see the example below). It is true if  $A$  has finite measure. In fact, suppose that  $A$  is an atom of finite measure and let  $E$  be a

measurable set. If  $\mu(A \cap E) > 0$  then  $\mu(A) = \mu(A \cap E)$ , showing that  $\mu(A \setminus E) = 0$  since  $E \cap A$  has finite measure, and therefore  $A$  is an  $*$ -atom.

A measure is  $*$ -atomic if every measurable set  $A$  of positive measure contains an  $*$ -atom. A measure that has no  $*$ -atoms is said  $*$ -nonatomic. Here is an example of a  $*$ -nonatomic measure which is atomic in the standard sense. Consider  $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$  as a measurable space and

$$\mu(A) = \begin{cases} \infty & \text{if } A \text{ is uncountable} \\ 0 & \text{otherwise} \end{cases}$$

If  $\mu(A) > 0$ , then  $A$  is uncountable and can be splitted into two uncountable disjoint sets  $B$  and  $C$ . Then  $\mu(A \cap B) = \infty$  and  $\mu(A \setminus B) = \infty$ . So, there are no  $*$ -atoms and then according to the above definition  $\mu$  is  $*$ -nonatomic.

The other concept we need is the notion of  $*$ -singular. Two measures  $\nu$  and  $\lambda$  are said  $*$ -singular if for all measurable sets  $E$ , there exist two measurable sets  $F$  and  $G$  contained in  $E$  such that

$$\nu(F) = \nu(E), \lambda(F) = 0, \text{ and } \lambda(G) = \lambda(E), \nu(G) = 0.$$

The main theorem we need is the following.

**Theorem 2.1** in [9]. *Assume  $(\Omega, \mathcal{F}, \mu)$  is a measure space. Then  $\mu$  can be decomposed as  $\mu = \nu + \lambda$ , where  $\nu$  is  $*$ -atomic and  $\lambda$  is  $*$ -nonatomic. We can assume that  $\nu, \lambda$  are  $*$ -singular, in which case the decomposition is unique.*

We are now ready to prove the main result of this subsection:

*Proof of Theorem 4.7.* It is clear that  $(iii) \implies (ii) \implies (i)$ . Now suppose that  $(i)$  holds, that is,  $B_{L^p(\Omega, \mathcal{F}, \mu)}$  is UA. By Theorem 2.1 in [9], there is a unique decomposition  $\mu = \nu + \lambda$  where  $\nu$  is  $*$ -atomic measure,  $\lambda$  is  $*$ -nonatomic and  $\nu$  and  $\lambda$  are  $*$ -singular. Consider

$$\mathcal{C} = \{[A] : A \text{ is an } * \text{-atom for } \nu \text{ of finite } \nu \text{-measure} \}$$

where  $[A]$  is the equivalence class of measurable sets  $B$  such that  $\nu(A \Delta B) = 0$ . Notice that  $[A] \in \mathcal{C}$  if and only if  $A$  is an atom of finite  $\nu$ -measure. Therefore, if  $[A] \neq [B] \in \mathcal{C}$  then  $\nu(A \cap B) = 0$ , that is,  $A$  and  $B$  are disjoint up to  $\nu$ -measure 0.

If  $\mathcal{C}$  is infinite, we take a countable collection  $(E_n)_n$  of atoms for  $\nu$ , which are disjoint up to  $\nu$ -measure zero, and each one has finite and positive  $\nu$ -measure. For every  $n$  there exists  $F_n \subset E_n$ , such that  $\nu(F_n) = \nu(E_n)$  and  $\lambda(F_n) = 0$ . Clearly,  $(F_n)_n$  is a countable class of disjoint atoms for  $\nu$ , which have positive and finite measure. The measurable set  $A = \cup_{n=1}^{\infty} F_n$  satisfies  $\lambda(A) = 0$ . This shows that  $\mu|_A = \nu|_A$ , so  $L^p(A, \mathcal{F}|_A, \mu|_A)$  and  $L^p(A, \mathcal{F}|_A, \nu|_A)$  can be identified.

On the other hand, the measure  $\nu|_A$  is atomic. Indeed, assume that  $D \subset A$  has positive measure. Then for some  $n$  it holds  $\nu(D \cap F_n) > 0$  and then  $D \cap F_n$  contains an  $*$ -atom  $H$  of  $\nu$ , which has finite measure, and therefore it is an atom for  $\nu$ . We can apply Proposition 4.10 to conclude that  $B_{L^p(A, \mathcal{F}|_A, \nu|_A)}$  is not UA, and a fortiori  $B_{L^p(\Omega, \mathcal{F}, \mu)}$  is not UA, which is a contradiction.

The conclusion is that  $\nu$  has a finite number of atoms  $(A_n)_{n \in J}$  of finite measure, up to measure 0, where  $J$  is a finite (eventually empty) set. Therefore, if  $B = \Omega \setminus \cup_{n \in J} A_n$ , then any measurable  $C \subset B$  has 0 or infinite  $\nu$ -measure.

On the other hand, there exists  $G \subset B$  such that  $\nu(G) = 0$  and  $\lambda(G) = \lambda(B)$ . If there exists  $H \subset B$  a measurable set such that  $0 < \lambda(H) < \infty$ , then we arrive to a contradiction. Indeed, consider  $K \subset H$  such that  $\nu(K) = 0$  and  $\lambda(K) = \lambda(H)$ . Since  $\lambda(H)$  is finite, this means that  $\lambda(H \setminus K) = 0$ . Now,  $\lambda|_K$  is a diffuse measure, because if there exists  $L \subset K$  an atom for  $\lambda$ , then this atom has finite measure and therefore it is an  $*$ -atom for  $\lambda$ , which is not possible. The contradiction is obtained because  $B_{L^p(K, \mathcal{F}|_K, \mu|_K)}$  and  $B_{L^p(K, \mathcal{F}|_K, \lambda|_K)}$  can be identified and the latter is not UA, according to Corollary 4.9.



The conclusion is that  $\lambda(H)$  is 0 or infinite for every  $H \subset B$ . Since  $\lambda(B^c) = 0$ , we conclude that  $\lambda(H)$  is either 0 or infinite for every measurable set  $H$ . Also,  $\mu(H)$  is 0 or infinite, for any  $H \subset B$  and  $\mu = \nu$  on  $A = B^c$ . Therefore,  $\mu$  is an atomic measure and it has a finite collection of disjoint atoms with finite measure, up to measure zero.

The last part of the Theorem follows from either Corollary 4.9 or Proposition 4.10.  $\square$

#### 4.4 Stability of the class of UA sets

In this subsection, we study the image of a UA set under classical operations. We start with the following easy proposition:

**Proposition 4.11.** *Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space and  $p \in [1, +\infty]$ . Let  $\mathcal{A}, \mathcal{B} \subset L^p(\Omega, \mathcal{F}, \mu)$  and  $\varepsilon > 0$ . Then:*

- (i) if  $\mathcal{A} \subset \mathcal{B}$  then  $N_{p,\varepsilon}(\mathcal{A}) \leq N_{p,\varepsilon}(\mathcal{B})$ ;
- (ii)  $N_{p,\varepsilon}(\mathcal{A}) = N_{p,\varepsilon}(\overline{\mathcal{A}})$ ;
- (iii)  $N_{p,|\lambda|\varepsilon}(\lambda\mathcal{A}) = N_{p,\varepsilon}(\mathcal{A})$  for all  $\lambda \in \mathbb{R}$ ;
- (iv)  $N_{p,\varepsilon}(\mathcal{A} + \mathcal{B}) \leq \min_{t,s>0,t+s \leq \varepsilon} N_{p,t}(\mathcal{A})N_{p,s}(\mathcal{B})$ ;

In particular if  $\mathcal{A}$  and  $\mathcal{B}$  are UA then  $\overline{\mathcal{A}}$ ,  $\lambda\mathcal{A}$  and  $\mathcal{A} + \mathcal{B}$  are UA.

*Proof.* The proof is left to the reader.  $\square$

In the next result, we prove that the closed convex hull of a bounded UA set is still UA.

**Theorem 4.12.** *Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space and  $p \in (1, +\infty)$ . If  $\mathcal{A} \subset L^p(\Omega, \mathcal{F}, \mu)$  is a UA set, then  $\mathcal{A}_K = \{f \in \overline{\text{co}}(\mathcal{A}) \mid \forall g \in \mathcal{A} \ \|f - g\|_p \leq K\}$  is also UA for all  $K \geq 0$ . More precisely, we have that*

$$N_{p,\varepsilon}(\mathcal{A}_K) \leq \min_{\eta \in (0,1)} (N_{p,(1-\eta)\varepsilon}(\mathcal{A}))^{s(\eta)}$$

for all  $\varepsilon > 0$ , where  $s(\eta) = \left\lceil \frac{CK}{\eta\varepsilon} \right\rceil^{\frac{\min\{p,2\}}{\min\{p,2\}-1}} + 1$  and  $C$  is a constant depending on  $(\Omega, \mathcal{F}, \mu)$  and  $p$ . In particular, if  $\mathcal{A}$  is bounded then  $\overline{\text{co}}(\mathcal{A})$  is UA.

*Proof.* Fix  $K \geq 0$ . For  $n \in \mathbb{N}$ , define

$$\text{co}_n(\mathcal{A}) = \left\{ \sum_{i=1}^n a_i f_i \mid a_i \geq 0, \sum_{i=1}^n a_i = 1, f_i \in \mathcal{A} \right\}.$$

Remember that  $L^p(\Omega, \mathcal{F}, \mu)$  has non-trivial Rademacher type  $r = \min\{p, 2\}$  (see Theorem 6.2.14 in [1]). By Theorem 2.5 of [6], one has that

$$d(\text{co}_n(\mathcal{A}), f) \leq \frac{CK}{n^{1-\frac{1}{r}}}$$

for all  $f \in \mathcal{A}_K$  and  $n \in \mathbb{N}$  where  $C$  is a constant depending on  $(\Omega, \mathcal{F}, \mu)$  and  $p$ . Therefore, if we take  $\varepsilon > 0$ ,  $\eta \in (0, 1)$  and  $n_0 = \left\lceil \frac{CK}{\eta\varepsilon} \right\rceil^{\frac{r}{r-1}} + 1$ , we will have that  $d(\text{co}_{n_0}(\mathcal{A}), f) < \eta\varepsilon$ , for all  $f \in \mathcal{A}_K$ .

Thus, if  $f_0 \in \mathcal{A}_K$  there exists  $g_0 = \sum_{i=1}^{n_0} a_i f_i \in \text{co}_{n_0}(\mathcal{A})$  such that  $\|f_0 - g_0\|_p < \eta\varepsilon$ . On the other hand, since  $\mathcal{A}$  is UA, there exists  $h_i \in \mathcal{G}_{p,k}$  where  $k = N_{p,(1-\eta)\varepsilon}(\mathcal{A})$  such that

$$\|f_i - h_i\|_p \leq (1 - \eta)\varepsilon$$

for all  $i \in \{1, \dots, n_0\}$ . One can deduce that

$$\left\| f - \sum_{i=1}^{n_0} a_i h_i \right\|_p \leq \|f_0 - g_0\|_p + \left\| \sum_{i=1}^{n_0} a_i f_i - \sum_{i=1}^{n_0} a_i h_i \right\|_p \leq \eta\varepsilon + \sum_{i=1}^{n_0} a_i(1 - \eta)\varepsilon = \varepsilon$$

with  $\sum_{i=1}^{n_0} a_i h_i \in \mathcal{G}_{p, k^{n_0}}$ . We conclude that  $N_{p, \varepsilon}(\mathcal{A}_K) \leq k^{n_0}$ .  $\square$

**Remark.** Note that if  $\mathcal{A}$  is an unbounded UA set then  $\overline{\text{co}}(\mathcal{A})$  may not be UA. In fact,  $\mathcal{G}_{p, 2}$  is UA but  $\overline{\text{co}}(\mathcal{G}_{p, 2}) = L^p(\Omega, \mathcal{F}, \mu)$  (since  $\text{co}(\mathcal{G}_{p, 2})$  is the set of simple functions) is not UA in general for any  $p \in [1, \infty]$ . Remark that the previous theorem is not interesting if  $p = \infty$  since any bounded set is UA by Proposition 4.6.

**Remark.** The previous theorem is false if  $p = 1$ . In fact remember that  $B_{\ell_1} = \overline{\text{co}}(\text{Ext}(B_{\ell_1}))$  and  $\text{Ext}(B_{\ell_1}) = \{\pm\delta_n\}_{n \in \mathbb{N}}$ , where  $\text{Ext}(B_{\ell_1})$  is the set of extreme points of  $B_{\ell_1}$ . It follows that  $\text{Ext}(B_{\ell_1})$  is UA but we have seen that  $B_{\ell_1}$  is not UA (see Theorem 4.7). More generally, using the previous result, it is easy to show that there exists a UA set  $\mathcal{A} \subset \ell_p$  such that  $B_{\ell_p} = \overline{\text{co}}(\mathcal{A})$  if and only if  $p \in \{1, \infty\}$ .

In the next result we study stability properties of UA classes under Hölder transformations. Recall that a real function  $\Psi$  is uniformly  $\alpha$ -Hölder if there exists a constant  $K$ , such that

$$|\Psi(x) - \Psi(y)| \leq K|x - y|^\alpha.$$

With this definition, the identity function is not uniformly  $\alpha$ -Hölder for  $\alpha < 1$ . To enlarge the class of uniformly  $\alpha$ -Hölder functions we consider the following classes of Hölder functions, denoted  $\mathbb{H}(K, \alpha)$  for  $0 < \alpha \leq 1$ , which consists of real functions  $\Psi$  such that for all  $x, y$ , it holds

$$|\Psi(x) - \Psi(y)| \leq K(|x| + |y| + 1)^{1-\alpha}|x - y|^\alpha.$$

We can assume without loss of generality that  $K \geq 1$ . We notice that  $\mathbb{H}(K, 1)$  is the set of  $K$ -Lipschitz functions. If  $0 < \beta \leq \alpha$ , then  $\mathbb{H}(K, \alpha) \subset \mathbb{H}(K, \beta)$ . Also  $\mathbb{H}(K, \alpha)$  contains the class of uniformly Hölder functions.

**Proposition 4.13.** Assume that  $(\Omega, \mathcal{F}, \mu)$  is a finite measure space and  $\mathcal{A}$  is UA in  $L^q(\Omega, \mathcal{F}, \mu)$ , for some  $q \in [1, \infty]$ . Consider  $\alpha \in (0, 1]$ , and we assume further that  $\mathcal{A}$  is bounded in  $L^q(\Omega, \mathcal{F}, \mu)$  when  $\alpha < 1$ . Then, the  $\mathbb{H}(K, \alpha)$ -transform of  $\mathcal{A}$  given by

$$\mathbb{H}(K, \alpha)(\mathcal{A}) = \{\Psi(f) : f \in \mathcal{A}, \Psi \in \mathbb{H}(K, \alpha)\}$$

is UA in  $L^p(\Omega, \mathcal{F}, \mu)$  for any  $1 \leq p \leq q$ . Moreover, for  $\varepsilon \in (0, 1]$

$$N_{p, \varepsilon}(\mathbb{H}(K, \alpha)(\mathcal{A})) \leq N_{q, (\varepsilon/\Gamma)^{\frac{1}{\alpha}}}(\mathcal{A}),$$

where

$$\Gamma = \Gamma(\alpha, p, q) = K\mu(\Omega)^{-r} \begin{cases} \left(2B + 1 + \mu(\Omega)^{\frac{1}{q}}\right)^{1-\alpha} & \text{if } \alpha < 1 \\ 1 & \text{if } \alpha = 1 \end{cases}$$

with  $B$  a bound for  $\mathcal{A}$  in  $L^q(\Omega, \mathcal{F}, \mu)$  and  $r = \frac{1}{p} - \frac{1}{q}$ .

*Proof.* The case  $\alpha = 1$  is straightforward so, we assume  $\alpha < 1$ . We assume first that  $p = q$ . Consider  $\varepsilon \in (0, 1]$ ,  $k = N_{q, \varepsilon}(\mathcal{A})$ ,  $f \in \mathcal{A}$ ,  $g \in \mathcal{G}_{q, k}$  such that  $\|f - g\|_q \leq \varepsilon$  and  $\Psi \in \mathbb{H}(K, \alpha)$ . We have

$$\int |\Psi(f(x)) - \Psi(g(x))|^q d\mu(x) \leq K^q \int (|f(x)| + |g(x)| + 1)^{q(1-\alpha)} |f(x) - g(x)|^{q\alpha} d\mu(x).$$

Now, we apply Hölder's inequality for  $s = \frac{1}{\alpha}$  and its conjugated index  $t = \frac{1}{1-\alpha}$  to get

$$\int |\Psi(f(x)) - \Psi(g(x))|^q d\mu(x) \leq K^q \left( \int (|f(x)| + |g(x)| + 1)^q d\mu(x) \right)^{1-\alpha} \left( \int |f(x) - g(x)|^q d\mu(x) \right)^\alpha$$

which implies

$$\|\Psi(f) - \Psi(g)\|_q \leq K (\|f\|_q + \|g\|_q + \|1\|_q)^{(1-\alpha)} \|f - g\|_q^\alpha.$$

If  $B$  is a bound for  $\mathcal{A}$ , we conclude that  $\|g\|_q \leq B + 1$ , which shows

$$\|\Psi(f) - \Psi(g)\|_q \leq K \left( 2B + 1 + \mu(\Omega)^{\frac{1}{q}} \right)^{(1-\alpha)} \varepsilon^\alpha = \Gamma \varepsilon^\alpha.$$

Since  $\Psi(g) \in \mathcal{G}_{q,k}$ , we deduce that

$$N_{q,\Gamma\varepsilon^\alpha}(\mathbb{H}(K, \alpha)(\mathcal{A})) \leq N_{q,\varepsilon}(\mathcal{A}),$$

and the result is shown in this case. The case  $p < q$  follows from Corollary 4.4.  $\square$

We point out that under the hypothesis of the Theorem, we have  $\mathcal{A} \subset \mathbb{H}(K, \alpha)(\mathcal{A})$ .

**Acknowledgements:** G. Grelier was supported by the Grants of Ministerio de Economía, Industria y Competitividad MTM2017-83262-C2-2-P; Fundación Séneca Región de Murcia 20906/PI/18; and by MICINN 2018 FPI fellowship with reference PRE2018-083703, associated to grant MTM2017-83262-C2-2-P. J. San Martín was supported in part by BASAL ANID FB210005.

## References

- [1] F. Albiac, N. J. Kalton, *Topics in Banach space theory*, Second edition. Graduate Texts in Mathematics 233, Springer, Cham, 2016.
- [2] M. Anthony, P. L. Bartlett, *Neural network learning: Theoretical foundations*, Cambridge : cambridge university press, 1999.
- [3] S. Arora, J. Lee, A. Naor, *Euclidean distortion and the sparsest cut*, Journal of the American Mathematical Society **21** (2008), 1–21.
- [4] V. I. Bogachev, M. A. S. Ruas, *Measure theory* (Vol. 1), Berlin: Springer, 2007.
- [5] J. Bourgain, *On Lipschitz embedding of finite metric spaces in Hilbert space*, Israel Journal of Mathematics **52** (1985), 46–52.
- [6] M. J. Donahue, C. Darken, L. Gurvits and E. Sontag, *Rates of convex approximation in non-Hilbert spaces*, Constructive Approximation **13** (1997), 187–220.
- [7] G.A. Hunt, *Martingales et Processus de Markov*, Paris, Dunod, 1966.
- [8] R. Huotari, W. Li *Continuities of Metric Projection and Geometric Consequences*, journal of approximation theory **90** (1970), 319–339.
- [9] R. A. Johnson, *Atomic and nonatomic measures*, Proceedings of the American Mathematical Society **25** (1970), 650–655.
- [10] W. B. Johnson, J. Lindenstrauss *Extensions of Lipschitz mappings into a Hilbert space*, Con-temp. Math **26** (1984), 189–206.

- [11] J. Matoušek, *On the distortion required for embedding finite metric spaces into normed spaces*, Israel Journal of Mathematics **93** (1996), 333–344.
- [12] W. Sierpiński, *Sur les fonctions d'ensemble additives et continues*, Fundamenta Mathematicae **1** (1922), 240–246.
- [13] D. W. Stroock *Probability theory: an analytic view*, Cambridge university press, 2010.
- [14] L. P. Vlasov, *Chebyshev sets and approximately convex sets*, Mathematical notes of the Academy of Sciences of the USSR **2** (1967), 600–605.
- [15] M. J. Wainwright, *High-dimensional statistics: A non-asymptotic viewpoint*, Vol. 48, Cambridge University Press, 2019.