Function Spaces defined by Means of Oscillation, with Applications

Francesca Angrisani

Contents

Preface

Lebesgue spaces $L^p(\Omega, \mu)$ and many of their generalizations (Grand Lebesgue, Lorentz, Orlicz, Marcinkiewicz spaces) are defined by conditions that, in some intuitive sense, limit the size of the function, but in a way such that the distribution of the values of the function is irrelevant.
In other words, one measures the mass of the function irrespective of where it is allocated. This is formalized by saying that these spaces are *rearrangement invariant*, that is, given two functions f and g such that

$$|\{x \in \Omega : |f(x)| \geq \lambda\}| = |\{x \in \Omega : |g(x)| \geq \lambda\}|, \quad \forall \lambda \geq 0,$$

then f belongs to the space if and only if g does; two functions satisfying the above property are called equimeasurable. In the aforementioned Lebesgue spaces, for example, this is made clear by the so-called Cavalieri Principle.
If a function f belongs to a given rearrangement invariant function space X, then the same can be said of its *non-increasing rearrangement*

$$f^* : t \in [0, |\Omega|] \mapsto \inf\left\{\lambda : |\{x \in \Omega : |f(x)| \geq \lambda\}| \leq t\right\} \in \mathbb{R}^+,$$

which is a non-increasing function that is equimeasurable to the starting function f.
Not all function spaces, though, are rearrangement invariant. The space of Lipschitz functions, to make an example that students of mathematics encounter early on in their education, is not a rearrangement invariant space. One (valid, but maybe not the simplest) way to think intuitively of a counterexample would be to think of a function having violent oscillation, such as

$$f(t) = t \in [0, 1] \mapsto \int_0^t \sum_{n=1}^{\infty} (-1)^{n+1}(n^2 + n)\chi_{\left[1-\frac{1}{n}, 1-\frac{1}{n+1}\right]}(s)\, ds \in [0, 1]$$

which is a piece-wise linear function doing a zig-zag to connect all points of the type

$$(0, 0) - \left(\frac{1}{2}, 1\right) - \left(\frac{2}{3}, 0\right) - \left(\frac{3}{4}, 1\right) - \left(\frac{4}{5}, 0\right) - \dots$$

and so on which is clearly not Lipschitz, but then look at is non-increasing rearrangement $f^*(t)$, which can be computed to be $f^*(t) = 1 - t$, which is clearly Lipschitz.
This informal argument suggests that an interesting aspect of functions, that is the way they oscillate, which is neglected in rearrangement invariant spaces, is crucial to the definition of other interesting function spaces. In the mathematical literature, a

plethora of different meanings and formal definitions have been associated to the word "oscillation".
In this text we will explore some of the function spaces defined by means of oscillation, in many different senses of the word, fitting into two modes: spaces in which oscillation is bounded and spaces in which oscillation is vanishing, i.e. arbitrarily small when measured on a sufficiently small set. This general framework will be made precise in a diversity of ways.

The outline of the thesis is the following. In the first chapter a very large space of functions introduced by Brezis, Bourgain and Mironescu (see [32]) in 2015 and often denoted as B is introduced. It is a function space that caught the interest of both my former advisor Prof. Carlo Sbordone and my current advisor Prof. Gioconda Moscariello (see [59],[71] and [72]). Even if its introduction is very recent and inspired by other previously defined function spaces, the logic for its placement at Chapter 1 is that B is a space containing all of those who follow, which will be seen as special subspaces of B.
The space B was introduced to answer a question about functions with discrete range

$$f : \Omega \to \mathbb{Z}.$$

In many function spaces classical to the mathematical analysis, such as Sobolev spaces (with sufficiently large integrability exponent) these functions are necessarily almost everywhere constant functions: this is easily proven via the Sobolev-Poincarè inequality.
What is the largest space we can think of where functions with discrete range are necessarily almost everywhere equal to a constant?
That is the question that was on the table when the space B was introduced. As a matter of fact, a very natural subspace B_0 which will be introduced in Chapter 1 has the desired property and contains all function spaces that were known to have the desired property.
The couple (B, B_0) of Banach function spaces has many interesting functional properties.
Since it fits into a very general and abstract framework for couples of Banach function spaces introduced by K.M. Perfekt in [120], which we will refer to as an (o, O) pair or (o, O)–type structure and make us of multiple times throughout this text, we will present, with respect to this, results from a very recent paper [59]. Moreover, we will explore an interesting relationship that is between some quantities involved in defining the norm on B and the perimeter of a measurable set: these are the topics of [71] and [72], by Fusco, Moscariello and Sbordone.

In Chapter 2 we explore the space of Lipschitz functions, in the most general setting of an arbitrary compact metric space and together with other function spaces from the family of Holder spaces. The space of Lipschitz function is possibly the most famous example of a space defined by means of oscillation. As a matter of fact, one defines the pointwise oscillation of a function f over a ball B as

$$\omega_f(B) := \sup_{x,y \in B} |f(x) - f(y)|.$$

That is the most natural and naive definition of oscillation, i.e. the largest distance there is between two values of f in B. The Holder spaces Lip_α, $\alpha \in (0,1]$, of which the space Lip_1 is a special case, can be seen as the spaces of functions whose oscillation $\omega_f(B)$ is bounded in terms of α-th power $\mathrm{Diam}^\alpha(B)$ of the diameter of the ball, uniformly with respect to all balls contained in the domain of f.
Again in the general context of a compact metric space (K, ρ), we will also introduce the less famous subspace lip_α consisting of functions such that

$$\frac{|f(x) - f(y)|}{\rho^\alpha(x,y)}$$

is not only uniformly bounded for all $x, y \in K$, $x \neq y$, but vanishing as $\rho(x,y) \to 0$, i.e. such that

$$\lim_{\rho(x,y) \to 0} \frac{|f(x) - f(y)|}{\rho^\alpha(x,y)} = 0.$$

Again making use of the framework from [120] and some additional arguments from a recent paper [59], we explore the relationship between the smaller space lip_α and Lip_α, explaining exactly under which conditions we can say that the bidual of the smaller one is isometrically isomorphic to the larger one, as done with Ascione, D'Onofrio and Manzo in [16].
This type of functional relationship between the two spaces was already explored by many, but fitting the couple of spaces in the abovementioned framework by K.M. Perfekt gives new functional analytic information as, for example, an equivalent formulation for the distance of a given function from the smaller space lip_α, and also allows us to infer some interesting properties about the predual of Lip_α which is a Banach space of Borel measures on K, suitably normed. In particular, we present the results of [16] and [19].
More precisely, regarding the latter, if the space $M(K)$ of finite signed Borel measures on K is equipped with the Kantorovich-Rubinstein norm (see Chapter 2 for the definition), then we provide an atomic decomposition result for its completion $M(K)^c$. Its elements can be decomposed as an infinite series of measures whose support is of cardinality at most 2, as we will show with Theorem 2.23.
The last section of the Chapter 2 is devoted to some results obtained in the field of Optimal Control Theory. These results are not yet submitted for publication.
As a matter of fact, it is an interesting application, where almost everywhere differentiability of Lipschitz functions (Rademacher's Theorem) and some weaker notions of differentiability for Lipschitz functions (Clarke's Generalized Jacobian and Quasi Differential Quotient) play a crucial role.
The Pontryagin Maximum Principle (PMP) has been generalized to a wide variety of optimal control problems to provide necessary conditions for optima. In a work in preparation with Prof. Franco Rampazzo (see [20]), we first introduce a problem of

the following type:

$$\begin{cases} \frac{dx}{dt}(t) = f(x(t), a(t)) + \sum_{i=1}^{m} g_i(x(t))u^i(t), & \text{a.e. } t \in [0,T], \\ \frac{d\nu}{dt}(t) = |u(t)|, \\ (x,\nu)(0) = (\hat{x}, 0), \quad \nu(T) \le K, \quad (T, x(T)) \in \mathfrak{T}, \\ \text{minimize:} \quad \Psi(T, x(T)) + \int_0^T l(x(t), u(t), a(t))\, dt \end{cases}$$

where u is the control function with values in a cone C, $f : \mathbb{R}^n \times A \to \mathbb{R}^n$ is usually called "drift", $A \subseteq \mathbb{R}^q$, g_i, $i \le n$ are vector fields on $\mathbb{R}^n$, $\mathfrak{T} \subseteq \mathbb{R}^* \times \mathbb{R}^n$ is a set called "target", representing admissible final times and locations, t, K are positive real numbers and $\Psi : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}$ and $l : \mathbb{R}^n \times C \times A \to \mathbb{R}$ are the final and the lagrangian cost functions respectively.
If the fields g_i are $\mathscr{C}^1$, some second order conditions on optima can be found, expressed in the language of Lie brackets (see, for example [21]). Using a definition from [128] that generalized Lie Brackets to Lipschitz fields, and the concept of a Quasi Differential Quotient (QDQ) from [119], we will be able to find a second order condition for optima even if the fields are Lipschitz: it will be expressed in terms of set-valued Lie Brackets and QDQ's.

In Chapter 3 we introduce Orlicz spaces

$$L^\Psi(\Omega) = \left\{ u \in L^1(\Omega) : \ \exists \lambda > 0 \text{ such that} \int_\Omega \Psi\left(\frac{|u|}{\lambda}\right) < +\infty \right\}.$$

for any choice of a Young function Ψ and the closure of L^∞ in L^Ψ, also known as the Morse space M^Ψ and characterized as

$$M^\Psi(\Omega) = \left\{ u \in L^1(\Omega) : \ \forall \lambda > 0 \ \int_\Omega \Psi\left(\frac{u}{\lambda}\right) dx < +\infty \right\}.$$

We first show how in [18] we were able to individuate a large subfamily of Orlicz spaces for which (L^Ψ, M^Ψ) fits into the aforementioned mathematical framework by K.M. Perfekt, deriving many functional properties of the couple.
In the last section we introduce the Orlicz-Sobolev spaces $W^{1,\Psi}$, which are the spaces of functions in L^Ψ such that a weak derivative exists and is also in L^Ψ, and, in line with the general topic of the text, we present a nice and peculiar result by Tuominen [134] characterizing it as the space of L^Ψ functions f such that there exists a function g also in L^Ψ bounding the oscillation of f in the following way

$$|f(x) - f(y)| \le |x - y| \left[g(x) + g(y) \right].$$

This result is a generalization of one obtained by Hajlasz for $\Psi(t) = t^p$ and hence characterizing the classical Sobolev spaces. These types of results, showing that Orlicz-Sobolev spaces are defined by means of oscillation, also provide a way to define these spaces in the wider context of a generic compact metric space, as the concept of weak

derivative is no longer needed in the alternative definition.
Moreover, in the last section I present some results (see [12]) on Lipschitz regularity for minima of some functionals in the calculus of variations, a field I got a deeper interest in, since collaborating to [17].
We prove that for suitable Young functions φ and ψ with the Δ_2 growth condition and autonomous functionals

$$\mathscr{F} = \int f(Du)\,dx$$

with ψ-growth, assuming $W^{1,\varphi}$-quasiconvexity of $f = f(z)$ only asymptotically is enough to prove local Lipschitz regularity for minima u on a dense subset of the domain.
To be more precise, we begin by making use of asymptotic quasiconvexity to prove partial $C^{1,\alpha}$ regularity of minima in points where the gradient Du of the minimum is large enough, say larger than M, adapting arguments used for globally quasiconvex lagrangians.
As a corollary, we infer that minima are locally lipschitz on a dense subset of points. The idea behind it is that any x_0 such that all points in a neighbourhood of x_0 are not points of local holder regularity for Du, is such that Du is essentially bounded by M around x_0, hence u is locally Lipschitz around x_0.

Chapters 4 and 5 are dedicated to the spaces BMO and VMO and their respective subcones BLO and VLO, concluding the text with discussion of what are arguably the most natural "children" of the very large space B discussed at the beginning (on the contrary, they were introduced decades earlier than B was introduced).
More precisely, the fourth chapter regards the BMO and BLO which are function classes defined by imposing a bound on two different kinds of oscillation, the so called mean oscillation of f over I, i.e.

$$\frac{1}{|I|}\int_I \left| f(x) - \frac{1}{|I|}\int_I f(y)\,dy \right| \,dx.$$

and the so called lower oscillation of f over I, i.e.

$$\frac{1}{|I|}\int_I f(x)\,dx - \inf_{x\in I} f(x)$$

respectively, where I is an interval if we are dealing with functions of a single variable, while it is a hypercube with sides parallel to the coordinate hyperplanes if we wish to define the function class in $\mathbb{R}^n$.
In an introductory section, we present many properties of BMO and its subcone BLO and explore the relationship between them and among them and the Muckenhoupt function classes A_p.
Then, we address the problem of determining a distance formula for functions in BMO from the subspace L^∞ of essentially bounded functions, solved by Garnett and Jones in 1978 in their paper [74], and, subsequently, my transposition of their result to BLO in [14].

In the section that follows, I introduce a new equivalent norm in BLO, inspired by some characterizations of BLO by Coifman and Rochberg in 1980 (see [51]), that is much more natural when addressing the aforementioned distance problem. This section contains results not yet submitted for publishing.
We then discuss the behaviour of BLO functions with respect to truncation.
As the norm on BLO is defined by means of oscillation, we may expect it to have strange behaviour with respect to truncation, which is a procedure acting on the size of the function.
That is exactly what happens: we show that bounded functions, functions that can be BLO-approximated by their truncations and functions that can be BLO-approximated by bounded functions are three completely different concepts, forming subspaces $L^\infty, T, \overline{L^\infty}^{BLO}$ in strict consecutive inclusion

$$L^\infty \subsetneq T \subsetneq \overline{L^\infty}^{BLO}.$$

Examples are provided in the section. This section also contains original results that have not been submitted yet for publishing.
In the last section of Chapter 4 the characterizazion

$$BMO = BLO - BLO,$$

where of course $-$ stands for Minkowski set subtraction here, is presented. It was obtained by Coifman and Rochberg in [51].

The last chapter introduces and studies VMO and VLO. They are subsets of BMO and BLO respectively, defined by the condition that the mean (lower, respectively) oscillation of the function tends uniformly to 0 as the measure of the interval over which it is computed goes to 0.
One of the results contained in [120] by K.M. Perfekt is that (BMO, VMO) is an (o, O)-type structure.
A result by Donald Sarason from 1975 (see [130]) is then presented, regarding an equivalent formula for the BMO-distance of a function from VMO, followed by my trasposition of this result to VLO, contained in [14].
To draw parallelism between the couple (BMO, BLO) and (VMO, VLO) a result by M.B. Korey (see [108]) is then stated, showing that

$$VMO = VLO - VLO,$$

in analogy to the abovementioned result from [51].
Lastly, in the second to final section we display a result from Leibov showing that for functions in VMO a special interval I^* can be found such that the mean oscillation of f over I^* is the absolute largest, i.e.

$$⨍_{I^*} |f - f_{I^*}|\, dt \geq ⨍_I |f - f_I|\, dt, \quad \forall I,$$

providing example that this is not at all true for the generic BMO function.

The result is presented together with some key insights that allowed its proof and some of the consequences Leibov obtained from this result, which was solely a lemma in its paper [110] from 1990.
This result is then generalized, in the last section, to BLO and VLO, discussing how the approach of Leibov could not exactly apply and that, in a way different from Leibov result, the existence of this peculiar interval is not logically equivalent to the VLO property.
This is the content of paper [15] by me and Giacomo Ascione: we specify that for simplicity of notation and visual interpretation of the result, we restrict the attention to $n = 1$, as the argument would otherwise become inherently geometrical, considering all possible cases for the symmetric difference of two cubes in $\mathbb{R}^n$.

To summarize, there are numerous ways of interpreting formally the concept of oscillation and those give rise to many different function spaces in which the focus is not on the size of the function but on the variability of its values.
Here we present a collection of result stemming from a functional analytic approach (with the exceptions of a few applicative results) to studying and understanding some examples of such function spaces, mostly fitting into the often recurring framework by K.M. Perfekt which we will properly introduce in Chapter 2 and in large part fitting under the wide cape of B functions, as defined in Chapter 1.
I feel that the *fil rouge* for this text is the sometimes surprising and counterintuitive nature of function spaces in which the norm is not interpreted as a measure of size but of variability of the function values, in a way that cannot be trivially reduced to commenting the size of the derivatives, as we will explore many contexts in which derivatives fail to exist in the classical and even in the weak Sobolev or distributional sense. On the contrary, we will put many different interpretations of the concept of "oscillation" at the center of everything.

Chapter 1

The Bourgain-Brezis-Mironescu Space

In this chapter, we define and discuss a very large functional space that contains all of the other function spaces defined by means of oscillation discussed in this thesis. This space was introduced in 2015 by Bourgain, Brezis, and Mironescu in [32]. They had noticed that for functions in the fractional Sobolev space $W^{\frac{1}{p},p}$, $p \geq 1$ (see page 23 for a definition) or in VMO (see page 109 for a definition), if the range of the function is discrete, say $f : \Omega \to \mathbb{Z}$ then the function is almost everywhere equal to a constant. They wanted to identify a wider space X of functions containing all of these function spaces and such that the above property holds for functions in X having discrete range, i.e. they are necessarily almost everywhere equal to a constant function.
They found that a sufficient condition for this is to have a bound on a form of mean oscillation, which we are about to define.
To do so, it is more convenient to introduce new notation gradually. Throughout this chapter, $Q_\varepsilon(a)$ will denote a hypercube with sides parallel to the coordinate hyperplanes, whose sidelength is ε and whose center is a. For any function that is summable on such a cube Q, f_Q will denote its integral average over Q and

$$M(f,Q) = \fint_Q |f - f_Q|$$

will denote what we will call the mean oscillation of f over Q.
For any fixed $\varepsilon \in (0,1)$, we will denote by $\mathscr{G}_\varepsilon$ any family of pairwise disjoint cubes of type $Q_\varepsilon(a)$ such that the cardinality of the family does not exceed ε^{1-n}. We then define

$$[f]_\varepsilon := \varepsilon^{n-1} \sup_{\mathscr{G}_\varepsilon} \sum_{Q_\varepsilon \in \mathscr{G}_\varepsilon} M(f, Q_\varepsilon)$$

and finally we can define the space B as the space of functions such that

$$[f]_B = \sup_{0<\varepsilon<1} [f]_\varepsilon < +\infty.$$

Effectively, we are asking that the sums of mean oscillations over families of cubes whose sidelenght is ε and whose collective measure is not larger than ε are bounded whatever the value of $\varepsilon \in (0,1)$.
Associating $[f]_B$ to each f in B induces a seminorm, which is null on almost everywhere

constant functions. To obtain a normed space, which can then be proved to be a Banach space, we can either consider B modulo almost everywhere constant functions, or add another seminorm to $[f]_B$, like the value of $|f|$ at some fixed point in Ω.
The space B clearly contains (and is inspired by) the space BMO of functions of bounded mean oscillation. Moreover, it can be proven the two function spaces coincide if and only if we are considering functions of a single variable, otherwise we have strict inclusion.
With some effort, starting from the notorious Poincarè inequality and its generalization to BV functions or functions in fractionary Sobolev spaces, one can also prove that B contains $W^{\frac{1}{p},p}$ (see the much more general Definition 2.5) for any p and BV (for the definition and properties of BV, the space of functions of bounded variation, see [8]).
An important subspace of B is the space

$$B_0 = \{f \in B : [f] := \limsup_{\varepsilon \to 0^+} [f]_\varepsilon = 0\}.$$

It was shown very recently in [59] that the couple of spaces (B, B_0) fits into a framework for describing couples of functional Banach spaces introduced by K.M. Perfekt (see for example [120] and page 10 of this text) and many interesting functional properties of B_0 with respect to B can be deduced from this, for example:

1. $(B_0)^{**} = B$ and
2. for any other space E such that $B = E^*$, E is isometrically isomorphic to $(B_0)^*$.
3. The distance
$$d_B(f, B_0) = \inf_{g \in B_0} [f - g]_B$$
induced by the $[\cdot]_B$ seminorm of a generic function in B from the subspace B_0 is equivalent to the quantity $[f]$, i.e. whenever $[f]$ is not 0, it can be reasonably interpreted as the distance of f from the subspace of B where $[\cdot] = 0$.

In particular, in [59], the authors proved the following result obtaining atomic decomposition for the predual of B (compare with the result at page 22).

Theorem 1.1 (D'Onofrio, Greco, Perfekt, Sbordone, Schiattarella, 2020, [59]). *Consider the Brezis-Bourgain-Mironescu space $B([0,1]^n)$ for functions defined on the unit hypercube of $\mathbb{R}^n$. Let $\varphi \in B_*$. Then for each natural number r there is a positive real number $\varepsilon_r \in (0,1)$, a family $\mathscr{G}_{\varepsilon_r}$ of disjoint hypercubes of sidelength ε_r and a summable function g_r satisfying:*

1. $|\mathscr{F}_r| \leq \varepsilon_r^{1-n}$
2. *$supp(g_r) \subseteq \bigcup \mathscr{G}_{\varepsilon_r}$*
3. $|g_r|\chi_Q \leq \varepsilon_r^{n-1}\frac{1}{|Q|}$ *for every cube $Q \in \mathscr{G}_{\varepsilon_r}$*
4. $\int_Q g_r = 0$ *for every cube $Q \in \mathscr{G}_{\varepsilon_r}$*

such that φ can be decomposed as

$$\varphi = \sum_r \lambda_r g_r$$

for a suitable sequence $\{\lambda_r\}_{r\in\mathbb{N}} \in \ell^1(\mathbb{R})$.
The action of $f \in B$ on φ is then the following

$$f(\varphi) = \sum_r \lambda_r \int f g_r.$$

Moreover, the quantity

$$\inf \sum_r |\lambda_r|$$

where the infimum is taken over all such representations of φ is equivalent to $\|\varphi\|_{B_*}$.

The space B can be continuously embedded in the Lorentz/Marcinkiewicz space $L^{\frac{n}{n-1},\infty}$ of functions f such that

$$\sup_{t>0} t^{\frac{n-1}{n}} \fint_{[0,t]} f^*(s)\,ds < +\infty$$

but B cannot be embedded in the smaller Lebesgue space $L^{\frac{n}{n-1}}$.
On the other hand, for the characteristic function χ_E of a measurable set E, the following inequality holds

$$\left\| \chi_E - \fint_Q \chi_E \right\|_{L^{\frac{n}{n-1}}(Q)} \leq c(n)\|\chi_E\|_B.$$

On the other hand it is also trivial to notice that

$$\left\| \chi_E - \fint_Q \chi_E \right\|_{L^{\frac{n}{n-1}}(Q)} \leq 2$$

and we can deduce

$$\left\| \chi_E - \fint_Q \chi_E \right\|_{L^{\frac{n}{n-1}}(Q)} \leq cP(E,Q)$$

from the known isoperimetric inequality, concluding

$$\left\| \chi_E - \fint_Q \chi_E \right\|_{L^{\frac{n}{n-1}}(Q)} \leq c\min\{1, P(E,Q)\}.$$

Motivated by these two inequalities estimating $\left\| \chi_E - \fint_Q \chi_E \right\|_{L^{\frac{n}{n-1}}(Q)}$ one could ask if there is some kind of relation between $\|\chi_E\|_B$ and $\min\{1, P(E,Q)\}$. To address this matter, it is convenient to introduce another quantity that is inspired by the seminorm on B but has the advantage of isotropy:

$$I_\varepsilon(f) := \varepsilon^{n-1} \sup_{\mathscr{F}_\varepsilon} \sum_{Q' \in \mathscr{F}_\varepsilon} \fint_{Q'} \left| f(x) - \fint_{Q'} f \right| dx$$

where $\mathscr{F}_\varepsilon$ is a collection of hypercubes of sidelength ε with arbitrary orientation (i.e. not necessarily with sides parallel to the coordinate hyperplanes) and $|\mathscr{F}_\varepsilon| < \varepsilon^{1-n}$.
In [7], the authors proved

Theorem 1.2 (Ambrosio, Bourgain, Brezis, Figalli, 2016, [7]). *For every measurable set E,*

$$\lim_{\varepsilon \to 0} I_\varepsilon(\chi_E) = \frac{1}{2} \min\{1, P(E)\}$$

where $P(E) = \infty$ if E is not a set of finite perimeter.

The two quantities $I_\varepsilon(f)$ and $[f]_\varepsilon$ are not at all uncorrelated. As a matter of fact, the following two inequalities hold for any function $f \in L^1(Q)$ and $\varepsilon \in \left(0, n^{-\frac{1}{2}}\right)$:

$$[f]_\varepsilon \leq I_\varepsilon(f) \leq [f]_{\sqrt{n}\varepsilon}.$$

Moreover, something more general than

$$\lim_{\varepsilon \to 0} \varepsilon^{n-1} \sup_{\mathscr{F}_\varepsilon} \sum_{Q' \in \mathscr{F}_\varepsilon} \fint_{Q'} \left| \chi_E(x) - \fint_{Q'} \chi_E \right| dx = \frac{1}{2} \min\{1, P(E)\}$$

holds if we replace the condition on the cardinality of $\mathscr{F}_\varepsilon$ with a weaker condition. As a matter of fact, it can be proven that for every integer M, one has

$$\lim_{\varepsilon \to 0} \varepsilon^{n-1} \sup_{\mathscr{F}_{M,\varepsilon}} \sum_{Q' \in \mathscr{F}_{M,\varepsilon}} \fint_{Q'} \left| \chi_E(x) - \fint_{Q'} \chi_E \right| dx = \frac{1}{2} \min\{M, P(E)\}$$

where $\mathscr{F}_{M,\varepsilon}$ is a family of arbitrarily oriented hypercubes of sidelength ε whose cardinality is not larger than $M\varepsilon^{1-n}$.
In particular,

$$\lim_{\varepsilon \to 0} \varepsilon^{n-1} \sup_{\mathscr{F}_{\infty,\varepsilon}} \sum_{Q' \in \mathscr{F}_{\infty,\varepsilon}} \fint_{Q'} \left| \chi_E(x) - \fint_{Q'} \chi_E \right| dx = \frac{P(E)}{2}$$

where $\mathscr{F}_{\infty,\varepsilon}$ is an arbitrarily large family of arbitrarily oriented hypercubes of sidelength ε.
Even more than that: upon defining

$$K_\varepsilon(f) = \varepsilon^{n-1} \sup_{\mathscr{F}_{\infty,\varepsilon}} \sum_{Q' \in \mathscr{F}_{\infty,\varepsilon}} \fint_{Q'} \left| f(x) - \fint_{Q'} f \right| dx$$

the previous result can be extended to

$$\lim_{\varepsilon\to 0} K_\varepsilon(f) = \frac{1}{2}|Df|(\mathbb{R}^n)$$

for every function of bounded variation with values in $\mathbb{Z}$.
The authors in [7] conjectured that for functions in SBV, i.e. functions of bounded variation that have no Cantor part $D^c f$, the following equality would hold

$$\lim_{\varepsilon\to 0} K_\varepsilon(f) = \frac{1}{4}|D^a f|(\mathbb{R}^n) + \frac{1}{2}|D^s f|(|\mathbb{R}^n).$$

This conjecture was proven true by Fusco, Moscariello and Sbordone in [71], with the following theorem

Theorem 1.3 (Fusco, Moscariello, Sbordone, 2016, [71])**.** *Let $f \in SBV_{loc}(\mathbb{R}^n)$. Then*

$$\lim_{\varepsilon\to 0} K_\varepsilon(f) = \frac{1}{4}|D^a f|(\mathbb{R}^n) + \frac{1}{2}|D^s f|(|\mathbb{R}^n).$$

and its corollary

Corollary 1.4. *Let $f \in W^{1,1}_{loc}(\mathbb{R}^n)$. Then*

$$\lim_{\varepsilon\to 0} K_\varepsilon(f) = \frac{1}{4}|D^a f|(\mathbb{R}^n).$$

This formula does not hold for a generic BV function, and it is possible that $\lim_{\varepsilon\to 0} K_\varepsilon(f)$ does not exist as $K_\varepsilon(f)$ may tend to oscillate between

$$\frac{1}{4}|Df|(\mathbb{R}^n) \quad \text{and} \quad \frac{1}{2}|Df|(\mathbb{R}^n)$$

as $\varepsilon \to 0$.
As a matter of fact

Theorem 1.5. *Let $f \in L^1_{loc}(\mathbb{R}^n)$. Then:*

$$\frac{1}{4}|Df|(\mathbb{R}^n) \le \liminf_{\varepsilon\to 0} K_\varepsilon(f) \le \limsup_{\varepsilon\to 0} K_\varepsilon(f) \le \frac{1}{2}|Df|(\mathbb{R}^n)$$

and $f \in BV_{loc}(\mathbb{R}^n)$ if and only if $\liminf_{\varepsilon\to 0} K_\varepsilon(f) < \infty$.

The space BV could then be defined as the space of functions such that $\liminf_{\varepsilon\to 0} K_\varepsilon(f)$ is bounded but we remark here that the corresponding smaller space of functions such that

$$\limsup_{\varepsilon\to 0} K_\varepsilon(f) = 0$$

is trivial, as it only contains constant functions.
Similarly, we can define

$$K_\varepsilon(p, f) := \varepsilon^{n-p} \sup_{\mathscr{F}_{\infty,\varepsilon}} \sum_{Q'\in\mathscr{F}_{\infty,\varepsilon}} \fint_{Q'} \left| f(x) - \fint_{Q'} f \right|^p dx$$

for any $p > 1$.
In [72], the same authors from [71] proved the following theorem

Theorem 1.6. *For any function in* $W^{1,p}_{loc}(\mathbb{R}^n)$, $p \in [1, \infty)$,

$$\lim_{\varepsilon \to 0} K_\varepsilon(p, f) = \gamma \int_{\mathbb{R}^n} |Df|^p \, dx$$

where

$$\gamma(n,p) := \max_{\nu \in \mathbb{S}^{n-1}} \int_{\left[-\frac{1}{2}, \frac{1}{2}\right]^n} |x \cdot \nu|^p \, dx$$

Moreover, it was proven that for any function in $L^p_{loc}(\mathbb{R}^n)$ and p strictly larger than 1,

$$|Df| \in L^p_{loc}(\mathbb{R}^n) \iff \liminf_{\varepsilon \to 0} K_\varepsilon(p, f) < \infty$$

which is a very peculiar characterization of $W^{1,p}$ as the right hand side of the equivalence sign does not involve weak derivatives whatsoever. Sadly, as we stated before, this last assertion does not hold for $p = 1$, as finiteness of the infimum limit of $K_\varepsilon(f) = K_\varepsilon(1, f)$ is a characterization of BV and not of $W^{1,1}$.

Chapter 2

The space *Lip* and related function spaces

2.1 Definition of the space of Lipschitz functions on a compact metric space

Let (K, ρ) be a compact metric space and $f : K \to \mathbb{R}$ a function defined on it. Let $\mathfrak{B}(K, \rho)$ (or simply $\mathfrak{B}(K)$, if no ambiguity arises) be the set of all balls in (K, ρ). We will use the notation $B_r^{K,\rho}(x_0)$ to denote a ball of radius r, centered at x_0 in the metric space (K, ρ), omitting K and ρ from the notation whenever they are understood from the context.
We define the **oscillation** ω_f as a function on the set $\mathfrak{B}(K)$:

$$\omega_f : B \in \mathfrak{B}(K) \mapsto \sup_{x,y \in B} |f(x) - f(y)| \in \mathbb{R}.$$

For any fixed point $x_0 \in K$ and function $f : K \to \mathbb{R}$, the function $\omega_f(B_r(x_0))$ is decreasing with r and so we can define

$$\omega_f(x_0) := \lim_{r \to 0^+} \omega_f(B_r(x_0)). \tag{2.1}$$

Oscillation is an useful concept as it allows to quantify exactly many abstract concepts and unify a variety of different definitions. To make the simplest example, a function $f : K \to \mathbb{R}$ is continuous at x_0 if and only if $\omega_f(x_0) = 0$ and continuous in K if and only if $\omega_f \equiv 0$ but wherever a function f is not continuous, ω_f is a quantitative measure of the discontinuity at that point. For example, for the single variable real functions

$$f(x) = \begin{cases} 0 & \text{if } x = 0 \\ \sin(1/x) & \text{if } x \in (0, 1] \end{cases} \qquad \text{and} \qquad g(x) = \chi_{[\frac{1}{2}, 0]}(x)$$

defined from $[0, 1]$ to $\mathbb{R}$ we have $\omega_f(0) = 2$ and $\omega_g(\frac{1}{2}) = 1$.
Classically, a function f is said to be Lipschitz in K with Lipschitz constant L if

$$\sup_{\substack{x,y \in K \\ x \neq y}} \frac{|f(x) - f(y)|}{\rho(x, y)} = L < +\infty$$

Of course, though, in terms of oscillation, a function is Lipschitz if and only if

$$\frac{\omega_f(B_r(x))}{2r} \leq L' < \infty, \quad \forall x \in K, \quad \forall r > 0$$

(but the best constant L' is not necessarily equal to the Lipschitz constant L as defined earlier).
Any Lipschitz function f is of course uniformly continuous and bounded: the difference between its maximum and minimum on K is at most $L \cdot \mathrm{diam}(K)$, where L is its Lipschitz constant and $\mathrm{diam}(K)$ denotes the diameter of the compact set K. Even more so, it is possible to prove Lipschitz functions are absolutely continuous.
Lipschitzianity is a global property, but a local analogue can be given in the following way: we will say f is locally Lipschitz at $x \in K$ if there exists a neighbourhood U_x of x in K such that f is Lipschitz in U_x.
The space $Lip(K, \rho)$ consisting of all Lipschitz functions defined from the compact metric space (K, ρ) is a vector space that, when normed with the norm

$$\|f\|_{Lip} = \sup_{\substack{x,y \in K \\ x \neq y}} \frac{|f(x) - f(y)|}{\rho(x,y)} + \|f\|_\infty$$

or, equivalently, with the norm

$$\|f\|_1 = \max\left\{ \sup_{\substack{x,y \in K \\ x \neq y}} \frac{|f(x) - f(y)|}{\rho(x,y)}, \|f\|_\infty \right\}$$

is a Banach Space.
Whenever the compact metric space is chosen as (K, d), where K is a compact subset of an Euclidean space $\mathbb{R}^n$ and d is the Euclidean distance, the choice of the metric d will be often omitted from the notation and we have even higher knowledge of the properties Lipschitz functions: the following is a classical theorem by Rademacher, implying the Lipschitz condition is relevant to the existence of derivatives.

Theorem 2.1 (Rademacher)**.** *Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be a Lipschitz function with Lipschitz constant L. Then the set Diff$(f) \subseteq U$ consisting of points in which f is differentiable has full measure in U, i.e. is such that*

$$|U \setminus \mathit{Diff}(f)| = 0$$

and the gradient Df, defined almost everywhere, is essentially bounded by L.

From this theorem, one can prove that the space of Lipschitz functions $Lip(K)$, where $K \subseteq \mathbb{R}^n$ is nothing but the Sobolev space $W^{1,\infty}(K)$.

2.2 The spaces Lip_α and lip_α of Hölder functions

With the same notation as for the previous section and for any $\alpha \in (0,1]$ we now introduce the Hölder spaces

$$Lip_\alpha(K,\rho) := \left\{ f : [f]_\alpha := \sup_{x,y\in K} \frac{|f(x)-f(y)|}{\rho^\alpha(x,y)} < \infty \right\}. \tag{2.2}$$

with the norm

$$\|f\|_\alpha = \max\{[f]_\alpha, \|f\|_\infty\}$$

We remark that every Hölder space can be seen as a Lipschitz space as for any metric ρ and for any $\alpha \in (0,1]$ we have that ρ^α is still a metric and of course $Lip_\alpha(K,\rho) = Lip(K,\rho^\alpha)$. Of course any Lipschitz space is also a Hölder space with choice of $\alpha = 1$, but some Lipschitz spaces cannot be expressed as Hölder spaces for any choice of α strictly smaller than 1. This is the case, for example, for any compact metric space that is compact subset of an Euclidean space with Euclidean distance.
A relevant subspace of $Lip(K,\rho)$ is the subspace

$$lip(K,\rho) := \left\{ f \in Lip(K,\rho) : \limsup_{\rho(x,y)\to 0} \frac{|f(x)-f(y)|}{\rho(x,y)} = 0 \right\}. \tag{2.3}$$

and we can also define

$$lip_\alpha(K,\rho) := \left\{ f \in Lip_\alpha(K,\rho) : \limsup_{\rho(x,y)\to 0} \frac{|f(x)-f(y)|}{\rho^\alpha(x,y)} = 0 \right\}. \tag{2.4}$$

For some choices of a metric ρ, it may happen that $lip(K,\rho)$ is trivial, i.e. it only contains constant functions. This happens, again, for any compact metric space that is compact and connected subset of an Euclidean space with Euclidean distance. However, whenever $\alpha < 1$, the following proposition (see Example 3.8.3 of [50]) shows $lip_\alpha(K,\rho)$ is never, in this sense, trivial.

Proposition 2.2. *For any compact metric space (K,ρ) and $0 < \alpha < \beta \le 1$ we have*

$$Lip_\beta(K,\rho) \subseteq lip_\alpha(K,\rho)$$

Whatever the compact metric space (K,ρ), the space $Lip_\beta(K,\rho)$ will always contain at least one family of functions that is not constant, i.e. for any choice of x_0 in K the distance functions:

$$x \mapsto \rho^\beta(x_0,x).$$

Also α-Hölder functions for $\alpha < 1$ are necessarily uniformly continuous and bounded but, in the Euclidean case, they are not necessarily differentiabile. Extending an analogous one-dimensional result from De Leeuw, in [56], Wulbert proved the following theorem, isolating an important functional relationship between $lip_\alpha(K)$ and $Lip_\alpha(K)$, with K being a compact subset of some $\mathbb{R}^n$, provided $\alpha < 1$.

Theorem 2.3 (Wulbert, 1974, [139])**.** *Let K be a compact subset of $\mathbb{R}^n$ and $\alpha \in (0,1)$.*

Then

$$(lip_\alpha(K))^{**} \simeq Lip_\alpha(K),$$

i.e. there is an isometry of normed spaces between the two.

In this section we will reach a similar conclusion in a wider setting, making use of the concept of o–O structures, which we now define.

Definition 2.1 (o–O structures, K.M. Perfekt, 2013, [120])**.** We say that a pair of Banach spaces (E_0, E), where E_0 is a subspace of E, form a o–O structure if there exist

- a reflexive separable Banach space X;
- a Banach space Y;
- a family $\mathscr{L}$ of bounded linear operators from X to Y;
- a topology τ on $\mathscr{L}$ that is σ-compact locally compact[1] Hausdorff topology and such that the maps $T_x : \mathscr{L} \to Y$ given by $T_x L = Lx$ for any $x \in X$ are continuous;

such that E is given by

$$E = \left\{ x \in X : \sup_{L \in \mathscr{L}} \|Lx\|_Y < +\infty \right\},$$

it holds $\|x\|_E = \sup\limits_{L \in \mathscr{L}} \|Lx\|_Y$ and E_0 is given by

$$E_0 = \left\{ x \in E : \limsup_{L \to \infty} \|Lx\|_Y = 0 \right\}$$

where $L \to \infty$ is intended in the Alexandrov one point compactification[2] of $(\mathscr{L}, \tau)$. Moreover, we are interested in o–O pairs (E_0, E) satisfying

ASSUMPTION AP. For any $x \in E$ there exists a sequence $\{x_n\}_{n \in \mathbb{N}} \subseteq E_0$ such that $x_n \rightharpoonup x$ in X and $\sup_{n \in \mathbb{N}} \|x_n\|_E \leq \|x\|_E$.

Let us also give the definition of M-ideal.

Definition 2.2 (Harmand, 2006, [87])**.** Let X be a Banach space. A linear projection P is called an L-projection if for any $x \in X$

$$\|x\|_X = \|Px\|_X + \|x - Px\|_X .$$

A closed subspace $J \subseteq X$ is called an L-summand if it is the range of an L-projection. Given a subspace $J \subseteq X$, the subspace $J^\perp = \{x^* \in X^* : \; x^*(y) = 0 \; \forall y \in J\}$ of X^* is

[1] σ-compactness means that the topological space is union of a sequence of compact sets and local compactness means that any point of the topological space has compact neighbourhoods.

[2] For any topological space T, its Alexandrov compactification is the set $T \cup \{\infty\}$ with topology uniquely defined by the condition that any point in T keeps its base of neighbourhoods and a base of neighbourhoods for ∞ is given by the complements in $T \cup \{\infty\}$ of compact sets of T.

called annihilator of J.
A closed subspace $J \subseteq X$ is called M-ideal in X if its annihilator $J^{\perp}$ is an L-summand of X^*.

o–O structures are very rich in functional informations about the two Banach spaces they involve. As a matter of fact, K. M. Perfekt proved the following properties of this structure.

Theorem 2.4 (K.M. Perfekt, 2013, [120])**.** *Under* **Assumption AP**, *E_0^{**} is isometrically isomorphic to E.*
It can also be proved E_0^* is the ***strongly unique predual*** *of E, that is any other Banach space that has dual space isometric to E is itself isometric to E_0^*.*
Also, E_0 is an M-ideal in E and the following distance formula holds:

$$dist(x, E_0)_E = \limsup_{L \to \infty} \|Lx\|_Y.$$

In [16] we identified necessary and sufficient conditions to prove o–O structure for the pair $(lip(K, \rho), Lip(K, \rho))$. One of these conditions is what we called Assumption H.

Definition 2.3. Let (K, ρ) be any compact metric space. We will say that (K, ρ) satisfies **Assumption H** if, for any $f \in Lip(K, \rho)$, A a finite subset of K and $C > 1$ real constant, there exists a function $g \in lip(K, \rho)$ such that $g|_A \equiv f|_A$ and $\|g\|_1 \leq C\|f\|_1$.

The above assumption allows the proof of an approximation property in $Lip(K, \rho)$.

Theorem 2.5 (Angrisani, Ascione, D'Onofrio, Manzo, 2019, [16])**.** *Let us suppose* **Assumption H** *holds and consider $f \in Lip(K, \rho)$.*
Then, there is a sequence $\{f_n\}_{n \in \mathbb{N}} \subset lip(K, \rho)$ pointwise converging to f and such that $\sup_{n \in \mathbb{N}} \|f_n\|_1 \leq \|f\|_1$.

Proof. Since K is totally bounded, it can be covered by a finite number of balls of radius 1, so let us call A_0 the set of centers of these balls. We will now recursively define a sequence of sets $\{A_n\}_{n \in \mathbb{N}}$. Suppose now that we have defined the set A_n and consider the set $K_{n+1} := K \setminus \bigcup_{x \in A_n} B_{2^{-n-1}}(x)$. Since K_{n+1} is a compact and thus totally bounded subset of K, it can be covered by balls of radius 2^{-n-1}, so if we denote by B_{n+1} the corresponding set of centers, we can take $A_{n+1} := A_n \cup B_{n+1}$. This ensures that every point of K has distance less that 2^{-n} from the points in A_n. We also take $C_n := 1 + \frac{1}{n+1}$.
Let g_n be the function from **Assumption H** obtained by considering $A = A_n$ and $C = C_n$ and define $f_n := \frac{g_n}{C_n} \in lip(K, \rho)$. We have that $\|f_n\|_\alpha \leq \|f\|_\alpha$, so the only thing that's left to show is the pointwise convergence, which implies weak convergence in X. We notice that, by definition of f_n, it is enough to show that $g_n \to f$ pointwise. If we define $A_\infty := \bigcup_{n \in \mathbb{N}} A_n$ we see that A_∞ is dense and for all $x \in A_\infty$ the sequence $g_n(x)$ eventually becomes constantly equal to $f(x)$. By using the Lipschitz property we can easily extend the pointwise convergence to the whole K.

□

Another assumption needed to reach the main result of this section is a purely geometrical hypothesis that is needed on the metric space.

Definition 2.4. We say that a metric space (K, ρ) has the *doubling condition* if there exists a positive integer C such that any ball B can be covered by at most C balls having half the radius.
A Borel measure μ on a metric space (K, ρ) is said to have the *doubling condition* if

(i) there exist two balls B_1, B_2 such that $\mu(B_1) > 0$ and $\mu(B_2) < +\infty$;

(ii) there exists a constant $C > 0$ such that

$$\mu(B_{2r}(x)) \leq C\mu(B_r(x)) \tag{2.5}$$

for all $x \in K$ and all $r > 0$. The space (K, ρ, μ) is said to be a doubling metric measure space. A measure μ that satisfies (i) is said to be non-degenerate. Condition (ii) is called doubling condition and the constant C is called doubling constant.

First of all, let us observe that the role of the number 2 in (2.5) is easily replaced by any $c > 1$. Indeed, if we fix a constant $c > 1$, one can show that a non-degenerate measure μ is a doubling measure on (K, ρ) if and only if there exists a constant $C_c > 0$ such that

$$\mu(B_{cr}(x)) \leq C_c\mu(B_r(x)) \tag{2.6}$$

for all $x \in K$ and all $r > 0$.
This property implies that any doubling measure μ is fully supported. Indeed, let us consider a generic $x \in K$ and $r > 0$. Let us suppose by contradiction that $\mu(B_r(x)) = 0$. Then B_1 is not contained in $B_r(x)$. However, there exists a $c > 1$ such that $B_1 \subseteq B_{cr}(x)$, but by doubling condition (2.6) we have $\mu(B_{cr}(x)) = 0$, concluding the proof.
Moreover, if (K, ρ, μ) is a compact doubling metric measure space, then $\mu(K) < +\infty$. Indeed, if we consider $x \in K$ and $r > 0$ such that $B_2 = B_r(x)$, since K is compact, there exists a constant $c \geq 1$ such that $B_{cr}(x) = K$, concluding that $\mu(B_{cr}(x)) < +\infty$ by doubling condition (2.6).
It is easy to see that if a measure is doubling then the underlying metric space must be doubling [52], while the converse is not true in general. However in [111] it is shown that every complete (and in particular compact) doubling metric space can be given a doubling measure.
We will also work with compact metric spaces of the form (K, ρ^α) for some metric ρ. First of all, let us show the following easy lemma.

Lemma 2.6. *Fix $\alpha \in (0, 1)$. If (K, ρ) is a compact doubling metric space then (K, ρ^α) is a doubling metric space. Moreover, any doubling measure μ on (K, ρ) is also doubling on (K, ρ^α).*

Proof. First of all, since (K, ρ) is a doubling metric space, then there exists a doubling measure μ on (K, ρ). Using the doubling condition in the form (2.6) for $c = 2^\alpha$ and setting $C_{2^\alpha} = C_\alpha$ we have

$$\mu(B_{(2r)^\alpha}(x)) \leq C_\alpha\mu(B_{r^\alpha}(x))$$

hence μ is a doubling measure on (K,ρ^α). Finally, since we have a doubling measure on (K,ρ^α), (K,ρ^α) is a doubling metric space. $\square$

Moreover, let us observe, as a consequence of [81, Lemma 4.7], the following lemma.

Lemma 2.7. *Let (K,ρ) be a compact doubling metric space and μ a doubling measure on it. Then there exist two constants $C,Q>0$ such that*

$$\mu(B_r(x)) \geq Cr^Q.$$

In the following we will need the notion of Besov spaces on metric measure spaces. From now on, in this section, let us fix a doubling compact metric space (K,ρ) and a doubling measure μ on K.

Definition 2.5 (Gogatishvili, Koskela, Shanmugalingam, 2010, [78])**.** The **Besov** space of parameters $s\in(0,1)$ and $p,q\in[1,\infty)$ on (K,ρ,μ) is the space

$$\mathscr{B}^s_{p,q}(K,\rho,\mu)=\Big\{f:K\to\mathbb{R}: f\in L^p(K,\mu) \text{ and}$$
$$[f]_{\mathscr{B}^s_{p,q}} := \left[\int_0^{+\infty}\frac{dr}{r}\left[\int_K \fint_{B_r(x)} \frac{|f(x)-f(y)|^p}{r^{sp}}\,d\mu(y)\,d\mu(x)\right]^{q/p}\right]^{1/q} < +\infty\Big\}.$$

It is a Banach space when endowed with the norm

$$\|f\|_{\mathscr{B}^s_{p,q}} = \|f\|_{L^p} + [f]_{\mathscr{B}^s_{p,q}}.$$

The space $\mathscr{B}^s_{p,q}$ is obviously separable as it embeds continuously as a subspace of $L^p(K,\rho)$ which is separable because L^p spaces on separable metric measure spaces are separable.
With the additional assumption that $p=q>2$, a **Clarkson type inequality** (see [34],[79]) can be proved, in the sense that follows: as it is more convenient for technical reasons, introduce the equivalent norm

$$\|f\|'_{\mathscr{B}^s_{p,q}} = \left[\|f\|^p_{L^p} + ([f]_{\mathscr{B}^s_{p,q}})^p\right]^{1/p}$$

and then, in a way similar to how it is done in L^p spaces, prove:

$$\left\|\frac{f+g}{2}\right\|^p_{\mathscr{B}^s_{p,p}} + \left\|\frac{f-g}{2}\right\|^p_{\mathscr{B}^s_{p,p}} \leq \frac{1}{2}\left[\|f\|^p_{\mathscr{B}^s_{p,p}} + \|g\|^p_{\mathscr{B}^s_{p,p}}\right]$$

This allows to prove that $\mathscr{B}^s_{p,p}$ is **uniformly convex**, that is for every $\varepsilon>0$ there exists $\delta>0$ such that:

$$\|f\|_{\mathscr{B}^s_{p,p}}, \|g\|_{\mathscr{B}^s_{p,p}} = 1 \text{ and } \left\|\frac{f+g}{2}\right\|_{\mathscr{B}^s_{p,p}} > 1-\delta \quad \Longrightarrow \quad \|f-g\|_{\mathscr{B}^s_{p,p}} < \varepsilon.$$

Recalling that a theorem by Milman and Pettis shows that every uniformly convex Banach space is reflexive (see for instance [34]), we know that $\mathscr{B}^s_{p,p}$ is a reflexive and

separable Banach space.
In [78] it has been shown that the seminorm $[f]_{\mathscr{B}^s_{p,p}}$ is equivalent to the semi-norm

$$[f]_{\mathscr{B}^s_p} = \int_K \int_K \frac{|f(x)-f(y)|^p}{\rho(x,y)^{\alpha p}\mu(B_{\rho(x,y)}(x))} d\mu(x)d\mu(y).$$

In particular, if $K=\mathbb{R}^n$, ρ is the Euclidean distance and μ is the Lebesgue measure, then $[\cdot]_{\mathscr{B}^s_p}$ is the semi-norm characterizing the fractional Sobolev space $W^{s,p}(\mathbb{R}^n)$, as used in [120].
Let us give a definition of another interesting functional space on metric measure spaces.

Definition 2.6 (Karak, 2019, [102])**.** Let $f \in L^p(K,\mu)$. Then we say that g is a Hajłasz s-gradient if

$$|f(x)-f(y)| \leq \rho^s(x,y)|g(x)+g(y)|.$$

We denote with $\mathscr{D}_s(f)$ the set of all Hajłasz gradients of f and with $\mathscr{D}^p_s(f)$ the set of the Hajłasz gradients of f that are in $L^p(K,\mu)$. We say that f belongs to the fractional Hajłasz-Sobolev space $\mathscr{H}^{s,p}(K,\rho,\mu)$ if $\mathscr{D}^p_s(f) \neq \emptyset$. In particular $\mathscr{H}^{s,p}(K,\rho,\mu)$ is a Banach space when endowed with the norm

$$\|f\|_{\mathscr{H}^{s,p}} := \|f\|_{L^p} + [f]_{\mathscr{H}^{s,p}}$$

where

$$[f]_{\mathscr{H}^{s,p}} := \inf_{g\in\mathscr{D}^p_s(f)} \|g\|_{L^p}$$

Let us first observe that the fractional Hajłasz-Sobolev space $\mathscr{H}^{s,p}(K,\rho,\mu)$ coincides with the Hajłasz-Sobolev space $\mathscr{H}^{1,p}(K,\rho^s,\mu)$. Hence, in particular, a Morrey-type embedding theorem can be shown, as a direct consequence of [81, Theorem 8.7].

Theorem 2.8. *Let (K,ρ,μ) be a compact metric measure space such that $\mu(B_r(x)) \geq cr^Q$ for some constants $c>0, Q\geq 0$ and for all $x\in K, r>0$, and suppose $p>\frac{Q}{s}$. Then there exists a constant C such that for any function $u\in\mathscr{H}^{s,p}(K,\rho,\mu)$ and any $g\in\mathscr{D}^p_s(u)$ it holds*

$$|u(x)-u(y)| \leq Cd(x,y)^{s-\frac{Q}{p}} \|g\|_{L^p}, \ \forall x,y\in K.$$

Other embedding theorems can be shown, also for more general spaces (see, for instance, [102]).
From this punctual estimate we deduce that if $u\in\mathscr{H}^{s,p}(K,\rho,\mu)$ for $p>\frac{Q}{s}$ then u is $\left(s-\frac{Q}{p}\right)$-Hölder continuous. In Lemma 2.7 we have shown that doubling measures satisfy the previous condition for some $Q\geq 0$. From now on let Q be such a constant. We have

Proposition 2.9. *Let (K,ρ,μ) be a doubling compact metric measure space and $p>\frac{Q}{s}$. Then $\mathscr{H}^{s,p}(K,\rho,\mu)$ embeds with continuity in L^∞.*

Proof. Let us denote with $D=\operatorname{diam}(K)$. Fix $u\in\mathscr{H}^{s,p}(K,\rho,\mu)$, $g\in\mathscr{D}^p_s(u)$ and

$x, y \in K$. Observe that, by using the previous punctual estimate,

$$\begin{aligned}|u(x)| &\le |u(y)| + |u(x) - u(y)| \\ &\le |u(y)| + Cd(x,y)^{s-\frac{Q}{p}} \|g\|_{L^p} \\ &\le |u(y)| + CD^{s-\frac{Q}{p}} \|g\|_{L^p} .\end{aligned}$$

By using the p-homogeneity and convexity of the function $t \mapsto t^p$ for $t > 0$ we have

$$|u(x)|^p \le C_1(|u(y)|^p + C^p D^{ps-Q} \|g\|_{L^p}^p)$$

and then integrating in $d\mu(y)$, setting $M = \mu(K)$, we have

$$|u(x)|^p \le \frac{C_1}{M}(\|u\|_{L^p}^p + MC^p D^{ps-Q} \|g\|_{L^p}^p).$$

Now, by using the fact that there exists a constant C_p such that

$$(a^p + b^p)^{\frac{1}{p}} \le C_p(a+b)$$

for any $a, b > 0$, we have

$$|u(x)| \le C_2(\|u\|_{L^p} + \|g\|_{L^p}).$$

Now let us take the infimum over $\mathscr{D}_s^p(u)$ to achieve

$$|u(x)| \le C_2 \|u\|_{\mathscr{H}^{s,p}} ,$$

and then, taking the maximum on K, we have

$$\|u\|_{L^\infty} \le C_2 \|u\|_{\mathscr{H}^{s,p}} .$$

□

Concerning the relation between $\mathscr{B}^s_{p,p}$ and $\mathscr{H}^{s,p}$, one can show the following embedding theorem as a direct consequence of [78, Lemma 6.1].

Theorem 2.10. *Let (K, ρ, μ) be a doubling compact metric measure space. Then $\mathscr{B}^s_{p,p}(K, \rho, \mu)$ embeds with continuity in $\mathscr{H}^{s,p}(K, \rho, \mu)$.*

Actually, the statement of [78, Lemma 6.1] only refers to the inclusion of $\mathscr{B}^s_{p,p}(K, \rho, \mu)$ in $\mathscr{H}^{s,p}(K, \rho, \mu)$. However, in the proof, it is shown that there exists a constant $C > 0$, depending only on the doubling constant of μ, such that for any $u \in \mathscr{B}^s_{p,p}(K, \rho, \mu)$ there exists a $g \in \mathscr{D}^p_s(u)$ such that $[u]_{\mathscr{H}^{s,p}} \le \|g\|_{L^p} \le C[u]_{\mathscr{B}^s_{p,p}}$. Summing on both sides $\|u\|_{L^p}$ one has the continuous embedding.
As a corollary of the previous two embedding theorems we have the following

Corollary 2.11. *Let (K, ρ, μ) be a doubling compact metric measure space and $p > \frac{Q}{s}$. Then $\mathscr{B}^s_{p,p}(K, \rho, \mu)$ embeds continuously in L^∞.*

Now we are ready to show the main result of the section.

Theorem 2.12. *Let (K,ρ,μ) be a doubling compact metric measure space. Then the pair $(lip(K,\rho), Lip(K,\rho))$ exhibit a o–O structure if and only if* **Assumption H** *holds, in which case we have the following consequences:*

- $(lip(K,\rho))^{**} \simeq Lip(K,\rho)$ *isometrically;*
- *for $f \in \mathrm{Lip}(K,\rho)$ the following distance formula holds:*

$$\operatorname*{dist}_{Lip(K,\rho)}(f, lip(K,\rho)) = \limsup_{\rho(x,y)\to 0} \frac{|f(x)-f(y)|}{\rho(x,y)}; \tag{2.7}$$

- $lip(K,\rho)$ *is an M-ideal in $Lip(K,\rho)$, hence*

$$(\mathrm{Lip}(K,\rho))^* \simeq (lip(K,\rho))^* \oplus_1 (lip(K,\rho))^{\perp}; \tag{2.8}$$

- $(lip(K,\rho))^*$ *is the strongly unique predual of* $\mathrm{Lip}(K,\rho)$.

Proof. First of all we need a reflexive and separable Banach space X in which we can embed $\mathrm{Lip}(K,\rho)$. Thus fix $s\in(0,1)$ and $p>\frac{Q}{s}$ and consider

$$X = \overline{\mathrm{Lip}(K,\rho)}^{\mathscr{B}^s_{p,p}},$$

since we have shown that $\mathscr{B}^s_{p,p}$ is reflexive, separable and $\mathrm{Lip}(K,\rho)$ continuously embeds in it, so the closure of Lip in $\mathscr{B}^s_{p,p}$ also has these properties and will be our X.
As Banach space Y let us choose $\mathbb{R}\times\mathbb{R}$, endowed with the L^∞ norm, i.e.

$$\|(x,y)\|_{\mathbb{R}\times\mathbb{R}} = \max\{|x|,|y|\}.$$

Our family of operators will be the following:

$$\mathscr{L} = \left\{ L_{x,y,z} : f\in X \mapsto \left(\frac{f(x)-f(y)}{\rho(x,y)}, \frac{\rho(x,y)}{D} f(z) \right) \in \mathbb{R}^2, x,y,z\in K, x\neq y \right\}.$$

It is clear that these operators are linear.
If we set $V := K^2 \backslash \mathrm{Diag}(K^2)$, we can give $\mathscr{L}$ the product topology of $V\times K$, where on V we have the trace topology induced by the topology on K^2. In the following we will identify $\mathscr{L}$ with $W := V\times K$.
Since K is a compact metric space, it is σ-compact, locally compact, Hausdorff and separable and so is also V. These properties easily transfer to $\mathscr{L}$, being it a product space. In particular an exhaustive sequence K_n of compact subsets of $\mathscr{L}$ is given by

$$K_n = \left\{ (x,y)\in K^2 : \ \rho(x,y) \geq \frac{1}{n} \right\} \times K$$

hence taking the limit as $L\to\infty$ is equivalent to taking the limit as

$$\rho(x,y)\to 0.$$

Now we need to show the continuity of the maps $T_f : L\in\mathscr{L} \mapsto L(f) \in \mathbb{R}\times\mathbb{R}$ for

$f \in X$. We notice that it is enough to prove this for $f \in Lip(K,\rho)$, since we can use a diagonal argument, combined with the boundedness of the operators themselves, to extend this to the whole X.
This is easy because $(x_n, y_n, z_n) \to (x, y, z)$ as n goes to infinity implies

$$\rho(x_n, y_n) \to \rho(x, y) \text{ and } \rho(z_n, z) \to 0,$$

so using the continuity of f and ρ we easily obtain

$$\max\left\{\left|\frac{f(x_n)-f(y_n)}{\rho(x_n,y_n)} - \frac{f(x)-f(y)}{\rho(x,y)}\right|, \left|\frac{\rho(x_n,y_n)}{D}f(z_n) - \frac{\rho(x,y)}{D}f(z)\right|\right\} \to 0$$

proving that T_f is continuous for any $f \in \mathrm{Lip}(K,\rho)$.
It is easy to observe that

$$\sup_{(x,y,z)\in W} \|L_{x,y,z}f\|_{\mathbb{R}\times\mathbb{R}} = \sup_{(x,y,z)\in W} \max\left\{\left|\frac{f(x)-f(y)}{\rho(x,y)}\right|, \frac{\rho(x,y)}{D}|f(z)|\right\}$$
$$= \max\{[f]_1, \|f\|_\infty\} = \|f\|_1 ,$$

while the o-structure for $lip(K,\rho)$ follows from the inequality

$$\frac{|f(x)-f(y)|}{\rho(x,y)} \le \|L_{x,y,z}f\|_{\mathbb{R}\times\mathbb{R}} \le \frac{|f(x)-f(y)|}{\rho(x,y)} + \frac{\rho(x,y)}{D}\|f\|_{L^\infty} .$$

Concerning the continuity of $L_{x,y,z}$, let us recall, from Corollary 2.11, that there exists a constant C such that $\|f\|_{\mathscr{B}^s_{p,p}} \ge C\|f\|_{L^\infty}$. Hence we have

$$\frac{|f(x)-f(y)|}{\rho(x,y)\,\|f\|_{\mathscr{B}^s_{p,p}}} \le \frac{2}{C\rho(x,y)}$$

while

$$\frac{\rho(x,y)|f(z)|}{D\,\|f\|_{\mathscr{B}^s_{p,p}}} \le \frac{\rho(x,y)}{CD},$$

thus $L_{x,y,z} : X \to \mathbb{R}\times\mathbb{R}$ is a bounded linear operator.
Finally let us observe that we have shown that, supposed that **Assumption H** holds, for any $f \in Lip(K,\rho)$ there exists a sequence $\{f_n\}_{n\in\mathbb{N}} \subset lip(K,\rho)$ such that $f_n \to f$ point-wise and $\sup_{n\in\mathbb{N}} \|f_n\|_1 \le \|f\|_1$, hence, by Banach-Alaoglu theorem, we can extract a subsequence of f_n that weakly converges to f in X, concluding the proof of one implication.
Concerning the other implication, let us suppose that the o–O structure holds. Then we know that $(lip(K,\rho))^{**} \simeq Lip(K,\rho)$ isometrically. However, in [82] it is shown that such isometry is equivalent to **Assumption H**, concluding the proof. □

The following proposition by Hanin allows to prove that all of the sufficient conditions for exhibiting o–O structure are satisfied by $(lip_\alpha(K,\rho), Lip_\alpha(K,\rho))$ provided $\alpha < 1$.

Proposition 2.13 (Hanin, 1994, [82])**.** *Let (K,ρ) be a compact metric space, $\alpha \in (0,1)$, $f \in Lip_\alpha(K)$, A a finite subset of K and $C > 1$ a real constant. Then there exists a* *function* $g \in Lip_1(K)$ *such that* $g|_A \equiv f|_A$ *and* $\|g\|_\alpha \leq C\|f\|_\alpha$.

Corollary 2.14. *For any* $\alpha \in (0,1)$ *the pair*

$$(lip_\alpha(K,\rho), Lip_\alpha(K,\rho))$$

exhibits an o–O structure.

Proof. By Proposition 2.13, Assumption H is satisfied, hence the thesis. $\square$

2.3 Atomic decomposition of the space of Borel Measures: an application to measure theory

Let (K, ρ) be a compact metric space and let us denote by $M(K)$ the space of all finite signed Borel measures on K. Of course $M(K)$ is a vector space and it is usually normed with total variation, i.e.

$$\nu \mapsto |\nu|(K) := \nu^+(K) + \nu^-(K).$$

However, we will be interested in a different norm on $M(K)$, which has the advantage of being strictly related to the metric ρ and, as a matter of fact, we can even say it extends it, in some sense. Of course K can be embedded as a set of $M(K)$ by mapping each point x of K to the Dirac measure δ_x concentrated at x. This norm we are about to define, called the Kantorovich-Rubinstein norm in honour of the mathematicians that first thought of it, has the property that

$$\|\delta_x - \delta_y\|_\rho = \rho(x, y)$$

hence, in the sense specified above, extending the metric ρ from K to the set of finite signed Borel measures on K.
To define it on $M(K)$ we first have to define it on a subspace, $M_0(K)$, which is the space of all measures ν in $M(K)$ such that $\nu(K) = 0$. These measures are called **vanishing measures**.
To any such a measure ν we associate the family Ψ_ν of all non-negative measures $\psi \in M(K \times K)$ such that for any Borel $F \subset K$ *the balance condition*

$$\psi(K, F) - \psi(F, K) = \nu(F) = \nu^+(F) - \nu^-(F) \tag{2.9}$$

Remark 2.15. This balance condition has an interpretation in Optimal Transport Theory, as we can read Equation (2.9) to say that ψ is a transport plan from ν^+ to ν^-. As a matter of fact, finite non-negative Borel measures on K are the perfect mathematical object to describe (eventually degenerate, i.e. concentrated in null measure sets) distributions of mass in a set K, this is because any question of the type: *how much of the mass distributed according to ν is in the Borel subset B of K* is naturally answered by computing $\nu(B)$. In the framework of Optimal Transport Theory, a non-negative measure Ψ on $K \times K$ satisfying the balance condition is a proposed method of redistributing the mass from the way it is distributed according to ν^+ to the way it is distributed according to ν^-: with this interpretation, we read $\Psi(A, B)$ as how much of the mass from A is moved to B and we read the balance condition as saying that the mass that Ψ moves from all of K to F minus the mass that Ψ moves from K away to F is exactly equal to the difference of masses in F in the starting and target distributions. Integrating $\rho(x, y)$ over $K \times K$ with respect to the measure Ψ means computing the cost of the transport plan Ψ, as $\rho(x, y)d\psi(x, y)$ is the infinitesimal cost of moving mass from x to y, with the assumption that the cost is proportional to the distance.

The norm of $\nu \in M_0(K)$ is defined by:

$$||\nu||_\rho^0 = \inf_{\psi \in \Psi_\nu} \int\int_{K\times K} \rho(x,y) d\psi(x,y) \tag{2.10}$$

and its value with the corresponding optimal transfer gives the solution to the mass transfer problem from μ^- to μ^+ with cost ρ (see [9] and reference therein); moreover for generic $\mu \in M(K)$ by

$$||\mu||_\rho = \inf_{\nu \in M_0(K)} \{||\nu||_\rho^0 + |\mu - \nu|(K)\} \tag{2.11}$$

It is worth noting that the infimum in (2.10) is the same if taken over a smaller set $\bar{\Psi}_\nu$ of measures $\psi \in M_+(K \times K)$ such that

$$\psi(\cdot, K) = \nu_-, \qquad \psi(K, \cdot) = \nu_+$$

where ν_- and ν_+ are the negative and the positive variation of ν and the set of measures with finite support is dense in $M(K)$ in the KR-norm.
When equipped with this norm, $M(K)$ enjoys important properties which are well revealed by Dunford-Pettis Theorem [28]: in particular weakly compact sets are characterised as those consisting in uniformly bounded and uniformly additive (in other words: equiabsolutely continuous) sets of measures. Of course "weakly" means with respect to the duality $\sigma(M, M')$. Unfortunately there is not trasparent characterization of the dual M'.
The main property of the space $M(K)$ endowed with the KR norm is the following duality relation (see, for instance, [82, Theorem 0]):

$$(M(K))^* = Lip(K, \rho) \tag{2.12}$$

(alternatively, we also have $(M_0(K))^* = Lip(K, \rho)$, where in this case the functions in *Lip* are considered modulo a constant). What Hanin did in [82] is to find a necessary and sufficient condition for ρ to enjoy this other duality relation:

$$(lip(K, \rho))^* = (M(K))^c, \tag{2.13}$$

where $(\cdot)^c$ refers to the completion. These two results combined, in particular, would imply biduality for the pair $(lip(K, \rho), Lip(K, \rho))$.
A family of distances that satisfies the assumption given by Hanin is given by $\{\tilde{\rho}^\alpha\}_{0<\alpha<1}$, where $\tilde{\rho}$ is a given distance, so in particular we obtain that

$$(lip_\alpha(K, \tilde{\rho}))^{**} = \mathrm{Lip}_\alpha(K, \tilde{\rho}),$$

with the intermediate space being $M(K)$ endowed with the Kantorovich-Rubinstein norm $\|\cdot\|_{\tilde{\rho}^\alpha}$ corresponding in this case to the metric $\tilde{\rho}^\alpha$.
What we proved in Section 2.2 about o–O structure for $(lip_\alpha(K, \rho), Lip_\alpha(K, \rho))$ will have a consequence on the possibility to decompose any finite signed Borel measure on (K, ρ) as series of simpler finite signed Borel measures on (K, ρ) that we will call atoms as their support is of cardinality at most 3. This will be the main result of this section.

To be able to reach it, let us show the following abstract result on o–O structures, which deals with atomic decomposition of elements of the predual E_* of a given Banach space E. The first occurrence in the literature of the concept of atomic decompositions was in a paper of E. Stein when considering the Hardy space $E_* = H^1$ (see [66],[53]: in that case $E = BMO$, the space of functions of bounded mean oscillation introduced by John and Nirenberg in 1961 (see Chapter 4).

Theorem 2.16 (Angrisani, Ascione, D'Onofrio, Manzo, 2020, [16])**.** *Let (E_0, E) be an o-O pair such that the space of the operators $(\mathscr{L}, \tau)$ is separable. Then there exists a constant $C \in (0,1)$ such that for any $\Phi \in E_*$ there exist two sequences $(g_n)_{n\in\mathbb{N}} \subset E_*$ and $(\lambda_n)_{n\in\mathbb{N}} \in \ell^1(\mathbb{R}^+)$ such that $\|g_n\|_{E_*} = 1$,*

$$\Phi = \sum_{n\in\mathbb{N}} \lambda_n g_n$$

and

$$C\sum_{n=1}^{+\infty} \lambda_n \leq \|\Phi\|_{E_*} \leq \sum_{n=1}^{+\infty} \lambda_n.$$

Proof. Since we have that the map $T_x : L \in \mathscr{L} \mapsto Lx \in Y$ is continuous, we can also observe that

$$E = \left\{ x \in X : \sup_{L\in\mathscr{L}} \|Lx\|_Y < +\infty \right\}.$$

From [59, Theorem 3] we know that for $\Phi \in E_*$ there exists a sequence $(y_n^*)_{n\in\mathbb{N}} \in \ell^1(Y^*)$ such that

$$\Phi = \sum_{n=1}^{+\infty} L_n^* y_n^*$$

for some sequence $\{L_n\}_{n\geq 0} \subset \mathscr{L}$ independent from Φ, where $L_n^* \in \mathrm{Lin}(Y^*, E_*)$ is the adjoint operator of L_n. Now let us recall that $\|L_n\|_{\mathrm{Lin}(E,Y)} = \|L_n^*\|_{\mathrm{Lin}(Y^*,E_*)}$. By definition of E, we can use the Banach-Steinhaus theorem to assure that there exists a constant K_1 such that

$$\|L_n^*\|_{\mathrm{Lin}(Y^*,E_*)} = \|L_n\|_{\mathrm{Lin}(E,Y)} \leq K_1.$$

Let us then define

$$g_n = \frac{L_n^* y_n^*}{\|L_n^* y_n^*\|_{E_*}}$$

($g_n = 0$ if $L_n^* y_n^* = 0$) and $\lambda_n = \|L_n^* y_n^*\|_{E_*}$ to obtain

$$\Phi = \sum_{n=1}^{+\infty} \lambda_n g_n.$$

Now let us observe that, since $\|g_n\|_{E_*} = 1$,

$$\|\Phi\|_{E_*} \leq \sum_{n=1}^{+\infty} \lambda_n.$$

Let us also observe that

$$\|L_n^* y_n^*\|_{E_*} \le \|L_n^*\|_{\mathrm{Lin}(Y^*,E_*)} \|y_n^*\|_{Y^*} \le K_1 \|y_n^*\|_{Y^*}$$

so that

$$\|y_n^*\|_{Y^*} \ge \frac{1}{K_1}\lambda_n.$$

Since the predual is strongly unique, by using the isometry in [59, Theorem 3] we have that there exists a constant K_2 (independent from Φ) such that

$$\|\Phi\|_{E_*} \ge K_2 \sum_{n=1}^{+\infty} \|y_n^*\|_{Y^*} \ge \frac{K_2}{K_1} \sum_{n=1}^{+\infty} \lambda_n.$$

Pose $C = \frac{K_2}{K_1}$ to conclude the proof. □

Now we are ready to give an atomic decomposition of the space $M(K)$ endowed with the Kantorovich-Rubinstein norm induced by a metric ρ^α for some $\alpha < 1$.

Theorem 2.17 (Angrisani, Ascione, D'Onofrio, Manzo, 2020, [16]).
Fix $\alpha \in (0,1)$ and let $\mu \in M(K)$.
Then there exist a sequence of atomic measures $(\mu_n)_{n\in\mathbb{N}} \subset M(K)$ with

$$card(supp(\mu_n)) \le 3$$

and a sequence $(\gamma_n)_{n\in\mathbb{N}} \in \ell^1(\mathbb{R})$ with $\gamma_n \ge 0$ such that

$$\mu = \sum_{n=1}^{+\infty} \gamma_n \mu_n$$

where the convergence is intended in the Kantorovich-Rubinstein norm with respect to ρ^α. Moreover there is $C > 0$ such that

$$C \sum_{n=1}^{+\infty} \gamma_n \le \|\mu\|_{\rho^\alpha} \le \sum_{n=1}^{+\infty} \gamma_n \tag{2.14}$$

Proof. Since we have shown that the pair $(lip_\alpha(K,\rho), Lip_\alpha(K,\rho))$ admits a o–O structure whenever $\alpha \in (0,1)$, we know that $(lip_\alpha(K,\rho))^*$ is the strongly unique predual of $Lip_\alpha(K,\rho)$ and $(lip_\alpha(K,\rho))^* \simeq (M(K))^c$. Moreover, let us observe that the topology on $\mathscr{L}$ is the one induced by the natural topology on $V \times K \times (0,1]$, hence it is separable. Thus we know that there exist a sequence $(\gamma_n)_{n\in\mathbb{N}} \in \ell^1(\mathbb{R})$ with $\gamma_n \ge 0$ and a sequence $(\mu_n)_{n\in\mathbb{N}} \subset (M(K))^c$ such that

$$\mu = \sum_{n=1}^{+\infty} \gamma_n \mu_n$$

with $\|\mu_n\|_{\rho^\alpha} = 1$.
Now recall that $\mu_n = \frac{L_n^* g_n^*}{\gamma_n}$ by definition for some operators $L_n \in \mathscr{L}$ and for some

$g_n^* \in Y^* = \mathbb{R}^2$. Let us observe that for any $g \in \mathbb{R}^2$ and $f \in Lip_\alpha(K,\rho)$ we have

$$\langle L_n f, g\rangle = \frac{f(x_n) - f(y_n)}{\rho(x_n,y_n)^\alpha} g_1 + \frac{\rho(x_n,y_n)^\alpha}{D^\alpha} f(z_n) g_2$$

while

$$\langle f, L_n^* g\rangle = \int_K f dL_n^* g.$$

From these two relations we have for any $g \in \mathbb{R}^2$

$$L_n^* g = \frac{\delta_{x_n} - \delta_{y_n}}{\rho(x_n,y_n)^\alpha} g_1 + \frac{\rho(x_n,y_n)^\alpha}{D^\alpha} g_2 \delta_{z_n}$$

where δ_x is the Dirac delta measure concentrated in x. Hence we can conclude that μ_n is a purely atomic measure with $card(supp(\mu_n)) \leq 3$. □

Everything we said so far has the flaw of not being applicable to the most classic of all metric spaces: Euclidean spaces. With the choice of $K \subset \mathbb{R}^n$ and $\rho(x,y) = |x-y|$ we can still infer an interesting atomic decomposition for the space $M(K)$ of Borel measures on K and its completion $M(K)^c$ with respect to the Kantorovich-Rubinstein norm. In what follows, since the metric space is fixed, there is no use in specifying the metric as a subscript of the Kantorovich-Rubinstein norm, so that we will simply denote it with $\|\cdot\|_{KR}$ (or $\|\cdot\|_{KR_0}$ for the one defined on $M_0(K)$) honouring the initials of Kantorovich and Rubinstein. What is left of this section is dedicated to this question and its answer contained in [19]. To do so, let us introduce the notation

$$\mathrm{Lip}_0(K) = \mathrm{Lip}(K)/\,\mathbb{R},$$

i.e. the Lipschitz space $\mathrm{Lip}(K)$ modulo constant functions.
In $\mathrm{Lip}_0(K)$, to simplify the notation, we will identify any function $f : K \to \mathbb{R}$ with its equivalence class. If we endow $\mathrm{Lip}_0(K)$ with the norm

$$\|f\|_{\mathrm{Lip}_0(K)} = \sup_{\substack{(x,y)\in K^2 \\ x\neq y}} \frac{|f(x) - f(y)|}{|x-y|},$$

then this normed space is a Banach space.
In the following we will need to embed the spaces $\mathrm{Lip}(K)$ and $\mathrm{Lip}_0(K)$ into suitable reflexive Banach spaces, which can be also seen as a special subcase of the aforementioned Besov spaces. For our purposes, the natural candidates are fractional Sobolev spaces. An almost complete survey on such spaces is given in [57].

Definition 2.7. Let us denote by $W^{s,p}(\Omega)$ for $s \in (0,1)$ and $p > 1$ the fractional Sobolev space consisting of the functions $f \in L^p(\Omega)$ such that

$$\|f\|^p_{\dot{W}^{s,p}(\Omega)} := \int_\Omega \int_\Omega \frac{|f(x) - f(y)|^p}{|x-y|^{ps+n}} dxdy < +\infty.$$

If we endow $W^{s,p}(\Omega)$ with the norm

$$\|f\|_{W^{s,p}(\Omega)} = \|f\|_{\dot{W}^{s,p}(\Omega)} + \|f\|_{L^p(\Omega)}$$

it is a reflexive separable Banach space (since it is uniformly convex by means of a Clarkson-type inequality [79]). The homogeneous fractional Sobolev space $\dot{W}^{s,p}(\Omega)$ is defined as $\dot{W}^{s,p}(\Omega) = W^{s,p}(\Omega)/\mathbb{R}$ and if we endow this space with the norm $\|f\|_{\dot{W}^{s,p}(\Omega)}$ it is a reflexive separable Banach space (for the same reason as before).

Remark 2.18. Let us recall that if $ps > n$ then, by a fractional Morrey-type embedding theorem, we have that $W^{s,p}(\Omega) \hookrightarrow C(K)$ (this is true for any doubling compact metric-measure space as a consequence of the Morrey embedding for Hajłasz-Sobolev spaces [81, Theorem 8.7] and the continuous embedding of Besov spaces into them [78, Lemma 6.1]). In this case we will always consider the continuous version of a function in $W^{s,p}(\Omega)$.

Another characterization of $\dot{W}^{s,p}(\Omega)$ for $sp > n$ is given as the space of functions $f \in W^{s,p}(\Omega)$ such that $f(z) = 0$, for an a priori fixed point $z \in K$ (here we are implicitly using the embedding $W^{s,p}(\Omega) \hookrightarrow C(K)$). In particular we have (by using the same idea adopted for $\mathrm{Lip}(K)$) that the norm

$$\|f\|_{W^{s,p}(\Omega),z} = \|f\|_{\dot{W}^{s,p}(\Omega)} + |f(z)|$$

is equivalent to $\|\cdot\|_{W^{s,p}(\Omega)}$. By identifying $C(K)/\mathbb{R}$ in the same way we have $\dot{W}^{s,p}(\Omega) \hookrightarrow C(K)/\mathbb{R}$.
Moreover, let us make the definition of atoms and dipoles precise

Definition 2.8. We will call δ-**atom** any measure $\mu \in M(K)$ whose support is finite. Moreover, we call **dipole** any measure $\mu \in M_0(K)$ of the form $\mu = \alpha(\delta_x - \delta_y)$ for some $\alpha \in \mathbb{R}$ and $(x, y) \in K^2$, with $\alpha \neq 0$ and $x \neq y$.

To obtain a decomposition of elements of $M_0(K)^c$ - which will induce a decomposition of elements of $M_0(K)$ - we generalize the approach of [16], which relies on the o–O structure of $(\mathrm{lip}_\alpha, \mathrm{Lip}_\alpha)$, by using results contained in [59], which allow us to remove the dependence on the "little o" space, as it is trivial for Lip and Lip_0. We start by writing Lip_0 in a suitable way. Indeed we want to make use of [59, Theorem 3] and, to do this, we have to characterize Lip_0 by means of linear bounded operators $L : X \to Y$ where X is a reflexive Banach space containing Lip_0 and Y is some other Banach space. In particular, we want to find a countable family $\mathscr{F} = \{L_j\}_{j\in\mathbb{N}}$ of such kind of operators such that

$$\mathrm{Lip}_0(K) = \{f \in X : \sup_{j\in\mathbb{N}} \|L_j f\|_Y < +\infty\}.$$

As we will see from the following Lemma, the natural choice we have for Y is $\mathbb{R}$ and for X is $\dot{W}^{s,p}(K)$. Indeed, as we stated before, $\dot{W}^{s,p}(K)$ is separable and reflexive and contains $\mathrm{Lip}_0(K)$ by definition. Moreover, we can chose s and p in a suitable way to obtain $\dot{W}^{s,p}(K)$ continuously embedded into the quotient space $C(K)/\mathbb{R}$. This choice will be useful to show the boundedness of L_j. Here the compactness of K plays a prominent role, as in this case $C(K) \subset L^\infty(K)$ (that will be useful to show boundedness of L_j). In case we choose K to be not compact (for instance unbounded), then we need to find a different approach to show boundedness of the operators. For now, let us focus on the compact case.

Lemma 2.19. *There exists a sequence of functionals*

$$(L_j)_{j\in\mathbb{N}} : X = (\dot{W}^{s,p}(\Omega)) \to Y = \mathbb{R}$$

such that

$$\mathrm{Lip}_0(K) = \{f \in \dot{W}^{s,p}(\Omega) : \sup_{j\in\mathbb{N}} |L_j f| < +\infty\}$$

and

$$\|f\|_{\mathrm{Lip}_0(K)} = \sup_{j\in\mathbb{N}} |L_j f|.$$

Proof. First of all, let us fix $s \in (0,1)$ and $p > 1$ such that $ps > n$, so that $\dot{W}^{s,p}(\Omega) \hookrightarrow C(K)/\mathbb{R}$. Let us consider $D_1 \subset K$, a countable set such that $K = \overline{D_1}$, and $K_1 = K \setminus D_1$. Now let us consider $D_2 \subset K_1$, a countable set such that $K_1 = \overline{D_2}$. Finally, let us define $D = D_1 \times D_2$. Observe that $D_1 \cap D_2 = \emptyset$ so, for any $(x,y) \in D$, $x \neq y$. Moreover, D is countable, hence we can enumerate $D = \{(x_j, y_j)\}_{j\in\mathbb{N}}$. Finally $\overline{D} = K \times K$. Let us define

$$L_j : f \in \dot{W}^{s,p}(\Omega) \to \frac{f(x_j) - f(y_j)}{|x_j - y_j|} \in \mathbb{R}\,.$$

L_j is obviously linear. Moreover, since $\dot{W}^{s,p}(\Omega) \hookrightarrow C(K)/\mathbb{R}$ we have

$$\frac{|f(x_j) - f(y_j)|}{|x_j - y_j|} \le \frac{2}{|x_j - y_j|} \|f\|_{L^\infty(K)} \le C_j \|f\|_{\dot{W}^{s,p}(\Omega)}\,,$$

hence $L_j \in (\dot{W}^{s,p}(\Omega))^*$ for any $j \in \mathbb{N}$.
Finally, let us observe that by density of D in $K \times K$ and continuity of $f \in \dot{W}^{s,p}(\Omega)$,

$$\|f\|_{\mathrm{Lip}_0(K)} = \sup_{j\in\mathbb{N}} |L_j f|$$

concluding the proof. □

Now that we have this rewriting of the definition of $\mathrm{Lip}_0(K)$ we can use the techniques employed in [59] to obtain the desired atomic decomposition. Before giving the main result, let us make use of the ideas behind [59]. Indeed, in such case, one can define the operator $V : \mathrm{Lip}_0 \to \ell^\infty$ as, for any $f \in \mathrm{Lip}_0$, $Vf(j) = L_j f$ for any $j \in \mathbb{N}$. Thus, after obtaining that $V\,\mathrm{Lip}_0 \simeq \mathrm{Lip}_0$ (here we are using $Y = \mathbb{R}$ and $\mathbb{R}^{**} \simeq \mathbb{R}$) it is not difficult to check that a predual $(\mathrm{Lip}_0)_*$ is isometrically isomorphic to ℓ^1/P where $P = (V\,\mathrm{Lip}_0)^\perp \cap \ell^1$ (where with $\perp$ we denote the annihilator). This gives us a series representation of the elements of $M_0(K)^c$ viewed as a predual of $\mathrm{Lip}_0(K)$. This is an underlying reason for the following result.

Theorem 2.20 (Angrisani, Ascione, Manzo, 2021, [19])**.** *There exists a constant $C \in (0,1)$ such that for any choice of functional $\mu \in M_0(K)^c$ there exists a sequence*

$(\alpha_j)_{j\in\mathbb{N}} \in \ell^1(\mathbb{R})$ such that

$$\mu = \sum_{j=1}^{+\infty} \frac{\delta_{x_j} - \delta_{y_j}}{|x_j - y_j|}\alpha_j,$$

where the series converges in KR_0, and

$$C\sum_{j=1}^{+\infty} |\alpha_j| \leq \|\mu\|_{KR_0} \leq \sum_{j=1}^{+\infty} |\alpha_j|, \tag{2.15}$$

where the sequences $(x_j)_{j\in\mathbb{N}}$ and $(y_j)_{j\in\mathbb{N}}$ are defined in Lemma 2.19. Moreover, the sequence of δ-atoms $(\mu_j)_{j\in\mathbb{N}} \subset M_0(K)$ defined as

$$\mu_j = \frac{\delta_{x_j} - \delta_{y_j}}{|x_j - y_j|}$$

spans $M_0(K)^c$, with $\|\mu_j\|_{KR_0} = 1$ for any $j \in \mathbb{N}$. In particular, the δ-atoms μ_j are dipoles, hence have support of cardinality exactly 2.

Proof. By [59, Theorem 3] we know that there exists $C \in (0,1)$ such that for any $\mu \in M_0(K)^c$ there exists a sequence $(\alpha_j)_{j\in\mathbb{N}}$ such that

$$\mu = \sum_{j=1}^{+\infty} L_j^*\alpha_j,$$

where L_j^* is the adjoint operator of L_j, and

$$C\sum_{j=1}^{+\infty} \left\|L_j^*\alpha_j\right\|_{KR_0} \leq \|\mu\|_{KR_0} \leq \sum_{j=1}^{+\infty} \left\|L_j^*\alpha_j\right\|_{KR_0}.$$

Since one has

$$\langle f, L_j^*\alpha_j\rangle = \langle L_j f, \alpha_j\rangle = \frac{f(x_j) - f(y_j)}{|x_j - y_j|}\alpha_j,$$

then

$$L_j^*\alpha_j = \frac{\delta_{x_j} - \delta_{y_j}}{|x_j - y_j|}\alpha_j$$

concluding the proof. □

Remark 2.21. Let us remark that one could use any separable Banach space X such that $\mathrm{Lip}(K) \subset X \subset L^\infty(K)$, where the second inclusion is continuous, in place of $W^{s,p}(\Omega)$.
Moreover, let us observe that the previous theorem provides a ℓ^1/P-atomic decomposition of $M_0(K)^c$.

We now devote to a similar atomic decomposition in the larger space $M(K)^c$, with the help of the space $\mathrm{Lip}(K)$. This time we cannot use the same operators as in Lemma 2.19 since they define a seminorm on $\mathrm{Lip}(K)$. The following rewriting of $\mathrm{Lip}(K)$ relies on the fact that we can consider on $\mathbb{R}^2$ the ℓ^∞ norm.

Lemma 2.22 (Angrisani, Ascione, Manzo, 2021, [19]). *There exists a sequence of operators $(L_j)_{j\in\mathbb{N}} \in \mathrm{Lin}(W^{s,p}(\Omega), \mathbb{R}^2)$, where we equip $\mathbb{R}^2$ with the norm $\|(x,y)\|_{\ell^\infty} = \max\{|x|,|y|\}$, such that*

$$\mathrm{Lip}(K) = \{f \in W^{s,p}(\Omega) : \sup_{j\in\mathbb{N}} \|L_j f\|_{\ell^\infty} < +\infty\}$$

and

$$\|f\|_{\mathrm{Lip}(K)} = \sup_{j\in\mathbb{N}} \|L_j f\|_{\ell^\infty}.$$

Proof. First of all, let us fix $s \in (0,1)$ and $p > 1$ such that $ps > n$, so that $W^{s,p}(\Omega) \hookrightarrow C(K)$, and let us consider the set $D \subset K^2$ defined in Lemma 2.19. Let us define

$$L_j : f \in W^{s,p}(\Omega) \to \left(\frac{f(x_j) - f(y_j)}{|x_j - y_j|}, f(x_j)\right) \in \mathbb{R}^2.$$

L_j is obviously linear. Moreover, since $W^{s,p}(\Omega) \hookrightarrow C(K)$ we have

$$\max\left\{\frac{|f(x_j) - f(y_j)|}{|x_j - y_j|}, |f(x_j)|\right\} \le \max\left\{\frac{2}{|x_j - y_j|}, 1\right\} \|f\|_{L^\infty(K)} \le C_j \|f\|_{W^{s,p}(\Omega)},$$

hence $L_j \in \mathrm{Lin}(W^{s,p}(\Omega), \mathbb{R}^2)$ for any $j \in \mathbb{N}$.
Finally, let us observe that by density of D in $K \times K$, D_1 in K, and continuity of $f \in \dot{W}^{s,p}(\Omega)$ we have

$$\|f\|_{\mathrm{Lip}(K)} = \sup_{j\in\mathbb{N}} \|L_j f\|_{\ell^\infty}$$

concluding the proof. □

As we did earlier, we can now use the techniques of [59] to obtain the atomic decomposition of $M(K)^c$. Let us recall that the starting point of the following result is still the series decomposition that follows from [59, Theorem 3] that we discussed before Theorem 3.2. Moreover, let us recall that Remark 2.21 holds also for the following theorem.

Theorem 2.23 (Angrisani, Ascione, Manzo, 2021, [19]). *There exists a constant $C \in (0,1)$ such that for any functional $\mu \in M(K)^c$ there exists a sequence $((\alpha_j^1, \alpha_j^2))_{j\in\mathbb{N}} \in \ell^1(\mathbb{R}^2)$ such that*

$$\mu = \sum_{j=1}^{+\infty} \left(\frac{\delta_{x_j} - \delta_{y_j}}{|x_j - y_j|}\alpha_j^1 + \delta_{x_j}\alpha_j^2\right),$$

where the series converges in KR, and

$$C\sum_{j=1}^{+\infty}(|\alpha_j^1| + |\alpha_j^2|) \le \|\mu\|_{KR} \le \sum_{j=1}^{+\infty}(|\alpha_j^1| + |\alpha_j^2|), \tag{2.16}$$

where the sequences $(x_j)_{j\in\mathbb{N}}$ and $(y_j)_{j\in\mathbb{N}}$ are defined in Lemma 2.22. In particular, the *sequence of δ-atoms $(\mu_j)_{j\in\mathbb{N}} \subset M(K)$ defined as*

$$\mu_j = \begin{cases} \frac{\delta_{x_k}-\delta_{y_k}}{|x_k-y_k|} & j = 2k-1 \\ \delta_{x_k} & j = 2k \end{cases} \tag{2.17}$$

spans $M(K)^c$, and $\|\mu_j\|_{KR} \leq 1$ for any $j \in \mathbb{N}$.

Proof. By [59, Theorem 3] we know that there exist $\widetilde{C} \in (0,1)$ and $((a_j^1, a_j^2))_{j\in\mathbb{N}} \in \ell^1(\mathbb{R}^2)$ such that for any $\mu \in M(K)^c$

$$\mu = \sum_{j=1}^{+\infty} L_j^* \alpha_j,$$

where L_j^* is the adjoint operator of L_j, $\alpha_j = (\alpha_j^1, \alpha_j^2) \in \mathbb{R}^2$, and

$$\widetilde{C} \sum_{j=1}^{+\infty} \left\|L_j^*\alpha_j\right\|_{KR} \leq \|\mu\|_{KR} \leq \sum_{j=1}^{+\infty} \left\|L_j^*\alpha_j\right\|_{KR}. \tag{2.18}$$

As in the proof of Theorem 2.20, we have

$$L_j^*\alpha_j = \frac{\delta_{x_j} - \delta_{y_j}}{|x_j - y_j|}\alpha_j^1 + \delta_{x_j}\alpha_j^2.$$

Now let us determine some upper and lower bounds for $\left\|L_j^*\alpha_j\right\|_{KR}$. To do this, let us recall that

$$\|\delta_x - \delta_y\|_{KR} = \min\{|x-y|, 2\} \leq |x-y|, \qquad \|\delta_x\|_{KR} = 1 \;\forall x, y \in K.$$

Hence we have for the upper bound

$$\left\|L_j^*\alpha_j\right\|_{KR} \leq \frac{\left\|\delta_{x_j} - \delta_{y_j}\right\|_{KR}}{|x_j - y_j|}|\alpha_j^1| + \left\|\delta_{x_j}\right\|_{KR}|\alpha_j^2| \leq |\alpha_j^1| + |\alpha_j^2|. \tag{2.19}$$

Concerning the lower bound, let us recall (see [84, Section 4.1]) that

$$\left\|L_j^*\alpha_j\right\|_{KR} = \sup_{\|f\|_{\mathrm{Lip}(K)}\leq 1} \left(\frac{f(x_j) - f(y_j)}{|x_j - y_j|}\alpha_j^1 + f(x_j)\alpha_j^2\right). \tag{2.20}$$

Let $d = \mathrm{diam}(K)$ and let us define the functions

$$f_j(z; \alpha_j^1, \alpha_j^2) = \begin{cases} \frac{1-|x_j-z|}{d+1} & \alpha_j^1, \alpha_j^2 \geq 0 \\ \frac{1+|x_j-z|}{d+1} & \alpha_j^1 < 0 \text{ and } \alpha_j^2 \geq 0 \\ \frac{-1-|x_j-z|}{d+1} & \alpha_j^1 \geq 0 \text{ and } \alpha_j^2 < 0 \\ \frac{-1+|x_j-z|}{d+1} & \alpha_j^1, \alpha_j^2 < 0. \end{cases}$$

By using this function as test function in (2.20) we obtain

$$\left\|L_j^*\alpha_j\right\|_{KR} \geq \frac{1}{d+1}(|\alpha_j^1| + |\alpha_j^2|) \tag{2.21}$$

Using Equations (2.19) and (2.21) in Equation (2.18) and setting $C = \frac{\widetilde{C}}{d+1}$ we finally achieve Equation (2.16). □

Remark 2.24. Let us observe that the sequence of δ-atoms $(\mu_j)_{j\in\mathbb{N}}$ is composed by delta measures and dipoles. In particular, if j is even, then μ_j is a delta measure and then the cardinality of its support is exactly 1. On the other hand, if j is odd, then μ_j is a dipole and then the cardinality of its support is exactly 2. Thus we have that for any functional $\mu \in M(K)^c$ there exists a sequence $(\alpha_j)_{j\in\mathbb{N}} \in \ell^1(\mathbb{R})$ such that $\mu = \sum_{j=1}^{+\infty} \alpha_j\mu_j$ where μ_j are δ-atoms with support of cardinality at most 2.
We still have a ℓ^1/P-atomic decomposition of $M(K)^c$. However, in this case, the atoms μ_j are such that $\|\mu_j\|_{KR} \leq 1$. In particular, if $\operatorname{diam} K \leq 2$, we obtain again $\|\mu_j\|_{KR} = 1$ for any $j \in \mathbb{N}$, while, in general, this is true only for even j. Let us also observe that to obtain the lower bound in this case, Kantorovich-Rubinstein duality for the norm on $M(K)^c$ (see [84]) is actually the main tool.

Remark 2.25. Let us stress that both inequalities (2.15) and (2.16) hold true for respectively a certain sequence $(\alpha_j)_{j\in\mathbb{N}} \in \ell^1(\mathbb{R})$ and $((\alpha_j^1, \alpha_j^2))_{j\in\mathbb{N}} \in \ell^1(\mathbb{R}^2)$. In particular, setting $\mu \in M_0(K)^c$, inequality (2.15) is not necessarily valid for any sequence $(\alpha_j)_{j\in\mathbb{N}} \in \ell^1(\mathbb{R})$ such that $\mu = \sum_{j=1}^{+\infty} \frac{\delta_{x_j}-\delta_{y_j}}{|x_j-y_j|}\alpha_j$ in KR_0 with the same constant C. This is due to the fact that we have an isometric isomorphism between $(M_0(K))^c$ and $\ell^1(\mathbb{R})/P$, which is a quotient space with norm $\|[\alpha]\|_{\ell^1/P} = \inf_{\beta\in P} \|\alpha - \beta\|_{\ell^1}$ for any $\alpha = (\alpha_j)_{j\in\mathbb{N}} \in \ell^1(\mathbb{R})$, while the inequality is expressed in terms of the ℓ^1 norm of one of the representatives of the class $[\alpha]$ characterizing $\mu \in (M_0(K))^c$. The same holds for (2.16).

Remark 2.26. Let us observe that if $\mu \in M_0(K)^c$ and

$$\mu = \sum_{j=1}^{+\infty} \left(\frac{\delta_{x_j} - \delta_{y_j}}{|x_j - y_j|}\alpha_j^1 + \delta_{x_j}\alpha_j^2\right), \quad \alpha^1, \alpha^2 \in \ell^1(\mathbb{R})$$

then $\sum_{j=1}^{+\infty} \alpha_j^2 = 0$. This is a direct consequence of the fact that $\mu(K) = 0$.

With the same strategy exploited in the previous remark, we can prove a similar property for any $\mu \in M(K)^c$, as we can see from the following Proposition.

Proposition 2.27 (Angrisani, Ascione, Manzo, 2021, [19])**.** *Let $\mu \in M(K)^c$ and $((\alpha_j^1, \alpha_j^2))_{j\in\mathbb{N}} \in \ell^1(\mathbb{R}^2)$ be the sequence defined in Theorem 2.20. Suppose $((\beta_j^1, \beta_j^2))_{j\in\mathbb{N}} \in \ell^1(\mathbb{R}^2)$ is another sequence such that*

$$\mu = \sum_{j=1}^{+\infty} \left(\frac{\delta_{x_j} - \delta_{y_j}}{|x_j - y_j|}\beta_j^1 + \delta_{x_j}\beta_j^2\right)$$

and inequalities (2.16) *hold. Then*

$$\sum_{j=1}^{+\infty}(\alpha_j^2 - \beta_j^2) = 0$$

Proof. Let us define the following measures for $N \in \mathbb{N}$:

$$\begin{aligned}
\mu_N^\alpha &= \sum_{j=1}^{N} \frac{\delta_{x_j} - \delta_{y_j}}{|x_j - y_j|}\alpha_j^1 + \delta_{x_j}\alpha_j^2 \\
\mu_N^\beta &= \sum_{j=1}^{N} \frac{\delta_{x_j} - \delta_{y_j}}{|x_j - y_j|}\beta_j^1 + \delta_{x_j}\beta_j^2 \\
\nu_N = \mu_N^\alpha - \mu_N^\beta &= \sum_{j=1}^{N} \frac{\delta_{x_j} - \delta_{y_j}}{|x_j - y_j|}(\alpha_j^1 - \beta_j^1) + \delta_{x_j}(\alpha_j^2 - \beta_j^2).
\end{aligned}$$

First of all, let us observe that both μ_N^α and μ_N^β converge in KR norm towards μ. Now let us observe that

$$\begin{aligned}
\sum_{j=1}^{N} \left\| \frac{\delta_{x_y} - \delta_{y_j}}{|x_j - y_j|}(\alpha_j^1 - \beta_j^1) + \delta_{x_j}(\alpha_j^2 - \beta_j^2) \right\|_{KR} &\leq \sum_{j=1}^{N} (|\alpha_j^1 - \beta_j^1| + |\alpha_j^2 - \beta_j^2|) \\
&\leq \sum_{j=1}^{N} (|\alpha_j^1| + |\alpha_2^j|) + \sum_{j=1}^{N} (|\beta_j^1| + |\beta_j^2|).
\end{aligned}$$

Taking the limit as $N \to +\infty$ we obtain that the series in the left-hand side converges and in particular

$$\sum_{j=1}^{+\infty} \left\| \frac{\delta_{x_y} - \delta_{y_j}}{|x_j - y_j|}(\alpha_j - \beta_j) \right\|_{KR} \leq \frac{2}{C} \|\mu\|_{KR}.$$

Now let us consider $M > N > 0$ in $\mathbb{N}$ and observe that

$$\begin{aligned}
\|\nu_N - \nu_M\|_{KR} &= \left\| \sum_{j=N+1}^{M} \frac{\delta_{x_y} - \delta_{y_j}}{|x_j - y_j|}(\alpha_j^1 - \beta_j^1) + \delta_{x_j}(\alpha_j^2 - \beta_j^2) \right\|_{KR} \\
&\leq \sum_{j=N+1}^{M} \left\| \frac{\delta_{x_y} - \delta_{y_j}}{|x_j - y_j|}(\alpha_j^1 - \beta_j^1) + \delta_{x_j}(\alpha_j^2 - \beta_j^2) \right\|_{KR}.
\end{aligned}$$

In particular $(\nu_N)_{N \geq 0}$ is a Cauchy sequence in the Banach space $M(K)^c$, thus it admits a limit $\nu \in M(K)^c$ given by

$$\nu = \sum_{j=1}^{+\infty} \left(\frac{\delta_{x_y} - \delta_{y_j}}{|x_j - y_j|}(\alpha_j^1 - \beta_j^1) + \delta_{x_j}(\alpha_j^2 - \beta_j^2) \right).$$

Now we need to identify ν. To do this, let us just observe that

$$\nu = \lim_{N\to+\infty} \nu_N = \lim_{N\to+\infty} (\mu_N^\alpha - \mu_N^\beta) = \mu - \mu = 0,$$

and then we have

$$0 = \sum_{j=1}^{+\infty} \left(\frac{\delta_{x_y} - \delta_{y_j}}{|x_j - y_j|} (\alpha_j^1 - \beta_j^1) + \delta_{x_j} (\alpha_j^2 - \beta_j^2) \right). \tag{2.22}$$

However, we have, by [84, Equation 1.18]

$$0 = \|0\|_{KR} = \left\| \sum_{j=1}^{+\infty} \left(\frac{\delta_{x_y} - \delta_{y_j}}{|x_j - y_j|} (\alpha_j^1 - \beta_j^1) + \delta_{x_j} (\alpha_j^2 - \beta_j^2) \right) \right\|_{KR} \geq \left| \sum_{j=1}^{+\infty} (\alpha_j^2 - \beta_j^2) \right|,$$

concluding the proof. □

Let us observe that the same strategy does not lead to uniqueness of the coefficients. Indeed Equation (2.22) does not imply

$$\sum_{j=1}^{+\infty} (|\alpha_j^1 - \beta_j^1| + |\alpha_j^2 - \beta_j^2|) = 0,$$

in view of Remark 2.25.

2.4 Generalized Differentiability for Lipschitz Functions: an application to Optimal Control Theory

2.4.1 An introduction to the problem

On the one hand, necessary conditions for minimizers of nonsmooth optimal control problems are usually given in terms of a Maximum Principle involving some kind of generalized differentiation and connected approximating cones: since the early Seventies the vast and rich literature of Nonsmooth Analysis has addressesed this important issue (see [48], [116] and [137] for example). On the other hand, the so-called higher order necessary conditions require an higher regularity of the dynamics. This is the case, in particular, of necessary conditions encountered in Geometric Control Theory, which involve Lie brackets of the dynamics: the mere existence of these brackets requires that the involved vector fields have a sufficient degree of differentiation. Indeed, let us recall that the Lie bracket of two differentiable vector fields h_1, h_2 is defined, in any system of coordinates, as $[h_1, h_2] = Dh_2h_1 - Dh_1h_2$. Therefore, one needs h_1, h_2 to be at least differentiable, or even smoother when iterated brackets are considered. A natural question is: what to do if h_1, h_2 are just Lipschitz continuous? After all, the domains $\mathrm{Diff}(h_1)$ and $\mathrm{Diff}(h_2)$ where h_1 and h_2, respectively, are differentiable, have full measure by Rademacher's Theorem, so that the Lie bracket $[h_1, h_2]$ is classically defined on $\mathrm{Diff}(h_1) \cap \mathrm{Diff}(h_2)$, i.e. almost everywhere. So it is reasonable to wonder if, as it happens in the smooth case, such almost-everywhere-defined brackets would be capable to provide informations about the limit behavior of controlled trajectories having a vanishing first order part. Actually, in connection with *controllability* problems, this question has been addressed in [128] , where the following *everywhere-defined, set-valued* Lie bracket for Lipschitz continuous vector field has been introduced:

$$[f,g]_{set}(x) = \overline{\mathrm{co}}\Big\{ \lim_{n\to\infty} [f,g](x_n),\ x_n \to x,\ x_n \in \mathrm{Diff}(f) \cap \mathrm{Diff}(g) \Big\} \tag{2.23}$$

In [126], [127] this notion of set-valued Lie bracket has proved suitable for the generalization of some classical results of differential geometry –like the *commutativity criterion* or *Frobenius Theorem*– from the smooth case to the non-smooth one but see Simich, Montanari for different approaches. Let us point out that the utilization of some *larger* Lie brackets does not lead to satisfactory generalizations of the known results involving the classical Lie brackets.[3]

Now, controllability and optimality can be regarded (from the view-point of set-separation) as sort of *dual* problems. So, it is somehow natural to investigate the following question:

Q. *Can one add to a standard (non-smooth) Maximum Principle a necessary condition for minima which involves the set-valued Lie brackets defined in* (2.23)*?*

In the present section, which is a condensation of the results obtained in [20](still in state of preprint), we give a positive answer to this question, in relation with an optimal

[3]For instance, as a bracket larger than $[f,g]_{set}$ one could consider the set $\{L_2h_1 - L_1h_2, (L_1, L_2) \in \partial h_1 \times \partial h_2\}$, where ∂h denotes the Clarke's generalized Jacobian. On the other hand, one would obtain a bracket smaller than $[f,g]_{set}$ simply undoing convexification in (2.23).

control problem of the form[4]

$$(P)\qquad \begin{cases} \text{minimize } \Psi(T, x(T)) \quad \text{over processes } (T, u, a, x), \\ \text{where } x : [0, T] \to \mathbb{R}^n \text{ solves} \\ \dfrac{dx}{dt} = f(x) + \displaystyle\sum_{i=1}^{m} g_i(x) u^i \qquad x(0) = \hat{x} \quad \|u\|_1 \le K \quad (T, x(T)) \in \mathfrak{T}. \end{cases} \tag{2.24}$$

Here, the vector fields $g_i : \mathbb{R}^n \to \mathbb{R}^n$, $i = 1, \ldots, m$ *are Lipschitz continuous* and the controls u take values in $\mathbb{R}^m$. In particular, the controls $u(\cdot)$ are unbounded in L^∞ while they are costrained to have L^1-norm less or equal to K. The end-time T is not fixed, but the considered trajectories $x : [0, T] \mapsto \mathbb{R}^n$ verify $(T, x(T)) \in \mathfrak{T}$, where the *target* $\mathfrak{T}$ is a given subset of $\mathbb{R}^+ \times \mathbb{R}^n$. Let us point out that a minimizer for problem (P) rarely exists with the slow growth hypothesis we will make on the cost function. Actually, in order to guarantee existence of a minimizer, one densely embeds the original problem in the extended problem

$$(P_{ext})\qquad \begin{cases} \text{minimize } \Psi(t(S), y(S)) \quad \text{over processes } (S, w^0, w, t, y) \\ \text{where } (t, y) : [0, S] \to \mathbb{R}^{1+n} \text{ solves} \\ \dfrac{dt}{ds} = w^0(s) \\ \dfrac{dy}{ds} = f(y) w^0 + \displaystyle\sum_{i=1}^{m} g_i(y) w^i \qquad (t, y)(0) = (0, \hat{x}) \quad \|w\|_1 \le K, \quad (t, y)(S) \in \mathfrak{T} \end{cases}$$

where the controls (w^0, w) belong to the set

$$\bigcup_{S>0} \Big\{ (w^0, w) \in L^\infty([0, S], \mathbb{R}^+ \times \mathscr{C}) : w^0(s) + |w(s)| = 1 \Big\}.$$

Observe that problem (P_{ext}) is simply obtained from (P) by first reparametrizing time through

$$t(s) := \int_0^s w^0(\sigma) d\sigma, \qquad w^0 > 0, \qquad \begin{cases} w(s) := u(t(s)) w^0(s), \\ y(s) := x(t(s)) \end{cases},$$

and then allowing also *impulsive subintervals* $I \subseteq [0, S]$ (i.e. $w^0(s) \equiv 0 \, \forall s \in I$). Notice, in particular, that (P_{ext}) is a problem with controls which are bounded in L^∞.

As for necessary condition for a minimizer $(\overline{S}, \overline{w}^0, \overline{w}, \overline{y}^0, \overline{y})$ of the extended problem $(P)_{ext}$, intuition suggests that answering question **Q** should mean complementing the usual, non smooth, maximum principle (in one of the available versions) with conditions that tell something about the relation between the corresponding adjoint variable $p(\cdot)$

[4] Here a simplified form of the problem is presented. In the general form of the problem considered later in this section, the drift f is allowed to depend on *bounded control* a, and an additional current cost $\displaystyle\int_0^T l(x, u, a)\, dt$ is considered. Moreover, the controls u are allowed to range over a cone $\mathscr{C} = \mathscr{C}_1 \times \mathscr{C}_2$, where, $\mathscr{C}_1$ is a closed cone in $\mathbb{R}^{m_1}$ containing the coordinate axes, $\mathscr{C}_2$ is a closed cone in $\mathbb{R}^{m_2}$ not containing any straight line, and $m = m_1 + m_2$.

and set-valued Lie brackets $[g_i, g_j]_{set}$.

Our main result, which we state below in a simplified form –see Theorem 2.35 for a rigorous and more general statement–, actually says that at almost every $s \in [0, \overline{S}]$ a sort of generalized Goh condition holds true:

Theorem 2.28 (Angrisani, Rampazzo, 2021, preprint). *Let $(\overline{S}, \overline{w}^0, \overline{w}, \overline{y}^0, \overline{y})$ be a local minimizer for the extended problem (P_{ext}), and assume that $\|w\|_1 < K$. Then there exist multipliers $(p_0, p, \lambda) \in \mathbb{R}^* \times AC([0, \overline{S}]; (\mathbb{R}^n)^*) \times \mathbb{R}^*$ such that, besides the standard necessary conditions of Pontryagin Maximum Principle, we have the following new condition:*

- *For any $i, j \in \{1, \dots, m\}$ and for almost any $s \in [0, \overline{S}]$,*

$$\boxed{0 \in p(s) \cdot [g_i, g_j]_{set}(\overline{y}(s)).} \tag{2.25}$$

Clearly, the importance of such a result relies on the possibility that a trajectory allowed by the standard maximum principle is not a minimizer, in that it does not verify condition (2.25). An example in the last subsection of this section illustrates this circumstance is actually possible.

Let us mention that a crucial tool for the proof of Theorem 2.28 (in the more general version of Theorem 2.35), is represented by the Quasi Differential Quotient, a notion of generalized differentation (valid also for set-valued maps) introduced in [119] as a special case of Sussmann's Approximate Generalized Differential Quotients [4]. This tool and the corresponding notion of approximating multi-cone are flexible enough to allow an expression of variational cones generated by multiple set-valued Lie brackets as well as appying of a peculiar criterion (see Theorem 4.37, p. 265 in [4] or Theorem 2.3 in [119]) which connects set-separation and cone-separability. In turn, the latter property is equivalent to the existence of multipliers (p_0, p, λ).

We conclude this presentation by making some remarks concerning applications and desirable developments. Though the pervasiveness of control-affine control systems is such that no justification is needed for treating a low regularity issue involving them, the fact that we allow unbounded controls must be underlined. Let us just mention that they show up naturally e.g. in Classical Mechanics as soon one identifies the *control* with a moving part of a given mechanical system (but also in other technological applications, e.g. in neurological dynamics or aerospace navigation). However, another, fully mathematical, application maybe deserves some attention: we refer to case that when the control system is driftless, and, in particular, to sub-Riemannian geometry. Actually, in this event, due to rate-independence, any unbounded control can be replaced by a bounded control without changing the resulting trajectory (up to reparametrization). Hence our result might be regarded as a first step towards the investigation of some kind of non-smooth sub-Riemannian geometry.

Throughout this section the elements of an Euclidean space $\mathbb{R}^q$ will be thought as *column vectors*, while *row vectors* will represents the linear one-forms, i.e. the elements of the dual space $(\mathbb{R}^q)^*$. To save space, if $q_1 \dots, q_r$ are positive integers and if $x_i \in \mathbb{R}^{q_i}$ for all $i = 1, \dots, r$, $q := q_1 + \dots, q_r$, the notation $(x_1, \dots, x_r)$ will denote the q-dimensional column vector

$$\begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix}.$$

The space of all linear operators from a vector space X to a vector space Y will be denoted by $\mathrm{Lin}(X,Y)$. Whenever X and Y are Euclidean spaces, it will be understood as a space of matrices.

If $M(\cdot)$ is a $n\times n$-matrix-valued L^1 function on $[0,\bar S]$, let us use the notation $e^{\int_{s_1}^{s_2} M(s)\,ds}$, $s_1, s_2 \in [0,\bar S]$ to denote the fundamental matrix solution of the linear equation

$$\frac{dv}{ds}(s) = M(s)v(s). \tag{2.26}$$

Namely, for every $s_1, s_2 \in [0,\bar S]$, $\bar v \in \mathbb{R}^n$, $e^{\int_{s_1}^{s_2} M}\bar v = v(s_2)$, where $v(\cdot)$ is the solution to (2.26) such that $v(s_1) = \bar v$.

Moreover, we remark that the symbol $\mathrm{dist}\,(\cdot,\cdot)$ is used in this section both for the Euclidean distance between two points and for the distance of a point x from a set S, namely

$$\mathrm{dist}\,(x,S) = \inf_{s\in S} \mathrm{dist}\,(x,s)\,.$$

Lastly, for any subset S of an Euclidean space, by $\mathrm{co}(S)$ we mean the convex hull of S, i.e. the smallest convex set containing S, obtainable by intersection of all convex sets containing S. The symbol $\overline{\mathrm{co}}(S)$ denotes the closure of such convex hull, and it is also the smallest closed convex set containing S.

2.4.2 Set-valued Lie brackets, Quasi Differential Quotients, and Approximating Cones

Definition 2.9 (Clarke's Generalized Jacobian)**.** Let $F : y \in \mathbb{R}^N \mapsto F(y) \in \mathbb{R}^n$ be an almost everywhere differentiable map and $y_0 \in \mathbb{R}^N$. We say that the set $\partial_y F(y_0)$ is the Clarke's Generalized Jacobian of F at point y_0 if

$$\mathrm{Lin}(\mathbb{R}^N,\mathbb{R}^n) \supseteq \partial_y F(y_0) := \overline{\mathrm{co}}\Big\{ \lim_{n\to\infty} DF(y_n) :\ \ \mathrm{Diff}(F) \ni y_n \to y_0 \Big\}.$$

Definition 2.10 (Quasi Differential Quotients (QDQ))**.** Let $F : \mathbb{R}^N \rightrightarrows \mathbb{R}^n$ be a set-valued map, $(\bar\gamma,\bar y) \in \mathbb{R}^N \times \mathbb{R}^n$, $\Lambda \subset \mathrm{Lin}\{\mathbb{R}^N,\mathbb{R}^n\}$ be a compact set, and $\Gamma \subset \mathbb{R}^N$ be any subset. We say that Λ is a *Quasi Differential Quotient (QDQ) of F at $(\bar\gamma,\bar y)$* in the direction of Γ if there exists a modulus $\rho : (0,+\infty) \to (0,+\infty)$ and $\forall\delta > 0$ there is a continuous map

$$(L_\delta, h_\delta) : (\bar\gamma + B_\delta) \cap \Gamma \to \mathrm{Lin}(\mathbb{R}^N,\mathbb{R}^n) \times \mathbb{R}^n$$

such that

$$\min_{L'\in\Lambda} |L_\delta(\gamma) - L'| \le \rho(\delta), \quad |h_\delta(\gamma)| \le \delta\rho(\delta), \quad \text{and}$$

$$\bar y + L_\delta(\gamma)\cdot(\gamma - \bar\gamma) + h_\delta(\gamma) \in F(\gamma), \tag{2.27}$$

whenever $\gamma \in (\bar{\gamma} + B_\delta) \cap \Gamma$.

Remark 2.29. Whenever F is a single-valued continuous map from $\mathbb{R}^N$ to $\mathbb{R}^n$, the inclusion (2.27) is actually an equality. In this case, to show that Λ is a QDQ of F at $(\bar{\gamma}, F(\bar{\gamma}))$ it would be enough to find a family $\{L_\delta,\ \delta > 0\}$ of continuous maps

$$L_\delta : (\bar{\gamma} + B_\delta) \cap \Gamma \to \mathrm{Lin}(\mathbb{R}^N, \mathbb{R}^n)$$

satisfying,

$$\min_{L' \in \Lambda} |L_\delta(\gamma) - L'| \leq \rho(\delta), \quad \text{ and } \quad |F(\gamma) - F(\bar{\gamma}) - L_\delta(\gamma)(\gamma - \bar{\gamma})| \leq \delta\rho(\delta)$$

whenever $\delta > 0$ and $\gamma \in (\bar{\gamma} + B_\delta) \cap \Gamma$, as in this case the continuity of the error function $h_\delta := F(\gamma) - F(\bar{\gamma}) - L_\delta(\gamma)(\gamma - \bar{\gamma})$ would follow from the hypotheses.

Remark 2.30. To further specialize the previous remark, assume F is single-valued and continuous and assume one is able to find a positive real number $\tilde{\varepsilon} > 0$ and a single continuous map

$$L : B_{\tilde{\varepsilon}} \cap \Gamma \to \mathrm{Lin}(\mathbb{R}^N, \mathbb{R}^n)$$

satisfying

$$\lim_{\Gamma \ni \gamma \to 0} \operatorname{dist}(L(\gamma), \Lambda) = 0, \quad \text{ and } \quad \lim_{\Gamma \ni \gamma \to 0} \frac{|F(\gamma) - F(0) - L(\gamma)\gamma|}{|\gamma|} = 0.$$

This is enough to prove that Λ is a QDQ for F at $(0, F(0))$ in the direction of Γ. This is because one could define L_δ to be the restriction of L to $B_\delta \cap \Gamma$ and a valid choice for the modulus $\rho(\delta)$ would then be, trivially

$$\rho(\delta) := \max \left\{ \sup_{\gamma \in B_\delta \cap \Gamma} \operatorname{dist}(L(\gamma), \Lambda), \sup_{\gamma \in B_\delta \cap \Gamma} \frac{|F(\gamma) - F(0) - L(\gamma)\gamma|}{\delta} \right\},$$

which is well defined on $(0, \tilde{\varepsilon})$ (provided $\tilde{\varepsilon}$ was chosen suitably small), positive, non-decreasing, tends to 0 as $\delta \to 0^+$ and so Λ is proven a QDQ as prefigured.

In any Euclidean space $\mathbb{R}^n$, a subset $C \subseteq \mathbb{R}^n$ is called a cone if $\alpha v \in C, \forall \alpha \geq 0$ and $\forall v \in C$. A set $\mathscr{C}$ whose elements are cones is called a multicone.
For any given cone $C \subseteq \mathbb{R}^n$, the set

$$C^\perp = \{v \in \mathbb{R}^n,\ v \cdot c \leq 0\ \forall c \in C\}$$

is a closed cone in $\mathbb{R}^n$ called the polar cone of C. On the other hand, the polar cone of a multicone, defined by

$$\mathscr{C}^\perp = \bigcup_{C \in \mathscr{C}} C^\perp,$$

is not necessarily a closed subset of $\mathbb{R}^n$.
Two cones C_1 and C_2 are said to be transversal if $C_1 - C_2 = \mathbb{R}^n$, where $C_1 - C_2$ is the set of differences $c_1 - c_2$, with $c_1 \in C_1$ and $c_2 \in C_2$. In addition, they are strongly transversal if they also have non-trivial intersection. Trivially, this is equivalent to

saying there exists $\mu \in (\mathbb{R}^n)^*$ such that $\mu \cdot v > 0$ for some $v \in C_1 \cap C_2$
Two multicones $\mathscr{C}_1$ and $\mathscr{C}_2$ are transversal if the same can be said for any couple of cones $C_1 \in \mathscr{C}_1$ and $C_2 \in \mathscr{C}_2$. Two transversal multicones $\mathscr{C}_1$ and $\mathscr{C}_2$ are also strongly transversal, if there exists a single $\mu \in (\mathbb{R}^n)^*$ such that for any $C_1 \in \mathscr{C}_1$ and $C_2 \in \mathscr{C}_2$ an element $v \in C_1 \cap C_2$ such that $\mu \cdot v > 0$ can be found.
Since two cones C_1 and C_2 are linearly separated, i.e. $C_1^\perp \cap -C_2^\perp \neq \{0\}$, if and only if they are not transversal, two cones or even multicones that are not strongly transversal are sometimes termed weakly linearly separated.
It was shown by Hector Sussmann in [132] that two multicones $\mathscr{C}_1$ and $\mathscr{C}_2$ are weakly linearly separated if for every $0 \neq \mu \in (\mathbb{R}^n)^*$ there exist $C_1 \in \mathscr{C}_1$, $C_2 \in \mathscr{C}_2$, $\pi_1 \in C_1^\perp$, $\pi_2 \in C_2^\perp$ and a non-negative real number π_0 such that

$$(\pi_0, \pi_1, \pi_2) \neq (0,0,0) \quad \text{and} \quad \pi_1 + \pi_2 = \mu \cdot \pi_0.$$

This statement of weak linear separation becomes extremely useful whenever a non-zero functional $-\mu$ is chosen in the polar cone of every cone of the multicone $\mathscr{C}_2$, as it then follows from $\pi_2 - \mu \cdot \pi_0 = -\pi_1$, that two cones $C_1 \in \mathscr{C}_1$ and $C_2 \in \mathscr{C}_2$ can be found, that are linearly separated.
A convex multicone is simply a multicone consisting solely of convex cones.

Definition 2.11 (QDQ approximating multicones). Let S be any subset of $\mathbb{R}^n$ and $x \in S$. A convex multicone $\mathscr{C}$ is said to be a *QDQ* approximating multicone if there exists a function $F : (\mathbb{R}^+)^N \to \mathbb{R}^n$ such that $F((\mathbb{R}^+)^N) \subseteq S$, Λ is a *QDQ* for F at $(0,x)$ and

$$\mathscr{C} = \{L \cdot (\mathbb{R}^+)^N,\ L \in \Lambda\}.$$

Definition 2.12 (Local separation of sets). Two subsets S_1 and S_2 are locally separated at x if and only if there exists a neighbourhood U_x of x such that

$$S_1 \cap S_2 \cap U_x = \{x\}.$$

As a consequence of an open mapping theorem involving QDQ's, the following fact holds true (see Theorem 4.37, p. 265 in [4] where the lemma was proven in the more general context of $AGDQ$'s, of which QDQ's are a special case)

Lemma 2.31. *If two subsets S_1 and S_2 are locally separated at x and if $\mathscr{C}_1$ and $\mathscr{C}_2$ are QDQ approximating multicones for S_1 and S_2, respectively, at x, then $\mathscr{C}_1$ and $\mathscr{C}_2$ are not strongly transverse.*

The proof of the following lemma can then be recovered from [132].

Lemma 2.32. *Let $\mathscr{C}_1$ and $\mathscr{C}_2$ be two multicones that are not strongly transversal. If there is a linear functional μ that is in $C_2^\perp$ but not in $-C_2^\perp$ for all $C_2 \in \mathscr{C}_2$, then there are two cones $C_1 \in \mathscr{C}_1$ and $C_2 \in \mathscr{C}_2$ that are linearly separated.*

2.4.3 Lie brackets for Lipschitz vector fields

Given two Lipschitz vector fields $f, g : \mathbb{R}^n \to \mathbb{R}^n$, it is known by Rademacher's theorem that the set $\mathrm{Diff}(f) \cap \mathrm{Diff}(g)$ of points of differentiability of f and g is a has full measure. Basing on this fact, the authors of [128] introduce the following concept of a *set-valued Lie bracket*:

$$[f,g]_{set}(x) = \overline{\mathrm{co}}\Big\{ \lim_{n\to\infty} [f,g](x_n),\ \to x,\ x_n \in \mathrm{Diff}(f)\cap \mathrm{Diff}(g)\Big\},$$

where we mean that all existing limits along sequences $(x_n) \subset \mathrm{Diff}(f)\cap\mathrm{Diff}(g)$ are considered. Clearly, for every $x\in\mathbb{R}^n$, one has $[f,g]_{set}(x) \neq \emptyset$. Moreover $[f,g]_{set}(x) = \{[f,g](x)\}$ as soon as f,g are of class C^1 in a neighborhood od x. One trivially has that the relations $[f,f]_{set} = \{0\}$ and $[f,g]_{set} = -[g,f]_{set}$ keep holding for set-valued brackets, with the understanding that $-S$ is the set of opposites of elements in S. Furthermore, some basic results have been generalized to set-valued Lie brackets . For instance, the flow of f,g locally commute if and only if $[f,g]_{set} = 0$. Furthermore, by means of generalization of the involutivity condition involving set-valued brackets, a Frobenius-type result, holds true for Lipschitz distributions as well(see, for instance, [127]).

Finally, iterated set-valued Lie brackets for Lipschitz vector fields and their relations with local controllability have been introduced in [67].

2.4.4 The problem and the main result

Let us recall problem (P), which has been presented in the introductive subsection. Actually, the problem below (still labeled (P)) is more general, in that it includes a current cost, a more general set where the controls u take values, and the continuous dependence of the drift f on some bounded control a.

$$(P)\qquad \begin{cases} \text{minimize } \Psi(T,x(T)) + \displaystyle\int_0^T l(x(t),u(t),a(t))\,dt, \\ \text{over all feasible strict-sense processes } (T,u,a,x,\nu) \text{ of} \end{cases}$$

$$\begin{cases} \dfrac{dx}{dt} = f(x,a) + \displaystyle\sum_{i=1}^m g_i(x)u^i, \qquad \text{a.e. } t\in[0,T], \\ \dfrac{d\nu}{dt} = |u|, \\ (x,\nu) = (\hat{x},0), \quad \nu \le K, \quad (T,x)\in\mathfrak{T}, \qquad {}^5 \end{cases} \tag{2.28}$$

The hypotheses for problem (P) are as follows:

i) the state variable x belongs to $\mathbb{R}^n$, for some $n>0$;[6];
ii) the vector fields $g_i:\mathbb{R}^n\to\mathbb{R}^n$, $i=1,\dots,m$ are locally Lipschitz continuous vector fields on $\mathbb{R}^n$;
iii) the controls $u=(u^1,\dots,u^m)$ belong to (the closed cone) $\mathscr{C} = \mathscr{C}_1\times\mathscr{C}_2$, where, for some non negative integers m_1 and m_2 such that $m = m_1+m_2$, $\mathscr{C}_1 \subseteq \mathbb{R}^{m_1}$ is a closed cone containing the coordinate axes, and $\mathscr{C}_2\subseteq\mathbb{R}^{m_2}$ is a closed cone which does not contain any straight line;
iv) the control a takes values in a compact set $A\subset\mathbb{R}^q$;
v) f, which is sometimes called the *drift*, is continuous in (x,a), and, for any value $a\in A$, the vector field $f(\cdot,a):\mathbb{R}^n\to\mathbb{R}^n$ is locally Lipschitz continuous;

[5] Notice that we have rewritten the L^1 bound $\|u\|_1\le K$ in the equivalent form $\nu\le K$.
[6] See Remark 2.33.

vi) the real-valued *Lagrangian* $l := l(x,u,a)$ has the form $l(x,u,a) = l_0(x,a)+l_1(x,u)$ and is continuous; furthermore, for every (u,a), $x \mapsto l(x,u,a)$ is locally Lipschitz continuous with respect to x, uniformly with respect to a; moreover, the so-called *recess function* $\hat{l}_1(x,w^0,w) := \lim_{r\to w^0} r l_1\left(x, \frac{w}{r}\right)$ is well-defined and locally Lipschitz with respect to x, uniformly when (w^0,w) ranges on the bounded set $[0,1]\times(\mathscr{C}\cap B_1)$;

vii) the *final cost* $\Psi(t,x)$ is of class C^1, $0 \le K \le +\infty$, and the (time-dependent) *target* $\mathfrak{T} \subseteq \mathbb{R}_+ \times \mathbb{R}^n$ is a closed subset.

The elements of

$$\mathscr{U} = \bigcup_{T>0} \{T\} \times L^1([0,T], \mathscr{C}\times A)$$

will be called *strict-sense controls.* If $(T,u,a)\in\mathscr{U}$ is a strict-sense control and (x,ν) is the unique Carathéodory solution of the above system, then (T,u,a,x,ν) will be called a *strict-sense process.* Moreover, we will say that a strict-sense process (T,u,a,x,ν) is *feasible* if $(T,x(T),\nu(T))\in\mathfrak{T}\times[0,K]$.

Remark 2.33. This problem, our results and all the mathematical tools and ideas we make use of throughout this section could also be extended to the context of an n-dimensional manifold M, being intrinsic/chart-independent. We restrict ourselves to $\mathbb{R}^n$ for the sake of clarity and to avoid further complicating the notation.

Definition 2.13. We say that $(\overline{T},\overline{u},\overline{a},\overline{x},\overline{\nu})$ is a *local minimizer* for problem (P) if there exists $\delta>0$ such that

$$\Psi(\overline{T},\overline{x}(\overline{T})) + \int_0^{\overline{T}} l(\overline{x}(t),\overline{u}(t),\overline{a}(t))\,dt \le \Psi(T,x(T)) + \int_0^T l(x(t),u(t),a(t))\,dt$$

for all feasible processes (T,u,a,x,ν) such that $|T-\overline{T}|+\|(x,\nu)-(\overline{x},\overline{\nu})\|_\infty<\delta$ where, since (x,ν) and $(\overline{x},\overline{\nu})$ may have different domains, we first extend them to $\mathbb{R}^+$ so they stay constant to their final values $(x,\nu)(T)$ and $(\overline{x},\overline{\nu})(\overline{T})$, respectively.

As mentioned in the introductive subsection, since the control u are unbounded and the problem (P) has slow growth, in order to achieve existence of a minimizer one tries to consider some form of *closure*, or even *compactification.* The impulsive extension makes the job, so our necessary conditions will refer to it. To describe it, let us begin introducing the set

$$\mathscr{W} := \bigcup_{S>0} \{S\} \times \left\{(w^0,w,\alpha)\in L^\infty([0,S],\mathbb{R}^+\times\mathscr{C}\times A) : \operatorname{essinf}(w^0+|w|)>0\right\}$$

of *extended-sense controls.* For any $(S,w^0,w,\alpha)\in\mathscr{W}$, we say that $(S,w^0,w,\alpha,y^0,y,\beta)$

is an *extended-sense process* if (y^0, y, β) is the unique Carathéodory solution of

$$\begin{cases} \dfrac{dy^0}{ds}(s) = w^0(s), \\ \dfrac{dy}{ds}(s) = f(y(s), \alpha(s))w^0(s) + \displaystyle\sum_{i=1}^m g_i(y(s))w^i(s), & \text{a.e. } s \in [0, S], \\ \dfrac{d\beta}{ds}(s) = |w(s)|, \\ (y^0, y, \beta)(0) = (0, \hat{x}, 0). \end{cases}$$

Furthermore an extended-sense process $(S, w^0, w, \alpha, y^0, y, \beta)$ is said *feasible* if $(y^0(S), y(S), \beta(S)) \in \mathfrak{T} \times [0, K]$. The *extended problem* is defined as

$$(P_{ext}) \quad \begin{cases} \text{minimize } \Psi(y^0(S), y(S)) + \displaystyle\int_0^S l^e((y, w^0, w, \alpha)(s))\, ds \\ \text{over all feasible processes } (S, w^0, w, \alpha, y^0, y, \beta) \end{cases}$$

where the *extended lagrangian* l^e is defined by setting

$$l^e(x, w^0, w, \alpha) := l_0(x, \alpha)w^0 + \hat{l}_1(x, w^0, w) \quad \forall (x, w^0, w, \alpha) \in \mathbb{R}^+ \times \mathscr{C} \times A,$$

$\hat{l}_1$ being the above-defined recession function. For instance, if $l(x, u, a) = l_0(x, a) + \ell(x)|u|^r$ for some $r \in [0, 1]$ and some Lipschitz function ℓ, one has $l^e(x, w^0, w, \alpha) = l_0(x, \alpha)w^0 + \ell(x)|w|^r(w^0)^{1-r}$. In general, in view of the sublinearity of l in u, l^e is well-defined.

The notion of local minimizer in the extended problem (P_{ext}) is defined similarly to that of the original problem:

Definition 2.14. We say that $(\overline{S}, \overline{w}^0, \overline{w}, \overline{\alpha}, \overline{y}^0, \overline{y}, \overline{\beta})$ is a *local minimizer* for problem $(P)_{ext}$ if there exists $\delta > 0$ such that

$$\Psi(\overline{y}^0(\overline{S}), \overline{y}(\overline{S})) + \int_0^{\overline{S}} l^e(\overline{y}(s), \overline{w}^0(s), \overline{w}(s), \overline{\alpha}(s))\, ds \leq$$
$$\leq \Psi(y^0(S), y(S)) + \int_0^S l^e(y(s), w^0(s), w(s), \alpha(s))\, ds$$

for all feasible processes $(S, w^0, w, \alpha, y^0, y, \beta)$ such that $|S - \overline{S}| + \|(y^0, y, \beta) - (\overline{y}^0, \overline{y}, \overline{\beta})\|_\infty < \delta$, where, since (y^0, y, β) and $(\overline{y}^0, \overline{y}, \overline{\beta})$ may have different domains, we first extend them to $\mathbb{R}^+$ so they stay constant to their final values $(y^0, y, \beta)(T)$ and $(\overline{y}^0, \overline{y}, \overline{\beta})(\overline{T})$, respectively.

There is an obvious one-to-one correspondence between strict-sense processes and extended-sense processes such that $w_0(s) > 0$ for almost any $s \in (0, S)$. This correspondence preserves feasibility of a process, and minima of the strict sense problem correspond to minima for the space-time extended problem having $w^0(s)$ almost everywhere positive (see [21]).

Thanks to the rate-independence of the extended system [7] without loss of generality

[7] By, rate-independence we mean that if $\sigma : [0, \hat{S}] \to [0, S]$ is a diffeomorphism, then

we can make the following convention:

Convention: we assume that any given local minimizer $(\overline{S}, \overline{w}^0, \overline{w}, \overline{\alpha}, \overline{y}^0, \overline{y}, \overline{\beta})$ is *canonical*, meaning that $w^0(s) + |w(s)| = 1$ for almost every $s \in [0, \overline{S}]$.

Remark 2.34. By allowing $w^0 = 0$ and imposing the constraint $w^0(s) + |w(s)| = 1$ we get the sought compactness of the minimization domain.

For every process $(S, w^0, w, \alpha, y^0, y, \beta)$, we will set

$$\overline{y}^l(s) := \int_0^s l^e(\overline{y}(\sigma), \overline{w}^0(\sigma), \overline{w}(\sigma), \overline{\alpha}(\sigma))d\sigma$$

(so that $\overline{y}^l$ is the unique Carathéodory solution to the trivial differential equation $\frac{dy^l}{ds}(s) = l^e(\overline{y}(s), \overline{w}^0(s), \overline{w}(s), \overline{\alpha}(s))$ with initial condition $y^l(0) = 0$.)

Lastly, let us introduce the Hamiltonian function

$$H : \mathbb{R}^n \times \left(\mathbb{R}^{1+n+1+1}\right)^* \times \left(\mathbb{R}^+ \times \mathscr{C} \times A\right) \to \mathbb{R}$$

defined by

$$H(y, p_0, p, \lambda, \pi, w^0, w, a) := p_0 w^0 + p\left(f(y, a) + \sum_{i=1}^m g_i(y)w^i\right) - \lambda l^e(y, w^0, w, a) + \pi|w|.$$

We are now in the position of stating our main result in this section:

Theorem 2.35 (Maximum Principle). *Let $(\overline{S}, \overline{w}^0, \overline{w}, \overline{\alpha}, \overline{y}^0, \overline{y}, \overline{\beta})$ be a canonical local minimizer for the extended problem (P_{ext}), and let $\boldsymbol{\mathscr{T}}$ be any QDQ-multicone approximating the target set $\mathfrak{T}$ at $(\overline{y}^0, \overline{y})(\overline{S})$.*
Then there exist multipliers $(p_0, p, \lambda, \pi) \in \mathbb{R}^* \times AC([0, \overline{S}]; (\mathbb{R}^n)^*) \times \mathbb{R}^* \times \mathbb{R}^*$ such that $\pi \leq 0$ (with $\pi = 0$ as soon as $\overline{\beta}(\overline{S}) < K$) and the following conditions are satisfied:

i) $\lambda \geq 0$,

ii) $(p_0, p, \lambda) \neq 0$;

iii) *For almost all $s \in [0, \overline{S}]$, p verifies the adjoint differential inclusion*

$$\dot{p} \in -p\,\partial_y\left(f(\overline{y}, \overline{\alpha})\overline{w}^0 + \sum_{i=1}^m g_i(\overline{y})\overline{w}^i\right) + \lambda\partial_y l^e(\overline{y}, \overline{w}^0, \overline{w}, \overline{\alpha}) \ ;$$

iv) $(p_0, p(\overline{S})) + \lambda\left(\frac{\partial\Psi}{\partial y^0}, \frac{\partial\Psi}{\partial y}\right)\left((\overline{y}^0, \overline{y})(\overline{S})\right) \in -\overline{\bigcup_{\mathscr{T} \in \boldsymbol{\mathscr{T}}} \mathscr{T}^\perp}$;

$(S, w^0, w, \alpha, y^0, y, \beta)$ is a process if and only if $(\hat{S}, w^0 \circ \sigma, w \circ \sigma, \alpha \circ \sigma, y^0 \circ \sigma, y \circ \sigma, \beta \circ \sigma)$ is a process.

v) *For almost all* $s \in [0, \overline{S}]$,

$$\boldsymbol{max}\Big[H(\overline{y}(s), p_0, p(s), \lambda, \pi, w^0, w, a)\Big](1+\zeta) = H(\overline{y}(s), p_0, p(s), \lambda, \pi, \overline{w}^0(s), \overline{w}(s), \overline{\alpha}(s)).$$

where the maximization is performed among values $(w^0, w, a, \zeta) \in \mathbb{R}^+ \times \mathscr{C} \times A \times [-\rho, \rho]$.

If, in addition, we assume that that $\overline{\beta}(\overline{S}) < K$ *and* $\hat{l}_1(\cdot, 0) \equiv 0$, *the following* $\dfrac{m_1(m_1-1)}{2}$ *second order conditions hold true as well:*

vi) *For any* $i, j \in \{1, \dots, m_1\}$ *and for almost any* $s \in [0, \overline{S}]$, *one has*

$$0 \in p(s)\,[g_i, g_j]_{set}(\overline{y}(s)). \tag{2.29}$$

2.4.5 An equivalent fixed end-time problem

Let us begin with a standard reparametrization procedure which allows one to reduce problem (P) to a fixed end-time problem. To save space, we introduce the notation

$$\mathscr{F}(y, w^0, w, a) := \left(w^0, F^e(y, w^0, w, a), l^e(y, w^0, w, a)\right),$$

for all $(y, w^0, w, a) \in \mathbb{R}^{1+n+1} \times \mathbb{R}^+ \times \mathscr{C} \times A$, where

$$F^e(y, w^0, w, a) := f(y, a)w^0 + \sum_{i=1}^{m} g_i(y)w^i,$$

Let us fix $\bar{S} > 0$, $\rho > 0$. We say that $(\bar{S}, w^0, w, \alpha, \zeta, y^0, y, y^l, \beta)$ is a *rescaled space-time process* if $(\bar{S}, w^0, w, \alpha, \zeta)(\cdot) \in \mathscr{W} \times L^\infty([0, \overline{S}], [-\rho, \rho])$ and $\big((y^0, y, y^l), \beta\big)$ is the unique (Carathéodory) solution of *the rescaled Cauchy problem*

$$\begin{cases} \dfrac{d}{ds}((y^0, y, y^l), \beta) = \Big(\mathscr{F}(y, w^0, w, a), |w|\Big) \cdot \Big(1 + \zeta\Big) \\ \big((y^0, y, y^l), \beta\big)(0) = \big((0, \hat{x}, 0), 0\big) \end{cases} \tag{2.30}$$

Moreover, we call $(\bar{S}, w^0, w, \alpha, \zeta, y^0, y, y^l, \beta)$ *feasible* if $\big((y^0, y), \beta\big) \in \mathfrak{T} \times [0, K]$.
The *rescaled optimization problem* is defined as as

$$\begin{cases} \text{minimize } \Psi((y^0, y)(\overline{S})) + y^l(\overline{S}), \\ \text{ over feasible rescaled processes .} \end{cases} \tag{2.31}$$

Remark 2.36. It is common knowledge that, for small enough $\rho > 0$, a process $(\overline{S}, \overline{w}^0, \overline{w}, \overline{\alpha}, \overline{y}^0, \overline{y}, \overline{\beta})$ is a canonical local minimizer for the extended problem (P_{ext}) if and only if the rescaled space-time process $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, 0, \bar{y}^0, \bar{y}, \bar{y}^l, \bar{\beta})$ is a local minimizer for fixed-end-time problem (2.31). Therefore, in the proof of the Maximum Problem we are allowed to replace the hypothesis of the theorem with the assumption that $(\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, 0, \bar{y}^0, \bar{y}, \bar{y}^l, \bar{\beta})$ is a local minimizer of the rescaled problem (2.31). *Therefore, we posit the assumption that* $\bar{\zeta} \equiv 0$ throughout the proof.

2.4.6 Set separation

For some $\delta > 0$, let us consider the *δ-reachable set* $\mathfrak{R}_\delta \subseteq \mathbb{R}^{1+n+1+1}$ defined by:

$$\mathfrak{R}_\delta := \left\{ \begin{array}{l} \Big(y^0, y, y^l + \Psi \circ (y^0, y), \beta\Big)(\overline{S}): \ (\overline{S}, w^0, w, \alpha, \zeta, y^0, y, y^l, \beta) \text{ is a process} \\ \text{that verifies } |S - \overline{S}| + \|(y^0 - \overline{y}^0, y - \overline{y}, y^l - \overline{y}^l, \beta - \overline{\beta})\|_\infty < \delta \end{array} \right\}$$

and the *projected δ-reachable set*

$$\mathfrak{R}'_\delta := \mathfrak{pr}\Big(\mathfrak{R}_\delta\Big) \subseteq \mathbb{R}^{1+n+1},$$

where the projection operator $\mathfrak{pr}$ is defined by setting

$$\mathfrak{pr}(x^0, x, x^l, \beta) := (x^0, x, x^l), \quad \forall (x^0, x, x^l, \beta) \in \mathbb{R}^{1+n+1+1}.$$

Let us introduce also the *profitable set*

$$\mathfrak{P}_\delta := \left(\left(\mathfrak{T} \times \Big] - \infty, \overline{y}^l(\overline{S}) + \overline{\Psi}(\overline{S}) \Big[\right) \bigcup \left\{ \left(\overline{y}^0, \overline{y}, \overline{y}^l(\overline{S}) + \overline{\Psi}(\overline{S})\right) \right\} \right) \times [0, K]$$

and the **projected profitable set**

$$\mathfrak{P}'_\delta := \mathfrak{pr}\Big(\mathfrak{P}_\delta\Big) = \left(\mathfrak{T} \times \Big] - \infty, \overline{y}^l(\overline{S}) + \overline{\Psi}(\overline{S}) \Big[\right) \bigcup \left\{ \left(\overline{y}^0, \overline{y}, \overline{y}^l(\overline{S}) + \overline{\Psi}(\overline{S})\right) \right\}$$

Lemma 2.37. *Let us assume that $\overline{\beta}(\overline{S}) < K$. Then for any $\delta > 0$ sufficiently small, the projected profitable set $\mathfrak{P}'_\delta$ and the projected δ-reachable set $\mathfrak{R}'_\delta$ are locally separated at $\big(\overline{y}^0, \overline{y}, \overline{y}^l(\overline{S}) + \overline{\Psi}(\overline{S})\big)$.*

Proof. Indeed, by the definition of local minimizer it follows that the profitable set $\mathfrak{P}_\delta$ and the δ-reachable set $\mathfrak{R}_\delta$ are locally separated at $\big(\overline{y}^0, \overline{y}, \overline{y}^l(\overline{S}) + \overline{\Psi}(\overline{S}), \overline{\beta}(\overline{S})\big)$. From this one get the thesis trivially (see [21], Lemma 6.12). □

With the ultimate aim of applying a suitable separability criterion for approximating cones, we now build a family of QDQ approximating cones to the projected δ-reachable set $\mathfrak{R}'_\delta$ at $\big(\overline{y}^0, \overline{y}, \overline{y}^l(\overline{S}) + \overline{\Psi}(\overline{S})\big)$. Let us define the set $\mathfrak{V}$ of *variation generators* as the union $\mathfrak{V} := \mathfrak{V}_1 \bigcup \mathfrak{V}_2$, where $\mathfrak{V}_1$ and $\mathfrak{V}_2$ are the sets of *needle variation generators* and of *bracket-like variation generators* defined as $\mathfrak{V}_1 := \mathbb{R}^+ \times \mathscr{C} \times A \times [-\rho, \rho]$ and $\mathfrak{V}_2 := \left[\{1, \ldots, m_1\}^2 \setminus \operatorname{diag}\Big(\{1, \ldots, m_1\}^2\Big) \right]$, respectively.

Definition 2.15. Let $(0, \overline{S})_{Leb} \subset [0, \overline{S}]$ be the set of Lebesgue points of the function $s \mapsto (\overline{w}^0(s), \overline{F}^e(s), \overline{l}^e(s), |\overline{w}|(s))$ ì where $\overline{F}^e$ and $\overline{l}^e$ denote the functions F^e and l^e evaluated along the optimal process we are studying. For every variation generator

$\mathbf{c} \in \mathfrak{V}$ and every instant $\overline{s}$, let us define the variation $\big(v^0_{\mathbf{c},\overline{s}}, v_{\mathbf{c},\overline{s}}, v^l_{\mathbf{c},\overline{s}}\big)$ as follows:

$$\big(v^0_{\mathbf{c},\overline{s}}, v_{\mathbf{c},\overline{s}}, v^l_{\mathbf{c},\overline{s}}\big) := \begin{cases} \left\{\begin{pmatrix} w^0(1+\zeta) - \overline{w}^0(\overline{s}) \\ F^e(\overline{y}(\overline{s}), w^0, w, a)(1+\zeta) - \overline{F}^e(\overline{s}) \\ l^e(\overline{y}(\overline{s}), w^0, w, a)(1+\zeta) - \overline{l}^e(\overline{s}) \end{pmatrix}\right\} & \begin{array}{c}\text{if } \mathbf{c} = (w^0, w, a, \zeta) \in \mathfrak{V}_1, \\ \text{and } \overline{s} \in (0, \overline{S})_{Leb}\end{array} \\ \\ \{0\} \times [g_i, g_j]_{set}(\overline{y}(\overline{s})) \times \{0\} & \begin{array}{c}\text{if } \mathbf{c} = (i,j) \in \mathfrak{V}_2 \\ \text{and } \overline{s} \in (0, \overline{S}).\end{array} \end{cases}$$

Moreover, when $\mathbf{c} = (w^0, w, a, \zeta) \in \mathfrak{V}_1$, and $\overline{s} \in (0, \overline{S})_{Leb}$, we set $v^\nu_{\mathbf{c},\overline{s}} := |w|(1+\zeta) - |\overline{w}(\overline{s})|$.

Let us point out that, to retain uniformity of notation, we always regard $\big(v^0_{\mathbf{c},\overline{s}}, v_{\mathbf{c},\overline{s}}, v^l_{\mathbf{c},\overline{s}}\big)$ as a *subset of vectors of* $\mathbb{R}^{1+n+1}$, though it reduces to the singleton formed by that the usual needle variation vector as soon as $\mathbf{c} \in \mathfrak{V}_1$.

Definition 2.16. Let us fix an instant $\overline{s} \in (0, \overline{S})$ and a rescaled control $\mathbf{w} = (w^0, w, \alpha, \zeta) \in L^\infty\big([0, \overline{S}], \mathbb{R}^+ \times \mathscr{C} \times A \times [-\rho, \rho]\big)$ (with $\operatorname{essinf}(w^0 + |w|) > 0$).

- If $\mathbf{c} = (\hat{w}^0, \hat{w}, \hat{a}, \hat{\zeta}) \in \mathfrak{V}_1$, the family of controls $\big\{\mathbf{w}_{\varepsilon,\mathbf{c},\overline{s}}(s) : \varepsilon \in [0, \overline{s})\big\}$ defined setting
$$\mathbf{w}_{\varepsilon,\mathbf{c},\overline{s}}(s) = \begin{cases} \mathbf{w}(s) & \text{if } s \in [0, \overline{s} - \varepsilon) \cup (\overline{s}, \overline{S}] \\ (\hat{w}^0, \hat{w}, \hat{a}, \hat{\zeta}) & \text{if } s \in [\overline{s} - \varepsilon, \overline{s}] \end{cases}$$
will be called a *needle control approximation of* $\mathbf{w}(s)$ at $\overline{s}$ associated to $\mathbf{c}$.

- If $\mathbf{c} = (i,j) \in \mathfrak{V}_2$, the family of controls $\{\mathbf{w}_{\varepsilon,\mathbf{c},\overline{s}}(s) : 0 < 8\sqrt{\varepsilon} \le \overline{s}\}$ defined by setting
$$\mathbf{w}_{\varepsilon,\mathbf{c},\overline{s}}(s) = \begin{cases} \mathbf{w}(s) & \text{if } s \notin [\overline{s} - 8\sqrt{\varepsilon}, \overline{s}] \\ (2w^0, 2w, \alpha, \zeta) \circ \gamma^\varepsilon(s) & \text{if } s \in [\overline{s} - 8\sqrt{\varepsilon}, \overline{s} - 4\sqrt{\varepsilon}] \\ (0, \mathbf{e}_i, a, 0) & \text{if } s \in [\overline{s} - 4\sqrt{\varepsilon}, \overline{s} - 3\sqrt{\varepsilon}] \\ (0, \mathbf{e}_j, a, 0) & \text{if } s \in [\overline{s} - 3\sqrt{\varepsilon}, \overline{s} - 2\sqrt{\varepsilon}] \\ (0, -\mathbf{e}_i, a, 0) & \text{if } s \in [\overline{s} - 2\sqrt{\varepsilon}, \overline{s} - \sqrt{\varepsilon}] \\ (0, -\mathbf{e}_j, a, 0) & \text{if } s \in [\overline{s} - \sqrt{\varepsilon}, \overline{s}], \end{cases}$$
where $a \in A$ is arbitrarily chosen[8] and $\gamma^\varepsilon(s) := 2s - \overline{s} + 8\sqrt{\varepsilon}$, the *bracket-like approximation of* $\mathbf{w}$.

For any $(\overline{s}, \mathbf{c}) \in [0, S] \times \mathfrak{V}$ and any ε sufficiently small, consider the functional $\mathscr{A}_{\varepsilon,\mathbf{c},\overline{s}}$ (from the space of rescaled controls $\mathbf{w}$ into itself) defined by setting $\mathscr{A}_{\varepsilon,\mathbf{c},\overline{s}}(\mathbf{w}) := \mathbf{w}_{\varepsilon,\mathbf{c},\overline{s}}$. In addition, given N variation generators $\mathbf{c}_1, \dots, \mathbf{c}_N \in \mathfrak{V}$ and N instants $0 < s_1 < s_2 < \dots s_N \le \overline{S}$ for a $\tilde{\varepsilon} > 0$ sufficiently small, let us define the multiple perturbation

$$[0, \tilde{\varepsilon}]^N \ni \boldsymbol{\varepsilon} \mapsto \overline{\mathbf{w}}_{\boldsymbol{\varepsilon}} := \mathscr{A}_{\varepsilon_N,\mathbf{c}_N,s_N} \circ \dots \circ \mathscr{A}_{\varepsilon_1,\mathbf{c}_1,s_1}(\overline{\mathbf{w}}).$$

[8] Since $w^0 = 0$, the choice of a is irrelevant

Let us set $(\overline{w}^0_\varepsilon, \overline{w}_\varepsilon, \overline{a}_\varepsilon, \overline{\zeta}_\varepsilon) := \overline{\mathbf{w}}_\varepsilon$, and let us use $(y^0_\varepsilon, y_\varepsilon, y^l_\varepsilon, \beta_\varepsilon)$ to denote the solution (on $[0,\overline{S}]$) of the Cauchy problem [9]

$$\begin{cases} \dfrac{d}{ds}(y^0, y, y^l, \beta) = \left(\mathscr{F}\left(y, \overline{w}^0_\varepsilon, \overline{w}_\varepsilon, \overline{a}_\varepsilon\right), |w_\varepsilon|\right)(1+\overline{\zeta}_\varepsilon) \\ (y^0, y, y^l, \beta)(0) = (0, \hat{x}, 0, 0) \end{cases} \qquad (P_{\varepsilon,\mathbf{c},\overline{s}})$$

Definition 2.17. Let N be a natural number, and let us choose choose N variation generators $\mathbf{c}_1, \dots, \mathbf{c}_N \in \mathfrak{V}$ and N instants $0 < s_1 < s_2 < \dots \le s_N \le \overline{S}$, with $s_k \in [0,\overline{S}]_{Leb}$ as soon as $\mathbf{c}_k \in \mathfrak{V}_1$. For every $k = 1, \dots, N$ and any L^1-map $[0,\overline{S}] \ni s \mapsto (M,\omega)(s) \in \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^n) \times (\mathbb{R}^n)^*$, let us consider the $(1+n+1) \times (1+n+1)$ matrix

$$\mathscr{E}'_k(M,\omega) := \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{\int_{s_k}^{\overline{S}} M} & 0 \\ \dfrac{\partial \overline{\Psi}}{\partial y^0}(\overline{S}) & \dfrac{\partial \overline{\Psi}}{\partial y}(\overline{S}) e^{\int_{s_k}^{\overline{S}} M} + \displaystyle\int_{s_k}^{\overline{S}} \left(\omega(s) e^{\int_s^{s_k} M}\right) ds & 1 \end{pmatrix} \qquad (2.32)$$

where the exponential of a matrix is defined as in Subsection 2.4.1. Subsequently, define the subset $\Lambda'_N \subset \mathrm{Lin}(\mathbb{R}^N, \mathbb{R}^{1+n+1})$ as

$$\Lambda'_N := \left\{ \left(\mathscr{E}'_1(M,\omega) \begin{pmatrix} V^0_1 \\ V_1 \\ V^l_1 \end{pmatrix}, \dots, \mathscr{E}'_N(M,\omega) \begin{pmatrix} V^0_N \\ V_N \\ V^l_N \end{pmatrix} \right), \quad \begin{array}{c} (M,\omega)(s) \in \partial_y\left(\overline{F}^e, \overline{l}^e\right)(s),\ s \in [0,\overline{S}], \\ \text{is a measurable selection, and} \\ (V^0_k, V_k, V^l_k) \in (v^0_{\mathbf{c}_k,s_k}, v_{\mathbf{c}_k,s_k}, v^l_{\mathbf{c}_k,s_k}) \end{array} \right\}.$$

In the special case when $\boldsymbol{c}_k \in \mathfrak{V}_1$ for all $k \in \{1, \dots, N\}$, we can also define the $(1+n+1+1) \times (1+n+1+1)$ matrix

$$\mathscr{E}_k(M,\omega) := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{\int_{s_k}^{\overline{S}} M} & 0 & 0 \\ \dfrac{\partial \overline{\Psi}}{\partial y^0}(\overline{S}) & \dfrac{\partial \overline{\Psi}}{\partial y}(\overline{S}) e^{\int_{s_k}^{\overline{S}} M} + \displaystyle\int_{s_k}^{\overline{S}} \left(\omega(s) e^{\int_s^{s_k} M}\right) ds & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad (2.33)$$

[9] Of course, $\overline{\mathbf{w}}_\varepsilon$ and $(y_\varepsilon, \beta_\varepsilon)$ depend also on the parameters $\mathbf{c}_k$ and s_k, but we avoid writing them when possible in order to simplify the notation.

and the subset $\Lambda_N \subset \mathrm{Lin}(\mathbb{R}^N, \mathbb{R}^{1+n+1+1})$,

$$\Lambda_N := \left\{ \left(\mathscr{E}_1(M,\omega) \begin{pmatrix} V_1^0 \\ V_1 \\ V_1^l \\ V_1^\nu \end{pmatrix}, \dots, \mathscr{E}_N(M,\omega) \begin{pmatrix} V_N^0 \\ V_N \\ V_N^l \\ V_N^\nu \end{pmatrix} \right), \quad \begin{array}{l} (M,\omega)(s) \in \partial_y\big(\overline{F}^e, \overline{l}^e\big)(s),\ s \in [0,\overline{S}], \\ \text{is a measurable selection, and} \\ (V_k^0, V_k, V_k^l, V_k^\nu) \in \\ \in (v^0_{\mathbf{c}_k,s_k}, v_{\mathbf{c}_k,s_k}, v^l_{\mathbf{c}_k,s_k}, v^\nu_{\mathbf{c}_k,s_k}) \end{array} \right\}.$$

Theorem 2.38 and Corollary 2.39 represent the most important technical step of the proof of the Maximum Principle, and will be proved in Section 2.4.8.

Theorem 2.38. *Let $\left(\overline{y}^0, \overline{y}, \overline{y}^l, \overline{\beta}\right)$ and $\left(\overline{y}^0_{\boldsymbol{\varepsilon}}, \overline{y}_{\boldsymbol{\varepsilon}}, \overline{y}^l_{\boldsymbol{\varepsilon}}, \overline{\beta}_{\boldsymbol{\varepsilon}}\right)$ as in the previous subsections. If we assume the extra assumption $\hat{l}_1(\cdot,0,) \equiv 0$, then the set Λ'_N is a QDQ at $\mathbf{0}$ of the map*

$$(\mathbb{R}^+)^N \ni \boldsymbol{\varepsilon} \mapsto \left(y^0_{\boldsymbol{\varepsilon}}(\overline{S}), y_{\boldsymbol{\varepsilon}}(\overline{S}), y^l_{\boldsymbol{\varepsilon}}(\overline{S}) + \Psi\left(y^0_{\boldsymbol{\varepsilon}}(\overline{S}), y_{\boldsymbol{\varepsilon}}(\overline{S})\right) \right).$$

Moreover, in the special case when $\mathbf{c}_k \in \mathfrak{V}_1$ for all $k \in \{1,\dots,N\}$ (and $\hat{l}_1(\cdot,0,)$ is possibly non vanishing) , Λ_N is a QDQ at $\mathbf{0}$ of the map

$$(\mathbb{R}^+)^N \ni \boldsymbol{\varepsilon} \mapsto \left(y^0_{\boldsymbol{\varepsilon}}(\overline{S}), y_{\boldsymbol{\varepsilon}}(\overline{S}), y^l_{\boldsymbol{\varepsilon}}(\overline{S}) + \Psi\left(y^0_{\boldsymbol{\varepsilon}}(\overline{S}), y_{\boldsymbol{\varepsilon}}(\overline{S})\right), \beta_\varepsilon(S) \right).$$

Corollary 2.39. *Let us use the same notations as in Theorem 2.38 and let us assume that $\hat{l}_1(\cdot,0,) \equiv 0$. For any choice of $\delta > 0$, the family*

$$\Lambda'_N(\mathbb{R}^+)^N = \{\mathscr{L}'(\mathbb{R}^+)^N : \mathscr{L}' \in \Lambda'_N\}, \ ^{10}$$

is an approximating QDQ-multicone of the projected δ-reachable set $\mathscr{R}'_\delta$ at $\big(\overline{y}^0, \overline{y}, \overline{y}^l(\overline{S}) + \overline{\Psi}(\overline{S})\big)$. Moreover, in the special case when $\mathbf{c}_k \in \mathfrak{V}_1$ for all $k \in \{1,\dots,N\}$ (and $\hat{l}_1(\cdot,0,)$ is possibly non vanishing),

$$\Lambda_N(\mathbb{R}^+)^N = \{\mathscr{L}(\mathbb{R}^+)^N : \mathscr{L} \in \Lambda_N\}$$

is an approximating QDQ-multicone of δ-reachable set $\mathscr{R}_\delta$ at $\big(\overline{y}^0, \overline{y}, \overline{y}^l(\overline{S}) + \overline{\Psi}(\overline{S}), \overline{\beta}(\overline{S})\big)$.

Generalizing a classical procedure, we now use the fact that $\Lambda'_N(\mathbb{R}^+)^N$ is a *QDQ* approximating multi-cone of $\mathscr{R}'_\delta$ at $\big(\overline{y}^0, \overline{y}, \overline{y}^l(\overline{S}) + \overline{\Psi}(\overline{S})\big)$ to deduce a linear separability result at time $\overline{S}$, which in turn can be regarded as a Maximum Principle *at* $\overline{S}$.

Lemma 2.40. *Let N a positive integer and let Λ'_N be defined as in the previous subsection. Let us assume that $\overline{\beta}(\overline{S}) < K$ whenever $\mathbf{c}_k \in \mathfrak{V}_2$ for some $k \in \{1,\dots,N\}$. Then, for any approximating QDQ-multicone $\boldsymbol{\mathscr{T}}$ to the target $\mathfrak{T}$ at $\left(\overline{y}^0, \overline{y}\right)$, there exists $(\xi_0, \xi, \xi_c) \in (\mathbb{R} \times \mathbb{R}^n \times \mathbb{R})^* \backslash \{(0,0,0)\}$ such that*

(i) $\xi_c \leq 0$,

[10] One sets $\mathscr{L}(\mathbb{R}^+)^N = \{\mathscr{L}\overline{\varepsilon} : \overline{\varepsilon} \in (\mathbb{R}^+)^N\}$

(ii) $(\xi_0, \xi, \xi_c) \in \left(\mathscr{L}'(\mathbb{R}^+)^N\right)^\perp$, *for some* $\mathscr{L}' \in \Lambda'_N$,

(iii) $(\xi_0, \xi) \in -\mathscr{T}^\perp$ *for some* $\mathscr{T} \in \boldsymbol{\mathscr{T}}$.

Additionally, if $c_k \in \mathfrak{V}_1$ *for all* $k \in \{1, \dots, N\}$ *then there exists an additional multiplier* $\pi \leq 0$ *with the condition*

(iv) $(\xi_0, \xi, \xi_c, \pi) \in \left(\mathscr{L}(\mathbb{R}^+)^N\right)^\perp$, *for some* $\mathscr{L} \in \Lambda_N$,

Proof. By Lemma 2.37 we know that, for $\delta > 0$ sufficiently small, the projected profitable set $\mathfrak{P}'$ and the projected δ-reachable set $\mathfrak{R}'_\delta$ are locally separated.
Moreover, if $\boldsymbol{\mathscr{T}}$ is an approximating QDQ-multicone to the target at $(\overline{y}^0, \overline{y})$, then $\left\{\mathscr{T} \times (-\infty, 0) : \mathscr{T} \in \boldsymbol{\mathscr{T}}\right\}$ is an approximating QDQ-multicone to the projected profitable set $\mathfrak{P}'$ at $\left(\overline{y}^0, \overline{y}, \overline{y}^l(\overline{S}) + \overline{\Psi}(\overline{S})\right)$.
It follows approximating QDQ-multicones $\left\{\mathscr{T} \times (-\infty, 0) : \mathscr{T} \in \boldsymbol{\mathscr{T}}\right\}$ and $\Lambda'_N(\mathbb{R}^+)^N$ are not strongly transverse, from Lemma 2.31. Now, since $\left\{\mathscr{T} \times (-\infty, 0) : \mathscr{T} \in \boldsymbol{\mathscr{T}}\right\}$ is a multicone such that all its cones are contained in the semispace $\mathbb{R}^{1+n} \times (-\infty, 0]$ we can apply Lemma 2.32 and infer the existence of $(\xi_0, \xi, \xi_c) \in (\mathbb{R} \times \mathbb{R}^n \times \mathbb{R})^* \backslash \{(0,0,0)\}$, $\mathscr{L} \in \Lambda_N$, and $\mathscr{T} \in \boldsymbol{\mathscr{T}}$, such that

$$(\xi_0, \xi, \xi_c) \in \left(\mathscr{L}(\mathbb{R}^+)^N\right)^\perp, \quad (\xi_0, \xi, \xi_c) \in \left(\mathscr{T} \times \mathbb{R}^-\right)^\perp.$$

In particular, $\xi_c \leq 0$ and $(\xi_0, \xi) \in -\mathscr{T}^\perp$, which concludes the first part of the proof. The existence of a $\pi \leq 0$ such that property (iv) holds follows from a similar reasoning addressing the local separation of the profitable set and the δ-reachable set in the augmented space $\mathbb{R}^{1+n+1+1}$. □

By using propagation due to the adjoint inclusion, as a direct consequence of Lemma 2.40 we get a *maximum principle for the instants* $s_1, \dots, s_N$ *and the variation generators* $\mathbf{c}_1, \dots, \mathbf{c}_N$*:*

Lemma 2.41. *Let* N *a positive integer. Assume* $\overline{\beta}(\overline{S}) < K$ *and* $\hat{l}_1(\cdot, 0, \cdot) \equiv 0$ *whenever* $\mathbf{c}_k \in \mathfrak{V}_2$ *for some* $k \in \{1, \dots, N\}$*. Then, for any approximating QDQ-multicone* $\boldsymbol{\mathscr{T}}$ *to the target* $\mathfrak{T}$ *at* $(\overline{y}^0, \overline{y})$ *, there exist*

$$(p_0, p, \lambda) \in \mathbb{R}^* \times AC\left([0, \overline{S}], (\mathbb{R}^n)^*\right) \times \mathbb{R}^* \qquad \text{and} \qquad \mathscr{T} \in \boldsymbol{\mathscr{T}}$$

such that $\lambda \geq 0$ *and:*

i)

$$\dot{p} \in -p \ \partial_y \left(f(\overline{y}, \overline{\alpha})\overline{w}^0 + \sum_{i=1}^m g_i(\overline{y})\overline{w}^i \right) + \lambda \ \partial_y l^e(\overline{y}, \overline{w}^0, \overline{w}, \overline{\alpha}) \ ; \tag{2.34}$$

ii)

$$(p_0, p(\overline{S})) + \lambda \left(\frac{\partial \Psi}{\partial y^0}, \frac{\partial \Psi}{\partial y} \right) \left((\overline{y}^0, \overline{y})(\overline{S}) \right) \in -\mathscr{T}^\perp \ ; \tag{2.35}$$

iii) *if* $\mathbf{c}_k = (w^0_k, w_k, a_k, \zeta_k) \in \mathfrak{V}_1$,

$$H(\overline{y}(s_k), p_0, p(s_k), \lambda, 0, w^0_k, w_k, a_k) \leq \\ \leq H(\overline{y}(s_k), p_0, p(s_k), \lambda, 0, \overline{w}^0(s_k), \overline{w}(s_k), \overline{\alpha}(s_k)) \quad (2.36)$$

iv) *if* $\mathbf{c}_k = (i_k, j_k) \in \mathfrak{V}_2$,

$$\min_{V \in [g_{i_k}, g_{j_k}]_{set}} p(s_k)\ V \leq 0. \quad (2.37)$$

If, instead, $\overline{\beta}(\overline{S}) = K$ *and all* $\mathbf{c}_k$ *are in* $\mathfrak{V}_1$ *there exists* $\pi \leq 0$ *such that*

$$H\Big(\overline{y}(s_k), p_0, p(s_k), \lambda, \pi, w^0_k, w_k, a_k\Big) \leq \\ \leq H\Big(\overline{y}(s_k), p_0, p(s_k), \lambda, \pi, \overline{w}^0(s_k), \overline{w}(s_k), \overline{\alpha}(s_k)\Big) \quad (2.38)$$

holds in place of 2.36.

Proof. Let us observe that, by the elementary theory of affine ODEs, we can rephrase (ii) from Lemma 2.40 by saying that there exists a linear form $(\xi_0, \xi, \xi_c) \in (\mathbb{R} \times \mathbb{R}^n \times \mathbb{R})^* \backslash \{(0,0,0)\}$, a measurable selection $M(s) \in \partial_y \left(f(\overline{y}, \overline{\alpha})\overline{w}^0 + \sum_{i=1}^m g_i(\overline{y})\overline{w}^i \right)$, a.e. $s \in (0, \overline{S})$, a measurable selection $\omega(s) \in \partial_y l^e(\overline{y}, \overline{w}^0, \overline{w}, \overline{\alpha})$ a.e. $s \in (0, \overline{S})$, and a choice of

$$\left(V^0_j, V_j, V^l_j\right) \in \left(v^0_{\mathbf{c}_j, s_j}, v_{\mathbf{c}_j, s_j}, v^l_{\mathbf{c}_j, s_j}\right), \quad \forall j = 1, \dots, N,$$

such that $\xi_0 \leq 0$ and, $\forall k = 1, \dots, N$,

$$\begin{aligned} &\xi_0 V^0_k + \xi e^{\int_{s_k}^{\overline{S}} M(s)} V_k + \xi_c \Bigg[\frac{\partial \overline{\Psi}}{\partial y^0}(\overline{S}) V^0_k + \\ &+ \frac{\partial \overline{\Psi}}{\partial y}(\overline{S}) e^{\int_{s_k}^{\overline{S}} M(s)} V_k + \int_{s_k}^{\overline{S}} \omega(s) e^{\int_s^{s_k} M(\sigma)\, d\sigma}\, ds V_k + V^l_k \Bigg] \leq 0. \end{aligned} \quad (2.39)$$

Setting $\lambda := -\xi_c$, $p_0 := \xi_0 - \lambda \frac{\partial \overline{\Psi}}{\partial y^0}(\overline{S})$ and, for all $s \in [0, \overline{S}]$,

$$p(s) := \left(\xi - \lambda \frac{\partial \overline{\Psi}}{\partial y}(\overline{S}) \right) e^{\int_s^{\overline{S}} M(\sigma)\, d\sigma} - \lambda \int_s^{\overline{S}} \omega(\sigma) e^{\int_\sigma^s M(\tau)\, d\tau}\, d\sigma,$$

$p(\cdot)$ satisfies the adjoint differential equation $\dot{p}(s) = -p(s)M(s) + \lambda\omega(s)$. In particular, $p(\cdot)$ verifies the differential inclusion (2.34). Therefore, inequality (2.39) can be written as

$$p_0 V^0_k + p(\overline{s}_k) V_k - \lambda V^l_k \leq 0, \quad (2.40)$$

while (iii) of Lemma 2.40 now reads as (2.35) Specializing (2.40) to bracket-like variations $\mathbf{c}_k = (i_k, j_k) \in \mathfrak{V}_2$, we obtain (2.37), whereas, when $\mathbf{c}_k = (w^0_k, w_k, a_k, \zeta_k) \in \mathfrak{V}_1$ is a needle variation generator, we get (2.36).

The case $\overline{\beta}(\overline{S}) = K$ and all needle variations is proved similarly, by making use of

(iv) instead of (ii) from Lemma 2.40.

□

2.4.7 Infinitely many variations

To complete the proof of Theorem 2.35, we now combine some classical non-empty intersection arguments with the crucial fact that the set-valued brackets are convex-valued.
We will do so only in the case when $\overline{\beta}(\overline{S}) < K$, but the same ideas can be used to deal with the case $\overline{\beta}(\overline{S}) = K$ (which however allows only for the 'first order' part of our maximum principle). As a matter of fact, the maximum condition is the easiest to deal with, and its proof is not that different from classical proofs in the smooth case. Instead, a special care is needed to prove the set-valued Lie-Bracket higher order condition.

By Lusin's Theorem, there exists subsets $E_q \subset [0, \bar{S}]s$, $q = 0, 1, 2, ...$, such that

i) E_0 has null measure,

ii) for every $q > 0$ E_q is a compact set such that the restriction to E_q of the map

$$s \mapsto \left(\overline{w}^0(s), \left(f(\overline{y}(s), \overline{\alpha})(s), \overline{w}^0(s) + \sum_{i=1}^{m} g_i(\overline{y}(s))\overline{w}^i(s)\right), l^e(\overline{y}(s), \overline{w}^0(s), \overline{w}(s)), |\overline{w}|(s)\right)$$

is continuous, and

iii) $(0, \overline{S})_{Leb} = \bigcup_{q=0}^{+\infty} E_q$.

For every $q > 0$ let use $D_q \subseteq E_q$ to denote the set of all density points of E_q. By the Lebesgue density theorem we have that, for every integer $q \geq 1$, D_q has the same Lebesgue measure as E_q, so the set $D := \bigcup_{q=1}^{+\infty} D_q$ is a set of measure equal to $\overline{S}$.

Definition 2.18. Let $X \subseteq D \times \mathfrak{V}$ be any subset of time-generator pairs. We will say that a triple $(p_0, p, \lambda) \in \mathbb{R} \times AC([0, \overline{S}]; \mathbb{R}^n) \times \mathbb{R}^+$ *satisfies property* (P_X) if the following conditions (1)-(3) are verified:

(1) p satisfies the differential inclusion

$$\dot{p} \in -p\ \partial_y \left(f(\overline{y}, \overline{\alpha})\overline{w}^0 + \sum_{i=1}^{m} g_i(\overline{y})\overline{w}^i\right) + \lambda \partial_y l^e(\overline{y}, \overline{w}^0, \overline{w}, \overline{\alpha}) \quad \text{for a.e. } s \in [0, \overline{S}]; \tag{2.41}$$

(2) one has

$$(p_0, p(\overline{S})) + \lambda \left(\frac{\partial \Psi}{\partial y^0}, \frac{\partial \Psi}{\partial y}\right)\left((\overline{y}^0, \overline{y})(\overline{S})\right) \in -\overline{\bigcup_{\mathcal{T} \in \mathscr{T}} \mathcal{T}^{\perp}}; \tag{2.42}$$

(3) for every $(s,\mathbf{c}) \in X$, if $\mathbf{c} = (w^0, w, a, \zeta)$, then

$$p_0 w^0(1+\zeta)+p(s)\left(f(\overline{y}(s),a)w^0+\sum_{i=1}^m g_i(\overline{y}(s))w^i\right)(1+\zeta)-\lambda l^e(\overline{y}(s),w^0,w,a) \leq p_0\overline{w}^0+p(s)\left(f(\overline{y}(s),\overline{\alpha}(s))\overline{w}^0(s)+\sum_{i=1}^m g_i(\overline{y}(s))\overline{w}^i(s)\right)- -\lambda l^e(\overline{y}(s),\overline{w}^0(s),\overline{w},\overline{\alpha}(s)), \quad (2.43)$$

while, if $\mathbf{c} = (i,j)$, then

$$\min_{V\in[g_i,g_j]_{set}(\overline{y}(s))} p_n(s)V \leq 0. \quad (2.44)$$

For any given subset $X \subseteq D \times \mathfrak{V}$, let us set

$$\Theta(X) := \left\{ \begin{array}{l} (p_0,p,\lambda) \in \mathbb{R}\times AC([0,\overline{S}];\mathbb{R}^n)\times\mathbb{R} \,:\, |(p_0,p(\overline{S}),\lambda)| = 1, \\ (p_0,p,\lambda) \text{ verifies the property } (P_X) \end{array} \right\}.$$

Lemma 2.42. *For any subset $X \subseteq D\times\mathfrak{V}$, $\Theta(X)$ is a compact subset of $\mathbb{R}\times AC([0,\overline{S}];\mathbb{R}^n)\times\mathbb{R}$, when the latter is endowed with the norm $\|(p_0,p(\cdot),\lambda)\| := |p_0|+|\lambda|+\|p\|_\infty$.*

Proof. Consider a sequence $(p_{0,n},p_n(s),\lambda_n) \in \Theta(X)$. The set-valued maps

$$s \mapsto \partial_y\left(f(\overline{y}(s),\overline{\alpha}(s))\overline{w}^0(s)+\sum_{i=1}^m g_i(\overline{y}(s))\overline{w}^i(s)\right)$$

$$s \mapsto \partial_y l^e\Big(\overline{y}(s),\overline{w}^0(s),\overline{w}(s),\overline{\alpha}(s)\Big)$$

have uniformly bounded closed convex values as they are Clarke Jacobians of functions that are globally Lipschitz when cut off to a compact set containing a tubular neighbourhood of the set $\overline{y}([0,\overline{S}])$. Furthermore, the quantities $|p_n(\overline{S})|$, λ_n and $p_{0,n}$ are bounded in norm by 1, so that we are in the position to use the following fact:

- *Let $C(s) : [0,\overline{S}] \rightrightarrows \mathbb{R}^n$ and $B(s) : [0,\overline{S}] \rightrightarrows \mathbb{R}^n$ be a measurable multivalued function with compact convex non-empty values. Moreover, assume that the sets $B(s)$ (respectively $C(s)$) are all contained in a fixed ball centered at $0 \in \mathbb{R}^n$ and of radius B^* (respectively C^*).*
  Let $p_n(s)$ be a sequence of solutions to the differential inclusion

  $$\dot{p}(s) \in p(s)B(s)+C(s), \qquad \text{for almost all } s \in [0,\overline{S}] \quad (2.45)$$

  all satisfying $|p_n(\overline{S})| \leq 1$. Then there is a subsequence of $p_n(s)$ that uniformly converges to a function $p(s)$, and $p(s)$ is also a solution to the differential inclusion (2.45).

The proof of this fact can be deduced, e.g., from Theorem 1 in Chapter 2 of [22].

Therefore, modulo thrice extracting subsequences from our sequence, we can assume

$$\lambda_n \to \lambda \geq 0,\ p_{0,n} \to p_0 \text{ and } p_n \to p \in AC \text{ uniformly for } s \in [0, \overline{S}],$$

with $p(s)$ still satisfying the differential inclusion (2.41). Since the paths p_n converges uniformly to p, properties (2.43) and (2.42) are inherited by $p(s)$ from the sequence $p_n(s)$ by passing to the limit. Finally, passing to the limit we get that (2.44) holds true as well. □

Our main result from this section is concluded showing that $\Theta(D \times \mathfrak{V}) \neq \emptyset$ with a non-empty intersection argument. This was of course taken care of in [20] but the proof is rather standard in its classical components while rather complicated and technical in its parts concerning the set-valued Lie brackes, so that we do not fully report it here for the sake of readability and brevity.
Next subsection, however, deals with the proof of crucial Theorem 2.38, which we had postponed.

2.4.8 Proof of Theorem 2.38

In order to conclude the proof of Theorem 2.35, we have to prove Theorem 2.38. For the sake of brevity and readability, here, we will not be slavish on the technical details. Let us recall its statement:

Theorem 2.38. (Angrisani, Rampazzo, 2021, preprint) *Let $\left(\overline{y}^0, \overline{y}, \overline{y}^l, \overline{\beta}\right)$ and $\left(\overline{y}^0_{\boldsymbol{\varepsilon}}, \overline{y}_{\boldsymbol{\varepsilon}}, \overline{y}^l_{\boldsymbol{\varepsilon}}, \overline{\beta}_{\boldsymbol{\varepsilon}}\right)$ as in the previous subsection. If we assume the extra assumption $\hat{l}_1(\cdot, 0,) \equiv 0$, then the set Λ'_N is a QDQ at $\mathbf{0}$ of the map*

$$(\mathbb{R}^+)^N \ni \boldsymbol{\varepsilon} \mapsto \left(y^0_{\boldsymbol{\varepsilon}}(\overline{S}), y_{\boldsymbol{\varepsilon}}(\overline{S}), y^l_{\boldsymbol{\varepsilon}}(\overline{S}) + \Psi\left(y^0_{\boldsymbol{\varepsilon}}(\overline{S}), y_{\boldsymbol{\varepsilon}}(\overline{S})\right)\right).$$

Moreover, in the special case when $\mathbf{c}_k \in \mathfrak{V}_1$ for all $k \in \{1, \ldots, N\}$ (and $\hat{l}_1(\cdot, 0,)$ is possibly non vanishing) , Λ_N is a QDQ at $\mathbf{0}$ of the map

$$(\mathbb{R}^+)^N \ni \boldsymbol{\varepsilon} \mapsto \left(y^0_{\boldsymbol{\varepsilon}}(\overline{S}), y_{\boldsymbol{\varepsilon}}(\overline{S}), y^l_{\boldsymbol{\varepsilon}}(\overline{S}) + \Psi\left(y^0_{\boldsymbol{\varepsilon}}(\overline{S}), y_{\boldsymbol{\varepsilon}}(\overline{S})\right), \beta_{\varepsilon}(S)\right).$$

Proof. Let us introduce a mollifier (in the variable $y \in \mathbb{R}^n$), namely a C^∞ function $\varphi : \mathbb{R}^n \to \mathbb{R}^+$, with support contained in the unit ball, and such that $\displaystyle\int_{\mathbb{R}^n} \varphi(y)\, dy = 1$. For every $\eta > 0$, let us set $\varphi_\eta(y) := \dfrac{1}{\eta^n}\varphi\left(\dfrac{y}{\eta}\right)$, and let us define the mollified vector field

$$\mathscr{F}_\eta\Big(y, w^0, w, a\Big) := \int_{\mathbb{R}^n} \mathscr{F}\Big(y + h, w^0, w, a\Big)\varphi_\eta(h)\, dh.$$

Observe that the control vector field $\mathscr{F}$ is continuous, and, in addition, it is locally Lipschitz continuous in the variable y. Moreover, we can apply a cut off tehcnique and make $\mathscr{F}$ equal to zero outside a compact set containing a small neighbourhood of our local minimizer, so that global Lipschitz continuity (in y) can be assumed as well. It follows that $\mathscr{F}_\eta$ converges uniformly to $\mathscr{F}$ as η goes to 0. For any fixed $\boldsymbol{\varepsilon} \in \mathbb{R}^+$ with a suitably small norm, let us introduce the mollified perturbed Cauchy problem

$$\begin{cases} \dfrac{d}{ds}\left((y^0, y, y^l), \beta\right) = \Big(\mathscr{F}_\eta(y, w^0_{\boldsymbol{\varepsilon}}, w_{\boldsymbol{\varepsilon}}, a_{\boldsymbol{\varepsilon}}), |w|\Big) \cdot \Big(1 + \zeta_{\boldsymbol{\varepsilon}}\Big) \\ \big((y^0, y, y^l), \beta\big)(0) = \big((0, \hat{x}, 0), 0\big) \end{cases} \tag{2.46}$$

and let us use $\left(y^0_{\eta,\boldsymbol{\varepsilon}}, y_{\eta,\boldsymbol{\varepsilon}}, y^l_{\eta,\boldsymbol{\varepsilon}}, \beta_{\boldsymbol{\varepsilon}}\right)$ to denote its unique Carathéodory solution.

Let us define the function $z_{\eta,\boldsymbol{\varepsilon}}(s) := \left|(y^0_{\eta,\boldsymbol{\varepsilon}}, y_{\eta,\boldsymbol{\varepsilon}}, y^l_{\eta,\boldsymbol{\varepsilon}})(s) - (y^0_{\boldsymbol{\varepsilon}}, y_{\boldsymbol{\varepsilon}}, y^l_{\boldsymbol{\varepsilon}})(s)\right|$, $s \in [0, \bar{S}]$, and let us observe that, from the inequality

$$z_{\eta,\boldsymbol{\varepsilon}}(s) \leq \int_0^s \left|\mathscr{F}_\eta(y_{\eta,\boldsymbol{\varepsilon}}, w^0_{\boldsymbol{\varepsilon}}, w_{\boldsymbol{\varepsilon}}, a_{\boldsymbol{\varepsilon}}) - \mathscr{F}(y_{\boldsymbol{\varepsilon}}, w^0_{\boldsymbol{\varepsilon}}, w_{\boldsymbol{\varepsilon}}, a_{\boldsymbol{\varepsilon}})\right|\Big(1 + \zeta_{\boldsymbol{\varepsilon}}\Big) d\sigma \leq$$
$$\int_0^s \left|\mathscr{F}_\eta(y_{\eta,\boldsymbol{\varepsilon}}, w^0_{\boldsymbol{\varepsilon}}, w_{\boldsymbol{\varepsilon}}, a_{\boldsymbol{\varepsilon}}) - \mathscr{F}(y_{\eta,\boldsymbol{\varepsilon}}, w^0_{\boldsymbol{\varepsilon}}, w_{\boldsymbol{\varepsilon}}, a_{\boldsymbol{\varepsilon}})\right|\Big(1 + \zeta_{\boldsymbol{\varepsilon}}\Big) d\sigma + L(1+2\rho)\int_0^s z_{\eta,\boldsymbol{\varepsilon}}(\sigma) \leq$$
$$2K(1+2\rho)\overline{S}\eta + L(1+2\rho)\textstyle\int_0^s z_{\eta,\boldsymbol{\varepsilon}}(\sigma)d\sigma,$$ [11]

and Gronwall's Lemma, we deduce

$$\left|(y^0_{\eta,\boldsymbol{\varepsilon}}, y_{\eta,\boldsymbol{\varepsilon}}, y^l_{\eta,\boldsymbol{\varepsilon}})(s) - (y^0_{\boldsymbol{\varepsilon}}, y_{\boldsymbol{\varepsilon}}, y^l_{\boldsymbol{\varepsilon}})(s)\right| = z_{\eta,\boldsymbol{\varepsilon}}(s) \leq C\eta, \quad \forall s \in [0, \bar{S}], \tag{2.47}$$

where C is a positive constant depending only on $\overline{S}$, K, and L. Therefore

$$\left\|(y^0_{\eta,\boldsymbol{\varepsilon}}, y_{\eta,\boldsymbol{\varepsilon}}, y^l_{\eta,\boldsymbol{\varepsilon}}) - (y^0_{\boldsymbol{\varepsilon}}, y_{\boldsymbol{\varepsilon}}, y^l_{\boldsymbol{\varepsilon}})\right\|_\infty \to 0$$

as $\eta \to 0$, uniformly with respect to $\boldsymbol{\varepsilon}$.

Assume, for a start, that $N = 1$. By the definition of $\overline{w}_{\varepsilon,\mathbf{c},s_1}(s)$ and by standard estimates on needle variations (see e.g. [21]), one gets the following fact:[12]

- if $\mathbf{c} = (w^0, w, a, \zeta)$ is a needle variation generator, then

$$\begin{pmatrix} y^0_{\eta,\varepsilon} \\ y_{\eta,\varepsilon} \\ y^l_{\eta,\varepsilon} \end{pmatrix}(s_1) - \begin{pmatrix} y^0_\eta \\ y_\eta \\ y^l_\eta \end{pmatrix}(s_1) = \varepsilon \begin{pmatrix} w^0(1+\zeta) - \overline{w}^0(s_1) \\ F^e_\eta\left(y_\eta(s_1), w^0, w, a)\right)(1+\zeta) - \overline{F}^e_\eta(s_1) \\ l^e_\eta\left(y_\eta(s_1), w^0, w, a)\right)(1+\zeta) - \overline{l}^e_\eta(s_1) \end{pmatrix} + \phi_1(\eta, \varepsilon), \tag{2.48}$$

where we have used the notation

$$\Big(\overline{F}^e_\eta(s), \overline{l}^e_\eta(s)\Big) := \Big(F^e_\eta, l^e_\eta\Big)\left(y_\eta(s), \overline{w}^0(s), \overline{w}(s), \overline{a}(s)\right) \qquad \forall s \in [0, S],$$

and, for every $\eta \geq 0$, $\phi_1(\eta, \cdot)$ is a continuous function verifying $\phi_1(\eta, \varepsilon) = o(\varepsilon)$, uniformly with respect to η. In other words, $\phi_1(\eta, \varepsilon)/|\varepsilon| \to 0$ as $\varepsilon \to 0$, uniformly with respect to η. This is because it can be proven that $|\phi_1(\eta, \varepsilon)| < C\varepsilon^2$, with C only depending on the Lipschitz constant and L^∞ norm of $\mathscr{F}$: such an inequality is, in turn, proven by expressing $(y^0_{\eta,\varepsilon}, y_{\eta,\varepsilon}, y^l_{\eta,\varepsilon})(s_1)$ and $(y^0_\eta, y_\eta, y^l_\eta)(s_1)$ as Caratheodory solutions of their respective dynamic systems.

[11] K is a bound for the maps $\mathscr{F}$ and $\mathscr{F}_\eta$ and L is a Lipschitz constant for the maps $(y^0, y, y^l) \mapsto \mathscr{F}(y^0, y, y^l, w^0, w, a)$, independent of (w^0, w, a)

[12] We write $\mathbf{c}, \varepsilon, s_1$ instead of $\mathbf{c}_1, \boldsymbol{\varepsilon}, s_1$, respectively.

To write down a similar estimate for bracket-like variation generator, we will need an exact integral formula for commutator-like multiflows that was found in [128]. More precisely we make use of the following fact that the authors proved: for any two smooth fields that are bounded and with bounded derivatives, such as $g_{i,\eta}$ and $g_{j,\eta}$ in our situation, one has

$$\left(e^{-\sqrt{\varepsilon}g_{j,\eta}} \circ e^{-\sqrt{\varepsilon}g_{i,\eta}} \circ e^{\sqrt{\varepsilon}g_{j,\eta}} \circ e^{\sqrt{\varepsilon}g_{i,\eta}}\right)(y_\eta(s_1)) =$$
$$= y_\eta(s_1) + \int_0^{\sqrt{\varepsilon}} \int_0^{\sqrt{\varepsilon}} [g_{i,\eta}, g_{j,\eta}](\theta(y_\eta(s_1), \sigma, \sqrt{\varepsilon}, \tau))\, d\tau\, d\sigma + o(\varepsilon) \quad (2.49)$$

where $\theta(x, \sigma, t, \tau) := \left(e^{(\tau-t)g_{i,\eta}} \circ e^{\sigma g_{j,\eta}} \circ e^{tg_{i,\eta}}\right)(x)$ and $e^{tf}(x)$ denotes the value at t of the solution to the Cauchy Problem $\{y' = f(y), y(0) = x\}$.
Using the integral average theorem, we have

$$\left(e^{-\sqrt{\varepsilon}g_{j,\eta}} \circ e^{-\sqrt{\varepsilon}g_{i,\eta}} \circ e^{\sqrt{\varepsilon}g_{j,\eta}} \circ e^{\sqrt{\varepsilon}g_{i,\eta}}\right)(y_\eta(s_1)) =$$
$$= y_\eta(s_1) + \varepsilon \cdot [g_{i,\eta}, g_{j,\eta}](\theta^*(y_\eta(s_1), \varepsilon)) + o(\varepsilon) \quad (2.50)$$

where $\theta^*(y_\eta(s_1), \varepsilon)$ is an abbreviation of $\theta(y_\eta(s_1), \sigma^*, \sqrt{\varepsilon}, \tau^*)$ for some suitable σ^* and τ^* in $(0, \sqrt{\varepsilon})$. The precise expression of $\theta^*(y_\eta(s_1), \varepsilon)$ is not going to be relevant in future calculations: we only remark that $|\theta^*(y_\eta(s_1), \varepsilon) - y_\eta(s_1)| < C\sqrt{\varepsilon}$ for some constant $C > 0$.

- if $\mathbf{c} = (i, j)$ is a bracket-like variation generator, then

$$\begin{pmatrix} y^0_{\eta,\varepsilon} \\ y_{\eta,\varepsilon} \\ y^l_{\eta,\varepsilon} \end{pmatrix}(s_1) - \begin{pmatrix} y^0_\eta \\ y_\eta \\ y^l_\eta \end{pmatrix}(s_1) = \varepsilon \begin{pmatrix} 0 \\ [g_{i,\eta}, g_{j,\eta}](\theta^*(y_\eta(s_1), \varepsilon)) + o(\varepsilon) \\ 0 \end{pmatrix}, \quad (2.51)$$

  where, for every $\ell = 1, \dots, N, \eta > 0$, the mollified vector field $g_{\ell,\eta}$ is defined as $g_{\ell,\eta}(y) := \int_{\mathbb{R}^n} g_\ell(y + h)\varphi_\eta(h)\, dh$,

To deal with the general case $N \geq 1$, for every $k = 1, \dots, N$, let us set

$$\begin{pmatrix} v^0_{k,\eta,\varepsilon} \\ v_{k,\eta,\varepsilon} \\ v^l_{k,\eta,\varepsilon} \end{pmatrix} := \begin{cases} \begin{pmatrix} w^0(1+\zeta) - \overline{w}^0(s_k) \\ F^e_\eta\left(y_\eta(s_k), w^0, w, a)\right)(1+\zeta) - \overline{F}^e_\eta(s_k) \\ l^e_\eta\left(y_\eta(s_k), w^0, w, a)\right)(1+\zeta) - \overline{l}^e_\eta(s_k) \end{pmatrix} & \text{if, } \boldsymbol{c}_k \in \mathfrak{V}_1, \\ \begin{pmatrix} 0 \\ [g_{i,\eta}, g_{j,\eta}](\theta^*(y_\eta(s_1), \varepsilon_k)) \\ 0 \end{pmatrix}. & \text{if } \boldsymbol{c}_k \in \mathfrak{V}_2 \end{cases}$$

Using induction, by (2.48), (2.51) one easily gets (see e.g. [21])

$$\begin{pmatrix} y^0_{\eta,\varepsilon} \\ y_{\eta,\varepsilon} \\ y^l_{\eta,\varepsilon} \end{pmatrix}(\overline{S}) - \begin{pmatrix} y^0_{\eta} \\ y_{\eta} \\ y^l_{\eta} \end{pmatrix}(\overline{S}) = \sum_1^N \varepsilon_k \, e^{\int_{s_k}^{\bar S} \frac{\partial \mathscr{F}_\eta}{\partial y}} \begin{pmatrix} v^0_{k,\eta,\varepsilon} \\ v_{k,\eta,\varepsilon} \\ v^l_{k,\eta,\varepsilon} \end{pmatrix} + o(|\boldsymbol{\varepsilon}|), \tag{2.52}$$

where $o(|\boldsymbol{\varepsilon}|)$ is independent of η.

from which we obtain

$$\begin{pmatrix} y^0_{\eta,\varepsilon}(\overline{S}) \\ y_{\eta,\varepsilon}(\overline{S}) \\ y^l_{\eta,\varepsilon}(\overline{S}) + \Psi\left(y^0_{\eta,\varepsilon}(\overline{S}), y_{\eta,\varepsilon}(\overline{S})\right) \end{pmatrix} = \begin{pmatrix} y^0_{\eta}(\overline{S}) \\ y_{\eta}(\overline{S}) \\ y^l_{\eta}(\overline{S}) + \Psi\left(y^0_{\eta}(\overline{S}), y_{\eta}(\overline{S})\right) \end{pmatrix} + \mathscr{L}(\eta,\boldsymbol{\varepsilon})\boldsymbol{\varepsilon} + o(|\boldsymbol{\varepsilon}|),$$

where we have set, for every small enough $\eta > 0$, [13]

$$\mathscr{L}(\eta,\boldsymbol{\varepsilon}) = \left(\mathscr{E}_1'\left(\frac{\partial \overline{F}^e_\eta}{\partial y}, \frac{\partial \bar{l}^e_\eta}{\partial y}\right) \cdot \begin{pmatrix} v^0_{1,\eta,\varepsilon} \\ v_{1,\eta,\varepsilon} \\ v^l_{1,\eta,\varepsilon} \end{pmatrix}, \dots, \mathscr{E}_N'\left(\frac{\partial \overline{F}^e_\eta}{\partial y}, \frac{\partial \bar{l}^e_\eta}{\partial y}\right) \cdot \begin{pmatrix} v^0_{N,\eta,\varepsilon} \\ v_{N,\eta,\varepsilon} \\ v^l_{N,\eta,\varepsilon} \end{pmatrix} \right) \in \mathrm{Lin}\left(\mathbb{R}^N, \mathbb{R}^{1+n+1}\right)$$

Therefore, using (2.47), one gets

$$\begin{aligned}
&\left(y^0_{\boldsymbol{\varepsilon}}(\overline{S}), y_{\boldsymbol{\varepsilon}}(\overline{S}), y^l_{\boldsymbol{\varepsilon}}(\overline{S}) + \Psi\left(y^0_{\boldsymbol{\varepsilon}}(\overline{S}), y_{\boldsymbol{\varepsilon}}(\overline{S})\right)\right)^\top = \\
&\qquad \left(y^0_{\eta,\varepsilon}(\overline{S}), y_{\eta,\varepsilon}(\overline{S}), y^l_{\eta,\varepsilon}(\overline{S}) + \Psi\left(y^0_{\eta,\varepsilon}(\overline{S}), y_{\eta,\varepsilon}(\overline{S})\right)\right)^\top - \\
&\left[\left(y^0_{\eta,\varepsilon}(\overline{S}), y_{\eta,\varepsilon}(\overline{S}), y^l_{\eta,\varepsilon}(\overline{S}) + \Psi\left(y^0_{\eta,\varepsilon}(\overline{S}), y_{\eta,\varepsilon}(\overline{S})\right)\right)^\top - \left(y^0_{\boldsymbol{\varepsilon}}(\overline{S}), y_{\boldsymbol{\varepsilon}}(\overline{S}), y^l_{\boldsymbol{\varepsilon}}(\overline{S}) + \Psi\left(y^0_{\boldsymbol{\varepsilon}}(\overline{S}), y_{\boldsymbol{\varepsilon}}(\overline{S})\right)\right)^\top\right] = \\
&\quad \left(y^0_{\eta}(\overline{S}), y_{\eta}(\overline{S}), y^l_{\eta}(\overline{S}) + \Psi\left(y^0_{\eta}(\overline{S}), y_{\eta}(\overline{S})\right)\right)^\top + \mathscr{L}(\eta,\boldsymbol{\varepsilon})\boldsymbol{\varepsilon} + o(|\boldsymbol{\varepsilon}|) + o(\sqrt{\eta}) = \\
&\quad \left(y^0(\overline{S}), y(\overline{S}), y^l(\overline{S}) + \Psi\left(y^0(\overline{S}), y(\overline{S})\right)\right)^\top + \mathscr{L}(\eta,\boldsymbol{\varepsilon})\boldsymbol{\varepsilon} + o(|\boldsymbol{\varepsilon}|) + o(\sqrt{\eta}) - \\
&\left[\left(y^0(\overline{S}), y(\overline{S}), y^l(\overline{S}) + \Psi\left(y^0(\overline{S}), y(\overline{S})\right)\right)^\top - \left(y^0_{\eta}(\overline{S}), y_{\eta}(\overline{S}), y^l_{\eta}(\overline{S}) + \Psi\left(y^0_{\eta}(\overline{S}), y_{\eta}(\overline{S})\right)\right)^\top\right] = \\
&\quad \left(y^0(\overline{S}), y(\overline{S}), y^l(\overline{S}) + \Psi\left(y^0(\overline{S}), y(\overline{S})\right)\right)^\top + \mathscr{L}(\eta,\boldsymbol{\varepsilon})\boldsymbol{\varepsilon} + o(|\boldsymbol{\varepsilon}|) + o(\sqrt{\eta})
\end{aligned} \tag{2.53}$$

Now, our intention is to choose η as a function of $\boldsymbol{\varepsilon}$, namely $\eta = \eta(\boldsymbol{\varepsilon}) = |\boldsymbol{\varepsilon}|^2$, so that we reduce to just one vector-valued parameter $\boldsymbol{\varepsilon}$, we have $o(\sqrt{\eta}) = o(|\boldsymbol{\varepsilon}|)$ and we denote, more simply $\mathscr{L}(\boldsymbol{\varepsilon}) = \mathscr{L}(\eta(\boldsymbol{\varepsilon}), \boldsymbol{\varepsilon})$.
Clearly, as soon as $\boldsymbol{c}_k \in \mathfrak{V}_1$, one has

$$\mathrm{dist}\left(\left(v^0_{k,\eta(\varepsilon),\varepsilon}, v_{k,\eta(\varepsilon),\varepsilon}, v^l_{k,\eta(\varepsilon),\varepsilon}\right) \ , \ \left(v^0_{\mathbf{c}_k,s_k}, v_{\mathbf{c}_k,s_k}, v^l_{\mathbf{c}_k,s_k}\right) \right) \to 0 \quad \text{as } \boldsymbol{\varepsilon} \to \mathbf{0}. \tag{2.54}$$

This holds true if $\boldsymbol{c}_k \in \mathfrak{V}_2$ as well, but we postpone the technical and lenghty proof of this to Lemma 2.43 below, for the sake of clarity.

[13] See (2.32) for the definition of the matrices $\mathscr{E}_k(\cdot,\cdot)$

Given a convex compact set $Q \subset \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^n)$, let us use $P_Q : \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^n) \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^n)$ to denote the projection of x on Q. For any $\eta > 0$, let us consider the selection $M_\eta : [0, \overline{S}] \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}) \in \partial_y \overline{F}^e$ defined by setting

$$M_\eta(s) := P_{\partial_y \overline{F}^e(s)} \left(\frac{\partial \overline{F}^e_\eta}{\partial y}(s) \right) \quad \forall s \in [0, \overline{S}],$$

where we have used the overlined notation

$$\frac{\partial \overline{F}^e_\eta}{\partial y}(s) := \frac{\partial F^e_\eta}{\partial y}\big(y_\eta(s), \overline{w}^0(s), \overline{w}(s), \overline{\alpha}(s)\big) \qquad \partial_y \overline{F}^e(s) := \partial_y F^e\big(\overline{y}(s), \overline{w}^0(s), \overline{w}(s), \overline{\alpha}(s)\big) \quad \forall s \in [0, \overline{S}].$$

Now, consider the function

$$s \mapsto d_\eta(s) := \mathrm{dist}\left(\frac{\partial \overline{F}^e_\eta}{\partial y}(s) \ , \ \partial_y \overline{F}^e(s) \right) = \left| \frac{\partial \overline{F}^e_\eta}{\partial y}(s) - M_\eta(s) \right|,$$

which is measurable (see e.g. Proposition 5.8 in [4]) and essentially bounded in terms of the Lipschitz constant of F. Furthermore, through an argument akin to that utilized to prove Lemma 2.43 below , one proves that the sequence of measurable, equi-bounded functions $d_\eta(s)$ pointwise converges to 0 as $\eta \to 0$, for almost every s in the compact set $[0, \overline{S}]$. Therefore $\|d_\eta\|_1 \to 0$ converges to zero in L^1. Now, the exponential map $M(\cdot) \mapsto e^{\int_{(\cdot)}^{\overline{S}} M}$ is Lipschitz continuous (from L^1 to C^0) when restricted to a family of matrices uniformly bounded in L^1 , i.e., for some constant C,

$$\max_{s \in [0, \overline{S}]} \left| e^{\int_s^{\overline{S}} M_\eta} - e^{\int_s^{\overline{S}} \frac{\partial \overline{F}^e_\eta}{\partial y}} \right| \leq C \|d_\eta\|_1.$$

Hence,

$$\max_{s \in [0, \overline{S}]} \mathrm{dist}\left(e^{\int_s^{\overline{S}} \frac{\partial \overline{F}^e_\eta}{\partial y}}, \left\{ e^{\int_s^{\overline{S}} M}, \ \ M(\cdot) \in \partial \overline{F}^e_\eta(\cdot) \right\} \right) = \max_{s \in [0, \overline{S}]} \left| e^{\int_s^{\overline{S}} M_\eta} - e^{\int_s^{\overline{S}} \frac{\partial \overline{F}^e_\eta}{\partial y}} \right| \to 0$$

as $\eta \to 0$. Moreover, in the same spirit, $\frac{\partial \overline{l}^e}{dy}$ has vanishing distance from $\partial_y \overline{l^e}$ and the integral of $\frac{\partial \overline{l}^e}{dy}$ on any time interval has vanishing distance from the set of integrals of measurable selections of $\partial_y \overline{l}^e$ over the same time interval. Therefore, as $(\mathbb{R}^+)^N \ni \boldsymbol{\varepsilon} \to 0$ implies $\eta = \eta(\boldsymbol{\varepsilon}) \to 0$, we can conclude:

$$\mathrm{dist}\,(\mathscr{L}(\boldsymbol{\varepsilon}), \Lambda'_N) \to 0 \quad \text{as} \ \ (\mathbb{R}^+)^N \ni \boldsymbol{\varepsilon} \to 0.$$

The proof of the first statement in Theorem 2.38 is then concluded as we are precisely in the situation described by Remark 2.30.

The second statement is trivially obtained by repeating the same calculations also keeping track of β, which does not depend on y, hence it does not call for the above mollification argument and the procedure is straightforward noticing:

$$\beta_\varepsilon(\overline{S}) - \overline{\beta}(\overline{S}) := \sum_{k=1}^{N} \int_{s_k-\varepsilon_k}^{s_k} |w_k|(1+\zeta_k) - |\overline{w}(\sigma)|\, d\sigma.$$

□

As it is rather technical, we omit here the proof of

Lemma 2.43. *If $\mathbf{c}_k$ is a bracket-like variation generator and $\eta = \eta(\boldsymbol{\varepsilon}) = |\boldsymbol{\varepsilon}|^2$, then*

$$\operatorname{dist}\Big(\big(v^0_{k,\eta,\boldsymbol{\varepsilon}}, v_{k,\eta,\boldsymbol{\varepsilon}}, v^l_{k,\eta,\boldsymbol{\varepsilon}}\big) \ , \ \big(v^0_{\mathbf{c}_k,s_k}, v_{\mathbf{c}_k,s_k}, v^l_{\mathbf{c}_k,\overline{s}}\big) \Big) \to 0 \quad \textit{as } \boldsymbol{\varepsilon} \to 0 \tag{2.55}$$

2.4.9 An example

In this simple example, by direct computation one can check that, for any value of the parameter $r \in]-2,-1]$ the control $(\bar{w}^0_r, \bar{w}^1_r, \bar{w}^2_r, a_r)$ defined below is not optimal, while for $r = -2$ it is optimal. Let us see that, while this fact was not deductible from the First Order Maximum Principle, the the non-optimality of the controls $(\bar{w}^0_r, \bar{w}^1_r, \bar{w}^2_r, a_r)$ for any $r \in]-2,-1]$ is immediately established by the Higher Order Maximum Principle.

$$\begin{cases} \dfrac{dx}{dt} = f(x,a) + g_1(x)u^1 + g_2(x)u^2 \\ \frac{d\nu}{dt} = |(u^1,u^2)| \\ (x(0),\nu(0) = \big(1,0,2,0\big) \qquad (x(1),\nu(1) \in \mathfrak{T} \times [0,K] \end{cases}$$

$\mathfrak{T} = \mathbb{R}_+ \times \Big\{x \in \mathbb{R}^3,\ |x-(0,0,0.5)| \le 0.5\Big\}$, $K = 2$.

$$f(x,a) = \begin{pmatrix} 0 \\ 0 \\ a \end{pmatrix}, \qquad g_1(x) = \begin{pmatrix} 1 \\ 0 \\ -x_2 + |x_2| \end{pmatrix} \qquad g_2(x) = \begin{pmatrix} 0 \\ 1 \\ x_1 + |x_1| \end{pmatrix}$$

$(a,u) \in A \times \mathscr{C} := [1,2] \times \mathbb{R}^2$

$$\begin{cases} \left(\dfrac{dy^0}{ds}, \dfrac{dy^1}{ds}, \dfrac{dy^2}{ds}, \dfrac{dy^3}{ds}\right) = \big(w^0\,,\, w^1\,,\, w^2\,,\, (-y^2+|y^2|)\,w^1 + (y^1+|y^1|)\,w^2 + rw^0\big) \\ (y^0,y) = \big(0,1,0,2\big) \end{cases}$$

$$\Psi(t,x) := |x|^2 + (t-1)^2$$

For every $r \in [-2,-1]$ let us consider the space-time control $(\bar{w}^0_r, \bar{w}^1_r, \bar{w}^2_r, a_r)$ defined by

$$\big(\bar{w}^0_r, \bar{w}^1_r, \bar{w}^2_r, a_r\big)(s) := \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 0, r\right) \qquad \forall s \in [0,\sqrt{2}].$$

$$\big(\bar{y}^0_r, \bar{y}^1_r, \bar{y}^2_r, \bar{y}^3_r\big)(s) := \left(\frac{\sqrt{2}}{2}s, 1 - \frac{\sqrt{2}}{2}s, 0, 2 + r\frac{\sqrt{2}}{2}s\right)$$

$$\left(\frac{dp_0^r}{ds},\frac{dp_1^r}{ds},\frac{dp_2^r}{ds},\frac{dp_3^r}{ds}\right)(s)\in$$
$$\big(0,-p_3^r(s)\big(1+\tilde{\mathrm{sgn}}(\bar y_r^1(s))\big)\bar w^2(s),-p_3^r(s)\big(-1+\tilde{\mathrm{sgn}}(\bar y_r^2(s))\big)\bar w^1(s),0\big)$$

where we have used the notation

$$\tilde{\mathrm{sgn}}(\eta):=\{\mathrm{sgn}(\eta)\}\ \forall\eta\neq 0,\tilde{\mathrm{sgn}}(\bar 0):=[-1,1],$$
$$\{1+\tilde{\mathrm{sgn}}(\bar 0)\}:=[0,2]\,,\{-1+\tilde{\mathrm{sgn}}(\bar 0)\}:=[-2,0]\,.$$ [14]

Hence
$$\left(\frac{dp_0^r}{ds},\frac{dp_1^r}{ds},\frac{dp_2^r}{ds},\frac{dp_3^r}{ds}\right)(s)\in\left(0\,,\,0\,,\,\frac{\sqrt2}{2}p_3^r(s)[-2,0],0\right)$$

If $r=-1$
$$\big(p_0^r,p_1^r,p_2^r,p_3^r\big)\,(\sqrt2)\in\lambda\,\big(0,0,0,2\big)+\big(0,0,0,(-\infty,0]\big)$$

for some $\lambda\in(-\infty,0]$. Notice that $p_3^r(\sqrt2)<0$: indeed, if it were $p_3^r(\sqrt2)=\rho+2\lambda\geq 0$ for some $\rho\leq 0$, this would imply $p_3^r(\sqrt2)=\rho=\lambda=0$, hence it would follow $\big((p_0^r,p_1^r,p_2^r,p_3^r)(\cdot),\lambda\big)=\Big((0,0,0,0),0\Big)$, which would violate the nontriviality condition i) of the Maximum Principle in Theorem 2.35. So, we can definitely assume[15] that

$$\big(p_0^r,p_1^r,p_2^r,p_3^r\big)\,(\sqrt2)\equiv(0,0,0,-1)\tag{2.56}$$

If $r=(-2,-1)$

$$\big(p_0^r,p_1^r,p_2^r,p_3^r\big)\,(\sqrt2)\in\lambda\,\big(0,0,0,2(r+2)\big)+\big(0,0,0,0\big)$$

for some $\lambda\in(-\infty,0]$. One has that $p_3^r(\sqrt2)=2(r+2)\lambda<0$: indeed, if it were $p_3^r(\sqrt2)=2(r+2)\lambda\geq 0$, this would imply $\lambda\geq 0$, i.e. $\lambda=0$, thence $\big((p_0^r,p_1^r,p_2^r,p_3^r)(\cdot),\lambda\big)=\Big((0,0,0,0),0\Big)$, which would violate the nontriviality condition i). So, also in this case, we can definitely assum that

$$\big(p_0^r,p_1^r,p_2^r,p_3^r\big)\,(\sqrt2)\equiv(0,0,0,-1)\tag{2.57}$$

If $r=-2$
$$\big(p_0^r,p_1^r,p_2^r,p_3^r\big)\,(\sqrt2)\in\lambda\,\big(0,0,0,0\big)+\big(0,0,0,[0,+\infty]\big)$$

So we can choose

$$\big(p_0^r,p_1^r,p_2^r,p_3^r\big)\,(\sqrt2)=(0,0,0,0)\quad\text{and}\quad\lambda=-1\tag{2.58}$$

[14] Also, we set $aB:=\{ab,\ b\in B\}$, for any natural number q, any $a\in\mathbb{R}$, and any subset $B\subset\mathbb{R}^q$. Furthermore, for any natural number q and any collection of subsets $B_1,\ldots,B_q\subset\mathbb{R}$, we use the notation $(B_1,\ldots,B_r):=\{(b_1,\ldots,b_r)\,,\ b_1\in B_1,\ldots,b_r\in B_r\}\,.$

[15] For multipliers are defined up to multiplication by a positive number

so that

$$\left(p_0^r, p_1^r, p_2^r, p_3^r\right)(s) \equiv \left(0,0,0,0\right)$$

is a solution of the adjoint equation verifying the end-point condition (2.58). Let us notice that, for any value of the parameter $r \in [-2,-1]$, the adjoint path $s \mapsto \left(p_0^r, p_1^r, p_2^r, p_3^r\right)(s)$ satisfies the Hamiltonian maximization

$$\max_{(w^0,w,a,\zeta)\in W\times A\times[-\rho,\rho]} H\Big(\bar{y}(s), p_0^r, p(s), \lambda, w^0, w, a, \zeta\Big)$$
$$= H\Big(\bar{y}(s), p_0^r, p(s), \lambda, \bar{w}^0, \bar{w}(s), \bar{a}(s), 0\Big),$$

where (we have written y, w, and p for (y^1, y^2, y^3), (w^1, w^2), and (p_1, p_2, p_3), respectively, and) we have set

$$H(y, p_0^r, p^r, w^0, w, a) := p_0^r w^0 + p\left(f(y) + g_1(y)w^1 + g_2(y)w^2\right).$$

Therefore, upon setting $\bar{\beta}_r(s) := \frac{\sqrt{2}}{2}s$, $\forall s \in [0,\sqrt{2}]$, we have proved that, for any value of the parameter $r \in [-2,-1]$, the canonical process $\left(\sqrt{2}, \bar{w}_r^0, \bar{w}_r, \bar{y}_r^0, \bar{y}_r, \bar{\beta}_r\right)$ *verifies the First Order Maximum Principle* defined by **i)**-**v)** in Theorem 2.35 However, for all values of the parameter r belonging to $]-2,1]$ the canonical process $\left(\sqrt{2}, \bar{w}_r^0, \bar{w}_r, \bar{y}_r^0, \bar{y}_r, \bar{\beta}_r\right)$ *does not verify the bracket-involving condition* **vi)** in Theorem 2.35. Indeed, one has

$$\begin{array}{ll}
[g_1,g_2]_{set}(x) = (0,0,2+\mathrm{sgn}(x_1)-\mathrm{sgn}(x_2)) & \text{if } x_1, x_2 \neq 0,\\
[g_1,g_2]_{set}(x) = \Big(0,0,\Big[1-\mathrm{sgn}(x_2), 3-\mathrm{sgn}(x_2)\Big]\Big) & \text{if } x_1 = 0,\ x_2 \neq 0,\\
[g_1,g_2]_{set}(x) = \Big(0,0,\Big[1+\mathrm{sgn}(x_1), 3+\mathrm{sgn}(x_1)\Big]\Big) & \text{if } x_1 \neq 0,\ x_2 = 0,\\
[g_1,g_2]_{set}(x) = \Big(0,0,\Big[0,4\Big]\Big) & \text{if } x_1 = x_2 = 0.
\end{array}$$

Therefore, for all $s \in [0,\sqrt{2})$,

$$p^r(s)[g_1,g_2]_{set}(\bar{y}_r(s)) = (0,0,-1)\Big(0,0,\Big[2,4\Big]\Big) = [-4,-2] \not\ni 0.$$

Now let $r = -2$. By direct computation we already know that the control

$$\left(w_r^0, w_r^1, w_r^2, r\right) \equiv \Big(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 0, r\Big)$$

is optimal (indeed $(y_r^0, y_r^1, y_r^2, y_r^2, y_r^3) = (1,0,0,0)$), so the Higher Order Maximum Principle applies. Actually, with the above choice of the multipliers [16] one has

$$p^r(s)[g_1,g_2]_{set}(\bar{y}_r(s)) = 0 \quad \forall s \in [0,\sqrt{2}].$$

[16] i.e. $\left(p_0^r, p_1^r, p_2^r, p_3^r\right) \equiv (0,0,0,0)$, $\lambda = -1$.

Chapter 3

Orlicz Spaces and related function spaces

3.1 Orlicz Spaces: definition and functional analytic properties

In this section we will define Orlicz spaces L^Ψ, which are a generalization of the idea of Lebesgue L^p spaces. They are defined as spaces of functions satisfying some integrabilty conditions, and it can also be shown they are rearrangement invariant function spaces, hence they do not fit into the family of spaces defined by means of oscillation we are exemplifying in this text, but they are introduced for a reason nonetheless. We will see a beautiful result in the next section stating that functions whose derivatives belong to Orlicz spaces can be characterized in terms of a suitable kind of oscillation. In addition to that, in this section, we report a result we recently obtained (see [18]) with Giacomo Ascione and Gianluigi Manzo for a family of Orlicz spaces L^Ψ, stating that the closure M^Ψ of L^∞ in them forms an o–O structure (M^Ψ, L^Ψ). As we know from Theorem 2.4 in Section 2.2 and from Theorem 2.16 in Section 2.3, this will imply a series of interesting functional properties satisfied by L^Ψ.
Moreover at the end of the chapter we will devote a section to presenting an application of these concepts to the theory of regularity in Calculus of Variations.

Definition 3.1. We say that a function $\Psi : [0, +\infty) \to [0, +\infty)$ is a *Young function* if

$$\Psi(t) = \int_0^t \psi(\tau)\, d\tau \text{ for } t > 0$$

where $\psi : [0, +\infty) \to [0, +\infty)$ is a non-decreasing function such that:

- $\psi(0) = 0$;
- $\psi(t) > 0$ for any $t > 0$;
- ψ is right-continuous;
- $\lim\limits_{t\to+\infty} \psi(t) = +\infty$.

Definition 3.2. Let Ψ be a Young function and Ω a measurable space. The *Orlicz space* $L^\Psi(\Omega)$ is defined as

$$L^\Psi(\Omega) = \left\{ u \in L^1(\Omega) : \ \exists \lambda > 0 \text{ such that} \int_\Omega \Psi\left(\frac{|u|}{\lambda}\right) < +\infty \right\}.$$

The quantity $\|u\|_{L^\Psi(\Omega)} = \inf \left\{ \lambda > 0 : \ \int_\Omega \Psi\left(\frac{|u|}{\lambda}\right) dx \leq 1 \right\}$ is called *Luxemburg norm* of u. Orlicz spaces are Banach function spaces with respect to the Luxemburg norm [25].

From now on we will deal mainly with $\Omega = Q_0 := [0,1]^n \subset \mathbb{R}^n$: if $n = 1$ we will equivalently denote it with I_0, with the I standing for "interval". Moreover, whenever it is not necessary to specify the domain of integration, we will denote the Luxemburg norm as $\|\cdot\|_\Psi$. Of course, with the choice $\Psi(t) = t^p$ we obtain precisely Lebesgue spaces L^p with their usual L^p norm. Another famous example of an Orlicz space is the space Exp of exponentially summable functions that one obtains if he chooses $\Psi(t)$ to be $\Psi(t) = e^t - 1$.
It can be shown that two Orlicz spaces are set-theoretically equal and with equivalent norms if the two functions Ψ_1 and Ψ_2 defining them are such that $\Psi_1(t)/\Psi_2(t)$ is bounded away from 0 and $+\infty$ as t goes to $+\infty$, showing that it is only the definitive growth rate of the Young function that determines the corresponding Orlicz space.
We will denote by Φ the Young conjugate function of Ψ, that is the only Young function Φ such that $\Phi' = (\Psi')^{-1}$. In other words, if

$$\Psi(t) = \int_0^t \psi(s)\, ds \quad \text{ then } \Phi(t) = \int_0^t \psi^{-1}(s)\, ds,$$

where the inverse function ψ^{-1} is intended in the generalized sense, i.e.

$$\psi^{-1}(s) = \inf\{\sigma : \psi(\sigma) \geq s\}$$

as it is not obvious that ψ is bijective, as it is taken as a non-decreasing but not necessarily strictly increasing function.
Conjugate (or complementary) Young functions induce Orlicz spaces L^Ψ and L^Φ with a very strong interplay stemming from the Young inequality

$$|ab| \leq \Psi(a) + \Phi(b)$$

which allows the proof of a Holder inequality

$$\int |f(x)g(x)| \leq \|f\|_\Psi \|g\|_\Phi.$$

In turn, this is a starting point to understand any possible duality relationship that there is between complementary Orlicz spaces.
As a matter of fact, whenever $\Psi = t^p/p$ with $p \in (1, +\infty)$, we have that the conjugate function is $\Phi = t^q/q$ where q is such that $\frac{1}{p} + \frac{1}{q} = 1$ and the corresponding Orlicz spaces L^p and L^q are a couple of Lebesgue spaces that is such that any of them represents the

dual space of the other.
But Young functions give rise to Orlicz spaces with very different functional properties based on their growth rate: for example the space Exp is not reflexive like (most of) the Lebesgue spaces L^p. To explore exactly what it is that determines functional properties like reflexivity, we give the following definitions.

Definition 3.3. We say a Young function Φ satisfies the Δ_2 *condition* if there exists a constant $C > 1$ such that for any $t \in [0, +\infty)$

$$\Phi(2t) \leq C\Phi(t).$$

([109]) We say that a Young function Ψ satisfies the Δ^0 *condition* if there is a $k > 1$ such that

$$\lim_{t \to +\infty} \frac{\Psi(kt)}{\Psi(t)} = +\infty.$$

It can be shown that this second Δ_0 growth condition is a strong negation of the Δ_2 condition for Ψ and implies that the conjugate function Φ of Ψ satisfies the Δ_2 condition.

Proposition 3.1 (Chen, 1996, [46])**.** *The following properties are equivalent:*

- *Φ satisfies the Δ_2 condition;*
- *$L^\infty(Q_0)$ is dense in $L^\Phi(Q_0)$.*
- *$L^\Phi(Q_0)$ is separable.*

Proposition 3.2 (Lefevre, 2008, [109])**.** *If Ψ satisfies the Δ^0 condition then its conjugate Φ satisfies the Δ_2 condition and for all $p > 1$*

$$\lim_{t \to +\infty} \frac{\Psi(t)}{t^p} = +\infty.$$

Moreover, it can be shown that

Proposition 3.3. *The space L^Ψ is reflexive and in duality with the space L^Φ if and only if Φ is the conjugate of Ψ and they both satisfy the Δ_2 condition.*

Definition 3.4. Let Ψ be a Young function. The Morse space M^Ψ is defined as the closure of L^∞ in L^Ψ:

$$M^\Psi(Q_0) = \overline{L^\infty(Q_0)}^{L^\Psi(Q_0)}$$

or, equivalently

$$M^\Psi(Q_0) = \left\{ u \in L^1(Q_0) : \ \forall \lambda > 0 \ \int_{Q_0} \Psi\left(\frac{u}{\lambda}\right) dx < +\infty \right\}.$$

With respect to this, we specify that equivalent formulas for the distance of a function from L^∞ (that is the same as the distance from its closure M^Ψ) have been found in a paper by Carozza and Sbordone [44].
The aim of this section is to identify a large class of Orlicz spaces L^Ψ which have a

$o-O$ structure with their subspace M^Ψ. In particular we will show that an equivalent norm of the form

$$|||u|||_\Psi := \sup_{0<t<1} \frac{u^{**}(t)}{\Psi^{-1}(1/t)}$$

where u^{**} will be defined later, can be introduced for Young functions Ψ such that the function $t \mapsto \Psi^{-1}(1/t)$ belongs to $L^\Psi(0,1)$. In these cases, when it is possible to introduce this new norm and prove it equivalent to the Luxemburg norm, we will be able to prove the o–O structure of (M^Ψ, L^Ψ).
Let us recall an important notion that is needed in order to define u^{**} and reach a proof of the above statement.

Definition 3.5. The *non-increasing rearrangement* of $u \in L^1(\Omega)$ where $|\Omega| = 1$ is the function $u^* : [0,1] \to [0, +\infty]$ defined as

$$u^*(t) = \sup_{|A|=t} \inf_A |u|.$$

We recall that two functions $u_1 \in L^1(\Omega_1)$ and $u_2 \in L^1(\Omega_2)$ are said to be *equimeasurable* if $u_1^* = u_2^*$. Orlicz spaces have the useful property that if $u \in L^\Psi(Q_0)$ and $v \in L^1(Q_0)$ is equimeasurable to u then $v \in L^\Psi(Q_0)$ and $\|u\|_\Psi = \|v\|_\Psi$. We recall also that a Banach space with this property is said to be *rearrangement invariant* (see also [25]).
We would like to use u^* to define an equivalent norm. However, the map $u \mapsto u^*$ is not sub-additive. For this reason we have to use another function, linked to the non-increasing rearrangement.

Definition 3.6. Given a function $u : Q_0 \to \mathbb{R}$ the *maximal function* of u^* is the function

$$u^{**}(t) = \fint_0^t u^*(\tau)\, d\tau.$$

For the maximal function, one can easily show the following result.

Proposition 3.4 (Bennett, 1988, [25])**.** *Consider $u, v \in L^1(Q_0)$. Then*

- $u^{**} \equiv 0$ *if and only if* $u = 0$ *almost everywhere;*
- *If* $|u| \le |v|$ *a.e. then* $u^{**} \le v^{**}$*;*
- $u^* \le u^{**}$*;*
- $(au)^{**} = |a|u^{**}$ *for any* $a \in \mathbb{R}$*;*
- $(u+v)^{**} \le u^{**} + v^{**}$.

In particular u^{**} is sub-additive and thus it can be used to define a norm. For a rearrangement invariant Banach function space X it is possible to define a function $\phi_X(t) := \|\chi_E\|_X$, where $|E| = t$, called the *fundamental function*. It satisfies a useful property.

Proposition 3.5. *Let Ψ be a finite Young function and $u \in L^1(Q_0)$. Then, for any $t > 0$:*

$$\frac{u^{**}(t)}{\Psi^{-1}(1/t)} \leq \|u\|_\Psi \,.$$

Proof. Given a Banach space X we denote its associate ([124, Definition 6.2.4]) with X'. It can be shown ([124, Lemma 7.9.13]) that

$$\int_0^t u^*(\tau)\, d\tau \leq \|u\|_X \, \varphi_{X'}(t) \tag{3.1}$$

and ([124, Theorem 7.9.6])

$$\varphi_X(t)\varphi_{X'}(t) = t, \tag{3.2}$$

so combining (3.1) and (3.2) we obtain that

$$u^{**}(t)\varphi_X(t) \leq \|u\|_X$$

Since an easy calculation shows that for the Orlicz space $L^\Psi(Q_0)$ endowed with the Luxemburg norm $\varphi_{L^\Psi}(t) = \dfrac{1}{\Psi^{-1}(1/t)}$, the proof is complete.

□

Definition 3.7 (Kaminsksa, 2004, [97]). Let $A : [0, +\infty) \to [0, +\infty)$, $A(0) = 0$, A be increasing, and $A(t) > 0$ for $t > 0$. Then the *Marcinkiewicz space* $L^{A,\infty}$ (also called *weak Lorentz space*) is the collection of all measurable functions f such that

$$\|f\|_{L^{A,\infty}} = \sup_{t>0} \frac{1}{A(t)} \int_0^t f^*(\tau)\, d\tau < +\infty$$

Corollary 3.6. *Let Ψ be a finite Young function. Then*

$$L^\Psi \hookrightarrow L^{A,\infty} \text{ where } A(t) = t\Psi^{-1}(1/t).$$

We are now ready to study the quantity

$$|||u|||_{L^\Psi} = \sup_{0<t<1} \frac{u^{**}(t)}{\Psi^{-1}(1/t)}.$$

From Proposition 3.4 and 3.5, one can easily prove that this is a norm on L^Ψ and if $|u| \leq |v|$ we have $|||u|||_\Psi \leq |||v|||_\Psi$.

Theorem 3.7 (Angrisani, Ascione, Manzo, 2019, [18]). *Let us define the quantity* $N_*(u) = \sup\limits_{0<t<1} \dfrac{u^*(t)}{\Psi^{-1}(1/t)}$.
The following statements are equivalent:

i. $\Psi^{-1}(1/t) \in L^\Psi(I_0)$ *i.e. there exists $k > 0$ such that*

$$\int_{I_0} \Psi\left(\frac{1}{k}\Psi^{-1}\left(\frac{1}{t}\right)\right) dt < \infty;$$

ii. $u \in L^\Psi(Q_0) \iff |||u|||_\Psi < +\infty$;

iii. $u \in L^\Psi(Q_0) \iff N_*(u) < +\infty$;

iv. There exist two constants $K_1, K_2 > 0$ *such that for all* $u \in L^1(Q_0)$

$$K_1 N_*(u) \le K_1 \, |||u|||_\Psi \le \|u\|_\Psi \le K_2 N_*(u) \le K_2 \, |||u|||_\Psi \, .$$

Proof. To show that i.⇒ii. we first notice that the forward implication in ii. is just Proposition 3.5. To show the other implication we notice that since $u^{**}(t) \le |||u|||_\Psi \, \Psi^{-1}\left(\frac{1}{t}\right)$ we obtain

$$\int_0^1 \Psi\left(\frac{u^{**}(t)}{\lambda}\right) dt \le \int_0^1 \Psi\left(\frac{|||u|||_\Psi}{\lambda} \Psi^{-1}\left(\frac{1}{t}\right)\right) dt$$

so that ii. implies $u^{**} \in L^\Psi(I_0)$, but this implies $u^* \in L^\Psi(I_0)$ and by Luxemburg representation theorem [124, Theorem 7.8.3] $u \in L^\Psi(Q_0)$.
The implication ii.⇒iii. follows from the inequality $u^* \le u^{**}$.
The implication iii.⇒i. is trivial if $n = 1$ (so that $Q_0 = I_0$) since if

$$\Psi^{-1}(1/t) \notin L^\Psi \text{ and } \left|\left|\left|\Psi^{-1}\right|\right|\right|_{L^\Psi(I_0)} = 1$$

we get a negation of iii., while in the general case we choose a function f such that $u^*(t) = \Psi^{-1}(1/t)$ (we can get such a function as the limit of a suitable sequence of simple functions).
Let's now show that i., iii.⇒iv. From iii. and from $N_*(u) \le |||u|||_\Psi$ we obtain that if one of the quantities involved is infinite then so are the others. From i. we can choose $K_2 = \|\Psi^{-1}(1/t)\|_{L^\Psi(I_0)} < +\infty$ because in this case

$$\int_0^1 \Psi\left(\frac{u^*(t)}{N_*(u) K_2}\right) dt \le \int_0^1 \Psi\left(\frac{1}{K_2} \Psi^{-1}\left(\frac{1}{t}\right)\right) dt \le 1$$

and $N_*(u) \le |||u|||_\Psi$. Finally, Proposition 3.5 shows that $K_1 = 1$.
Finally, iv. trivially implies ii. so the proof is concluded.

□

Corollary 3.8. *Let* Ψ *be a finite Young function. The embedding of* L^Ψ *in* $L^{A,\infty}$ *as in Corollary 3.6 has norm* 1.
Also, it is an isomorphism if and only if $\Psi^{-1}(1/t) \in L^\Psi$.

Remark 3.9. Observe that under the hypotheses of Theorem 3.7 we have that $u \in L^\Psi(Q_0)$ if and only if $u^{**} \in L^\Psi(I_0)$ since we know that if $u^{**} \in L^\Psi(I_0)$ then, by monotonicity of the norm, $u^* \in L^\Psi(I_0)$ and then $u \in L^\Psi(Q_0)$ by Luxemburg representation theorem [124, Theorem 7.8.3], while:

$$\|u^{**}\|_{L^\Psi(I_0)} \le K N_*(u^{**}) = K \, |||u|||_\Psi < +\infty.$$

We now consider, for $u \in L^\Psi(Q_0)$, the quantity

$$[u] := \limsup_{t\to 0} \frac{u^{**}(t)}{\Psi^{-1}(1/t)}.$$

Proposition 3.10. *Let Ψ be a Young function. Then $\forall u_1, u_2 \in L^\Psi(Q_0)$ and $\forall v \in L^\infty(Q_0)$*

- $[v] = 0$;
- $[u_1 + u_2] \le [u_1] + [u_2]$;
- $[u_1 - v] = [u_1]$.

Now we want to show a measure of the distance between a generic function $u \in L^\Psi(Q_0)$ and $L^\infty(Q_0)$ which follows from the quantities $[u]$ that is equivalent from the usual distance induced by the Luxemburg norm.

Theorem 3.11 (Angrisani, Ascione, Manzo, 2019, [18])**.** *Let Ψ be a Young function such that $\Psi^{-1}(1/t) \in L^\Psi(I_0)$. Then there exist two constants $D_1, D_2 > 0$ such that*

$$D_1[u] \le \inf_{v\in L^\infty} \|u - v\|_\Psi \le D_2[u].$$

Proof. Since, by Theorem 3.7, we have that

$$[u] = [u - v] \le |||u - v|||_\Psi \le \|u - v\|_\Psi\,,$$

we can take $D_1 = 1$. On the other hand, for any $\varepsilon > 0$, by definition of $[u]$, there exists a $\delta_\varepsilon < 1$ such that for any $t \in [0, \delta_\varepsilon]$ we have

$$u^*(t) \le u^{**}(t) \le ([u] + \varepsilon)\Psi^{-1}\left(\frac{1}{t}\right).$$

The function

$$u_\varepsilon(x) = \begin{cases} u(x) & u(x) \le u^*(\delta_\varepsilon) \\ 0 & u(x) > u^*(\delta_\varepsilon) \end{cases}$$

is in $L^\infty(Q_0)$, and we have that $(u - u_\varepsilon)^* = u^*\chi_{[0,\delta_\varepsilon]}$.
If we now take $D_2 = \|\Psi^{-1}(1/t)\|_{L^\Psi(I_0)}$ we have

$$\begin{aligned} \int_0^1 \Psi\left(\frac{(u - u_\varepsilon)^*(t)}{D_2([u] + \varepsilon)}\right) dt &= \int_0^{\delta_\varepsilon} \Psi\left(\frac{u^*(t)}{D_2([u] + \varepsilon)}\right) dt \\ &\le \int_0^{\delta_\varepsilon} \Psi\left(\frac{1}{D_2}\Psi^{-1}\left(\frac{1}{t}\right)\right) dt \le 1, \end{aligned}$$

hence, taking the limit as $\varepsilon \to 0$ we have

$$\inf_{v\in L^\infty} \|u - v\|_\Psi \le D_2 \cdot [u].$$

□

From this equivalent distance we obtain also the following norm-attaining property.

Corollary 3.12. *For any $v \in M^\Psi$ there exists a $\tilde{t} \in [0,1]$ such that*

$$|||v|||_\Psi = \frac{v^{**}(\tilde{t})}{\Psi^{-1}(1/\tilde{t})}.$$

Proof. From Theorem 3.11 we have for $v \in M^\Psi$ that $\lim_{t \to 0} \frac{v^{**}(t)}{\Psi^{-1}(1/t)} = 0$. Hence if we consider the function $g : t \in (0,1] \to \frac{v^{**}(t)}{\Psi^{-1}(1/t)}$, it can be extended with continuity in 0 by posing $g(0) = 0$. The existence of such $\tilde{t} \in [0,1]$ is then assured by Weierstrass theorem.

□

We can finally prove the announced main result of this section

Theorem 3.13 (Angrisani, Ascione, Manzo, 2019, [18])**.** *Let Ψ be a Young function such that $\Psi^{-1}(1/t) \in L^\Psi(I_0)$. Suppose there exists a $p > 1$ such that*

$$\lim_{t \to +\infty} \frac{\Psi(t)}{t^p} = +\infty$$

(in particular if Ψ is Δ^0).
Then $(M^\Psi(Q_0), L^\Psi(Q_0))$ is a (E_0, E) pair satisfying Assumption AP.
As a consequence, under this assumption,

- *$(M^\Psi(Q_0))^{**}$ is isometric to $L^\Psi(Q_0)$*
- *$(M^\Psi(Q_0))^*$ is the strongly unique predual of $L^\Psi(Q_0)$*
- *$M^\Psi(Q_0)$ is a M-ideal in $L^\Psi(Q_0)$ with respect to $|||\cdot|||_\Psi$*

Proof. The assumption on the growth of Ψ allows us to choose $X = L^p(Q_0)$ [45]. The choice for Y and $\mathscr{L}$ however is not that straightforward.
Let $\mathscr{K}$ be the set of all functions $\widetilde{\mathrm{E}} : [0,1] \to \mathscr{F}$, where $\mathscr{F}$ is the σ-algebra of measurable sets of Q_0, that satisfy the following conditions

$$|\widetilde{\mathrm{E}}(t)| = t, \ \forall t \in [0,1] \tag{3.3}$$

and

$$s < t \Rightarrow \widetilde{\mathrm{E}}(s) \subset \widetilde{\mathrm{E}}(t) \tag{3.4}$$

and we define the space Y to be

$$Y = \left\{ \mathbf{u} = \{u_{\widetilde{\mathrm{E}}}\}_{\widetilde{\mathrm{E}} \in \mathscr{K}} \in \left(L^1(Q_0)\right)^{\mathscr{K}} : \|\mathbf{u}\|_Y := \sup_{\widetilde{\mathrm{E}} \in \mathscr{K}} \left\|u_{\widetilde{\mathrm{E}}}\right\|_{L^1(Q_0)} < +\infty \right\}.$$

If $\{\mathbf{u}_j\}_{j\in\mathbb{N}}$ is a Cauchy sequence, by definition, for every $\varepsilon > 0$ there is a $\nu \in \mathbb{N}$ such that for all $j, k \geq \nu$ we have that

$$\sup_{\widetilde{\mathrm{E}}\in\mathscr{K}} \left\| u_{j,\widetilde{\mathrm{E}}} - u_{k,\widetilde{\mathrm{E}}} \right\|_{L^1(Q_0)} = \|\mathbf{u}_j - \mathbf{u}_k\|_Y < \varepsilon,$$

so that if we take $\mathbf{u} = \{u_{\widetilde{\mathrm{E}}}\}_{\widetilde{\mathrm{E}}\in\mathscr{K}}$ such that $u_{\widetilde{\mathrm{E}}} = \lim_{j\to\infty} u_{j,\widetilde{\mathrm{E}}}$ we easily see that $\mathbf{u} \in Y$ and $\mathbf{u}_j$ converges to $\mathbf{u}$ in Y, thus Y is a Banach space. We finally choose our collection of operators to be $\mathscr{L} = \{L_t : t \in (0,1]\}$, endowed with the topology induced from the natural topology on $(0,1]$, where

$$L_t u = \left\{ \frac{u\chi_{\widetilde{\mathrm{E}}(t)}}{t\,\Psi^{-1}(1/t)} \right\}_{\widetilde{\mathrm{E}}\in\mathscr{K}}.$$

Let us proceed to show that our choices for X, Y and $\mathscr{L}$ give us the spaces $L^\Psi(Q_0)$ and $M^\Psi(Q_0)$.
In [25] we can find the following alternative expression for u^{**}:

$$u^{**}(t) = \frac{1}{t} \sup_{|E|=t} \int_E u\,dx. \tag{3.5}$$

Suppose that for every measurable $E \subset Q_0$ and for all $0 \leq a < |E| < b \leq 1$ there is a function $\widetilde{\mathrm{E}} : [a,b] \to \mathscr{F}$ satisfying (3.3) and (3.4) such that $\widetilde{\mathrm{E}}(|E|) = E$. Then we would easily have that $\|L_t u\|_Y = \dfrac{u^{**}(t)}{\Psi^{-1}(1/t)}$ by Equation (3.5), and this implies that E is $L^\Psi(Q_0)$ endowed with our equivalent norm $|||\cdot|||_\Psi$.
Let us now show this claim. We first prove that for all measurable and bounded $U \subset \mathbb{R}^n$ there is a function $\widetilde{\mathrm{E}} : [0, |U|] \to \mathscr{F}$ satisfying (3.3) and (3.4).
Let $Q_s = \left[-\dfrac{s}{2}, \dfrac{s}{2}\right]^n$ and suppose $U \subset Q_M$. The function

$$m : s \in [0, M] \mapsto |U \cap Q_s| \in [0, |U|]$$

is continuous, increasing and takes the values 0 in $t = 0$ and $|U|$ in $t = M$, so it is surjective and we can choose $\widetilde{\mathrm{E}}(t) = U \cap Q_s$, where $m(s) = t$.
Properties (3.3) and (3.4) are easy to show. We can use this to prove our claim. We consider the two functions

$$\widetilde{\mathrm{E}}_1 : [0, |E|] \to \mathscr{F}\,|_E \text{ and } \widetilde{\mathrm{E}}_2 : [0, 1-|E|] \to \mathscr{F}\,|_{Q_0\setminus E},$$

so we can take

$$\widetilde{\mathrm{E}}(t) := \begin{cases} \widetilde{\mathrm{E}}_1(t) & \text{if } 0 \leq t < |E| \\ E & \text{if } t = |E| \\ E \cup \widetilde{\mathrm{E}}_2(t - |E|) & \text{if } |E| < t \leq 1, \end{cases}$$

and by restricting $\widetilde{\mathrm{E}}$ to the interval $[a,b]$ we can conclude the proof of this claim.
All we need to show now are the topological properties of $\mathscr{L}$. Since $\mathscr{L}$ is homeomorphic to $(0,1]$ we immediately get the local compactness, σ-compactness and the Hausdorff

property. The boundedness of the operators L_t easily follows from Hölder inequality, so all we need to prove is the continuity of the operator $T_u : L \in \mathscr{L} \mapsto Lu \in Y$. Let us fix a value $\widetilde{\mathrm{E}}$ for the index of Lu. Since $L_s \to L_t$ iff $s \to t$, we can write for $h > 0$ (the case $h < 0$ is similar)

$$\begin{aligned} \|L_{t+h}u - L_t u\|_{L^1} \leq & \int_{\widetilde{\mathrm{E}}(t)} \left| \frac{1}{(t+h)\Psi^{-1}(1/(t+h))} - \frac{1}{t\,\Psi^{-1}(1/t)} \right| |u| + \\ & + \int_{\widetilde{\mathrm{E}}(t+h)\setminus \widetilde{\mathrm{E}}(t)} \frac{|u|}{|(t+h)\Psi^{-1}(1/(t+h))|} \leq \\ & \leq \left| \frac{1}{(t+h)\Psi^{-1}(1/(t+h))} - \frac{1}{t\,\Psi^{-1}(1/t)} \right| \|u\|_{L^1} + \\ & + \left| \frac{1}{(t+h)\Psi^{-1}(1/(t+h))} \right| \|u\|_{L^p} \, h^{\frac{p-1}{p}} \end{aligned}$$

which, when $h \to 0^+$, converges to 0 uniformly with respect to $\widetilde{\mathrm{E}}$ due to the continuity of $\dfrac{1}{t\,\Psi^{-1}(1/t)}$.

Hence, with these operators we can describe $L^\Psi(Q_0)$ and $M^\Psi(Q_0)$ as

$$\begin{aligned} L^\Psi(Q_0) &= \left\{ u \in L^p : \sup_{L \in \mathscr{L}} \|Lu\|_Y < +\infty \right\} \\ M^\Psi(Q_0) &= \left\{ u \in L^p : \limsup_{L \to \infty} \|Lu\|_Y = 0 \right\}. \end{aligned}$$

Now consider $u \in L^\Psi(Q_0) \subseteq L^p(Q_0)$. Recall that for any positive function in $L^p(Q_0)$ there exists an increasing sequence of $L^\infty(Q_0)$ functions converging to it in $L^p(Q_0)$. Hence let us consider u_j^+ and u_j^- converging respectively to u^+ and u^-.

Thus $u_j = u_j^+ - u_j^-$ converges in $L^p(Q_0)$ to u and $|u_j|$ is an increasing sequence converging to $|u|$. Hence $|u_j| \leq |u|$ and by monotonicity of the norm $|||\cdot|||_\Psi$

$$|||u_j|||_\Psi \leq |||u|||_\Psi .$$

Finally taking the supremum over j we have

$$\sup_{j \in \mathbb{N}} |||u_j|||_\Psi \leq |||u|||_\Psi .$$

□

Let us now see some examples where our theory gives meaningful results.

First of all, it is well known that the space $EXP(Q_0)$, which is an Orlicz space with $\Psi(t) = e^t - 1$, admits $|||\cdot|||_\Psi$ as equivalent norm (see [25]). Indeed it is easy to check that $\Psi^{-1}(1/t) = \log(1 + (1/t)) \in L^\Psi(I_0)$. Moreover, it is also easy to check that Ψ satisfies the Δ^0 condition.

Let us now show a more general result.

Proposition 3.14. *Let Ψ be a Young function of the form $e^{\nu(t)} - 1$ such that ν is a convex function.*

Then Ψ satisfies the Δ^0 condition and $\Psi^{-1}(1/t) \in L^{\Psi}(I_0)$.

Proof. The Δ^0 condition easily follows from the inequality

$$\Psi(kt) = e^{\nu(kt)} - 1 \geq e^{k\nu(t)} - 1$$

for any $k > 1$ and $\nu(t) \to +\infty$ as $t \to +\infty$.
Now observe that, being Ψ a Young function, $\nu(0) = 0$ and ν is strictly increasing. Moreover for any $r > 0$

$$e^{\nu(r/2)} \leq \sqrt{e^{\nu(r)}} \leq \frac{e^{\nu(r)}}{\sqrt{e^{\nu(r)} - 1}},$$

since $\nu(r/2) \leq \nu(r)/2$. If $r > 0$ is such that $\nu'(r)$ and $\nu'(r/2)$ exists, then we also have $\nu(r/2) \leq \nu'(r)$. Combining these two inequalities we have

$$\frac{1}{2}\nu'\left(\frac{r}{2}\right)e^{\nu(r/2)} \leq \frac{1}{2}\frac{\nu'(r)e^{\nu(r)}}{\sqrt{e^{\nu(r)} - 1}}.$$

Integrating this inequality in the interval $[0, r]$, applying first the logarithm and then ν^{-1} we obtain

$$\frac{r}{2} \leq \nu^{-1}\left(\log\left(\sqrt{e^{\nu(r)} - 1} + 1\right)\right).$$

Now, for any $t > 0$, consider $r = \Psi^{-1}(1/t) = \nu^{-1}(\log(1 + 1/t))$ to obtain

$$\frac{1}{2}\Psi^{-1}\left(\frac{1}{t}\right) \leq \nu^{-1}\left(\log\left(\sqrt{1 + \frac{1}{t}}\right)\right) = \Psi^{-1}\left(\sqrt{\frac{1}{t}}\right)$$

and then applying Ψ we finally have

$$\Psi\left(\frac{1}{2}\Psi^{-1}\left(\frac{1}{t}\right)\right) \leq \sqrt{\frac{1}{t}}$$

hence $t \mapsto \Psi^{-1}(1/t) \in L^{\Psi}(0, 1)$.

□

Observe that $\Psi(t) = e^t - 1$ (which is the Young function defining EXP) fits in this case, together with functions like $\Psi(t) = e^{t^\alpha} - 1$ for $\alpha > 1$, $\Psi(t) = e^{e^t - 1} - 1$, $\Psi(t) = e^{e^{e^t - 1} - 1} - 1$ and so on.
However, Young functions of the form $\Psi(t) = e^{\log^{1+\varepsilon}(t+1)} - 1$ for $\varepsilon > 0$ do not fit in the previous cases. For these we can show the following proposition.

Proposition 3.15. *Let Ψ be the Young function $\Psi(t) = e^{\log^{1+\varepsilon}(1+t)} - 1$, with $\varepsilon > 0$. Then Ψ satisfies the Δ^0 condition and $\Psi^{-1}(1/t) \in L^{\Psi}(I_0)$.*

Proof. It is easy to prove that Ψ satisfies the Δ^0 conditon. To study this case, one can consider the change of variable $t = 1/s$, by which the condition $\Psi^{-1}(1/t) \in L^{\Psi}(I_0)$ is shown to be equivalent to the existence of a $k > 1$ such that

$$\int_0^1 \Psi\left(\frac{1}{k}\Psi^{-1}\left(\frac{1}{s}\right)\right) ds = \int_1^{+\infty} \frac{1}{t^2}\Psi\left(\frac{1}{k}\Psi^{-1}(t)\right) dt < +\infty.$$

Let us study the case in which $\varepsilon \in (0,1)$. We have

$$\Psi^{-1}(t) = e^{\log^{\frac{1}{1+\varepsilon}}(1+t)} - 1$$

and then, after some calculations

$$\begin{aligned}\Psi\left(\frac{1}{k}\,\Psi^{-1}(t)\right) &= \exp\left(\log^{1+\varepsilon}\left(1 + \frac{1}{k}\,e^{\log^{\frac{1}{1+\varepsilon}}(1+t)} - \frac{1}{k}\right)\right) - 1 \\ &\simeq \frac{t}{e^{(1+\varepsilon)\log^{\frac{\varepsilon}{1+\varepsilon}} t}},\end{aligned}$$

where the symbol $\simeq$ means we are considering functions with the same asymptotic behaviour at $+\infty$. This asymptotic equivalence follows from Taylor expansion of $t \mapsto (1+t)^{1+\varepsilon}$ near 0 up to the second order term.
Since for t large enough

$$e^{(1+\varepsilon)\log^{\frac{\varepsilon}{1+\varepsilon}} t} \geq \log^p t$$

for any $p > 1$, we obtain

$$\int_1^{+\infty} \frac{1}{t^2}\,\Psi\left(\frac{1}{k}\,\Psi^{-1}(t)\right) dt < +\infty.$$

A similar argument works for $\varepsilon > 1$ by considering a Taylor expansion of higher order.

$\square$

3.2 Orlicz-Sobolev Spaces as defined by means of oscillation

Let Ψ be any Young function and $L^{\Psi}(\Omega)$ the corresponding induced Orlicz space over an open subset Ω of an Euclidean space $\mathbb{R}^n$. We will say that a function $f \in L^{\Psi}(\Omega)$ is in the Orlicz-Sobolev space $W^{1,\Psi}(\Omega)$ if there is a function $g \in L^{\Psi}(\Omega; \mathbb{R}^n)$ acting as weak gradient of f, i.e.

$$\int_{\Omega} f(x) D\varphi(x)\, dx = -\int_{\Omega} g(x)\varphi(x)\, dx, \quad \forall \varphi \in C_c^{\infty}(\Omega),$$

where by $C_c^{\infty}(\Omega)$ we denote the space of test functions, i.e. indefinitely differentiabile functions with compact support, but it can be proven by density arguments that the role of $C_c^{\infty}(\Omega)$ can be equivalently played, a posteriori, by any function in a suitable Orlicz-Sobolev space $W^{1,\Phi}$, with Φ being the conjugate of Ψ.
Exactly in the same way as Lebesgue spaces are what one obtains by choosing $\Psi_p(t) = t^p$ and looking at the induced Orlicz space L^{Ψ}, classical Sobolev spaces are a special case of Orlicz-Sobolev spaces. The fact that all Lebesgue spaces are rearrangement-invariant extends to all Orlicz spaces while on the other hand examples to show that $W^{1,\Psi}$ is not rearrangement invariant can be found for all choices of a Young function Ψ.
In this section we will present a recent result by Heli Tuominen (see [134]) characterizing Orlicz-Sobolev spaces in terms of oscillation. To understand the result a few definitions are needed. We already looked at the Hardy-Littlewood maximal operator M: we now define a restricted version.

Definition 3.8. Let $f(x)$ be any locally integrable function, $f \in L^1_{loc}(\mathbb{R}^n)$, and $R > 0$. The restricted Hardy-Littlewood maximal operator M_R maps f to a function $M_R f$ defined by the position

$$M_R f(x) = \sup_{R > r > 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)|\, dy$$

where $B_r(x)$ is a ball of center x and radius r.

In [81], citing, in turn, some results from [29], Hajlasz remarked that a function $f \in L^p$ is in the Sobolev space $W^{1,p}$ if and only if

$$|f(x) - f(y)| \leq C(n)|x - y| \Big[M_{2|x-y|}(|Df|)(x) + M_{2|x-y|}(|Df|)(x) \Big], \quad \text{for a.e. } x, y.$$

As the restricted Hardy-Littlewood maximal operator preserves the p integrability of the function if $p > 1$, this can be expressed more neatly in other words: a function $f \in L^p$ (with $p > 1$) is in the Sobolev space $W^{1,p}$ if and only if there is another function $g \in L^p$ bounding the Lipschitz oscillation in the following way

$$\frac{|f(x) - f(y)|}{|x - y|} \leq g(x) + g(y),$$

for almost every $x, y \in \mathbb{R}^n$ with $x \neq y$.

This result is extremely interesting as it provides a very natural way of extending the definition of Sobolev spaces to compact metric measure spaces, by substituting $|x-y|$ for another metric $\rho(x,y)$. This definition, first proposed by Hajlasz, of a Sobolev space for functions of metric measure space, is at the core of [81]: provided a few reasonable assumptions on behaviour of the measure on balls, but we do not go into further technical detail, it gives rise to a function space for which many properties of the Sobolev space are preserved. This function spaces are often referred to as Hajlasz-Sobolev spaces and denoted by $M^{1,p}(K,\rho,\mu)$. The reader is invited to compare with Section 2.2
What Tuominen did in [134] was precisely an extension of this result to Orlicz-Sobolev spaces. He proved the following theorem.

Theorem 3.16. *Let Ψ be a Young function and Φ be its conjugate and assume Ψ and Φ both satisfy the Δ_2 condition. Then a function $f \in L^\Psi(\mathbb{R}^n)$ is in the Orlicz-Sobolev space $W^{1,\Psi}(\mathbb{R}^n)$ if and only if there is a function $0 \leq g \in L^\Psi$, real constants $C > 0$ and $\sigma > 1$ and a null set E, $|E| = 0$, such that for all $x, y \in \mathbb{R}^n \setminus E$, the following inequality holds*

$$|f(x) - f(y)| \leq C|x-y|\Big[\Psi^{-1}(M_{\sigma|x-y|}\Psi(g)(x)) + \Psi^{-1}(M_{\sigma|x-y|}\Psi(g)(y))\Big].$$

Sadly this result cannot be expressed as neatly as the one by Hajlasz as the one in $W^{1,p}$ with $p > 1$, because $\Psi(g)$ is a function in L^1, and the Hardy-Littlewood maximal operator is not expected to preserve simple summability ($p = 1$), so that we cannot rewrite the result substituting $\Psi^{-1}(M_{\sigma|x-y|}\Psi(g)(x))$ for a generic function g in L^Ψ, but this result further motivates and justifies the introduction of Orlicz-Sobolev spaces even in the context of compact metric measure spaces: a construction that was done by Aïssaoui in [5] at about the same time as the paper by Tuominen appeared. These spaces are often referred to as Musielak-Orlicz-Sobolev spaces and generalize the Hajlasz-Sobolev spaces by imposing conditions of type:

$$\frac{|f(x) - f(y)|}{\rho(x,y)} \leq g(x) + g(y)$$

for almost all $x \neq y$ in the metric space and for L^Ψ functions f and g.

3.3 Orlicz and Hölder spaces in regularity theory

Hölder spaces are absolutely pervasive in regularity theory, starting from an extremely famous result by Ennio de Giorgi (see [55]), proving Hölder regularity for solutions of elliptic equations in divergence form with measurable bounded coefficients. As a matter of fact, this result is immediately traslated into Hölder regularity for gradients of minima for strictly convex integral functionals on $W^{1,p}_{loc}(\Omega, \mathbb{R})$.
It is known that this, in turn, is the starting point of a bootstrap argument proving higher regularity for minima of functionals with more regular lagrangian functions, hence giving a positive answer to the famous XIX Hilbert problem.
Such a clean result, however, is only possible in the scalar case. Ennio de Giorgi himself found a counterexample in the vectorial case, i.e. a functional

$$\mathscr{F}(u) = \int_\Omega f(x, Du(x))\, dx$$

defined on $W^{1,2}(\Omega, \mathbb{R}^N)$ and a minimum $u : \Omega \to \mathbb{R}^N$ such that, for any choice of $N > 2$, u fails to have gradient Du belonging to the Hölder space $\mathrm{Lip}_\alpha(\Omega)$ for any choice of $\alpha \in (0, 1)$.
On the other hand, in the multidimensional case $N > 1$, it is often possible to prove partial regularity, i.e. pointwise regularity on a set of points whose complement is negligible, in some sense. Some assumption on f under which this is usually done is an integral form of convexity called $W^{1,p}$-quasiconvexity. Whenever the functional $\mathscr{F}$ is autonomous, i.e. not depending directly from x (and, in this section we will even assume $f : z \in \mathbb{R}^{nN} \to f(z) \in \mathbb{R}$ depends exclusively on the gradient Du), $W^{1,p}$-quasiconvexity is the assumption that the lagrangian f satisfies the inequality

$$\fint_{B_1} f(z + D\phi(x))\, dx \geq f(z), \quad \forall z, \forall \phi \in C_0^\infty(B_1, \mathbb{R}^N).$$

This condition was introduced by Morrey in 1951 (see [117]) and it was shown to be a necessary and sufficient condition for the weak lower semicontinuity of the functional $\mathscr{F}$ in $W^{1,p}$. It is a condition strong enough to imply that affine functions are local minima. Stronger forms of quasiconvexity were later introduced, such as uniform or strict quasiconvexity, whose definition we will see later. In light of this, we mention the first regularity result for minima of strictly quasiconvex functionals due to Evans (see [62])
Another idea that is central to this section is the concept of (p, q) growth: we will give definitions further in the section, but we also refer the reader to [112] by Paolo Marcellini.
As a matter of fact, classically (see [62], [2] and [41]), $W^{1,p}$-quasiconvexity is paired with an assumption on the growth of f, of the type $|f(z)| \leq C(1 + |z|)^p$.
Marcellini (see [113]) replaced it with the more flexible (p, q) growth conditions, by considering a Lagrangian function f that is $W^{1,p}$-quasiconvex but satisfies $|f(z)| \leq C(1 + |z|)^q$ for different p and q. In this context he proved many regularity results and subsequently, several contributions were added to the theory by various authors.
In [43] and in [92], autonomous functionals $\mathscr{F}(u) = \int f(Du)$ of the calculus of variations are studied, with the hypothesis of quasiconvexity for the Lagrangian function $f = f(z)$

only holding asymptotically, i.e. for $z : |z| > M$. The idea is to show that the typical methods used to prove density of the set $Reg(u)$, i.e. the set of points of local Hölderianity for the gradient Du of a minimizer u, can be adapted to prove that the set of points x_0 in which $|Du|(x_0)$ is large and u is not locally Lipschitz around x_0 is no-where dense.
Thinking of Lipschitz functions as functions in $W^{1,\infty}$, i.e. admitting an essentially bounded weak derivative, allows to conclude that u is at least locally Lipschitz in a dense subset of its domain.
This section is devoted to a similar problem, but for functionals with a (φ, ψ) structure, where φ and ψ are more generally Young functions with the so-called Δ_2 condition that play the role that $t \mapsto t^p$ and $t \mapsto t^q$ play in the usual (p, q) theory. More precisely, we will assume that the lagrangian f does not grow faster than ψ, but is asymptotically φ-quasiconvex for different Young functions φ and ψ (see further in the section for a precise formulation of the problem).
The problem is approached with inspiration to some methods used by T. Schmidt in his paper [131], but also looking at [92] and adapting those leading to the announced conclusion.
In general, a very good survey of the topic was done by Mingione in [115].
Lastly, with respect to Hölder spaces, we remark that, in this context, it is also very useful to rethink of them via Campanato's integral characterization. As a matter of fact, in [35], Campanato introduced the spaces

$$L^{\lambda,p}(\Omega) := \left\{ u : [u]^p_{\lambda,p} := \sup_{B_r(x_0)\subseteq\Omega} \frac{1}{r^\lambda} \int_{B_r(x_0)} |u(x) - u_{B_r(x_0)}|\, dx < +\infty \right\}$$

sometimes called Campanato or Morrey-Campanato spaces. He then proved that whenever λ in $(n, n+p)$ these spaces coincide with Lip_α, with $\alpha = \frac{\lambda - n}{p}$.

3.3.1 The problem

In this section we study multidimensional integrals of the type

$$\mathscr{F}(u) = \int_\Omega f(Du(x))\, dx \qquad \text{for } u : \Omega \to \mathbb{R}^N$$

where Ω is an open bounded set in $\mathbb{R}^n$, $n \geq 2$, $N \geq 1$. We consider Young functions φ, $\psi \in C^2([0, +\infty))$ and a Lagrangian function f s.t. the following assumptions hold:

(H.1) **Regularity**- $f \in C^2(\mathbb{R}^{nN}, \mathbb{R})$.

(H.2) The following inequalities hold

$$p_\varphi \frac{\varphi'(t)}{t} \leq \varphi''(t) \leq q_\varphi \frac{\varphi'(t)}{t}$$

$$p_\psi \frac{\psi'(t)}{t} \leq \psi''(t) \leq q_\psi \frac{\psi'(t)}{t}$$

for all $t \geq 0$, where $p_\varphi, q_\varphi, p_\psi, q_\psi > 0$ are positive constants.

(H.3) **Asymptotical $W^{1,\varphi}$-quasiconvexity**- There exists $M >> 0$, $\gamma > 0$ and a continuous function g such that

$$f(z) = g(z), \quad \forall z: |z| > M$$

and such that g is strictly $W^{1,\varphi}$-quasiconvex, i.e. satisfies

$$\fint_{B_1} g(z + D\phi) \geq g(z) + \gamma \fint_{B_1} \varphi_{1+|z|}(|D\phi|), \quad \forall z, \forall \phi \in C_0^\infty(B_1, \mathbb{R}^N)$$

where $\varphi_a(t)$ is defined for any $0 < a \in \mathbb{R}$ via the following equality

$$\varphi_a'(t) = \varphi'(a+t)\frac{t}{a+t} \quad \text{and} \quad \varphi_a(0) = 0$$

and it was shown in [92] to satisfy the property

$$\varphi_a(t) \sim t^2 \varphi''(a+t).$$

(H.4) **Growth conditions**- The following inequalities hold

$$\Gamma'\varphi(|z|) \leq f(z) \leq \Gamma''(1 + \psi(|z|))$$

$$|D^2 f(z)| \leq \Gamma''(1 + \psi''(|z|))$$

for all $z \in \mathbb{R}^{nN}$ for some positive constants $\Gamma', \Gamma'' > 0$.

(H.5) **Range of anisotropy** We assume that, for any $a > M$, the function $\mathcal{N}_a = \varphi_a \circ (\psi_a')^{-1}$ is a Young function and the following inequality regarding its complementary Young function $\mathcal{N}_a^*$ holds

$$[\mathcal{N}_a]^*(t) \leq c\varphi_a^\beta(t)$$

for all $t >> 1$ and some $1 \leq \beta < \frac{n}{n-1}$

It can be shown (see Lemma 3.24 in the next subsection) that $(H.5)$ implies the inequality

$$\psi(t) \leq c\varphi^\beta(t), \quad \forall t >> 1$$

. If $\varphi(t) = t^p$ and $\psi(t) = t^q$, $(H.5)$ is equivalent to $q < p + \frac{1}{n}$.
In [92], it was proven that $(H.3)$ is equivalent to:

(H.3') There exists $M >> 0$, $\gamma > 0$ such that

$$\forall z: |z| > M$$

$$\fint_{B_1} f(z + D\phi) \geq f(z) + \gamma \fint_{B_1} \varphi_{1+|z|}(|D\phi|), \quad \forall \phi \in C_0^\infty(B_1, \mathbb{R}^N)$$

under our hypotheses. In particular, it follows from the fact that f is locally bounded from below.

We will study local $W^{1,\varphi}$-minimizers of $\mathscr{F}$, i.e. functions u such that

$$\mathscr{F}(u) \leq \mathscr{F}(u+\phi) \quad \forall \phi \in W_0^{1,\varphi}(\Omega, \mathbb{R}^N).$$

In the case of a globally quasiconvex functional and with the superquadratic hypothesis $\varphi(t) > a(t^2-1)$, D. Breit and A. Verde in [33] proved that if u is a local minimizer of $\mathscr{F}$, then u is in $C^{1,\alpha}$ in a open dense subset of Ω. Adapting and generalizing their arguments, we prove partial $C^{1,\alpha}$ regularity of a local minimizer u of $\mathscr{F}$ around points in whose neighbourhoods $|Du|$ is large, without the assumption of superquadraticity and with quasiconvexity only in its asymptotical formulation. More precisely, we will obtain the following result

Theorem 3.17 (Angrisani, 2021, preprint). *Let f, φ, ψ satisfy hypotheses $(H.1)$,$(H.2)$,$(H.3)$,$(H.4)$ and $(H.5)$ and let u be a local minimizer of the corresponding functional $\mathscr{F}$. Let $z_0 \in \mathbb{R}^{nN}$ such that $|z_0| > M+1$ and assume there is a $x_0 \in \mathbb{R}^n$ with the property that*

$$\lim_{\rho \to 0^+} \fint_{B_\rho(x_0)} |V(Du(x)) - V(z_0)|^2 = 0,$$

then u is locally C^1 in a neighbourhood of x_0, where $V(z)$ is defined in section 3.3.2.

This theorem has the following corollary

Corollary 3.18. *In the hypotheses and notation of Theorem 3.17, the set of points of local Lipschitzianity of u is a dense open subset of Ω.*

3.3.2 Technical lemmas and definitions

The following lemma will be useful while obtaining our Caccioppoli estimate. The proof can be found in [69].

Lemma 3.19. *Let $-\infty < r < s < +\infty$ and a continuous nondecreasing function $\Xi : [r, s] \to \mathbb{R}$ be given. Then there are $\tilde{r} \in [r, \frac{2r+s}{3}]$ and $\tilde{s} \in [\frac{r+2s}{3}, s]$, for which hold:*

$$\frac{\Xi(t) - \Xi(\tilde{r})}{t - \tilde{r}} \leq 3 \frac{\Xi(s) - \Xi(r)}{s - r}$$

and

$$\frac{\Xi(\tilde{s}) - \Xi(t)}{\tilde{s} - t} \leq 3 \frac{\Xi(s) - \Xi(r)}{s - r}$$

for every $t \in (\tilde{r}, \tilde{s})$.
In particular, we have $\frac{s-r}{3} \leq \tilde{s} - \tilde{r} \leq s - r$.

The following lemma is about Young functions satisfying hypothesis $(H.2)$

Lemma 3.20. *Let h be a Young function satisfying $(H.2)$. Then we have*

(a) h satisfies $\Delta_2(h) < +\infty$ and $\Delta_2(h^*) < +\infty$

(b) For all $t > 0$ the following inequality holds:

$$h(1)(t^p - 1) \le h(t) \le h(1)(t^q + 1)$$

where $p = p_\varphi + 1$ and $q = q_\varphi + 1$ if $h = \varphi$ and $p = p_\psi + 1$ and $q = q_\psi + 1$ if $h = \psi$

(c) For all $t > 0$, $h'(t)t$ is equivalent to $h(t)$

For the proof, see Lemma 3.1 in [68].
We will also make use of Lemma 3.2 from [92], which we rewrite here for the convenience of the reader.

Lemma 3.21. *Let φ be a Young function such that φ and φ^* both enjoy the Δ_2 condition. For $z_1, z_2 \in \mathbb{R}^{nN}$ and $\theta \in [0,1]$, denote $z_\theta = z_1 + \theta(z_2 - z_1)$. Then, uniformly in $z_1, z_2 \in \mathbb{R}^{nN}$ with $|z_1| + |z_2| > 0$ and in $\mu > 0$, it holds*

$$\frac{\varphi'(\mu + |z_1| + |z_2|)}{\mu + |z_1| + |z_2|} \simeq \int_0^1 \frac{\varphi'(\mu + |z_\theta|)}{\mu + |z_\theta|} d\theta$$

Definition 3.9 (Excess)**.** For any $z \in \mathbb{R}^{nN}$ let us define the quantity

$$V(z) := \sqrt{\frac{\varphi'(|z|)}{|z|}}z$$

and let us notice that, with our hypotheses, we have

$$|V(z_1) - V(z_2)|^2 \simeq \varphi_{|z_1|}(|z_1 - z_2|).$$

We also define the excess function

$$\Phi_\varphi(u, x_0, \rho, z) := \fint_{B_\rho(x_0)} |V(Du) - V(z)|^2 \, dx$$

and

$$\Phi_\varphi(u, x_0, \rho) := \fint_{B_\rho(x_0)} |V(Du) - V[(Du)_{B_\rho(x_0)}]|^2 \, dx$$

where by putting a set as a pedix to a function we refer to the integral average of the function over the set, i.e. $(Du)_{B_\rho(x_0)} = \fint_{B_\rho(x_0)} Du(x)\, dx$. We immediately notice that

$$\Phi_\varphi(u, x_0, \rho, z) \simeq \fint \varphi_{|z|}(|Du - z|)\, dx$$

In [33], D. Breit and A. Verde proved that if u is a $W^{1,\varphi}$ minimizer of $\mathscr{F}$ on $B_\rho(x_0)$, for all $L > 0$ and $\alpha \in (0,1)$ there exists $\varepsilon_0 > 0$ such that if

$$\Phi_\varphi(u, x_0, \rho) \le \varepsilon_0 \quad \text{and} \quad \left| \fint_{B_\rho(x_0)} Du \right| \le \frac{L}{2}$$

then $u \in C^{1,\alpha}_{loc}(B_\rho(x_0); \mathbb{R}^n)$.

We will now replicate their reasoning in the weaker hypothesis of asymptotic φ-quasiconvexity and without assumption of superquadratic growth behaviour of φ. With this aim, we start with the following lemmas.

Lemma 3.22. *If there exists z_0, $|z_0| > M+1$ and x_0 such that:*

$$\fint_{B_\rho(x_0)} |V(Du) - V(z_0)|^2 \to 0 \text{ as } \rho \to 0^+$$

then there exists $r_1 = r_1(x_0, z_0)$such that for all $r < r_1$

$$\left| \fint_{B_r(x_0)} Du \right| > M+1.$$

Proof. Let $|z_0| = M+1+\varepsilon$. Then there must be a r_1 (of course this depends on the specific values of x_0 and z_0) such that for all $r < r_1$ we have:

$$\fint_{B_r(x_0)} \varphi_{|z_0|}(|Du - z_0|) \le \varphi_{|z_0|}\left(\frac{\varepsilon}{2}\right), \quad \forall r < r_1$$

where we have also made use of our previous remark that $|V(Du) - V(z_0)|^2$ is equivalent to $\varphi_{|z_0|}(|Du - z_0|)$.
which, by Jensen inequality means:

$$\left| \fint_{B_r(x_0)} Du - z_0 \right| \le \frac{\varepsilon}{2}, \quad \forall r < r_1.$$

This in turn gives:

$$\left| \fint_{B_r(x_0)} Du \right| \ge |z_0| - \frac{\varepsilon}{2} = M+1+\varepsilon - \frac{\varepsilon}{2} > M+1, \quad \forall r < r_1.$$

□

The following is a rewriting of Lemma 2.5 from [33] that will be useful to us.

Lemma 3.23. *Let $0 < r < s$ and $\beta < \frac{n}{n-1}$. Then there exists a linear operator*

$$T_{r,s} : W^{1,\varphi}(\Omega, \mathbb{R}^N) \to W^{1,\varphi}(\Omega, \mathbb{R}^N)$$

defined as

$$T_{r,s}u(x) = \fint_{B_1(0)} u(x + \xi(x)y)\, dy, \qquad \text{with } \xi(x) := \frac{\max\{0, \min\{|x| - r, s - |x|\}\}}{2}$$

such that

(a) $T_{r,s}u = u$ *on* B_r *and outside* $\overline{B}_s$;

(b) $T_{r,s}u \in u + W_0^{1,\varphi}(B_s \setminus \overline{B}_r, \mathbb{R}^n)$;

(c) $|DT_{r,s}u| \leq cT_{r,s}|Du|$

(d) The following estimates holds:

$$\int_{B_s \setminus B_r} \varphi(|T_{r,s}u|)\,dx \leq c \int_{B_s \setminus B_r} \varphi(|u|)\,dx$$

$$\int_{B_s \setminus B_r} \varphi(|DT_{r,s}u|)\,dx \leq c \int_{B_s \setminus B_r} \varphi(|Du|)\,dx$$

$$\int_{B_s \setminus B_r} \varphi^\beta(|T_{r,s}u|)\,dx \leq c(s-r)^{-n\beta+n+\beta} \left[\sup_{r \leq t \leq s} \frac{\theta(t)-\theta(r)}{t-r} + \sup_{r \leq t \leq s} \frac{\theta(s)-\theta(t)}{s-t} \right]^\beta$$

$$\int_{B_s \setminus B_r} \varphi^\beta(|DT_{r,s}u|)\,dx \leq c(s-r)^{-n\beta+n+\beta} \left[\sup_{r \leq t \leq s} \frac{\Theta(t)-\Theta(r)}{t-r} + \sup_{r \leq t \leq s} \frac{\Theta(s)-\Theta(t)}{s-t} \right]^\beta$$

where

$$\theta(t) := \int_{B_t} \varphi(|u|)\,dx, \quad \Theta(t) := \int_{B_t} \varphi(|Du|)\,dx.$$

The following result is contained in [34].

Lemma 3.24. *Choose any positive constant L larger than M and let $a \in (M, L)$. Then for any $t > 0$ we have*

$$\psi_a(t) \leq K \cdot H(\varphi_a(t))$$

where $K = K(M, L, \beta, \psi, \varphi)$ is a positive real constant depending on L and $H(t) := t + t^\beta$.

Proof. Let us start with the case $t \leq 1$. In this case, by $(H.2)$,

$$\psi_a(t) \simeq \psi''(a+t)t^2 \simeq \psi(a+t)\frac{t^2}{(a+t)^2} \leq \max_{[M,L+1]} \psi \cdot \frac{t^2}{(a+t)^2} \leq$$
$$\leq K_1 \min_{[M,L+1]} \varphi \cdot \frac{t^2}{(a+t)^2} \leq K_1 \varphi(a+t)\frac{t^2}{(a+t)^2} \simeq K_1 \varphi_a(t)$$

where

$$K_1 = \frac{\max_{[M,L+1]} \psi}{\min_{[M,L+1]} \varphi} \in (0, +\infty)$$

depends only on M, L, ψ and φ.
On the other hand, if $t > 1$,

$$\psi_a(t) \simeq \psi''(a+t)t^2 \simeq \psi(a+t)\frac{t^2}{(a+t)^2} \leq \varphi^\beta(a+t)[\frac{t^2}{(a+t)^2}]^\beta \cdot \left(1 + \frac{a}{t}\right)^{2\beta-2} \leq$$
$$\leq K_2 \left[\frac{\varphi(a+t)t^2}{(a+t)^2} \right]^\beta \simeq K_2 \varphi_a(t)^\beta$$

where $K_2 = (1+L)^{2\beta-2}$.
The thesis follows with $K = \max\{K_1, K_2\}$. $\square$

3.3.3 Caccioppoli Inequality

This subsection is dedicated to the Caccioppoli inequality, which will be the main tool of the proof of partial regularity.

Lemma 3.25 (Angrisani, 2021, preprint). *Let the assumptions $(H.1)-(H.5)$ hold for a given M. Consider any positive constant $L > M > 0$ and a consider $W^{1,\varphi}$-minimizer $u \in W^{1,\varphi}(B_\rho(x_0); \mathbb{R}^N)$ of $\mathscr{F}$ on a ball $B_\rho(x_0)$ contained in Ω. For all $z \in \mathbb{R}^{nN}$ with $M < |z| < L+1$, let $q(x)$ be an affine function with gradient z and $v(x) = u(x) - q(x)$. Then*

$$\fint_{B_{\frac{\rho}{2}}} \varphi_{|z|}(|Dv|)\,dx \leq c \fint_{B_\rho} \varphi_{|z|}\left(\frac{|v|}{\rho}\right) dx+$$

$$+ c\left\{\fint_{B_\rho} \left[\varphi_{|z|}(|Dv|) + \varphi_{|z|}\left(\frac{|v|}{\rho}\right)\right] dx\right\}^\beta . \tag{3.6}$$

Proof. Assume for simplicity $x_0 = 0$ and choose

$$\frac{\rho}{2} \leq r < s \leq \rho.$$

Define

$$\Xi(t) := \int_{B_t} \left[\varphi_{|z|}(|Dv|) + \varphi_{|z|}\left(\left|\frac{v}{\tilde{s}-\tilde{r}}\right|\right)\right] dx.$$

We choose in addition $r \leq \tilde{r} < \tilde{s} \leq s$ as in Lemma 3.19. Let η denote a smooth cut-off functions with support in $B_{\tilde{s}}$ satisfying $\eta \equiv 1$ in $\overline{B_{\tilde{r}}}$ and $0 \leq \eta \leq 1$, $|\nabla\eta| \leq \frac{2}{\tilde{s}-\tilde{r}}$ on B_ρ. Using the operator from Lemma 3.23, we set

$$\zeta := T_{\tilde{r},\tilde{s}}[(1-\eta)v] \text{ and } \xi := v - \zeta.$$

By quasiconvexity we have

$$\gamma \int_{B_{\tilde{s}}} \varphi_{|z|}(|D\xi|) \leq \int_{B_{\tilde{s}}} f(z + D\xi) - f(z) =$$

$$= \int_{B_{\tilde{s}}} f(z + D\xi) - f(Du) + f(Du) - f(Du - D\xi) + f(Du - D\xi) - f(z) \leq \dots$$

Now, of course $\int f(Du) - f(Du - D\xi) \leq 0$ and $Du = z + D\xi + D\zeta$, so

$$\dots \leq \int_{B_{\tilde{s}}} f(z + D\xi) - f(z + D\xi + D\zeta) + \int_{B_{\tilde{s}}} f(z + D\xi) - f(z) \leq$$

$$\leq \int_{B_{\tilde{s}}} \int_0^1 |Df(z + D\xi + \theta D\zeta) - Df(z)||D\zeta|\,d\theta, dx+$$

$$+ \int_{B_{\tilde{s}}} \int_0^1 |Df(z + \theta D\zeta) - Df(z)||D\zeta|\,d\theta, dx =: \mathscr{I}_1 + \mathscr{I}_2.$$

We now estimate $\mathscr{I}_1$, recalling our growth hypotheses $(H.4)$ and also making use of

Lemma 3.21. We deduce:

$$\begin{aligned}\mathscr{I}_1 \leq \int_{B_{\tilde{s}}} \int_0^1 \int_0^1 |D^2 f(tz + (1-t)(z+\theta D\zeta))||\theta D\zeta|D\zeta|\, dt\, d\theta, dx \leq \\ \leq \Gamma'' \int_{B_{\tilde{s}}} \int_0^1 \int_0^1 |\psi''(|tz + (1-t)(z+\theta D\zeta|))||D\zeta|^2\, dt\, d\theta, dx \leq \\ \leq c \int_{B_{\tilde{s}}} \frac{\psi'(2|z| + |z + D\zeta|)}{2|z| + |z+D\zeta|} |D\zeta|^2\, dx \leq \\ \leq c \int_{B_{\tilde{s}}} \psi_{|z|}(|D\zeta|)\, dx.\end{aligned}$$

Regarding $\mathscr{I}_2$ we can deduce:

$$\begin{aligned}\mathscr{I}_2 \leq \int_{B_{\tilde{s}}} \int_0^1 \int_0^1 |D^2 f(t(z + D\xi + \theta D\zeta) + (1-t)z)||D\xi + \theta D\zeta||D\zeta|\, dt\, d\theta\, dx \leq \\ \leq c \int_{B_{\tilde{s}}} \int_0^1 \int_0^1 \psi''(|t(z + D\xi + \theta D\zeta) + (1-t)z|)|D\xi + \theta D\zeta||D\zeta|\, dt\, d\theta\, dx \leq \\ \leq c \int_{B_{\tilde{s}}} \psi''(|z| + |D\xi| + |D\zeta|)(|D\xi| + |D\zeta|)|D\zeta|\, dx \leq \\ \leq \int_{B_{\tilde{s}}} \psi'_{|z|}(|D\xi| + |D\zeta|)|D\zeta|\, dx \leq \\ \leq c \int_{B_{\tilde{s}}} \psi'_{|z|}(|D\xi|)|D\zeta| + c \int_{B_{\tilde{s}}} \psi'_{|z|}(|D\zeta|)|D\zeta| \leq \\ \leq c \int_{B_{\tilde{s}}} \psi'_{|z|}(|D\xi|)|D\zeta| + c \int_{B_{\tilde{s}}} \psi_{|z|}(|D\zeta|)\, dx\end{aligned}$$

Combining the two, we obtain

$$\gamma \int_{B_{\tilde{s}}} \varphi_{|z|}(|D\xi|) \leq c \int_{B_{\tilde{s}}} \psi'_{|z|}(|D\xi|)|D\zeta| + c \int_{B_{\tilde{s}} \setminus B_{\tilde{r}}} \psi_{|z|}(|D\zeta|)\, dx.$$

Now we use a Young inequality on $\psi'_{|z|}(|D\xi)|D\zeta|$, namely

$$\psi'_{|z|} \leq c[\mathscr{N}_{|z|}(\psi'_{|z|}(|D\xi)) + \mathscr{N}^*_{|z|}(|D\zeta|)].$$

This, together with the definition of $\mathscr{N}_{|z|}$, our anisotropy assumption $(H.5)$ and Lemma 3.24 lets us deduce

$$\begin{aligned}\gamma \int_{B_{\tilde{s}}} \varphi_{|z|}(|D\xi|) \leq c \int_{B_{\tilde{s}}} H[\varphi_{|z|}(|D\zeta|)]\, dx + \\ + c \left[\int_{B_{\tilde{s}} \setminus B_{\tilde{r}}} \varphi_{|z|}(|D\xi|)\, dx + \int_{B_{\tilde{s}} \setminus B_{\tilde{r}}} \varphi^{\beta}_{|z|}(|D\zeta|)\, dx \right] \leq\end{aligned}$$

$$\leq \quad c\left[\int_{B_{\tilde{s}}\setminus B_{\tilde{r}}} \varphi_{|z|}(|D\zeta|)\,dx + \int_{B_{\tilde{s}}\setminus B_{\tilde{r}}} \varphi^{\beta}_{|z|}(|D\zeta|)\,dx + \int_{B_{\tilde{s}}\setminus B_{\tilde{r}}} \varphi_{|z|}(|D\xi|)\right] \quad \leq$$

$$\leq c\int_{B_{\tilde{s}}\setminus B_{\tilde{r}}} \varphi_{|z|}(|DT_{\tilde{r},\tilde{s}}[(1-\eta)v]|)\,dx + c\int_{B_{\tilde{s}}\setminus B_{\tilde{r}}} \varphi^{\beta}_{|z|}(|DT_{\tilde{r},\tilde{s}}[(1-\eta)v]|) +$$
$$+ c\int_{B_{\tilde{s}}\setminus B_{\tilde{r}}} \varphi_{|z|}(|Dv|) \leq$$
$$\leq c\int_{B_{\tilde{s}}\setminus B_{\tilde{r}}} \varphi_{|z|}(|D\eta||v|+|Dv|)+c(\tilde{s}-\tilde{r})^{-n\beta+n+\beta}\left[\sup_{[\tilde{r},\tilde{s}]}\frac{\Xi(t)-\Xi(\tilde{r})}{t-\tilde{r}}+\sup_{[\tilde{r},\tilde{s}]}\frac{\Xi(\tilde{s})-\Xi(t)}{\tilde{s}-t}\right]^{\beta} +$$
$$+ c\int_{B_{\tilde{s}}\setminus B_{\tilde{r}}} \varphi_{|z|}(|Dv|)\,dx \leq$$
$$\leq c'\int_{B_{\tilde{s}}\setminus B_{\tilde{r}}} \varphi_{|z|}\left(\left|\frac{v}{\tilde{s}-\tilde{r}}\right|\right) + \varphi_{|z|}(|Dv|)\,dx +$$
$$+ c(s-r)^{-n\beta+n}[\Xi(s)-\Xi(r)]^{\beta}$$

where an estimate from Lemma 3.23, (d) was used and Lemma 3.19 was also used. Now, starting again from the left,

$$\int_{B_r} \varphi_{|z|}(|Dv|)\,dx \leq c\int_{B_\rho} \varphi_{|z|}\left(\frac{|v|}{\tilde{s}-\tilde{r}}\right) +$$
$$+ c'\int_{B_s\setminus B_r} \varphi_{|z|}(|Dv|)\,dx + c(s-r)^n\left[\frac{\Xi(\rho)}{(s-r)^n}\right]^{\beta}$$

Using the hole-filling method we have

$$\int_{B_r} \varphi_{|z|}(|Dv|)\,dx \leq \frac{c'}{1+c'}\int_{B_s} \varphi_{|z|}(|Dv|) +$$
$$+ c(s-r)^n\left[(s-r)^{-n}\int_{B_\rho} \varphi_{|z|}(|Dv|) + \varphi_{|z|}\left(\frac{|v|}{\tilde{s}-\tilde{r}}\right)\right]^{\beta} +$$
$$+ c\int_{B_\rho} \varphi_{|z|}\left(\frac{|v|}{s-r}\right)\,dx$$

A well-known lemma (see [58], Lemma 3.1) concludes the proof. □

3.3.4 $\mathscr{A}$-harmonicity

Consider a bilinear form $\mathscr{A}$ on $\mathbb{R}^{nN}$. We assume that the upper bound

$$|\mathscr{A}| \leq \Lambda \tag{3.7}$$

with $\Lambda > 0$ holds and that the Legendre-Hadamard condition

$$\mathscr{A}(yx^T, yx^T) \geq \lambda|x|^2|y|^2 \quad \text{for all } x \in \mathbb{R}^n, y \in \mathbb{R}^N \tag{3.8}$$

with ellipticity constant $\lambda > 0$ is satisfied.
We say that $h \in W^{1,1}_{loc}(\Omega, \mathbb{R}^N)$ is $\mathscr{A}$-harmonic on Ω iff

$$\int_\Omega \mathscr{A}(Dh, D\phi)\, dx = 0$$

holds for all smooth $\phi : \Omega \to \mathbb{R}^N$ with compact support in Ω.
The following lemma will ensure that, for large z, the bilinear form $\mathscr{A} = D^2 f(z)$ satisfies the Legendre-Hadamard condition.

Lemma 3.26. *Let f satisfy $(H.1)$ and $(H.3')$ for a given $M > 0$. Then, for any given z such that $|z| > M$, we have that $\mathscr{A} = D^2 f(z)$ satisfies the Legendre-Hadamard condition*

$$\mathscr{A}(\zeta x^T, \zeta x^T) \geq \lambda |x|^2 |\zeta|^2 \qquad \text{for all } x \in \mathbb{R}^n \text{ and } \zeta \in \mathbb{R}^N$$

with ellipticity constant $\lambda = 2\gamma$.

Proof. Let u be the affine function $u(x) = zx$ with z such that $|z| > M$. Quasiconvexity in z ensures that u is a $W^{1,\varphi}$-minimizer of the functional $\mathscr{F}$ induced by f and that the function:

$$G(t) = G_\Phi(t) := \mathscr{F}_{B_1}(u + t\Phi) - \gamma \int_{B_1} \varphi_{1+|z|}(|tD\Phi|)\, dx$$

has a minimum in $t = 0$ for any $\phi \in W^{1,\varphi}_0(B_1, \mathbb{R}^N)$ and, in the same way as it is done in ([77], Prop. 5.2), from $G'_\Phi(0) = 0$ and $G''_\Phi(0) \geq 0$ the Legendre-Hadamard condition will follow.
As a matter of fact, from $G''(0) \geq 0$, we obtain:

$$\int_{B_1} \frac{\partial^2 F}{\partial z_k^\alpha \partial z_j^\beta}(z_0) D_k\phi^\alpha D_j\phi^\beta\, dx \geq 2\gamma \int_{B_1} |D\phi^2|\, dx \tag{3.9}$$

for every $\phi \in C^1_c(B_1, \mathbb{R}^N)$. Let us $\phi = \nu + i\mu$ and write (3.9) for ν and for μ, i.e.:

$$\int_{B_1} \frac{\partial^2 F}{\partial z_k^\alpha \partial z_j^\beta}(z_0) D_k\nu^\alpha D_j\nu^\beta\, dx \geq 2\gamma \int_{B_1} |D\nu^2|\, dx \tag{3.10}$$

and

$$\int_{B_1} \frac{\partial^2 F}{\partial z_k^\alpha \partial z_j^\beta}(z_0) D_k\mu^\alpha D_j\mu^\beta\, dx \geq 2\gamma \int_{B_1} |D\mu^2|\, dx \tag{3.11}$$

we obtain:

$$\int_{B_1} \frac{\partial^2 F}{\partial z_k^\alpha \partial z_j^\beta}(z_0) \left[D_k\nu^\alpha D_j\nu^\beta + D_k\mu^\alpha D_j\mu^\beta\right] dx \geq 2\gamma \int_{B_1} |D\nu^2| + |D\mu^2|\, dx \tag{3.12}$$

and hence:

$$\mathrm{Re} \int_{B_1} \frac{\partial^2 F}{\partial z_k^\alpha \partial z_j^\beta}(z_0) D_k\phi^\alpha D_j\overline{\phi}^\beta\, dx \geq 2\gamma \int_{B_1} |D\phi|^2\, dx$$

Now, consider any $\xi \in \mathbb{R}^n$, $\eta \in \mathbb{R}^N$, $\tau \in \mathbb{R}$ and $\Psi(x) \in C^\infty_c(B_1, \mathbb{R})$ and take ϕ to be

$\phi(x) = \eta e^{i\tau(\xi\cdot x)}\Psi(x)$. Since $\phi^\alpha(x) = \eta^\alpha \Psi(x) e^{i\tau\xi\cdot x}$, we have

$$\int_{B_1} \frac{\partial^2 F}{\partial z_k^\alpha \partial z_j^\beta}(z_0)\eta^\alpha\eta^\beta[\tau^2\xi_k\xi_j\Psi^2 + D_k\Psi D_j\Psi]\,dx \geq 2\gamma|\eta|^2 \int_{B_1} (|D\Psi|^2 + \tau^2|\xi|^2|\Psi(x)|^2)\,dx.$$

Dividing by τ^2 and letting $\tau \to \infty$ we get:

$$\int_{B_1} \frac{\partial^2 F}{\partial z_k^\alpha \partial z_j^\beta}(z_0)\xi_k\xi_j\eta^\alpha\eta^\beta\Psi^2(x)\,dx \geq 2\gamma|\eta|^2|\xi|^2 \int_{B_1} \Psi^2(x)\,dx$$

and since this holds for all $\Psi \in C_c^\infty(B_1, \mathbb{R})$ the proposition is proved. □

Remark 3.27. Assume $f \in C^2_{\text{loc}}(\mathbb{R}^{nN})$. Then for each $L > 0$, there is a modulus of continuity $\omega_L : [0, +\infty[\to [0, +\infty[$ satisfying $\lim_{z\to 0} \omega_L(z) = 0$ such that for all $z_1, z_2 \in \mathbb{R}^{nN}$ we have:

$$|z_1| \leq L,\ |z_2| \leq L+1 \Rightarrow |D^2 f(z_1) - D^2 f(z_2)| \leq \omega_L(|z_1 - z_2|^2).$$

Moreover, ω_L can be chosen such that the following properties hold:

1. ω_L is non-decreasing,
2. ω_L^2 is concave,
3. $\omega_L^2(z) \geq z$ for all $z \geq 0$.

The following lemma will allow approximation by $\mathscr{A}$-harmonic functions.

Lemma 3.28. *Let f satisfy (H.1)–(H.5) for a given $M > 0$. Choose any $L > M > 0$ and take $u \in W^{1,\varphi}(\Omega, \mathbb{R}^N)$ to be a $W^{1,\varphi}$-minimizer of $\mathscr{F}$ on some ball $B_\rho(x_0)$. Then for all $z:\ M < |z| \leq L$ and $\phi \in C_c^\infty(B_\rho(x_0))$ we have*

$$\left| \fint_{B_\rho(x_0)} D^2 f(z)(Du - z, D\phi)\,dx \right| \leq c\sqrt{\Phi_\varphi}\,\omega_L(\Phi_\varphi) \sup_{B_\rho(x_0)} |D\phi|. \tag{3.13}$$

where $\Phi_\varphi := \Phi_\varphi(u, x_0, \rho, z)$, the constant c depends only on n,N,Γ',Γ'',L and ω_L is the modulus of continuity of the above Remark (see also [131]).

Proof. Setting $v(x) := u(x) - zx$, the Euler equation of $\mathscr{F}$ gives

$$\left| \fint_{B_\rho} D^2 f(z)(Dv, D\phi)\,dx \right| =$$

$$= \left| \fint_{B_\rho} \left[D^2 f(z)(Dv, D\phi) + Df(z)D\phi - Df(Du)D\phi \right] dx \right|.$$

If $|Dv| \leq 1$ we have

$$
\begin{aligned}
&|D^2 f(z)(Dv, D\phi) + Df(z)D\phi - Df(Du)D\phi| \leq \\
&\qquad \leq \int_0^1 \left|D^2 f(z) - D^2 f(z + tDv)\right| \, dt |Dv| \|D\phi\|_\infty \leq \\
&\qquad\qquad \leq \omega_L(|Dv|^2)|Dv|\|D\phi\|_\infty \leq \\
&\qquad\qquad\qquad \leq c\omega_L(\varphi_{|z|}(|Dv|))\varphi_{|z|}(|Dv|)\|D\phi\|_\infty
\end{aligned}
$$

where in the last step we used $(H.2)$ to infer

$$
|Dv|^2 \lesssim \inf_{t\in[M,L+1]} \varphi''(t)|Dv|^2 \leq \varphi''(|z| + |Dv|)|Dv|^2 \simeq \varphi_{|z|}(|Dv|).
$$

If $|Dv| > 1$, we use the $M \leq |z| \leq L$ and $\varphi(t) > ct$ (see Lemma 3.20) on $t > 1$, together with $(H.4)$ to obtain:

$$
\begin{aligned}
&|D^2 f(z)(Dv, D\phi) + Df(z)D\phi - Df(Du)D\phi| \leq \\
&\qquad \leq c\left(|Dv| + |Dv| \int_0^1 D^2 f(|z + t(Du - z)|) \, dt\right) \|D\phi\|_\infty \leq \\
&\qquad c\left(|Dv| + \int_0^1 \frac{\psi'(|z + t(Du - z)|)}{|z + t(Du - z)|} \, dt\right) \|D\phi\|_\infty \leq \\
&\qquad \leq c\left[\varphi(|Dv|) + \frac{\psi'(|z| + |Du - z|)}{|z| + |Du - z|}\right] \|D\phi\|_\infty \leq \\
&\qquad \leq c\left[\varphi(|Dv|) + \frac{\psi'(2|z|)}{|z|} + \frac{\psi'(2|Du - z|)}{2|Du - z|}\right] \|D\phi\|_\infty \leq \\
&\qquad \leq c\left[\varphi(|Dv|) + 1 + \psi'(|Dv|)\right] \|D\phi\|_\infty \leq \\
&\qquad\qquad \leq c\varphi(|Dv|)\|D\phi\|_\infty \leq c\varphi_{|z|}(|Dv|)\|D\phi\|_\infty.
\end{aligned}
$$

In the second to last step we also used $\psi'(t) < c(\varphi(t) + 1)$ for $t > 1$ which follows by the Young inequality

$$
\psi'(t) \cdot 1 \leq \mathscr{N}(\psi'(t)) + \mathscr{N}^*(1)
$$

where $\mathscr{N} = \varphi \circ (\psi')^{-1}$, and the fact that $\varphi(|Dv|) > c|Dv| > c$.
Now, by the fact that $\omega_L^2(t) \geq t$ for $t \geq 0$ we get

$$
\left| \fint_{B_\rho} D^2 f(z)(Dv, D\phi) \, dx \right| \leq c\|D\phi\|_\infty \fint_{B_\rho} \omega_L(\varphi_{|z|}(|Dv|))\sqrt{\varphi_{|z|}(|Dv|)} \, dx
$$

and since ω_L is non-decreasing, using Cauchy-Schwartz and Jensen inequalities we get

$$
\left| \fint_{B_\rho} D^2 f(z)(Dv, D\phi) \, dx \right| \leq c\sqrt{\Phi_\varphi}\omega_L(\Phi_\varphi)\|D\psi\|_\infty
$$

which concludes the proof. □

The following lemma is from a paper by Teresa Isernia, Chiara Leone and Anna

Verde (see [91]).

Lemma 3.29 (Isernia, Leone, Verde, 2021, submitted)**.** *Let* $0 < \lambda \leq \Lambda < \infty$ *and* $\varepsilon > 0$*. Then there is a* $\delta(n, N, \varphi, \varphi^*, \Lambda, \lambda, \varepsilon) > 0$ *such that the following assertion holds: For all* $\kappa > 0$*, for all* $\mathscr{A}$ *satisfying* (3.7) *and* (3.8) *and for each* $u \in W^{1,\varphi}(B_\rho(x_0); \mathbb{R}^N)$ *satisfying*

$$\left| \fint_{B_\rho(x_0)} \mathscr{A}(Du, D\phi)\, dx \right| \leq \delta\kappa \sup_{B_\rho(x_0)} |D\phi|$$

for all smooth $\phi : B_\rho(x_0) \to \mathbb{R}^N$ *with compact support in* $B_\rho(x_0)$ *there is an* $\mathscr{A}$*-harmonic function* $h \in C^\infty_{loc}(B_\rho(x_0), \mathbb{R}^N)$ *with*

$$\sup_{B_{\rho/2}(x_0)} |Dh| + \rho \sup_{B_{\rho/2}(x_0)} |D^2 h| \leq c^* \varphi_{|z|}^{-1} \left(\fint_{B_\rho(x_0)} \varphi_{|z|}(|Du|) \right)$$

and

$$\fint_{B_{\rho/2}(x_0)} \varphi_{|z|} \left(\frac{|u-h|}{\rho} \right) dx \leq \varepsilon \left[\fint_{B_\rho(x_0)} \varphi_{|z|}(|Du|) + \varphi(\gamma) \right].$$

Here c^* *denotes a constant depending only on* $n, N, q_1, \Lambda, \lambda$*.*

3.3.5 Excess decay estimate

Proposition 3.30. *Let* z_0 *be s.t.* $|z_0| > M + 1$ *and* x_0 *be s.t.*

$$\lim_{\rho \to 0} \fint_{B_\rho(x_0)} |V(Du(x)) - V(z_0)|^2 = 0$$

then

$$\Phi_p(u, x_0, \rho) \to 0 \quad \text{as} \quad \rho \to 0.$$

Proof. Let $(Du)_\rho := \fint_{B_\rho(x_0)} |Du|$. We have, by triangular inequality

$$\begin{aligned} \Phi_p(u, x_0, \rho) &= \fint_{B_\rho(z_0)} |V(Du(x)) - V((Du)_\rho)|^2\, dx \leq \\ &\leq c \left[\fint_{B_\rho(x_0)} |V(Du(x)) - V(z_0)|^2\, dx + \fint_{B_\rho(x_0)} |V(Du_\rho) - V(z_0)|^2\, dx \right]. \end{aligned}$$

First summand of the right side is going to 0 by hypothesis. The second summand is equivalent to

$$\varphi_{|z_0|} \left(\left| \fint_{B_{\rho(x_0)}} Du(x) - z_0\, dx \right| \right)$$

which, by Jensen inequality, is less than

$$\fint_{B_{\rho(x_0)}} \varphi_{|z_0|}(|Du(x) - z_0|)\, dx.$$

This is, again, in turn, equivalent to

$$\fint_{B_\rho(x_0)} |V(Du(x)) - V(z_0)|^2$$

which is vanishing by hypothesis. □

Finally, we can prove

Lemma 3.31. *Assume f, φ and ψ satisfy hypotheses $(H.1)$ – $(H.5)$ for given $p_\varphi, q_\varphi, p_\psi, q_\psi$ and M.*
Let $L > M + 1 > 0$, $\alpha \in (0,1)$, $x_0 \in \Omega$ and $z_0 \in \mathbb{R}^{nN}$ such that $L > |z_0| > M + 1$.
Then there are constants $\varepsilon_0 > 0$, $\theta \in (0,1)$ and a radius $\rho^* > 0$ such that the following holds [1].
Let u a $W^{1,\varphi}$-minimizer of $\mathscr{F}$ on $B_\rho(x_0)$, with $\rho < \rho^*$ and $x_0 \in \mathbb{R}^n$ satisfying

$$\lim_{\rho \to 0} \fint_{B_\rho(x_0)} |V((Du(x)) - V(z_0)|^2 = 0.$$

If

$$\Phi_\varphi(u, x_0, \rho) \le \varepsilon_0 \tag{3.14}$$

then

$$\Phi_\varphi(u, x_0, \theta\rho) \le \theta^{2\alpha} \Phi_\varphi(u, x_0, \rho).$$

Proof. Let z_0 be such that $|z_0| > M + 1$ and x_0 any point such that

$$\lim_{\rho \to 0} \fint_{B_\rho(x_0)} |V(Du(x)) - V(z_0)|^2 = 0.$$

In what follows, for simplicity of notation, we assume that $x_0 = 0$ and we abbreviate

$$z = (Du)_\rho := \fint_{B_\rho} Du\, dx$$

and

$$\Phi_\varphi(\cdot) := \Phi_\varphi(u, 0, \cdot).$$

where $\rho > 0$ is any positive value small enough (smaller than a ρ^* that will be determined throughout the proof).
As the claim is trivial if $\exists \rho$ s.t. $\Phi_\varphi(\rho) = 0$ we can assume $\Phi_\varphi(\rho) \neq 0$.
Setting

$$w(x) := u(x) - zx$$

[1] ε_0, θ, ρ^* depending on $n, N, L, p_\varphi, q_\varphi, p_\psi, q_\psi, \Gamma, \alpha, \gamma, x_0, z_0$ and $\Lambda_L := \max_{B_{L+2}} |D^2 f|$

we have by our equivalent definition of $\Phi_\varphi(\rho)$ that

$$\fint_{B_\rho} \varphi_{|z|}(|Dw|)\,dx = \Phi_\varphi(\rho)$$

Next we will approximate by $\mathscr{A}$–harmonic functions, where $\mathscr{A} := D^2 f(z)$.
If ρ is chosen sufficiently small we have $L > |z| > M+1$, hence, from $|\mathscr{A}| \leq \max_{B_{L+2}} |D^2 f| =: \Lambda_L$ and from Lemma 3.23 we deduce that $\mathscr{A}$ satisfies (3.8) with ellipticity constant 2γ. Lemma 3.28 yields the estimate:

$$\left| \fint_{B_\rho} \mathscr{A}(Dw, D\phi)\,dx \right| \leq C_2 \sqrt{\Phi_\varphi(\rho)}\,\omega_L\left(\Phi_\varphi(\rho)\right) \sup_{B_\rho} |D\phi|$$

for all $\rho < \rho^*$ and for all smooth functions $\phi : B_\rho \to \mathbb{R}^N$ with compact support in B_ρ, where C_2 is a positive constant depending on $n, N, p_1, q_1, \Gamma, L, \Lambda_L$.
For $\varepsilon > 0$ to be specified later, we fix the corresponding constant $\delta(n, N, \varphi, \Lambda_L, \gamma, \varepsilon) > 0$ from Lemma 3.29.
Now, let $\varepsilon_0 = \varepsilon_0(n, N, \varphi, \Lambda_L, \gamma, \varepsilon)$ be small enough so that (3.14) implies:

$$C_2 \omega_L(\Phi_\varphi(\rho)) \leq \delta \tag{3.15}$$

$$\kappa = \sqrt{\Phi_\varphi(\rho)} \leq 1. \tag{3.16}$$

We apply Lemma 3.29 obtaining an $\mathscr{A}$-harmonic function $h \in C^\infty_{loc}(B_\rho; \mathbb{R}^N)$ such that

$$\sup_{B_{\rho/2}} |Dh| + \rho \sup_{B_{\rho/2}} |D^2 h| \leq c^* \varphi^{-1}_{|z|}\left(\Phi_\varphi(\rho)\right)$$

where $c^* = c^*(n, N, \varphi, \Lambda_L, \gamma)$ and

$$\fint_{B_{\rho/2}} \varphi_{|z|}\left(\frac{|w-h|}{\rho}\right) dx \leq \varepsilon\left[\Phi_\varphi(\rho) + \varphi_{|z|}(\kappa)\right] \leq c\varepsilon\Phi_\varphi(\rho). \tag{3.17}$$

where this last step follows by noticing that $\varphi_{|z|}(t) \simeq t^2$ when $t < 1$.
Now fix $\theta \in (0, 1/4]$. Taylor expansion implies the estimate:

$$\sup_{x \in B_{2\theta\rho}} |h(x) - h(0) - Dh(0)x| \leq$$

$$\leq \frac{1}{2}(2\theta\rho)^2 \sup_{x \in B_{\rho/2}} |D^2 h| \leq 2c^* \theta^2 \rho \varphi^{-1}_{|z|}\left(\Phi_\varphi(\rho)\right).$$

It follows:

$$
\begin{aligned}
&\fint_{B_{2\theta\rho}} \varphi_{|z|}\left(\frac{|w(x)-h(0)-Dh(0)x|}{2\theta\rho}\right) dx \leq \\
&\qquad \leq c\Big[\theta^{-q_\varphi-1-n} \fint_{B_{\rho/2}} \varphi_{|z|}\left(\frac{|w-h|}{\rho}\right) dx + \\
&\qquad + \fint_{B_{2\theta\rho}} \varphi_{|z|}\left(\frac{|h(x)-h(0)-Dh(0)x|}{2\theta\rho}\right) dx\Big] \leq \\
&\qquad \leq c\left[\theta^{-q_\varphi-1-n}\varepsilon\Phi_\varphi(\rho) + \varphi_{|z|}(\theta\kappa)\right] \leq \\
&\qquad \leq c\left[\theta^{-q_\varphi-1-n}\varepsilon\Phi_\varphi(\rho) + \theta^2\Phi_\varphi(\rho)\right] \leq c\theta^2\Phi_\varphi(\rho)
\end{aligned}
$$

where the last step is obtained by choosing $\varepsilon := \varepsilon(\theta) = \theta^{q_\varphi+n+3}$ (so, remember that ε and hence δ and ε_0 depend on whatever θ is) and recalling the definition of w we have:

$$
\fint_{B_{2\theta\rho}} \varphi_{|z|}\left(\frac{|u(x)-zx-(h(0)+Dh(0)x)|}{2\theta\rho}\right) dx \leq c\theta^2\Phi_\varphi(\rho). \tag{3.18}
$$

On the other hand, we remark that, using the properties of h:

$$
|Dh(0)| \leq c^*\varphi_{|z|}^{-1}\left(\Phi_\varphi(\rho)\right) \tag{3.19}
$$

We can take ε_0 small enough such that (3.14) implies

$$
|Dh(0)| \leq 1. \tag{3.20}
$$

Using this fact together with (3.19) we get

$$
\begin{aligned}
&\Phi_\varphi(2\theta\rho, z+Dh(0)) \leq \\
&\qquad \leq c\left[(2\theta)^{-n}\left(\fint_{B_\rho} |V(Du(x))-V(z)|^2\,dx + \varphi_{|z|}(|Dh(0)|)\right)\right] \leq \\
&\qquad \leq c\left[\theta^{-n}\left(\Phi_\varphi(\rho)+\Phi_\varphi(\rho)\right)\right] \leq c\theta^{-n}\Phi_\varphi(\rho).
\end{aligned} \tag{3.21}
$$

Now we need to use Caccioppoli inequality (3.6) with $q(x) = h(0) + [z + Dh(0)]x$ and $z + Dh(0)$ playing the role of z (we can do so because $|z + Dh(0)| > M$), and we get

$$
\Phi_\varphi(\theta\rho, z+Dh(0)) \leq c\left[\theta^2\Phi_\varphi(\rho) + \theta^{2\beta}\Phi_\varphi(\rho)^\beta + \theta^{-n\beta}\Phi_\varphi(\rho)^\beta\right]. \tag{3.22}
$$

Thereby the condition $|z + Dh(0)| \leq L + 1$ of Lemma 3.25 can be deduced from the smallness of $|Dh(0)|$.
Now, if ε_0 is chosen small enough, depending on θ, (3.14) implies the following:

$$
\theta^{-n\beta}\Phi_\varphi(\rho)^{\beta-1} \leq \theta^2, \tag{3.23}
$$

and from the fact that $\theta \leq 1$ we have

$$\Phi_\varphi(\theta\rho, z + Dh(0)) \leq c\theta^2 \Phi_\varphi(\rho).$$

Adapting Lemma 6.2 in [131] (it just uses simples ideas like the ones from Proposition 3.30) we deduce, from (3.23):

$$\Phi_\varphi(\theta\rho) \leq C_3\theta^2 \Phi_\varphi(\rho), \tag{3.24}$$

where $C_3 > 0$ depends on $n, N, \varphi, \Gamma, \gamma, \Lambda_L, L$.
Finally, we choose $\theta \in (0, \frac{1}{4}]$ (depending on α and whatever C_3 depends on) small enough such that

$$C_3\theta^2 \leq \theta^{2\alpha} \tag{3.25}$$

holds, and ε_0 small enough such that (3.15), (3.16), (3.23) follow from (3.14). Taking into account (3.24) and (3.25) the proof of the proposition is complete. □

Applying this last lemma in iteration we get the following (see also [131], Lemma 7.10)

Lemma 3.32. *Assume f, φ and ψ satisfy hypotheses $(H.1)$–$(H.5)$ for given $p_\psi, q_\psi, p_\varphi, q_\varphi$ and M.*
Let $L > 2M + 2 > 0$, $\alpha \in (0, 1)$, $x_0 \in \Omega$ and $z_0 \in \mathbb{R}^{nN}$ such that $\frac{L}{2} > |z_0| > M + 1$.
Then there is a constant $\tilde{\varepsilon}_0 > 0$ and a radius $\rho^* > 0$ [2] such that the following holds.
Let u a $W^{1,\varphi}$-minimizer of $\mathscr{F}$ on $B_\rho(x_0)$, with $\rho < \rho^*$ and $x_0 \in \mathbb{R}^n$ satisfying

$$\lim_{\rho \to 0} \fint_{B_\rho(x_0)} |V(Du(x)) - V(z_0)|^2 = 0.$$

If

$$\Phi_\varphi(u, x_0, \rho) \leq \tilde{\varepsilon}_0 \tag{3.26}$$

then there is a constant c depending on $n, N, L, p_1, q_1, \Gamma, \alpha, \gamma, x_0, z_0$ such that

$$\Phi_\varphi(u, x_0, r) \leq c \left(\frac{r}{\rho}\right)^{2\alpha} \Phi_\varphi(u, x_0, \rho)$$

for any $r < \rho$.

The theorem announced in the introduction follows from Campanato's integral characterization of Hölder continuity.

[2] depending on $n, N, L, p_\varphi, q_\varphi, p_\psi, q_\psi, \Gamma, \alpha, \gamma, x_0, z_0$ and $\Lambda_L := \max_{B_{L+2}} |D^2 f|$

Chapter 4

Bounded mean or lower oscillation: BMO and BLO

4.1 Definition of BMO and BLO

In this chapter we will deal with some function spaces and function classes defined by bounding the value of different types of oscillation. Our definitions will be given in the context of functions of a single variable. Many properties of the resulting spaces are shared by functions of several variables, but attention to some technical geometrical details needs to be made.
First, we will give a definition based on the concept of mean oscillation, that is defined by considering the average distance between a function f and its integral average over an interval I, i.e., the quantity

$$\frac{1}{|I|}\int_I \left| f(x) - \frac{1}{|I|}\int_I f(y)\,dy \right| \,dx.$$

Definition 4.1. A real valued locally integrable function $f(x) \in L^1_{loc}(\mathbb{R})$ is said to have *Bounded Mean Oscillation* ($f(x) \in BMO(\mathbb{R})$) if:

$$\sup_I \fint_I |f(x) - f_I|dx = \|f\|_{BMO} < \infty \tag{4.1}$$

where f_I denotes $\fint_I f(x)dx$ and I spans the set of all compact intervals.

One can prove that $BMO(\mathbb{R})$ is a vector space and, modulo the set of functions that are almost everywhere equal to a constant, the already defined quantity $\|\cdot\|_{BMO}$ is a norm on it that makes it a Banach Space. This is because any function that is almost everywhere equal to a constant will have zero mean oscillation over any interval. Another option to define a norm on it, other than to consider the quotient with respect to almost everywhere costant functions, would be to add to $\|f\|_{BMO}$ the L^1-norm of the function over $[0,1]$, for example.
To define the BMO property for functions of several variables one could equivalently employ hypercubes Q or balls B contained in $\mathbb{R}^n$, obtaining exactly the same space of functions. However, the two resulting norms are not exactly the same and hypercubes

with sides parallel to the coordinate axes are what are conventionally used.
This space was introduced by John and Nirenberg in 1961 ([63]). In this paper, the two authors proved a strong inequality for BMO functions, proving that the set of points such that the function deviates more than λ from its own integral average decays exponentially with λ. This is an important "self-improving" property of elements of BMO.

Theorem 4.1 (John, Nirenberg, 1961, [63]). *Let f be a function in $BMO(\mathbb{R}^n)$. There exist constants $c_1, c_2 > 0$ such that for any cube $Q \in \mathbb{R}^n$, we have*

$$|\{x \in Q : |f(x) - f_Q| > \lambda\}| < c_1 exp\left(-c_2 \frac{\lambda}{\|f\|_{BMO}}\right)|Q|$$

This inequality has a corollary implying some form of exponential summability of the difference between the function and its own integral average.

Corollary 4.2. *Let f be a function in $BMO(\mathbb{R}^n)$. Then there exists some $C > 0$ and some $A(f) \leq C\|f\|BMO$ such that*

$$\sup_{Q \subseteq \mathbb{R}^n} \int_Q e^{\frac{|f(x)-f_Q|}{A(f)}} dx < +\infty$$

In particular, from John-Nirenberg inequality one can also prove that BMO functions are locally p-summable for any finite $p \geq 1$.

Corollary 4.3. *Let f be a function in $BMO(\mathbb{R}^n)$. Then $f \in L^p_{loc}(\mathbb{R}^n)$ for any $p \in [1, +\infty)$.*

However, a function in BMO is not necessarily locally essentially bounded. An example of a function in BMO that is not bounded is the following:

$$f(x) = \log(|p(x)|) \in BMO \setminus L^\infty, \quad \text{for any polinomial } p(x) \not\equiv 0.$$

Of course, though, any function in L^∞ is also a BMO function and

$$L^\infty \hookrightarrow BMO$$

with a continuous embedding, as it is easy to prove that

$$\|f\|_{BMO} \leq 2\|f\|_\infty$$

always holds.

Definition 4.2. A real valued locally integrable function $f(x) \in L^1_{loc}(\mathbb{R})$ is said to have *Bounded Lower Oscillation* ($f(x) \in BLO(\mathbb{R})$) if

$$\sup_I ⨍_I [f(x) - \inf_I f] dx = \sup_I [f_I - \inf_I f] = \|f\|_{BLO} < \infty. \tag{4.2}$$

We remark here that inf denotes the essential infimum.

Of course, as in the definition of BMO, it is useful to think the class of $BLO(\mathbb{R})$ functions modulo the set of all functions which are almost everywhere equal to a constant.
Of course:

- $\forall f, g \in BLO, \quad f+g \in BLO$ and $\|f+g\|_{BLO} \leq \|f\|_{BLO} + \|g\|_{BLO}$
- $\forall \alpha > 0,\ \forall f \in BLO, \quad \alpha f \in BLO$ and $\|\alpha f\|_{BLO} = \alpha \|f\|_{BLO}$

but $BLO \neq -BLO$ so that BLO is not a vector space and an example of a function $f(x) = -\log(|x|) \in BLO$ but not in $-BLO$ was pointed out by Korey [108].
It is common among many authors to use the $\|\cdot\|_{BLO}$ notation (see [108]) and refer to it as a norm even if it is not defined on a vector space; we will also do so in this paper.
It is obvious, by the definitions, that:

$$\|f\|_{BMO} \leq 2\|f\|_{BLO}, \quad \forall f \in BLO. \tag{4.3}$$

As explained in the Introduction of [108], by intersecting BLO and $-BLO$ we exactly get the space of all essentially bounded functions L^∞.
In particular, of course, L^∞ is included in BMO and BLO and $\|\cdot\|_{BMO} \leq 2\|\cdot\|_{L^\infty}$ and $\|\cdot\|_{BLO} \leq 2\|\cdot\|_{L^\infty}$.

4.2 Muckenhoupt weights, maximal functions and their relation to BMO and BLO

In this section we will give definition of some function classes called Muckenhoupt weights and explain why they were introduced in connection to an operator called Hardy-Littlewood maximal operator, to then show the strong connection they have with BMO.

Definition 4.3. Let $p \in (1, \infty)$.
A weight $w : \mathbb{R}^n \to [0, +\infty[$ belongs to the A_p class of Muckenhoupt if

$$A_p(w) = \sup_{B \subset\subset \mathbb{R}^n} \fint_B w(x)dx \left(\fint_B w^{-\frac{q}{p}}(x)dx \right)^{\frac{p}{q}} < \infty \tag{4.4}$$

where q is the Hölder conjugate of p, i.e. the only real number such that $\frac{1}{p} + \frac{1}{q} = 1$, and B is any ball contained in $\mathbb{R}^n$.

It is possible to show that the same space of functions is defined upon substituting balls for cubes, but in this context balls are conventionally used.
We would like to remark that the case of $p = 2$ is peculiar as $p = q$:

Definition 4.4. A weight $w : \mathbb{R}^n \to [0, +\infty[$ belongs to the A_2 class of Muckenhoupt if

$$A_2(w) = \sup_{B \subset\subset \mathbb{R}^n} \fint_B w(x)dx \fint_B w^{-1}(x)dx < \infty \tag{4.5}$$

where $A_2(w)$ is sometimes called the A_2 constant of w

Also, we define Muckenhoupt classes of weights also for $p = 1$ as

Definition 4.5. A weight $w : \mathbb{R}^n \to [0, +\infty[$ belongs to the A_1 class of Muckenhoupt if

$$A_1(w) = \sup_{B \subset\subset \mathbb{R}^n} \frac{\fint_B w(x)dx}{\inf_B w(x)} < \infty \tag{4.6}$$

where $A_1(w)$ is sometimes called the A_1 constant of w

Lastly, we define the class even for $p = \infty$ as

$$A_\infty = \bigcup_{1 \le p} A_p.$$

Muckenhoupt weights are relevant to a useful operator called Hardy-Littlewood maximal operator, which we define here

Definition 4.6. Let $f(x)$ be any locally integrable function, $f \in L^1_{loc}(\mathbb{R}^n)$.
The Hardy-Littlewood maximal operator M maps f to a function Mf defined by the position

$$Mf(x) = \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)|\, dy$$

where $B_r(x)$ is a ball of center x and radius r

This new function $Mf(x)$ is measurable as the integral function is continuous as a function of x and r. It can be shown, though it is not trivial as it may seem, that $Mf(x)$ is almost everywhere finite. This last property is consequence of the so-called Hardy-Littlewood inequality, stated below.

Theorem 4.4. *(Hardy, Littlewood,1930,[86]) Let $f(x)$ be any integrable function, $f \in L^1(\mathbb{R}^n)$. There exists a constant $C_n > 0$, only depending on the dimension n, such that for any $\lambda > 0$ the following inequality, bounding the measures of the superlevel sets, holds*

$$|\{Mf \geq \lambda\}| < \frac{C_d}{\lambda}\|f\|_{L^1}$$

With some effort and the use of a Marcinkiewicz interpolation theorem, the following stronger estimate follows

Theorem 4.5. *(Hardy, Littlewood,1930,[86]) Let $f(x)$ be any function in $L^p(\mathbb{R}^n)$ for some $p \geq 1$. There exists a constant $C_{p,n} > 0$, depending on the dimension n and on p, such that the following estimate holds*

$$\|Mf\|_{L^p} < C_{p,n}\|f\|_{L^p}$$

The above theorem proves, in particular, that M is a bounded operator from the classic Lebesgue space L^p to itself for any $p \geq 1$.
The relationship between Muckenhoupt weights and the Hardy-Littlewood maximal operator is, in this context, made clear by the following theorem

Theorem 4.6. *(Muckenhoupt,1972,[118]) A weight $w \geq 0$ in $L^1_{loc}(\mathbb{R}^n)$ belongs to the Muckenhoupt class A_p for some p if and only if one of the following equivalent statements hold:*

1. *The Hardy–Littlewood maximal function is bounded on $L^p(w(x)dx)$, that is*

$$\int |M(f)(x)|^p \, w(x)\, dx \leq C \int |f|^p \, w(x)\, dx,$$

   for some C depending only on p and the supremum in the definition of A_p.

2. *There is a constant c such that for any locally integrable function f on $\mathbb{R}^n$, and all balls B:*

$$(f_B)^p \leq \frac{c}{w(B)} \int_B f(x)^p \, w(x)\, dx,$$

   where:

$$f_B = \frac{1}{|B|}\int_B f, \qquad w(B) = \int_B w(x)\, dx.$$

With clever use of the definition of A_p weights, Jensen inequality and John-Nirenberg inequality one can finally prove the abovementioned strong relationship between Muckenhoupt weights and BMO

Theorem 4.7 (Muckenhoupt, 1972, [118])**.** *The following relations hold:*

1. *If $0 \leq w \in A_p$ for some $p \geq 1$, then $\log w \in BMO$*

2. *If $f \in BMO$, then for sufficiently small δ, we have $e^{\delta f} \in A_p$ for some $p \geq 1$. In particular, it can be proven $e^{\frac{f}{\mu}} \in A_2$ if $\mu > \frac{\|f\|_{BMO}}{c_2}$ where c_2 is the constant from the John-Nirenberg inequality.*

Notice that the smallness assumption for δ is necessary as $-\log|x|$ is BMO but $e^{-\log|x|} = \frac{1}{|x|}$ is not A_p for any $p \geq 1$. For a suitable choice of $\delta = \delta(p) > 0$, however, $e^{-\delta \log|x|} = \frac{1}{|x|^\delta}$ is a Muckenhoupt weight in A_p.
This next theorem, due to Coifman and Rochberg ([51], Corollary 3), shows a similar connection between the Muckenhoupt class A_1, BLO and the Hardy-Littlewood maximal operator.

Theorem 4.8 (Coifman, Rochberg, 1980, [51])**.** *We have*

1. $w \in A_1 \Rightarrow \log w \in BLO$ *and* $\|\log w\|_{BLO} \leq \log A_1(w)$
2. $f \in BLO \Rightarrow \quad e^{f/\mu} \in A_1, \quad \forall \mu > \frac{\|f\|_{BMO}}{c_2}$ *where c_2 is the constant from the John-Nirenberg inequality.*

Lemma 4.9 (Coifman, Rochberg, 1980)**.** *Let $w \in L^1_{loc}$ be a weight and $\varepsilon \in [0, 1)$. Then:*

$$Mw(x)^\varepsilon \in A_1$$

with A_1 constant depending on ε but not on w.

4.3 Distance to the L^∞ subspace

We have already stated above that any L^∞ function is in BLO and, even more so, it is a BMO function. The position that L^∞ occupies in BMO and BLO will be further investigated in this section, making further use of the strong duality between Muckenhoupt classes, BMO and BLO.
The following theorem further explores this relation between Muckenhoupt classes and BMO and BLO, specializing to the Muckenhoupt classes A_2 and A_1.
It proves that the largest possible δ such that $e^{\delta f} \in A_2$ for a given function f is related to the distance of the function f in BMO from its subspace L^∞ of essentially bounded functions.

Theorem 4.10 (Garnett, Jones, 1978, [74])**.** *There exist two absolute constants $a_2, a_3 > 0$ depending only on n such that for every real valued function $f \in BMO(\mathbb{R}^n)$ the following inequalities hold:*

$$a_2\varepsilon(f) \leq \operatorname*{dist}_{BMO}(f, L^\infty) \leq a_3\varepsilon(f) \tag{4.7}$$

where

$$\varepsilon(f) = \inf\{\mu > 0 : \ e^{\frac{f}{\mu}} \in A_2\} \tag{4.8}$$

In [6] Alberico and Sbordone found sharp constants for these inequalities in the case $n = 1$. They proved that one can choose a_2 to be $\frac{2}{e}$.
In this section we restrict to $n = 1$ and prove a similar result underlining a connection between the smallest possible μ such that $e^{\frac{f}{\mu}}$ in A_1 and the distance of $f \in BLO$ to L^∞ measured with the (so-called) BLO norm. The result is the following

Theorem 4.11 (Angrisani, 2017, [14])**.** *There exist two absolute constants $d_1, d_2 > 0$ such that for every real valued function $f \in BLO(\mathbb{R})$ the following inequalities hold:*

$$d_1\sigma(f) \leq \operatorname*{dist}_{BLO}(f, L^\infty) \leq d_2\sigma(f) \tag{4.9}$$

where:

$$\sigma(f) = \inf\{\mu > 0 : \ e^{\frac{f}{\mu}} \in A_1\} \tag{4.10}$$

To be able to prove it, we will first need a lemma showing some relevant properties of the above defined quantity $\sigma(f)$.

Lemma 4.12. *Let $f_1, f_2 \in BLO$ and $g \in L^\infty$. We have:*

1. $\sigma(f_1) \leq \frac{2}{c_2}\|f_1\|_{BLO}$
2. $\sigma(f_1 + f_2) \leq \sigma(f_1) + \sigma(f_2)$
3. $\sigma(f) = \sigma(f - g)$

Proof. The first statement follows immediately from the second statement in Lemma 4.8 and the known inequality $\|\cdot\|_{BMO} \leq 2\|\cdot\|_{BLO}$.
To prove the second statement it will be enough to prove that for every μ_1 and μ_2 such

that $e^{\frac{f_1}{\mu_1}}$ and $e^{\frac{f_2}{\mu_2}}$ are in the Muckenhoupt class A_1, we have $e^{\frac{f_1+f_2}{\mu_1+\mu_2}} \in A_1$.
Observe that, if $\theta := \frac{\mu_1}{\mu_1+\mu_2}$, we have:

$$e^{\frac{f_1+f_2}{\mu_1+\mu_2}} = \left(e^{\frac{f_1}{\mu_1}}\right)^{\theta} \left(e^{\frac{f_2}{\mu_2}}\right)^{1-\theta}. \tag{4.11}$$

It is always the case, by use of the Holder inequality, that $w_1, w_2 \in A_1 \Rightarrow w_1^{\theta} w_2^{1-\theta} \in A_1$ so that the proof of the second statement is finished.
The third statement follows directly from the second with the observation that $\sigma(g) = 0, \ \forall g \in L^\infty$. □

We are now ready to prove the main result of this section

Proof of Theorem 4.11. To prove the left inequality, by Lemma 4.12, observe that for every $g \in L^\infty$:

$$\frac{c_2}{2}\sigma(f) = \frac{c_2}{2}\sigma(f-g) \leq \|f-g\|_{BLO} \tag{4.12}$$

and then take the infimum with respect to $g \in L^\infty$ over both sides, so that d_1 can be chosen to be $\frac{c_2}{2}$.
To prove the other inequality choose $\mu_0 \in (\sigma(f), 2\sigma(f))$ and define $w = e^{f/\mu_0} \in A_1$.
By Lemma 4.9 we have $(Mw)^{\frac{1}{2}} \in A_1$. This means that $\frac{1}{2}\log(Mw)$ is in BLO, by Theorem 4.8 and by the same theorem

$$\|\log(Mw)\|_{BLO} \leq 2\log\left(A_1((Mw)^{\frac{1}{2}})\right) =: K.$$

Since this last quantity does not depend on w, by what is stated in Lemma 4.9, we have that

$$\|\mu_0 \log(Mw)\|_{BLO} \leq \mu_0 K \text{ with } K \text{ an absolute constant.} \tag{4.13}$$

We observed that:

$$w \leq Mw \leq A_1(w) w \tag{4.14}$$

and this implies that:

$$\log\left(\frac{w}{Mw}\right) \in L^\infty. \tag{4.15}$$

We conclude that:

$$\begin{aligned} \operatorname*{dist}_{BLO}(f, L^\infty) &\leq \left\| f - \mu_0 \log\left(\frac{w}{Mw}\right) \right\|_{BLO} = \\ &= \|\mu_0 \log(Mw)\|_{BLO} \leq \mu_0 K < (2K)\sigma(f) \end{aligned} \tag{4.16}$$

so that d_2 can be chosen to be $2K$. □

4.4 A new norm to tackle the distance problem

In this section, inspired by a work (see [51]) of Coifman and Rochberg on a decomposition of A_1 weights through the Hardy-Littlewood maximal operator and by its consequences on BLO, we define a new norm on BLO that is equivalent to the usual one.
This new norm $\|\cdot\|'_{BLO}$ will have the advantage of making

$$\operatorname*{dist}_{BLO'}(f, L^\infty) := \inf_{g\in L^\infty} \|f-g\|'_{BLO}$$

explicitely and exactly computable in terms of $\sigma(f)$. The beauty of this result is that this norm is not artificially constructed for this scope, but rather naturally inspired from properties of BLO functions with respect to the Hardy-Littlewoood maximal operator.
As a matter of fact, we will prove the equality:

$$\operatorname*{dist}_{BLO'}(f, L^\infty) = \sigma(f) := \inf\{\mu > 0 : \ e^{\frac{f}{\mu}} \in A_1\} \tag{4.17}$$

showing that, in this new norm, the distance from L^∞ is exactly the reciprocal of the critical exponent for which the function $e^{\frac{f}{\mu}}$ is in A_1, reaching a result that is, notably, a quantitative improvement with respect to the one in the previous section.
To reach this objective, we quote another result by Coifman and Rochberg contained in [51], concerning powers of Mw for $w \in L^1_{loc}$. The two authors also proved that, modulo L^∞ functions, all A_1 functions arise as some power of Mw for a suitable $w \in L^1_{loc}$, i.e:

Theorem 4.13 (Coifman, Rochberg, 1980, [51]). *Assume $w \in A_1$.*
There are functions $0 < A < b(x) < 1 \in L^\infty$, $g \in L^1_{loc}$ and a number $\varepsilon \in [0,1)$ such that:

$$w(x) = b(x)Mg(x)^\varepsilon$$

For the convenience of the reader we will repeat the proof of this theorem paying particular attention to the value of A and how it depends on $A_1(w)$.
To do so, an inequality known as the reverse Holder inequality is needed:

Theorem 4.14. *If $w \in A_1$, then there is a sufficiently small value of $\eta > 0$ and a constant c_η depending on η but not on $I \subset \mathbb{R}$ such that:*

$$\left(\fint_I w^{1+\eta}(x)dx\right)^{\frac{1}{1+\eta}} \leq c_\eta \fint_I w(x)dx.$$

In particular, we have $w^{1+\eta} \in A_1$ and $A_1(w^{1+\eta}) \leq c_\eta^{1+\eta} A_1(w)^{1+\eta}$.
Furthermore, $c_\eta < c_3 := 2e, \ \forall \eta < 1$.

Using this theorem we have that for every $w \in A_1$ there is always a $\eta \in (0,1)$ such that $A_1(w^{1+\eta}) \leq c_3^{1+\eta} A_1(w)^{1+\eta}$.

Proof of Theorem 4.13. Choose a sufficiently small $\eta < 1$ as in the reverse Holder inequality and define

$$b(x) = \frac{w(x)}{[M(w^{1+\eta}(x))]^{\frac{1}{1+\eta}}} \text{ and } g(x) = w(x)^{1+\eta} \tag{4.18}$$

noticing that $g(x) \in L^1_{loc}$.
We have that $b(x) \leq 1 \iff w^{1+\eta}(x) \leq M(w^{1+\eta}(x))$ but this latter inequality holds a.e. since almost every point $x \in \mathbb{R}$ is a Lebesgue point for $w^{1+\eta}(x) \in L^1_{loc}$.
A lower bound on $b(x)$ can be obtained by observing that:

$$b(x) = \frac{w(x)}{[M(w^{1+\eta}(x))]^{\frac{1}{1+\eta}}} = \left[\frac{w(x)^{1+\eta}}{M(w^{1+\eta}(x))}\right]^{\frac{1}{1+\eta}} \geq$$

$$\geq \left[\frac{1}{A_1(w^{1+\eta})}\right]^{\frac{1}{1+\eta}} \geq \frac{1}{c_3 A_1(w)} > 0. \tag{4.19}$$

Now take $\varepsilon \in (0,1)$ to be $\frac{1}{1+\eta}$.
We have that $w(x) = b(x)\,[Mg(x)]^{\varepsilon}$ so that the theorem is proven and we can choose A to be $\frac{1}{c_3 A_1(w)}$. □

By combining theorems 4.9 and 4.13, in light of the known duality expressed by Lemma 4.8 we get that:

$$f(x) \in BLO \iff \exists\ \alpha > 0,\ b(x) \in L^{\infty},\ g(x) \in L^1_{loc},$$
$$f(x) = \alpha \log(Mg(x)) + b(x) \tag{4.20}$$

The logical equivalence in (4.20) leads us to define:

$$\|f\|'_{BLO} =$$
$$= \inf\left\{\alpha + \|b\|_{\infty} :\ \alpha > 0,\ b \in L^{\infty} \text{ s.t. } \exists g \in L^1_{loc},\ f = \alpha \log(Mg) + b\right\}$$

which has, of course, the same properties of $\|\cdot\|_{BLO}$ and we will now prove the equivalence of the two.

Theorem 4.15 (Angrisani, preprint)**.** *There exist two absolute constants $d_1, d_2 > 0$ such that, for every $f \in BLO$:*

$$d_1\|f\|_{BLO} \leq \|f\|'_{BLO} \leq d_2\|f\|_{BLO} \tag{4.21}$$

Proof. During this proof, c_1 and c_2 will still denote the constants in John-Nirenberg inequality.
We will divide the proof in steps, addressing the latter inequality first.
Step 1: Showing that $\exists c_4 > 0$ such that $\forall f \in BLO,\ A_1\left(e^{\frac{fc_2}{4\|f\|_{BLO}}}\right) \leq c_4$:
Take $\lambda = \frac{4\|f\|_{BLO}}{c_2}\log\zeta$ in John-Niremberg inequality.

We have:

$$|\{t \in I: \ |f(t) - f_I| > \frac{4\|f\|_{BLO}}{c_2} \log \zeta\}| \leq c_1 \zeta^{-\frac{4\|f\|_{BLO}}{\|f\|_{BMO}}} |I|. \tag{4.22}$$

Notice that:

$$|f(t) - f_I| > \frac{4\|f\|_{BLO}}{c_2} \log \zeta \iff e^{\frac{c_2|f(t)-f_I|}{4\|f\|_{BLO}}} > \zeta. \tag{4.23}$$

By Cavalieri's principle we have:

$$\begin{aligned}\int_I e^{\frac{c_2|f(t)-f_I|}{4\|f\|_{BLO}}} dt = \int_0^\infty |\{t \in I: \ e^{\frac{c_2|f(t)-f_I|}{4\|f\|_{BLO}}} > \zeta\}| d\zeta \leq \\ \leq |I| + \int_1^\infty |\{t \in I: \ e^{\frac{c_2|f(t)-f_I|}{4\|f\|_{BLO}}} > \zeta\}| d\zeta.\end{aligned} \tag{4.24}$$

But we can use (4.23) to bound the last integral so that (4.24) becomes:

$$\int_I e^{\frac{c_2|f(t)-f_I|}{4\|f\|_{BLO}}} \leq |I| + |I| \int_1^\infty c_1 \zeta^{-\frac{4\|f\|_{BLO}}{\|f\|_{BMO}}} \leq |I| + |I| \int_1^\infty c_1 \zeta^{-2} d\zeta = |I|(1 + c_1). \tag{4.25}$$

where $\|\cdot\|_{BMO} \leq 2\|\cdot\|_{BLO}$ was also used.
By dividing both sides of (4.25) by $|I|$ and using $x \leq |x|$, we get to

$$\fint_I e^{\frac{c_2(f(t)-f_I)}{4\|f\|_{BLO}}} \leq 1 + c_1 \tag{4.26}$$

so that:

$$\fint_I e^{\frac{c_2}{4\|f\|_{BLO}} f(t)} dt \leq (1 + c_1) e^{\frac{c_2 f_I}{4\|f\|_{BLO}}} \leq (1 + c_1) e^{\frac{c_2(\inf_I f + \|f\|_{BLO})}{4\|f\|_{BLO}}} \tag{4.27}$$

where we used $f \in BLO$. Lastly:

$$\fint_I e^{\frac{c_2}{4\|f\|_{BLO}} f(t)} dt \leq (1 + c_1) e^{\frac{c_2}{4}} \inf_I e^{\frac{c_2}{4\|f\|_{BLO}} f(t)} \tag{4.28}$$

and this first step is concluded by taking $c_4 = (1 + c_1) e^{\frac{c_2}{4}}$.
Step 2: Showing that $\|f\|'_{BLO} \leq d_2 \|f\|_{BLO}$ Take $f(x) \in BLO$. By the first step we have $w(x) = e^{\frac{f(x) c_2}{4\|f\|_{BLO}}} \in A_1$.
By theorem 4.13 we have $b \in L^\infty$, $0 < A < b(x) < 1$, $\varepsilon \in [0, 1)$, $g \in L^1_{loc}$ such that:

$$w(x) = b(x)[Mg(x)]^\varepsilon \tag{4.29}$$

and so

$$f(x) = \frac{4\|f\|_{BLO}}{c_2} \log(b(x)) + \frac{4\|f\|_{BLO}}{c_2} \varepsilon \log(Mg(x)). \tag{4.30}$$

In particular, define

$$B(x) = \tfrac{4\|f\|_{BLO}}{c_2} \log(b(x)) \text{ and } \alpha = \tfrac{4\|f\|_{BLO}}{c_2} \varepsilon.$$

Since $1 \geq b(x) \geq \frac{1}{c_3 A_1(w)}$, taking logarithms and using the first step we have that

$$0 < |B(x)| < [\log c_3 + \log c_4] \frac{4\|f\|_{BLO}}{c_2} =: c_5 \frac{4\|f\|_{BLO}}{c_2}. \tag{4.31}$$

Also, notice $\alpha < \frac{4\|f\|_{BLO}}{c_2}$ since $\varepsilon < 1$.
By combining this with (4.31) we get

$$\|f\|'_{BLO} \leq \alpha + \|B\|_\infty \leq \left[(1 + c_5)\frac{4}{c_2}\right] \|f\|_{BLO},$$

so that by denoting $d_2 = \left[(1 + c_5)\frac{4}{c_2}\right]$ we conclude the second step.
Step 3: Showing $d_1\|f\|_{BLO} \leq \|f\|'_{BLO}$.
Take any $\alpha > 0$, $g \in L^1_{loc}$ and $b \in L^\infty$ such that $f = \alpha \log(Mg) + b$.
We have:

$$\|f\|_{BLO} \leq \alpha \| \log(Mg)\|_{BLO} + \|b\|_{BLO} \leq 2\alpha \| \log((Mg)^{\frac{1}{2}})\|_{BLO} + 2\|b\|_\infty.$$

But by the first proposition in Lemma 4.8 we have:

$$\|f\|_{BLO} \leq 2\alpha \log \left[A_1((Mg)^{\frac{1}{2}})\right] + 2\|b\|_\infty.$$

Theorem 4.9 by Coifman and Rochberg stated that $A_1((Mg)^\varepsilon)$ does not depend on g, so that $A_1((Mg)^{\frac{1}{2}}) \leq c_6$, with c_6 absolute constant.
For this reason we have

$$\|f\|_{BLO} \leq 2\alpha \log c_6 + 2\|b\|_\infty. \tag{4.32}$$

for each possible decomposition of f, so that by taking the infimum and denoting $\frac{1}{d_1} = \max\{2, 2\log c_6\}$ we conclude the last step of the proof. □

We can now prove the announced main result of this section

Theorem 4.16 (Angrisani, preprint). *For every $f \in BLO$ we have that:*

$$\operatorname*{dist}_{BLO'}(f, L^\infty) = \inf_{h \in L^\infty} \|f - h\|'_{BLO} = \sigma(f)$$

Proof. First notice that:

$$\operatorname*{dist}_{BLO'}(f, L^\infty) = \inf_{h \in L^\infty} \inf\{\mu + \|b\|_\infty : \ f - h = \mu \log(Mg) + b\}$$

so that, for any μ and g, by a proper choice of h, b can always be chosen to be 0.
This means that

$$\operatorname*{dist}_{BLO'}(f, L^\infty) = \inf\{\mu > 0 : \ \exists g(x) \in L^1_{loc}, \ f(x) - \mu \log[Mg(x)] \in L^\infty\}. \tag{4.33}$$

On the other hand, since a weight $e^{\frac{f}{\mu}} \in L^1_{loc}$ is in the Muckenhoupt class A_1 if and only if it can be written as $b(x)[Mg(x)]^\varepsilon$ for suitable $0 < A < b(x) \in L^\infty$, $g \in L^1_{loc}$ and

$\varepsilon \in [0,1)$, we have:

$$\sigma(f) = \inf\left\{\mu > 0 :\ \exists b \in L^\infty,\ \varepsilon < 1,\ g \in L^1_{loc},\ e^{\frac{f(x)}{\mu}} = b(x)[Mg(x)]^\varepsilon\right\} \tag{4.34}$$

or, by taking logarithms:

$$\sigma(f) = \inf\Big\{\mu > 0 : \\ \exists b(x) \in L^\infty,\ \varepsilon < 1,\ g(x) \in L^1_{loc},\ \frac{f(x)}{\mu} = b'(x) + \varepsilon \log[Mg(x)]\Big\} \tag{4.35}$$

where we have denoted $b'(x) = \log(b(x)) \in L^\infty$.
Equivalently:

$$\sigma(f) = \inf\left\{\mu > 0 :\ \exists \varepsilon < 1,\ g(x) \in L^1_{loc},\ f(x) - \mu\varepsilon \log[Mg(x)] \in L^\infty\right\} \tag{4.36}$$

which is the same thing as:

$$\sigma(f) = \inf\left\{\mu > 0 :\ \exists g(x) \in L^1_{loc},\ f(x) - \mu \log[Mg(x)] \in L^\infty\right\} \tag{4.37}$$

Let us better explain this last step, between Equation (4.36) and (4.37).
A real number a satisfies

$$a < \inf\{\mu > 0 :\ \exists \varepsilon < 1,\ g(x) \in L^1_{loc},\ f(x) - \mu\varepsilon \log[Mg(x)] \in L^\infty\} \tag{4.38}$$

if and only if there is no $\varepsilon < 1$ and $g \in L^1_{loc}$ such that $f - a\varepsilon \log(Mg(x)) \in L^\infty$, which is equivalent to saying that:

$$\nexists a' < a :\ f - a' \log(Mg(x)) \in L^\infty \tag{4.39}$$

which happens if and only if:

$$a \leq \inf\{\mu > 0 :\ \exists g(x) \in L^1_{loc},\ f(x) - \mu \log[Mg(x)] \in L^\infty\}$$

The proof is concluded by comparing (4.33) e (4.37).

□

4.5 Approximation by truncation and mollification in BLO

Of course, the previous sections implies that the closure of L^∞ in BLO can be characterized as:

$$f \in \overline{L^\infty}^{BLO} \iff \forall \mu > 0,\ e^{\frac{f}{\mu}} \in A_1. \tag{4.40}$$

Let us start this section by providing an example of an unbounded function in $\overline{L^\infty}^{BLO}$, proving $L^\infty \subset \overline{L^\infty}^{BLO}$ is strict.
Throughout this example we restrict the attention to the interval $J = (0, \frac{1}{e})$ and consider the function $f(x) = \log(-\log x)$. This function is well defined but not $L^\infty(J)$. However, we will prove $e^{rf(x)} = [-\log x]^r$ is A_1 on $(0, \frac{1}{e})$ for arbitrarily large r, concluding through (4.40) that it is in the closure of L^∞ in BLO.
To prove this, observe that $g_r(x) = [-\log x]^r$ is decreasing and always greater than 1 in J: this helps with the estimate of $A_1(g_r)$.
Indeed:

$$A_1(g_r) = \sup_{I \subset\subset J} \frac{\fint_I g_r(x)dx}{\inf_I g_r} = \sup_{[0,b] \subset\subset J} \frac{\fint_0^b g_r(x)dx}{\inf_{[0,b]} g_r} = \sup_{b \le 1/e} \frac{1}{g_r(b)} \fint_0^b g_r(x)dx. \tag{4.41}$$

This is because $g_r(x)$ is decreasing and for every interval $I = [a, b]$, by considering the interval $I' = [0, b]$ we have the obvious inequalities:

- $(g_r)_{I'} \ge (g_r)_I$
- $\inf_{I'} g_r = g_r(b) = \inf_I g_r$.

By taking $b = e^{-t}$ we have

$$A_1(g_r) = \sup_{t \ge 1} \frac{e^t}{t^r} \int_0^{e^{-t}} [-\log x]^r dx. \tag{4.42}$$

We will now show via the induction principle on the integer r that the quantity of which we are taking the supremum in (4.41) is decreasing with respect to t for every positive integer r. This will help us with computing its supremum.
As a matter of fact, the right hand side equals $1 + \frac{1}{t}$ for $r = 1$ so the base case of the induction principle is satisfied.
We will now procede by induction and assume it is decreasing for $r - 1$ and let us call $f_r(t) = \int_0^{e^{-t}} [-\log x]^r dx$. Using integration by parts we get the recursive formula $f_r(t) = \frac{t^r}{e^t} + r f_{r-1}(t)$ and then $\frac{e^t}{t^r} f_r(t) = 1 + \frac{r}{t}\left[\frac{e^t}{t^{r-1}} f_{r-1}(t)\right]$.
As the product of two positive decreasing functions is still decreasing, we proved that the right hand side in Equation (4.42) is decreasing in t for every n and this implies:

$$A_1(g_r) = \fint_0^{1/e} [-\log x]^r dx$$

because the supremum in 4.41 is obtained when $t = 1$.
One could compute $A_1(g_r)$ exactly with this strategy, but for us it is enough to observe

it is finite for every r, i.e. $A_1(e^{rf}) < \infty$ for every $r \in \mathbb{N}$ and $f(x) \in \overline{L^\infty}^{BLO}$ by the aforementioned characterization.

Let us call T the set of BLO functions f such that, if we define the truncated functions,

$$\{f\}_k = \begin{cases} k & \text{if } f(x) \geq k \\ f(x) & \text{if } -k \leq f(x) \leq k \\ -k & \text{if } f(x) \leq -k \end{cases}$$

we get $\lim_{k\to\infty} \|f - \{f\}_k\|_{BLO} = 0$.

We will now also show that both inclusions in

$$L^\infty \subset T \subset \overline{L^\infty}^{BLO} \tag{4.43}$$

are strict.

Our example of an unbounded function that is approximable by truncation is exactly the one from the previous section: in fact we will show that $f(x) = \log(-\log x) \in T$ even if it is unbounded, concluding $L^\infty \subset T$ is strict.

To do so, notice:

$$f - \{f\}_k = \begin{cases} f(x) - k & \text{if } f(x) \geq k \\ 0 & \text{if } f(x) \leq k \end{cases}$$

and since $f - \{f\}_k$ is non-negative, when we compute $\|f - \{f\}_k\|_{BLO}$ we can ignore all intervals having intersection with $\{x : f(x) < k\} = (e^{-e^k}, e^{-1})$.

In other words, $\|f - \{f\}_k\|_{BLO(0,1/e)} = \|f - k\|_{BLO(0,1/e^{e^k})} = \|f\|_{BLO(0,1/e^{e^k})}$.

Let us call $J_k = (0, 1/e^{e^k})$ so that our goal is to show that $\|f\|_{BLO(J_k)} \to 0$ as $k \to \infty$.

In [73], Garcia showed that $e^{\|f\|_{BLO}} \leq A_1(e^f)$.

In this case, since $e^{f(x)} = -\log x$:

$$\|f\|_{BLO(J_k)} \leq \log[A_1(-\log x)]_{J_k}$$

where the subscript J_k denotes the fact that we are considering the A_1 constant of $-\log x$ over J_k.

With the same strategy as before, we get:

$$[A_1(-\log x)]_{J_k} = \sup_{t \geq e^k} \frac{e^t}{t} \int_0^{e^{-t}} [-\log x] dx = \frac{e^{e^k}}{e^k} \int_0^{e^{-e^k}} [-\log x] dx = \frac{1 + e^k}{e^k}$$

so that:

$$0 \leq \lim_{k\to\infty} \|f - \{f\}_k\|_{BLO} \leq \lim_{k\to\infty} \log[A_1(-\log x)]_{J_k} = \lim_{k\to\infty} \log(1 + \frac{1}{e^k}) = 0.$$

proving the strictness of the first inclusion in (4.43).

We now need to show that there is a function which is approximable by bounded functions but not by truncation, i.e. a function $f \in \overline{L^\infty}^{BLO}$ which is not in T.

To do so we will first show that any function of the type

$$\eta(x) = \sum_{n=0}^{\infty} n\chi_{[a_{n+1},a_n)} \quad \text{with } a_n \text{ strictly decreasing to } 0 \tag{4.44}$$

is not in T and then we carefully choose $\{a_n\}$ in such a way that $\eta(x)$ is in $\overline{L^\infty}^{BLO}$.
To prove the aforementioned claim notice that, for a function $f(x)$ of the type in (4.44):

$$\|f - \{f\}_k\|_{BLO} = \|f\|_{BLO(0,a_k)} \geq 1, \quad \forall k \in \mathbb{N}. \tag{4.45}$$

This is because the BLO norm of a jump function is at least equal to the maximum size of its jumps: one can prove that it is by choosing intervals of the tipe $[a_{n+1}, a_n + \varepsilon]$ with sufficiently small $\varepsilon > 0$.
Our example of a function in $\overline{L^\infty}^{BLO}$ which is not approximable by truncation will then be of the type described in (4.44) with the choice $a_n = 1/e^{e^n}$ (considered on the interval J).
As a matter of fact, taking f to be

$$f(x) = \sum_{n=0}^{\infty} n\chi_{[1/e^{e^{n+1}},1/e^{e^n})}$$

let us compute $A_1(e^{rf})$. Notice that f is decreasing and with the usual reasoning:

$$A_1(e^{rf}) = \sup_{b\leq 1/e} \frac{⨍_0^b e^{rf(x)}dx}{e^{rf(b)}} \leq \sup_{n\in\mathbb{N}} \frac{⨍_0^{a_{n+1}} e^{rf(x)}dx}{e^{rn}}$$

where the second inequality comes from considering $a_{n+1} \leq b < a_n$ and overestimating the integral average over $[0, b]$ with the one over $[0, a_{n+1}]$ since f is decreasing.
Explicitely computing the integral via the definition gets to:

$$A_1(e^{rf}) = \sup_{n\in\mathbb{N}} \frac{1}{a_{n+1}} \sum_{i=n+1}^{\infty} e^{r(i-n)}[a_i - a_{i+1}]$$

and then substituting a_n we get

$$A_1(e^{rf}) = \sup_{n\in\mathbb{N}} e^{e^{n+1}} \sum_{j=1}^{\infty} e^{rj}\left[\frac{1}{e^{e^{n+j}}} - \frac{1}{e^{e^{n+j+1}}}\right] = \sum_{j=1}^{\infty} \frac{1}{e^{[e^n(e^j-e)-rj]}}.$$

Noting that the series on the right hand side is dominated by a convergent one that does not depend on n, i.e:

$$\sum_{j=1}^{\infty} \frac{1}{e^{[e^n(e^j-e)-rj]}} \leq \sum_{j=1}^{\infty} \frac{1}{e^{[e^j-e-rj]}} = K(r) < +\infty, \quad \forall r, n \in \mathbb{N}$$

shows that $f(x)$ lies in $\overline{L^\infty}^{BLO}$, but, like every other function of this type, is not approximable by truncation.

4.6 Coifman and Rochberg's decomposition of BMO

In this section, we present a proof by Coifman and Rochberg of the fact that any BMO function can be written as the difference of two BLO functions. At the end of the section we will also say something about a different proof, made possible nowadays a consequence of a decomposition result for the Muckenhoupt class A_2 that was not known at the time of paper [51] by Coifman and Rochberg, making use, again, of the connection between Muckenhoupt classes of weights and BMO and BLO. The way Cofiman and Rochberg proved it in 1980 is as a consequence of the following representation theorem for BMO functions due to L. Carleson [40]

Theorem 4.17 (Carleson, 1976, [40]). *Let φ be a non-negative Lipschitz function supported in the unit ball of $\mathbb{R}^n$ with and satisfying $\int_{\mathbb{R}^n} \varphi(y)\,dy = 1$. There are constants c1 and c_2 such that if $\varepsilon(y)$ is any measurable function and b_1 and b_2 are bounded functions then the function*

$$f(x) = b_1(x) + \int_{\mathbb{R}^n} \frac{1}{\varepsilon(y)^n} \varphi\left(\frac{x-y}{\varepsilon(y)}\right) b_2(y)\,dy \tag{4.46}$$

is in BMO and

$$\|f\|_{BMO} < c_1(\|b_1\|_\infty + \|b_2\|_\infty).$$

Conversely, if f is in BMO then f can be written in the form (4.46) *with functions b_1 and b_2 which satisfy*

$$\|b_1\|_\infty + \|b_2\|_\infty \leq c_2 \|f\|_{BMO}$$

Theorem 4.18 (Coifman, Rochberg, 1980, [51]). *A function f in is BMO if and only if it can be written as difference of two functions f_1, f_2 in BLO, i.e.*

$$f(x) = f_1(x) - f_2(x)$$

in a way such that

$$\|f_1\|_{BLO} + \|f_2\|_{BLO} \leq c\|f\|_{BMO}$$

for some absolute constant $c > 0$.

Proof. We will prove the result by decomposing the function b_2 in (4.46) as the difference of two non-negative functions $b_2 = h_1 - h_2$ and then showing that for every non-negative function $0 \leq h(x) \leq k$, we have that

$$g(x) = \int_{\mathbb{R}^n} \frac{1}{\varepsilon(y)^n} \varphi\left(\frac{x-y}{\varepsilon(y)}\right) h(y)\,dy$$

is a BLO function. To do so we are going to need the difference between the integral average and the infimum of $g(x)$ over the generic cube Q.
Fix any cube Q and denote by $\overline{Q}$ the concentric cube whose sidelength is 5 times larger. We will again split $h(x) = h_1(x) + h_2(x)$, where $h_1(x) = h(x)\chi_{\overline{Q}}(x)$ and $h_2(x) = h(x)[1 - \chi_{\overline{Q}}(x)]$ and $0 \leq h_1(x), h_2(x) \leq k$ will still hold and this will induce a splitting $g(x) = g_1(x) + g_2(x)$. Since

$$g_Q - \inf_Q(g) \leq (g_1)_Q + (g_2)_Q - \inf_Q(g) \tag{4.47}$$

we need an estimate of $(g_1)_Q$:

$$\fint_Q \int_{\mathbb{R}^n} \frac{1}{\varepsilon(y)^n}\varphi\left(\frac{x-y}{\varepsilon(y)}\right) h_1(y)\,dy\,dx =$$
$$= \frac{1}{|Q|}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{1}{\varepsilon(y)^n}\varphi\left(\frac{x-y}{\varepsilon(y)}\right) dx h_1(y)\,dy =$$
$$= \frac{1}{|Q|}\int_{\mathbb{R}^n} h_1(y)\,dy \le 5^n k.$$

where $\int \varphi = 1$ was used.
On the other hand, let x and x' be in Q, and let us estimate

$$|g_2(x) - g_2(x')| \le \int_{\mathbb{R}^n\setminus\overline{Q}} \left|\frac{1}{\varepsilon(y)^n}\left[\varphi\left(\frac{x-y}{\varepsilon(y)}\right) - \varphi\left(\frac{x'-y}{\varepsilon(y)}\right)\right]\right| dy$$

Now, since y is outside the cube $\overline{Q}$, which is 5 times larger than the cube Q in which x and x' are taken, there is going to be a positive constant α such that if $\varepsilon(y) < \alpha|x-y|$ then $\varepsilon(y) < |x'-y|$ and so the integral is going to be 0. This means we can continue estimating assuming that $\varepsilon \ge \alpha|x-y|$. In what follows, let us call L the Lipschitz constant of φ and D the diameter of the original cube Q:

$$|g_2(x) - g_2(x')| \le$$
$$\le L\int_{\mathbb{R}^n\setminus\overline{Q}} \frac{1}{\varepsilon(y)^{n+1}}|x-x'|\,dy \le \frac{LD}{\alpha^{n+1}}\int_{\mathbb{R}^n\setminus\overline{Q}} \frac{1}{|x-y|^{n+1}}\,dy \le$$
$$\le \frac{LD}{\alpha^{n+1}}\int_{\{|x-y|>D\}} \frac{1}{|x-y|^{n+1}}\,dy \le \beta k$$

with some positive constant $\beta > 0$.
So, continuing from (4.47) and noticing $\inf_Q(g) = \inf_Q(g_2)$ we have:

$$g_Q - \inf_Q(g) \le (g_1)_Q + (g_2)_Q - \inf_Q(g_2) \le (5^n + \beta)k < +\infty$$

proving that g is in BLO. This shows that the BMO function we started with can be expressed as difference of two BLO functions. The bounds on the norms are given by the bounds by Carleson, together with the fact that BLO continuously embeds into BMO and L^∞ continuously embeds into BLO.

□

The same result can also be obtained avoiding Carleson theorem by making use of a more recent result about Muckenhoupt classes, proved by Peter Jones in [95] stating that any weight in the Muckenhoupt class A_2 can be written as the ratio of two weights in A_1, with bounds on the A_1 constants of the factors. Combining this result with the aforementioned strong interplay between BMO and A_2 or BLO and A_1, which is an exponential interplay, the result is effortlessly obtained. Of course this connection goes both ways, so that the independent proofs of these strongly correlated results both give some deep insights into the topic, but for the sake of brevity we stop at this proof.

Chapter 5

Vanishing mean or lower oscillation: VMO and VLO

5.1 Definition of VMO and VLO and their relation to BMO and BLO

In this section we give the definition of a couple of relevant subclasses of BMO and BLO respectively. We will start with the space of functions with vanishing mean oscillation, i.e. VMO functions. The idea defining this space is that, together with an upper bound on how big the mean oscillation can be, we also request that the mean oscillation on VMO functions vanishes uniformly as the Lebesgue measure of the interval I on which it is computed goes to 0. This is an integral version of the concept of uniform continuity.

Definition 5.1 (Sarason, 1975, [130]). A BMO function $f(x)$ is said to have *Vanishing Mean Oscillation* ($f \in VMO(\mathbb{R})$) if it also satisfies:

$$V(f) = \limsup_{|I| \to 0} ⨍_I |f(x) - f_I| dx = 0. \tag{5.1}$$

The same idea can be used in defining the class of functions of vanishing lower oscillation, which we will denote by VLO.

Definition 5.2 (Korey, 2001, [108]). A BLO function is said to have *Vanishing Lower Oscillation* ($f \in VLO(\mathbb{R})$) if it also satisfies:

$$W(f) = \limsup_{|I| \to 0} [f_I - \inf_I f] = 0. \tag{5.2}$$

In the same way as BLO is not a vector space (but instead we might think of it as a subcone of BMO) because it is not in general true that the opposite of a function of bounded lower oscillation is also a function of bounded lower oscillation, the same can be said of VLO: an example will follow after the next proposition.
The following proposition gives one possible interpretation of the property VLO by connecting it to a classical function property: uniform continuity. We already said that a function is in BLO and in $-BLO$ at the same time if and only if it is essentially bounded. The cone $-BLO$ is sometimes referred to as BUO (Bounded Upper

Oscillation).
Here we show a simple lemma proving that a function is at the same time in VLO and in $-VLO$, this latter one sometimes referred to as VUO (Vanishing Upper Oscillation), if and only if it is bounded and uniformly continuous.

Proposition 5.1. *Let $f \in L^1_{loc}(\mathbb{R})$ be a function such that f and $-f$ belong to $VLO(\mathbb{R})$. Then f is in the space $BUC(\mathbb{R})$ of bounded and uniformly continuous functions.*

Proof. We already now that f is essentially bounded. Let us now write the conditions

$$W(f) = \limsup_{|I|\to 0}[f_I - \inf_I f] = 0. \tag{5.3}$$

and

$$W(-f) = \limsup_{|I|\to 0}[(-f)_I - \inf_I(-f)] = 0. \tag{5.4}$$

The latter of the two can be rewritten as

$$W(-f) = \limsup_{|I|\to 0} \sup_I(f) - f_I = 0. \tag{5.5}$$

hence, by putting it together with the former, one deduces

$$\limsup_{|I|\to 0} \omega_f(I) = 0 \tag{5.6}$$

where ω_f is classical oscillation of f as defined in (2.1), proving f is uniformly continuous and essentially bounded. To conclude, of course notice that, for a continuous function, an essential bound is the same thing as a bound. □

Exactly in the same way one proves $BLO \subset BMO$, it is also true that $VLO \subset VMO$. The above inclusion is strict, as shown by the example of the function

$$f(x) = -\sqrt{\log\left(\frac{1}{|x|}\right)}$$

which is in VMO but it is in $VLO \setminus -VLO$, so that $-f(x) \in VMO \setminus VLO$.

5.2 Distance in BMO and BLO to VMO and VLO

In his paper [130] from 1975, Donald Sarason proved, among other results, the following theorem, expressing the distance of a function in BMO from the subspace VMO of functions of vanishing mean oscillation.

Theorem 5.2 (Sarason, 1975, [130]). *There exists an absolute constant $a_1 > 0$ such that for every real valued function $f \in BMO(\mathbb{R})$ the following inequalities hold:*

$$V(f) \leq \operatorname*{dist}_{BMO}(f, VMO) \leq a_1 V(f) \tag{5.7}$$

where

$$V(f) = \limsup_{|I|\to 0} ⨍_I |f(x) - f_I| dx. \tag{5.8}$$

The aim of this section is to prove

Theorem 5.3 (Angrisani, 2017, [14]). *There exists an absolute constant $a_4 > 0$ such that for every real valued function $f \in BLO(\mathbb{R})$ the following inequalities hold:*

$$W(f) \leq \operatorname*{dist}_{BLO}(f, VLO) \leq a_4 W(f) \tag{5.9}$$

where:

$$W(f) = \limsup_{|I|\to 0} ⨍_I [f(x) - \inf_I f] dx. \tag{5.10}$$

This theorem is evidently analogous to Sarason's Theorem 5.7.
We adapt the proof of a lemma by Sarason (see [130]) to show that we also have:

Lemma 5.4. *If $h \in BLO$ and $\varphi > 0$ is in C_c, then $g(x) = (h \star \varphi)(x)$ is in VLO.*

Proof. Since $BLO \cap UC \subseteq VLO$ and $BLO \subseteq BMO$, since Sarason already proved that such a convolution is uniformly continuous, we only need to prove that g is in BLO.
First, notice that:

$$g_I = ⨍_I \int_{\mathbb{R}} \varphi(y) h(x-y) dy dx = \int_{\mathbb{R}} ⨍_I \varphi(y) h(x-y) dx dy = \int_{supp(\varphi)} \varphi(y) h_{I-y} dy \tag{5.11}$$

and

$$\inf_I g = \inf_I \int_{supp(\varphi)} \varphi(y) h(x-y) dy \geq \int_{supp(\varphi)} \varphi(y) \inf_{I-y} h dy \tag{5.12}$$

where this last inequality requires $\varphi \geq 0$.
From these it follows that:

$$g_I - \inf_I g \leq \int_{supp(\varphi)} \varphi(y)[h_{I-y} - \inf_{I-y} h] dy \leq \|\varphi\|_{L^1} \|h\|_{BLO}$$

so that $g \in BLO$.

$\square$

We remind that, for every $f \in BMO$:

$$W(f) = \limsup_{|I|\to 0} \fint_I \Big(f(x) - \inf_I f \Big)\, dx.$$

The following properties are easy to check:

- $W(f_1 + f_2) \le W(f_1) + W(f_2), \quad \forall f_1, f_2 \in BLO.$
- $W(g) = 0, \quad \forall g \in VLO.$
- $W(f) \le W(f - g), \quad \forall f \in BLO, g \in VLO.$
- $W(f) \le \|f\|_{BLO}, \quad \forall f \in BMO$

We are now ready to prove (5.9), using also techniques from [43].

Proof of Theorem 5.3. Using properties of $W(\cdot)$:

$$W(f) \le W(f - g) \le \|f - g\|_{BLO}, \quad \forall g \in VLO,\ \forall f \in BLO. \tag{5.13}$$

The first inequality we have to prove follows from (5.13) taking the infimum over all possible $g \in VLO$.
We get:

$$W(f) \le \operatorname*{dist}_{BLO}(f, VLO) \tag{5.14}$$

Note that of course $\inf\limits_{g \in VLO} \|f - g\|_{BLO} < +\infty$, since $VLO \cap -BLO \ne \emptyset$.
We will now prove that:

$$\inf_{g \in VLO} \|f - g\|_{BLO} \le a_4 W(f), \quad \forall f \in BLO. \tag{5.15}$$

Pick $\lambda > W(f)$. Then:

$$\exists \varepsilon > 0, \quad \forall I, |I| \le \varepsilon \Rightarrow \fint_I f(x) - \inf_I f dx \le \lambda.$$

We define:

$$\mathscr{F} = \{I_n\}_{n \in \mathbb{Z}}, \quad I_n = [n\frac{\varepsilon}{3}, (n+1)\frac{\varepsilon}{3}], \quad I_n^+ = I_{n-1} \cup I_n \cup I_{n+1}, \quad \forall n \in \mathbb{Z}$$

and

$$h(x) = \sum_{n \in \mathbb{Z}} \chi_{I_n} f_{I_n}.$$

Our first goal is to prove that the size of the discontinuities of $h(x)$ is not bigger than 12λ.
This is done observing that, for $j \in \{n-1, n, n+1\}$, we have:

$$|I_j|^{-1} \int_{I_j} |f(x) - f_{I_n^+}| dx \le 3 |I_n^+|^{-1} \int_{I_n^+} |f(x) - f_{I_n^+}| dx \le 6\lambda.$$

so that:

$$|f_{I_j} - f_{I_n^+}| \leq 6\lambda$$

which in turn implies:

$$|f_{I_j} - f_{I_k}| \leq 12\lambda$$

for $j, k \in \{n-1, n, n+1\}$.

We now want to show that $f \in BLO \Rightarrow h \in BLO$.
If $|I| \leq \frac{\varepsilon}{3}$, we have $h_I - \inf_I h = ⨍_I [h(x) - \inf_I h] dx \leq 12\lambda$.
If $|I| \geq \frac{\varepsilon}{3}$ we choose a minimal $\mathscr{F}' \subset \mathscr{F}$ with the property that $I \subset I' = \bigcup_{J \in \mathscr{F}'} J$.
It is easily proven that $\frac{|I'|}{|I|} \leq 3$ from which it follows:

$$h_I - \inf_I h = \frac{1}{|I|'} \frac{|I|'}{|I|} \int_I h(x) - \inf_{I'} h dx \leq 3 ⨍_{I'} h(x) - \inf_{I'} h dx. \tag{5.16}$$

Note also that

$$h_{I'} = \frac{1}{|I'|} \int_{I'} h(x) dx = \frac{1}{|I'|} \sum_{J \in \mathscr{F}'} \int_J f_J dx = \frac{1}{|I'|} \sum_{J \in \mathscr{F}'} \int_J f(x) dx = f_{I'} \tag{5.17}$$

and that $\inf_{I'} f \leq \inf_{I'} h$. From that it follows that:

$$h_I - \inf_I h \leq 3[f_{I'} - \inf_{I'} f] \leq 3\|f\|_{BLO}$$

and so $h \in BLO$.
Now define $g(x) = (h \star \varphi)(x)$, with $\varphi > 0$ in C_c such that $\|\varphi\|_{L^1} = 1$ and $supp(\varphi) \subset (-\frac{\varepsilon}{6}, \frac{\varepsilon}{6})$.
We have that $g \in VLO$ for Lemma 5.4. Using that $|h(x) - h(x-y)| \leq 12\lambda$ if $y \in supp(\varphi)$ one can prove that $\|h - h \star \varphi\|_\infty \leq 12\lambda$ and so:

$$\|h - g\|_{BLO} \leq 2\|h - g\|_\infty \leq 2 \cdot 12\lambda = 24\lambda \tag{5.18}$$

The last thing to do is to prove that

$$\|f - h\|_{BLO} \leq 13\lambda \tag{5.19}$$

and then the proof will be finished because of triangular inequality.
We divide the proof of (5.19) according to the measure of I.
If $|I| \leq \frac{\varepsilon}{3}$ we have:

$$[(f-h)_I - \inf_I (f-h)] \leq [f_I - \inf_I f] + [\sup_I h - h_I] \leq \lambda + 12\lambda = 13\lambda.$$

If $|I| \geq \frac{\varepsilon}{3}$ we consider $\mathscr{F}'$ and I' as before and we have that:

$$[(f-h)_I - \inf_I (f-h)] \leq (f-h)_I + \sup_{J \in \mathscr{F}'} \sup_J [f_J - f] \leq (f-h)_I + \lambda.$$

We then notice that:

$$(f-h)_I = |I|^{-1} \sum_{J \in \mathscr{F}'} \int_{I \cap J} f(x) - h(x) dx \leq 3|I'|^{-1} \sum_{J \in \mathscr{F}'} \int_J |f(x) - f_J| dx$$

and:

$$3|I'|^{-1} \sum_{J \in \mathscr{F}'} \int_J |f(x) - f_J| dx \leq 3|I'|^{-1} \sum_{J \in \mathscr{F}'} |J| \fint_J |f(x) - f_J| dx \leq$$
$$\leq 6\lambda |I'|^{-1} \sum_{J \in \mathscr{F}'} |J| = 6\lambda.$$

By combining (5.18) and (5.19) through the triangular inequality we conclude the proof of the theorem:

$$\operatorname*{dist}_{BLO}(f, VLO) \leq \|f-g\|_{BLO} \leq \|f-h\|_{BLO} + \|h-g\|_{BLO} \leq 13\lambda + 24\lambda = 37\lambda.$$

$\square$

5.3 Korey's decomposition of VMO

In the section "Coifman and Rochberg's decomposition of BMO" we talked about the decomposition

$$BMO = BLO - BLO$$

i.e. the characterization of BMO as the space of functions that can be written as the difference of two BLO functions, with some bounds on the norm of the two BLO components in terms of the BMO norm of the decomposed function.
In 2001, in the paper [108], M.B. Korey proved a similar result for VMO, showing that

$$VMO = VLO - VLO.$$

More precisely he proved

Theorem 5.5. *Each VMO function is the difference of two VLO functions. That is, if $f \in VMO(Q_0)$ then there exist $VLO(Q_0)$ functions F and G such that $f = F - G$ and*

$$\|F\|_{BLO} + \|G\|_{BLO} \leq C\|f\|_{BMO}.$$

The constant C depends only on the dimension

This result was reached by direct use of a powerful tool called Calderon-Zygmund decomposition and, as it is also explained in [108], a direct approach like the one used for decomposing BMO was not possible and Carleson's result was not a useful starting point anymore.
Also, it is first proven in the dyadic case, i.e. looking at cubes that are obtained from Q_0 by partitioning cube into 2^n subcubes and iterating this partitioning procedure some finite number of times. For example, if $n = 2$ and $Q_0 = [0,1]^2$, every cube of the type

$$\left[\frac{a}{2^k}, \frac{a+1}{2^k}\right] \times \left[\frac{b}{2^k}, \frac{b+1}{2^k}\right]$$

with $0 \leq a, b < 2^k$ and $k \in \mathbb{N}$ is a dyadic cube.
In other words, the result was first proven by Korey in the context of VMO^d and VLO^d which are analogues of VMO and VLO but defined using only dyadic cubes, and then extended to VMO and VLO with geometrical techniques.

We remark here also that M.B. Korey could not make direct use of the aforementioned Peter Jones decomposition result for A_2, stating an A_2 weight w can be decomposed as the quotient of two A_1 weights w_1, w_2. With respect to this topic, he in fact improved the original result (see [107]). He showed Peter Jones estimates on the A_1 weights w_1 and w_2 are asymptotically optimal for ideal weights: i.e, if $A_2(w) < 1+\varepsilon$, then the w_i can be found such that $A_1(w_i) \leq 1 + C\varepsilon^{\frac{1}{2}}$ for an absolute constant $C > 0$, but this is not true for any exponent larger than 1/2 for a suitably small choice of $\varepsilon > 0$.

5.4 Leibov's norm-attaining intervals in VMO

This section is dedicated to another interesting property that is peculiar of VMO functions in the BMO space. It was proven by Leibov while reaching a structure theorem for closed subspaces of VMO and it has to do with the definition of the BMO norm: it is of course expressed as a supremum, but Leibov needed to individuate, whenever possible, an interval I^* attaining the supremum, i.e. such that

$$\fint_{I^*} |f - f_{I^*}|\, dt \geq \fint_I |f - f_I|\, dt$$

for any other interval I. In particular, in [110], Leibov was able to prove that this property holds for VMO functions

Lemma 5.6 (Leibov, 1990, [110])**.** *If $f \in VMO([0,1])$ then there exists and interval $I^* \subseteq [0,1]$ such that*

$$\|f\|_{BMO} = \fint_{I^*} |f - f_{I^*}| dt \tag{5.20}$$

In this section we will dissect and analyse the proof of this lemma, even if it is relatively straightforward, with the purpose of proving an analogue result in the next section.
To do so, it will be convenient to introduce some notation.
We will refer to an interval in terms of its center x and half-lenght h as in:

$$I_h^x := [x - h, x + h].$$

Let us fix $h \in \left(0, \frac{1}{2}\right]$ and define $S_h = [h, 1 - h]$: we have that $I_h^x \subseteq [0,1]$ if and only if $x \in S_h$. Thus we can define the set

$$T = \left\{(x,h) \in \mathbb{R}^2 :\ h \in \left(0, \frac{1}{2}\right],\ x \in S_h\right\} \tag{5.21}$$

such that $I_h^x \subseteq [0,1]$ if and only if $(x,h) \in T$. Finally, for a generic function $f \in BMO([0,1])$, let us define the function $F : T \to \mathbb{R}$ given by

$$F(x,h) = \fint_{I_h^x} |f - f_{I_h^x}| dt, \tag{5.22}$$

which is always a continuous function since $f \in L^1([0,1])$. Thus the norm on $BMO([0,1])$ can be also defined as:

$$\|f\|_{BMO} = \sup_{(x,h)\in T} F(x,h). \tag{5.23}$$

We can also restate the VMO property in terms of F. We have that:

$$f \in VMO \iff \lim_{h\to 0} \sup_{x \in S_h} F(x,h) = 0 \tag{5.24}$$

that is to say that F converges to 0 as $h \to 0$ uniformly with respect to x.
The idea of the proof of the lemma by Leibov is to notice that F is continuous, and

since f is in VMO, it can be extended by continuity to the closure of T, namely:

$$\tilde{T} = \left\{ (x,h) \in \mathbb{R}^2 : \ h \in \left[0, \frac{1}{2}\right], \ x \in S_h \right\} \tag{5.25}$$

by posing $F = 0$ on $[0,1] \times 0$.
A straightforward application of Weierstrass theorem concludes the proof. In particular three main ingredients emerge:

- The compactness of $\widetilde{T}$;
- The continuity of F on T;
- The fact that if $f \in VMO([0,1])$ then F can be extended with continuity of the whole $\widetilde{T}$.

5.5 Norm-attaining intervals in VLO

The aim of this section is to find an analogue of Lemma 5.6 by Leibov in the subclass of BLO-functions. In order to follow the same sketch of proof as described in the previous section and obtain the same result for functions in $VLO([0,1])$ with respect to the norm in $BLO([0,1])$ we will need to write the norm of a function f in such space in terms of a suitable two variables function F and then assure this three hypotheses. We will see that even the second one is not necessarly satisfied by functions in $BLO([0,1]) \setminus VLO([0,1])$.
By using the same notation in the previous section for I_h^x, S_h and T, for a generic function $f \in BLO([0,1])$ let us define the function $F : T \to \mathbb{R}$ to be

$$F(x,h) = \fint_{I_h^x} f dt - \inf_{I_h^x} f \tag{5.26}$$

Thus the norm on $BLO([0,1])$ can be also defined as

$$\|f\|_{BLO} = \sup_{(x,h)\in T} F(x,h). \tag{5.27}$$

In terms of F, we have that $f \in VLO([0,1])$ if and only if

$$\lim_{h\to 0} \sup_{x\in S_h} F(x,h) = 0 \tag{5.28}$$

that is to say that F converges to 0 as $h \to 0$ uniformly with respect to x.
However, this time it is not true that F is continuous in T for any $f \in BLO([0,1])$. In particular, let us observe that for any $f \in BLO([0,1])$ we can define:

$$G(x,h) = \fint_{I_h^x} f dt \tag{5.29}$$

$$H(x,h) = \inf_{I_h^x} f \tag{5.30}$$

to obtain that

$$F(x,h) = G(x,h) + H(x,h) \tag{5.31}$$

thus, since G is continuous in T (since $f \in L^1([0,1])$), then F is continuous if and only if H is continuous.
Let us consider for instance the function

$$f(x) = \begin{cases} 0 & 0 \le x \le \frac{1}{2} \\ 1 & \frac{1}{2} < x \le 1. \end{cases} \tag{5.32}$$

We have that $f \in BLO([0,1])$ since $f \in L^\infty$, but we also have that

$$H(x,h) = \begin{cases} 0 & h \le x < \frac{1}{2} + h \\ 1 & \frac{1}{2} + h \le x \le 1 - h \end{cases} \tag{5.33}$$

so H is not continuous and then also F is not continuous. In particular, let us observe that since f admits a discontinuity jump, $f \notin VLO([0,1])$.
Moreover, let us observe that

$$G(x,h) = \begin{cases} 0 & h \leq x \leq \frac{1}{2} - h \\ \frac{2x+2h-1}{4h} & \frac{1}{2} - h < x \leq \frac{1}{2} + h \\ 1 & \frac{1}{2} + h < x \leq 1 - h \end{cases} \tag{5.34}$$

and then

$$F(x,h) = \begin{cases} 0 & x \in \left[h, \frac{1}{2} - h\right) \cup \left[\frac{1}{2} + h, 1 - h\right] \\ \frac{2x+2h-1}{4h} & x \in \left[\frac{1}{2} - h, \frac{1}{2} + h\right) \end{cases} \tag{5.35}$$

for which $\sup_{(x,h)\in T} F = 1$ but $F(x,h) < 1$ for any $(x,h) \in T$.
Let us return to our aim, which is to show an analogue of Leibov's Lemma for VLO functions. To do so we need to prove that, at least for a function f in VLO, the corresponding F is continuous.

Proposition 5.7 (Angrisani, Ascione, 2018, [15])**.** *Let $f \in VLO([0,1])$. Then*

$$H(x,h) = \inf_{I_x^h} f$$

is continuous in T.

Proof. Let us prove this assertion by contradiction. Fix $(x,h) \in T$ and let us suppose there exists a $\bar{\varepsilon} > 0$ such that for any $\delta > 0$ there exists a point (y_δ, k_δ) such that

$$|x - y_\delta| + |h - k_\delta| \leq \delta \tag{5.36}$$

and

$$|H^f(x,h) - H^f(y_\delta, k_\delta)| \geq \bar{\varepsilon}. \tag{5.37}$$

To fix the ideas, let us consider $\delta_n = \frac{1}{n}$ and let us denote $y_n := y_{\delta_n}$, $k_n := k_{\delta_n}$, $I := I_h^x$ and $I_n := I_{k_n}^{y_n}$. Moreover, we can suppose n is big enough to have $I \cap I_n \neq \emptyset$.
Equation (5.37) assures that

$$H^f(x,h) = \inf_I f > \inf_{I_n \cup I} f =: m_n. \tag{5.38}$$

and that

$$\inf_{I \cap I_n} f \neq m_n \tag{5.39}$$

thus

$$m_n = \inf_{I \Delta I_n} f \tag{5.40}$$

where $I\Delta I_n = (I \setminus I_n) \cup (I_n \setminus I)$. Let us suppose that $I_n \setminus I \neq \emptyset$ and $m_n = \inf_{I_n \setminus I} f$. Now let us observe that $I_n \setminus I$ has at most two connected components $I_1 = [x+h, y_n + k_n]$ and $I_2 = [y_n - k_n, x - h]$. Thus let us suppose that $m_n = \inf_{I_1} f$.
Let us suppose that n is big enough to have $y_n + k_n < x + 3h$ and then let us consider the interval

$$I_1^* = [2(x+h) - y_n - k_n, y_n + k_n] \tag{5.41}$$

on which we have that $\inf_{I_1^*} f = m_n$. Moreover, let us observe that by construction $|I_1^* \cap I| = |I_1|$ and $|I_1^*| = 2|I_1|$. Now, since we have supposed that $m_n = \inf_{I_n \setminus I} f$, then

$$m_n = H(y_n, k_n) \le H(x,h) \tag{5.42}$$

and

$$H(x,h) \ge m_n + \overline{\varepsilon}. \tag{5.43}$$

By definition we have

$$\inf_{I_1^* \cap I} f \ge H(x,h) \ge m_n + \overline{\varepsilon}. \tag{5.44}$$

Now let us observe that

$$\int_{I_1^*} f dt = \int_{I_1^* \cap I} f dt + \int_{I_1} f dt \ge (2m_n + \overline{\varepsilon})|I_1| \tag{5.45}$$

and then

$$\fint_{I_1^*} f dt \ge m_n + \frac{\overline{\varepsilon}}{2} \tag{5.46}$$

so that

$$\fint_{I_1^*} f dt - \inf_{I_1^*} \ge \frac{\overline{\varepsilon}}{2}. \tag{5.47}$$

In such case, let us denote with $I_n^* := I_1^*$. If $m_n = \inf_{I_2} f$, we can construct in a similar way an interval I_2^* on which we have (5.47) and we can pose $I_n^* := I_2^*$. The construction of I_n^* can be done in the same way if $I \setminus I_n \neq \emptyset$ and $m_n = \inf_{I \setminus I_n} f$. Thus, for any n there exists an interval I_n^* with length $|I_n^*| = y_n + k_n - x - h$ such that

$$\fint_{I_n^*} f dt - \inf_{I_n^*} \ge \frac{\overline{\varepsilon}}{2} \tag{5.48}$$

that is a contradiction with the fact that $f \in VLO([0,1])$ since $|I_n^*| \to 0$. $\square$

Now we can prove the main result

Proposition 5.8 (Angrisani, Ascione, 2018, [15]). *If $f \in VLO([0,1])$ then there exists an interval $I^* \subseteq [0,1]$ such that*

$$\|f\|_{BLO} = \fint_{I^*} f dt - \inf_{I^*} f \tag{5.49}$$

Proof. Since $f \in VLO([0,1])$, H (and then F) is continuous on T. Moreover, since F converges to 0 as $h \to 0$ uniformly with respect to x, we can extend F with continuity to

$$\widetilde{T} = \left\{ (x,h) \in \mathbb{R}^2 : \ h \in \left[0, \frac{1}{2}\right], \ x \in S_h \right\} \tag{5.50}$$

by posing $F(x,0) = 0$. Thus, since F is continuous on $\widetilde{T}$ that is compact, there exists a point $(x^*, h^*) \in \widetilde{T}$ such that

$$\sup_{(x,h)\in T} F(x,h) = \sup_{(x,h)\in \widetilde{T}} F = F(x^*, h^*) \tag{5.51}$$

and in particular we have $I^* = I_{h^*}^{x^*}$. □

Let us conclude this section by showing that $f \in VLO([0,1])$ is a sufficient but not necessary condition. Consider

$$f(x) = \begin{cases} 4x-1 & x \in \left[0, \frac{1}{2}\right] \\ 0 & x \in \left(\frac{1}{2}, 1\right]. \end{cases} \tag{5.52}$$

First let us observe that $f \in BLO([0,1])$ since $f \in L^\infty([0,1])$. Then let us consider the intervals $I_h = \left[\frac{1}{2} - h, \frac{1}{2} + h\right]$ for $h < \frac{1}{4}$. Then

$$\lim_{h \to 0} \fint_{I_h} f dt - \inf_{I_h} f = \lim_{h \to 0} h - \frac{1}{2} = -\frac{1}{2} \tag{5.53}$$

so $f \notin VLO([0,1])$. Let us consider an interval $I \subseteq [0,1]$:

- If $I \subseteq \left[0, \frac{1}{2}\right]$ then, posing $I = [x-h, x+h]$,

$$\inf_I f = 4x - 4h - 1 \tag{5.54}$$

  and

$$\fint_I f dt = 4x - 1 \tag{5.55}$$

  hence

$$\fint_I f dt - \inf_I f = 4h. \tag{5.56}$$

  But if $I \subseteq [0, \frac{1}{2}]$, then $h \in \left[0, \frac{1}{4}\right]$ so we have on these intervals

$$\sup_I \fint_I f dt - \inf_I f = 1. \tag{5.57}$$

- If $I \subseteq \left[\frac{1}{2}, 1\right]$, then we have

$$\fint_I f dt - \inf_I f = 0. \tag{5.58}$$

- In the other cases, we can always suppose that $I \subseteq \left[\frac{1}{4}, 1\right]$, since if it is not true then we can construct $\widetilde{I} = I \cap \left[0, \frac{1}{2}\right]$ such that $\inf_I f = \inf_{\widetilde{I}} f$ and

$$\fint_I f dt \leq \fint_{\widetilde{I}} f dt \tag{5.59}$$

  and then I is not influential to determine the norm of f in BLO. If we consider $I \subseteq \left[\frac{1}{4}, 1\right]$ then $\inf_I f = 0$. Posing $I = [a, b]$, then we can consider $\widetilde{I} = \left[a, \frac{1}{2}\right]$ to obtain that

$$\fint_I f dt \leq \fint_{\widetilde{I}} f dt = 2a \leq 1. \tag{5.60}$$

With these considerations, we can conclude that $\|f\|_{BLO} = 1$.
However, posing $I^* = \left[0, \frac{1}{2}\right]$, we have

$$\fint_{I^*} f dt - \inf_{I^*} f = 0 + 1 = 1 = \|f\|_{BLO} . \tag{5.61}$$

Ringraziamenti

Innanzitutto voglio ringraziare il mio maestro, relatore di tesi triennale e magistrale e tutor di dottorato, il professor Carlo Sbordone, per avermi trasmesso la passione per l'analisi matematica fin da quando ero studentessa, per essere stato sempre il mio faro ed avermi guidato con grande professionalità, pazienza ed umanità.
Ringrazio poi la professoressa Gioconda Moscariello che è stata mia tutor di dottorato nell'ultima fase del mio percorso e mi ha dato molti preziosi consigli sulla stesura di questa tesi.
Ringrazio il professor Franco Rampazzo che, dopo aver catturato la mia attenzione con un corso introduttivo alla teoria controllo ottimo, mi ha offerto la possibilità di approfondire la conoscenza di questo settore sotto la sua guida e di collaborare con lui a lavori in questo ambito.
Un ringraziamento va anche al professor Carlo Mantegazza, un punto di riferimento importante per me da ormai 4 anni, per avermi ascoltata, seguita e consigliata in tanti contesti.
Ringrazio poi le professoresse Chiara Leone ed Anna Verde per avermi presentato la teoria della regolarità, seguita nei miei studi riguardanti quest'ambito ed incoraggiata in tante situazioni.
Un altro ringraziamento va alla professoressa Cristina Trombetti, che in qualità di direttrice del dipartimento, ha sempre tenuto in grande considerazione le esigenze mie e degli altri dottorandi ed è stata sempre pronta a darci utilissimi consigli.
Voglio ringraziare poi le professoresse Raffaella Giova ed Antonella Passarelli che sono state sempre estremamente disponibili nei miei confronti.
Un grazie va anche Giacomo e Gianluigi con i quali ho condiviso varie avventure, parlando di spazi di funzioni anche a Natale e a Ferragosto.
Infine ringrazio i referees per la loro collaborazione diretta in correzioni, consigli e suggerimenti.
Altri ringraziamenti vanno ai miei amici. In particolare ringrazio:
Leonardo, che per me è praticamente un fratello, per essermi stato sempre vicino nelle gioie e nelle difficoltà, avermi sempre compresa nel profondo e molte volte tirata su di morale.
Francesco perché negli ultimi due anni è diventato un amico importante, mi ha dato consigli preziosi e sinceri e mi ha mostrato un lato di sé che pochi conoscono.
Niccolò, Gloria, Mattia, Gianpaolo, Claudio, Domenico, Giovanni, Dario e di nuovo Leonardo per le serate scanzonate in compagnia.
Andrea, perché so, che anche se non ci vediamo più spesso come una volta, su di lui posso sempre contare.
Marisa, Rossella e Valeria per le mille avventure insieme e perché sono le migliori

amiche che potessi desiderare.
Eleonora, mia cugina ed amica, perché è il mio alter ego e quindi è mia complice nel bene e nel male.
Chiara perché le due volte all'anno che ci vediamo riesce sempre a farmi sentire come se fossi a casa ed è la mia amica psicologa di fiducia.
Rosamaria e Roberta perché ci sono sempre state ed hanno condiviso con me mille gioie ed ansie.
Livia perché ci sopportiamo a vicenda da 23 anni e mai smetteremo di farlo.
Tutti gli altri cari amici che non cito per brevità.
Ringrazio poi mia suocera Rosanna che mi tratta sempre come una figlia e il mio straordinario suocero Tonino che fino al suo ultimo giorno è stato sempre un punto di riferimento fondamentale ed è sempre riuscito a rassicurarmi come nessun altro. Ringrazio anche Gigi, Catia, Daniele, Miriam e le loro splendide bimbe perché sono la mia seconda famiglia.
Ringrazio anche la mia famiglia:
Papà perché è e sempre sarà il mio modello e perché nonostante sembri un po' burbero ha saputo starmi vicino in ogni brutto momento come nessun altro.
Mamma che si preoccupa sempre per me ben più di quanto sia necessario.
I miei fratelli Carlo, Alberto e Basilio che mi salvano sempre dai guai, i miei cugini Maria Teresa, Carlo e Claudia con cui condivido molto ed i loro simpaticissimi bambini. Mio fratello Carlo è anche uno dei miei amici più preziosi e pur avendo 7 anni meno di me ha tanto da insegnarmi.
La mia nonna Teresa che non perde occasione per dimostrare il suo affetto nei miei confronti ed il mio indimenticabile nonno Carlo per cui ero davvero, come lui diceva, una quarta figlia.
Il mio cane Bear e la mia gatta Birba che hanno riempito la mia quotidianità durante questo dottorato di gioia e tenerezza ed il mio dolcissimo gatto salernitano Chopin che amava impedirmi di studiare distendendosi sui miei libri di analisi.
Ringrazio infine il mio fidanzato Dario che, nonostante le asperità del mio carattere e le mie ansie, mi ha sempre profondamente amata. Da quasi sei anni è parte integrante della mia quotidianità, mi sostiene sempre e mi regala felicità autentica ogni giorno come nessun altro potrebbe fare. A lui dedico questo lavoro di tesi.

Bibliography

[1] W. Abu-Shammala and A. Torchinsky. The atomic decomposition in $L^1(\mathbb{R}^n)$. *Proceedings of the American Mathematical Society*, 135(9):2839–2843, 2007.

[2] E. Acerbi and N. Fusco. A regularity theorem for minimizers of quasiconvex integrals. *Arch. Rational Mech. Anal.*, 99(3):261–281, 1987.

[3] E. Acerbi and N. Fusco. Regularity for minimizers of nonquadratic functionals: the case $1 < p < 2$. *J. Math. Anal. Appl.*, 140(1):115–135, 1989.

[4] A. A. Agrachev, A. S. Morse, E. D. Sontag, H. J. Sussmann, and V. I. Utkin. *Nonlinear and optimal control theory*, volume 1932 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin; Fondazione C.I.M.E., Florence, 2008. Lectures given at the C.I.M.E. Summer School held in Cetraro, June 19–29, 2004, Edited by Paolo Nistri and Gianna Stefani.

[5] N. Aïssaoui. Another extension of Orlicz-Sobolev spaces to metric spaces. *Abstr. Appl. Anal.*, (1):1–26, 2004.

[6] T. Alberico and C. Sbordone. A precise interplay between BMO space and A_2 class in dimension one. *Mediterr. J. Math.*, 4(1):45–51, 2007.

[7] L. Ambrosio, J. Bourgain, H. Brezis, and A. Figalli. BMO-type norms related to the perimeter of sets. *Comm. Pure Appl. Math.*, 69(6):1062–1086, 2016.

[8] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems*, volume 1. Oxford Science Publications, 2000.

[9] L. Ambrosio and N. Gigli. *A User's Guide to Optimal Transport*, pages 1–155. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[10] L. Ambrosio and D. Puglisi. Linear extension operators between spaces of Lipschitz maps and optimal transport. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2016.

[11] L. Ambrosio and P. Tilli. *Topics on analysis in metric spaces*, volume 25 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2004.

[12] F. Angrisani. Autonomous and asymptotically quasiconvex functionals with general growth conditions. *preprint*.

[13] F. Angrisani. A new norm on BLO and matters of approximability. *preprint.*

[14] F. Angrisani. On the distance in $BLO(\mathbb{R})$ to $L^\infty(\mathbb{R})$ and $VLO(\mathbb{R})$. *Rendiconto dell'Accademia delle Scienze Fisiche e Matematiche di Napoli*, (84):75–86, 2017.

[15] F. Angrisani and G. Ascione. Note on VLO functions. *Rendiconto dell'Accademia delle Scienze Fisiche e Matematiche di Napoli*, (85):177–183, 2018.

[16] F. Angrisani, G Ascione, L. D'Onofrio, and G. Manzo. Duality and distance formulas in lipschitz-holder spaces. *Rend. Lincei*, 31(2):401–419, 2020.

[17] F. Angrisani, G. Ascione, C. Leone, and C. Mantegazza. *Appunti di Calcolo delle Variazioni.* Amazon, 2019.

[18] F. Angrisani, G. Ascione, and G. Manzo. Orlicz spaces with a o–O type structure. *Ricerche di Matematica*, pages 1–17, 2019.

[19] F. Angrisani, G. Ascione, and G. Manzo. Atomic decomposition of finite signed measures on compacts of $\mathbb{R}^n$.. *Ann. Acad. Sci. Fenn. Math.*, 2021.

[20] F. Angrisani and F. Rampazzo. Set-valued lie brackets and non-smooth higher order maximum principle. *preprint.*

[21] M. S. Aronna, M. Motta, and F. Rampazzo. A higher-order maximum principle for impulsive optimal control problems. *SIAM J. Control Optim.*, 58(2):814–844, 2020.

[22] J.P. Aubin and A. Cellina. *Differential inclusions: set-valued maps and viability theory*, volume 264. Springer, 1984.

[23] W.G. Bade, P.C. Curtis Jr, and H.G. Dales. Amenability and weak amenability for Beurling and Lipschitz algebras. *Proceedings of the London Mathematical Society*, 3(2):359–377, 1987.

[24] C. Bennett and K. Rudnick. *On Lorentz-Zygmund spaces.* 1980.

[25] C. Bennett and R.C. Sharpley. *Interpolation of operators*, volume 129. Academic press, 1988.

[26] H. Berninger and D. Werner. Lipschitz spaces and M-ideals. *Extracta mathematicae*, 18(1):33–56, 2003.

[27] N. .H Bingham, C.M. Goldie, and J.L. Teugels. *Regular variation*, volume 27. Cambridge university press, 1989.

[28] V.I. Bogachev. *Measure Theory*, volume 1. Springer, 2007.

[29] B. Bojarski. Remarks on some geometric properties of Sobolev mappings. In *Functional analysis & related topics (Sapporo, 1990)*, pages 65–76. World Sci. Publ., River Edge, NJ, 1991.

[30] F. F. Bonsall. A general atomic decomposition theorem and Banach's closed range theorem. *The Quarterly Journal of Mathematics*, 42(1):9–14, 1991.

[31] G. Bouchitté, T. Champion, and C. Jimenez. Completion of the space of measures in the Kantorovich norm. *Riv. Mat. Univ. Parma*, 7(4):127–139, 2005.

[32] J. Bourgain, H. Brezis, and P. Mironescu. A new function space and applications. *J. Eur. Math. Soc. (JEMS)*, 17(9):2083–2101, 2015.

[33] D. Breit and A. Verde. Quasiconvex variational functionals in Orlicz-Sobolev spaces. *Ann. Mat. Pura Appl. (4)*, 192(2):255–271, 2013.

[34] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Springer Science & Business Media, 2010.

[35] S. Campanato. Proprietà di hölderianità di alcune classi di funzioni. *Ann. Scuola Norm. Sup. Pisa*, 17:175–188, 1963.

[36] S. Campanato. Sistemi ellittici in forma di divergenza. regolarità all'interno. *Quaderni della Scuola Normale Superiore, Pisa*, 1980.

[37] C. Capone and M.R. Formica. The distance to L^∞ from the Morrey space $L^{p,\lambda}$. *Rend. Accad. Sci. Fis. Mat. Napoli (4)*, 62:291–299 (1996), 1995.

[38] C. Capone and M.R. Formica. A decomposition of the dual space of some Banach function spaces. *Journal of Function Spaces and Applications*, 2012, 2012.

[39] D. Carando and S. Lassalle. Duality, reflexivity and atomic decompositions in Banach spaces. *Studia Math*, 191(1):67–80, 2009.

[40] L. Carleson. Two remarks on H^1 and BMO. *Advances in Math.*, 22(3):269–277, 1976.

[41] M. Carozza, N. Fusco, and G. Mingione. Partial regularity of minimizers of quasiconvex integrals with subquadratic growth. *Ann. Mat. Pura Appl. (4)*, 175:141–164, 1998.

[42] M. Carozza and R. Mellone. The distance to L^∞ or to VMO for BMO functions. *Ricerche di Matematica*, 44(2):439–448, 1995.

[43] M. Carozza, A. Passarelli Di Napoli, T. Schmidt, and A. Verde. Local and asymptotic regularity results for quasiconvex and quasimonotone problems. *Q. J. Math.*, 63(2):325–352, 2012.

[44] M. Carozza and C. Sbordone. The distance to L^∞ in some function spaces and applications. *Differential and Integral Equations*, 10(4):599–607, 1997.

[45] A. Caruso. Two properties of norms in Orlicz spaces. *Le Matematiche*, 56(1):183–194, 2001.

[46] S. Chen. *Geometry of Orlicz spaces*. Polska Akademia Nauk, Instytut Matematyczny, 1996.

[47] S. Chen and H. Hudzik. E^ϕ and h^ϕ fail to be M-ideals in L^ϕ and l^ϕ in the case of the Orlicz norm. *Rendiconti del Circolo Matematico di Palermo*, 44(2):283–292, 1995.

[48] F.H. Clarke, Y.S. Ledyaev, R. J. Stern, and R. R. Wolenski. *Nonsmooth Analysis and Control Theory*, volume 178. Springer, 1998. Graduate Texts in Mathematics.

[49] J. A. Clarkson. Uniformly convex spaces. *Transactions of the American Mathematical Society*, 40(3):396–414, 1936.

[50] S. Cobzas, R. Micolescu, and A. Nicolae. *Lipschitz Functions*. Spinger, 2019.

[51] R. Coifman and R. Rochberg. Another characterization of bmo. *Proceedings of the American Mathematical Society*, 2(79):249–254, 1980.

[52] R. R. Coifman and G. Weiss. *Analyse harmonique non-commutative sur certains espaces homogènes: étude de certaines intégrales singulières*, volume 242. Springer, 2006.

[53] R.R. Coifman and G. Weiss. Extensions of hardy spaces and their use in analysis. *Bulletin of the American Mathematical Society*, 83(4):569–645, 1977.

[54] G. Dafni, T. Hytönen, R. Korte, and H. Yue. The space JN_p: nontriviality and duality. *Journal of Functional Analysis*, 275(3):577–603, 2018.

[55] E. De Giorgi. Sulla differenziabilit'a e l'analiticit'a delle estremali degli integrali multipli regolari. *Mem. Accad. Sci. Cl. Sci. Fis. Mat. Natur*, (3):25–43, 1957.

[56] K. De Leeuw. Banach spaces of Lipschitz functions. *Studia Mathematica*, 1(21):55–66, 1961.

[57] E. Di Nezza, G. Palatucci, and E. Valdinoci. Hitchhiker's guide to the fractional Sobolev spaces. *Bulletin des Sciences Mathématiques*, 136(5):521–573, 2012.

[58] L. Diening, D. Lengeler, B. Stroffolini, and A. Verde. Partial regularity for minimizers of quasiconvex functionals with general growth. *SIAM Journal on Mathematical Analysis*, 44(5):3594–3616, 2012.

[59] L. D'Onofrio, L. Greco, K. M. Perfekt, C. Sbordone, and R. Schiattarella. Atomic decompositions, two stars theorems, and distances for the Bourgain–Brezis–Mironescu space and other big spaces. In *Annales de l'Institut Henri Poincaré C, Analyse non linéaire*. Elsevier, 2020.

[60] L. D'Onofrio, C. Sbordone, and R. Schiattarella. Duality and distance formulas in Banach function spaces. *Journal of Elliptic and Parabolic Equations*, pages 1–23, 2018.

[61] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions*. Chapman and Hall/CRC, 2015.

[62] L.C. Evans. Quasiconvexity and partial regularity in the calculus of variations. *Archive for Rational Mechanics and Analysis*, (95):227–252, 1986.

[63] John F. and Nirenberg L. On functions of bounded mean oscillation. *Communications on pure and applied Mathematics*, 3(14):415–426, 1961.

[64] F. Farroni and R. Giova. Quasiconformal mappings and sharp estimates for the distance to L^∞ in some function spaces. *Journal of Mathematical Analysis and Applications*, 395(2):694–704, 2012.

[65] F. Farroni and R. Giova. The distance to L^∞ in the grand Orlicz spaces. *J. Funct. Spaces Appl.*, pages Art. ID 658527, 7, 2013.

[66] C. Fefferman and E. M. Stein. H^p spaces of several variables. *Acta Math.*, 129(3-4):137–193, 1972.

[67] E. Feleqi and F. Rampazzo. Iterated Lie brackets for nonsmooth vector fields. *NoDEA Nonlinear Differential Equations Appl.*, 24(6):Paper No. 61, 43, 2017.

[68] K. Fey and M. Foss. Morrey regularity results for asymptotically convex variational problems with (p, q) growth. *J. Differential Equations*, 246(12):4519–4551, 2009.

[69] I. Fonseca and J. Malý. Relaxation of multiple integrals below the growth exponent. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 14(3):309–338, 1997.

[70] M.R. Formica. The distance to Lip in the space $C^{0,\alpha}$ of Hölder continuous functions. *Ricerche Mat.*, 54(1):127–135 (2006), 2005.

[71] N. Fusco, G. Moscariello, and C. Sbordone. A formula for the total variation of *SBV* functions. *J. Funct. Anal.*, 270(1):419–446, 2016.

[72] N. Fusco, G. Moscariello, and C. Sbordone. *BMO*-type seminorms and Sobolev functions. *ESAIM Control Optim. Calc. Var.*, 24(2):835–847, 2018.

[73] J. García-Cuerva and J. L. Rubio de Francia. *Weighted norm inequalities and related topics*, volume 116 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1985. Notas de Matemática [Mathematical Notes], 104.

[74] J.B. Garnett and P.W. Jones. The distance in BMO to L^∞. *Ann. of Math. (2)*, 108(2):373–393, 1978.

[75] T. Gasymov and C. Hashimov. On an atomic decomposition in Banach spaces. *Sahand Communications in Mathematical Analysis*, 9(1):15–32, 2018.

[76] M. Giaquinta. *Multiple integrals in the calculus of variations and nonlinear elliptic systems*, volume 105 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1983.

[77] E. Giusti. *Direct methods in the calculus of variations*. World Scientific Publishing Co., Inc., River Edge, NJ, 2003.

[78] A. Gogatishvili, P. Koskela, and N. Shanmugalingam. Interpolation properties of Besov spaces defined on metric spaces. *Mathematische Nachrichten*, 283(2):215–231, 2010.

[79] A. Grigor'yan and L. Liu. Heat kernel and Lipschitz–Besov spaces. In *Forum Mathematicum*, volume 27, pages 3567–3613. De Gruyter, 2015.

[80] M. Hairer. An introduction to stochastic PDEs. *arXiv preprint arXiv:0907.4178*, 2009.

[81] P. Hajłasz. *Sobolev spaces on metric-measure spaces*, volume 338 of *Contemp. Math.*, pages 173–218. Amer. Math. Soc., Providence, RI, 2003.

[82] L. G. Hanin. Kantorovich-Rubinstein norm and its application in the theory of Lipschitz spaces. *Proceedings of the American Mathematical Society*, 115(2):345–352, 1992.

[83] L. G Hanin. Duality for general Lipschitz classes and applications. *Proceedings of the London Mathematical Society*, 75(1):134–156, 1997.

[84] L. G. Hanin. An extension of the Kantorovich norm. *Contemporary Mathematics*, 226:113–130, 1999.

[85] L.G. Hanin. On isometric isomorphism between the second dual to the "small" Lipschitz space and the "big" Lipschitz space. In *Nonselfadjoint Operators and Related Topics*, pages 316–324. Springer, 1994.

[86] G. H. Hardy and J. E. Littlewood. A maximal theorem with function-theoretic applications. *Acta Math.*, 54(1):81–116, 1930.

[87] P. Harmand, D. Werner, and W. Werner. *M-ideals in Banach spaces and Banach algebras.* Springer, 2006.

[88] D. Hebert and H. Lacey. On supports of regular Borel measures. *Pacific Journal of Mathematics*, 27(1):101–118, 1968.

[89] J. Heinonen, P. Koskela, N. Shanmugalingam, and J. T Tyson. *Sobolev spaces on metric measure spaces*, volume 27. Cambridge University Press, 2015.

[90] R. Hurri-Syrjänen, N. Marola, and A. V. Vähäkangas. Aspects of local-to-global results. *Bull. Lond. Math. Soc.*, 46(5):1032–1042, 2014.

[91] T. Isernia, C. Leone, and A. Verde. Partial regularity results for non-autonomous elliptic systems with general growth. *submitted.*

[92] T. Isernia, C. Leone, and A. Verde. Partial regularity results for asymptotic quasiconvex functionals with general growth. *Ann. Acad. Sci. Fenn. Math.*, 41(2):817–844, 2016.

[93] T. Ito. On Banach spaces with unique isometric preduals. *The Michigan Mathematical Journal*, 24(3):321–324, 1977.

[94] S. Jahan. Approximative K-atomic decompositions and frames in Banach spaces. *Arab Journal of Mathematical Sciences*, 2019.

[95] P. W. Jones. Factorization of A_p weights. *Ann. of Math. (2)*, 111(3):511–530, 1980.

[96] A. Jonsson. The duals of Lipschitz spaces defined on closed sets. *Indiana University Mathematics Journal*, pages 467–476, 1990.

[97] A. Kaminska and H.J. Lee. M-ideal properties in Marcinkiewicz spaces. *Comment. Math. Prace Mat. Tomus specialis in Honorem Juliani Musielak*, pages 123–144, 2004.

[98] A. Kamińska, H.J. Lee, and H.J. Tag. M-ideal properties in Orlicz-Lorentz spaces. *arXiv preprint arXiv:1705.10451*, 2017.

[99] L. V. Kantorovich. On mass transfer problem. *Dokl. Acad. Nauk SSSR37*, pages 199–201, 1942.

[100] L. V. Kantorovich and G. S. Rubinstein. On a functional space and certain extremum problems. In *Doklady Akademii Nauk*, volume 115, pages 1058–1061. Russian Academy of Sciences, 1957.

[101] L. V. Kantorovich and G. S. Rubinstein. On a space of completely additive functions. *Vestnik Leningrad. Univ*, 13(7):52–59, 1958.

[102] N. Karak. Measure density and embeddings of Hajłasz-Besov and Hajłasz-Triebel-Lizorkin spaces. *Journal of Mathematical Analysis and Applications*, 475(1):966–984, 2019.

[103] J. Karamata. Sur un mode de croissance régulière. théorèmes fondamentaux. *Bull. Soc. Math. France*, 61:55–62, 1933.

[104] J. Karamata. Some theorems concerning slowly varying functions. Technical report, Wisconsin Univ. Madison Math. Res. Center, 1962.

[105] J. Kinnunen and P. Shukla. The distance of L^∞ from BMO on metric measure spaces. *Advances in Pure and Applied Mathematics*, 5(2):117–129, 2014.

[106] A. Korenovskii. *Mean oscillations and equimeasurable rearrangements of functions*, volume 4 of *Lecture Notes of the Unione Matematica Italiana*. Springer, Berlin; UMI, Bologna, 2007.

[107] M.B. Korey. Ideal weights: Asymptotically optimal versions of doubling, absolute continuity, and bounded mean oscillation. *The Journal of Fourier Analysis and Application*, 4:491–519, 1998.

[108] M.B. Korey. A decomposition of functions with vanishing mean oscillation. In *Harmonic analysis and boundary value problems (Fayetteville, AR, 2000)*, volume 277 of *Contemp. Math.*, pages 45–59. Amer. Math. Soc., Providence, RI, 2001.

[109] P. Lefevre, D. Li, H. Queffélec, and L. Rodríguez-Piazza. A criterion of weak compactness for operators on subspaces of Orlicz spaces. *Journal of Function Spaces*, 6(3):277–292, 2008.

[110] M. V. Leĭbov. Subspaces of the space VMO. *Teor. Funktsiĭ Funktsional. Anal. i Prilozhen.*, (46):51–54, 1986.

[111] J. Luukkainen and E. Saksman. Every complete doubling metric space carries a doubling measure. *Proceedings of the American Mathematical Society*, 126(2):531–534, 1998.

[112] P. Marcellini. Approximation of quasiconvex functions, and lower semicontinuity of multiple integrals. *Manuscripta Math.*, 51(1-3):1–28, 1985.

[113] P. Marcellini. Regularity of minimizers of integrals of the calculus of variations with non standard growth conditions. *Archive for Rational Mechanics and Analysis*, (105):267–284, 1989.

[114] M. Milman. Marcinkiewicz spaces, Garsia-Rodemich spaces and the scale of John-Nirenberg self improving inequalities. *arXiv preprint arXiv:1508.05057*, 2015.

[115] G. Mingione. Regularity of minima: an invitation to the dark side of the calculus of variations. *Appl. Math.*, 51(4):355–426, 2006.

[116] B.S. Mordukhovich and H.J. Sussmann. Nonsmooth analysis and geometric methods in deterministic optimal control. In *The IMA Volumes in Mathematics and its Applications*, pages ix+246. Springer, 1996.

[117] C. B. Morrey, Jr. Quasi-convexity and the lower semicontinuity of multiple integrals. *Pacific J. Math.*, 2:25–53, 1952.

[118] B. Muckenhoupt. Weighted norm inequalities for the Hardy maximal function. *Trans. Amer. Math. Soc.*, 165:207–226, 1972.

[119] M. Palladino and F. Rampazzo. A geometrically based criterion to avoid infimum gaps in optimal control. *J. Differential Equations*, 269(11):10107–10142, 2020.

[120] K.M. Perfekt. Duality and distance formulas in spaces defined by means of oscillation. *Arkiv för matematik*, 51(2):345–361, 2013.

[121] K.M. Perfekt. Weak compactness of operators acting on o–O type spaces. *Bulletin of the London Mathematical Society*, 47(4):677–685, 2015.

[122] K.M. Perfekt. On M-ideals and o–O type spaces. *Mathematica Scandinavica*, 102(1):151–160, 2017.

[123] B.J. Pettis et al. A proof that every uniformly convex space is reflexive. *Duke Mathematical Journal*, 5(2):249–253, 1939.

[124] L. Pick, A. Kufner, O. John, and S. Fucík. *Function spaces, 1.* Walter de Gruyter, 2012.

[125] A. C. Ponce. On the distributions of the form $\sum_i(\delta_{p_i} - \delta_{n_i})$. *Journal of Functional Analysis*, 210(2):391–435, 2004.

[126] F. Rampazzo. Frobenius-type theorems for Lipschitz distributions. *J. Differential Equations*, 243(2):270–300, 2007.

[127] F. Rampazzo and H. J. Sussmann. Commutators of flow maps of nonsmooth vector fields. *J. Differential Equations*, 232(1):134–175, 2007.

[128] F. Rampazzo and H.J. Sussmann. Set-valued differentials and a nonsmooth version of Chow's theorem. In *Proceedings of the 40th IEEE Conference on Decision and Control (Cat. No.01CH37228)*, volume 3, pages 2613–2618, 2001.

[129] M.M. Rao and Z.D. Ren. *Theory of Orlicz spaces.* M. Dekker New York, 1991.

[130] D. Sarason. Functions of vanishing mean oscillation. *Trans. Amer. Math. Soc.*, 207:391–405, 1975.

[131] T. Schmidt. Regularity of minimizers of $W^{1,p}$-quasiconvex variational integrals with (p, q)-growth. *Calc. Var. Partial Differential Equations*, 32(1):1–24, 2008.

[132] H. J. Sussmann. High-order point variations and generalized differentials. In *Geometric control and nonsmooth analysis*, volume 76 of *Ser. Adv. Math. Appl. Sci.*, pages 327–357. World Sci. Publ., Hackensack, NJ, 2008.

[133] A. Torchinsky. *Real-variable methods in harmonic analysis*, volume 123 of *Pure and Applied Mathematics.* Academic Press, Inc., Orlando, FL, 1986.

[134] H. Tuominen. *Orlicz-Sobolev spaces on metric measure spaces.* ProQuest LLC, Ann Arbor, MI, 2004. Thesis (Dr.Phil.)–Jyvaskylan Yliopisto (Finland).

[135] A. M. Vershik. Long history of the Monge-Kantorovich transportation problem. *The Mathematical Intelligencer*, 35(4):1–9, 2013.

[136] C. Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.

[137] R. Vinter. Optimal control. In *Modern Birkhäuser Classics*, pages xx+507. Springer, 2010.

[138] D. Werner. New classes of Banach spaces which are M-ideals in their biduals. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 111, pages 337–354. Cambridge University Press, 1992.

[139] D. E. Wulbert. Representations of the spaces of Lipschitz functions. *Journal of Functional Analysis*, 15(1):45–55, 1974.