Descent modulus and applications

ARIS DANIILIDIS, LAURENT MICLO, DAVID SALAS

Abstract. The norm of the gradient $\|\nabla f(x)\|$ measures the maximum descent of a real-valued smooth function f at x. For (nonsmooth) convex functions, this is expressed by the distance $\mathrm{dist}(0, \partial f(x))$ of the subdifferential to the origin, while for general real-valued functions defined on metric spaces by the notion of metric slope $|\nabla f|(x)$. In this work we propose an axiomatic definition of descent modulus $T[f](x)$ of a real-valued function f at every point x, defined on a general (not necessarily metric) space. The definition encompasses all above instances as well as average descents for functions defined on probability spaces. We show that a large class of functions are completely determined by their descent modulus and corresponding critical values. This result is already surprising in the smooth case: a one-dimensional information (norm of the gradient) turns out to be almost as powerful as the knowledge of the full gradient mapping. In the nonsmooth case, the key element for this determination result is the break of symmetry induced by a downhill orientation, in the spirit of the definition of the metric slope. The particular case of functions defined on finite spaces is studied in the last section. In this case, we obtain an explicit classification of descent operators that are, in some sense, typical.

Key words Descent modulus, metric diffusion, critical theory, determination of a function.

AMS Subject Classification *Primary* 49J52, 58E05, 60G99; *Secondary* 30L15, 37C10, 58G32.

Contents

1 Introduction

In [6] the following surprising result was obtained: two $\mathcal{C}^2$-smooth convex bounded from below functions defined on a Hilbert space $\mathcal{H}$ are equal up to an additive constant, provided they have the same modulus of derivative at every point. In other words, for this class of functions, equality of moduli of derivatives ($\|\nabla f\| = \|\nabla g\|$) implies equality of the derivatives ($\nabla f = \nabla g$). An alternative way to announce this result is to say that the operator

$$f \mapsto \|\nabla f\| \tag{1}$$

determines the function f (modulo a constant) for the class of $\mathcal{C}^2$-smooth convex and bounded from below functions defined on the Hilbert space $\mathcal{H}$.

The above result has been extended in [18] to the class of convex continuous bounded from below functions on a Hilbert space $\mathcal{H}$. A further extension for functions defined on an arbitrary Banach space X has been achieved in [24]. In both cases, the operator

$$f \mapsto d(0, \partial f(x)) \quad \text{(remoteness of the subdifferential)} \tag{2}$$

determines the function f (modulo a constant) for the class of convex continuous and bounded from below functions on a Banach space X.

In [9] the authors worked on an arbitrary metric space (X, d). Using the notion of metric slope $|\nabla f|$, they established the following result: two continuous coercive functions $f, g : X \to \mathbb{R}$ are equal, provided they have the same metric slope ($|\nabla f| = |\nabla g|$) and coincide on the (common) critical set $S = |\nabla f|^{-1}(0) = |\nabla g|^{-1}(0)$. (We refer to Section 2 for notation and relevant definitions; see also Subsection 2.1 for a more detailed description of the above results.) Denoting by $\mathcal{K}(X)$ the class of continuous coercive functions on X (the exact definition of coercivity will be given in (4)), we consider the following equivalent relation: $f \sim g$ if and only if f and g have the same (metric) critical set and their values coincide there up to a constant, that is,

$$f \sim g \quad \Longleftrightarrow \quad S = |\nabla f|^{-1}(0) = |\nabla g|^{-1}(0) \text{ and } f\big|_S - g\big|_S = c, \text{ for some } c \in \mathbb{R}.$$

Then, the aforementioned result of [9] asserts that the operator

$$f \mapsto |\nabla f| \tag{3}$$

is injective on $\mathcal{K}(X)/\sim$, that is, it determines functions $f \in \mathcal{K}(X)$ modulo the equivalent relation $\sim$.

We refer to all above results as determination results on a specific class of function (modulo a natural equivalent relation). Although the last result is formulated in an abstract metric space and is quite general, we will show in this work that a deeper result is hidden. Namely, the metric structure is ostensibly required to define the determining operator, but is not really paramount: the quantities $\|\nabla f(x)\|$ (in the smooth case) and $|\nabla f|(x)$ (in a metric setting) express the steepest descent of f at a given point x, however, this is not the only possible choice to deal with descent properties. For instance, one can also consider average descent (based on some probability measure on the space X) and emancipate dependence from the metric structure. The above leads to a definition of an abstract descent operator (which does not rely on a distance or even a topology). This abstract scheme, developed in Section 3, encompasses

several instances of descent-type operators, in particular both paradigms of steepest descent and average descent. In Section 4 we study general diffusion processes in metric spaces and show that asymmetrization (via downhill orientation) is the key property to obtain a determination result, in a complete analogy to the asymmetric definitions of (2)–(3). In the last section we consider the particular case of descent operators in finite dimensional spaces and obtain an explicit classification of a broad subfamily of these operators.

2 Notation and Preliminaries

We set $\overline{\mathbb{R}} = \mathbb{R} \sqcup \{-\infty, +\infty\}$ and $\overline{\mathbb{R}}_+ = \mathbb{R}_+ \sqcup \{+\infty\}$. For any $a \in \mathbb{R}$ we set $a_+ = \max\{a, 0\}$. For two real numbers $r, s \in \mathbb{R}$, we write $r \wedge s := \min\{r, s\}$ and $r \vee s := \max\{r, s\}$. Given a nonempty set X and a function $f : X \to \overline{\mathbb{R}}$ we define the domain of f as follows:

$$\mathrm{dom}(f) := \{x \in X : f(x) < +\infty\}.$$

Given $\alpha \in \mathbb{R}$, we write

$$[f \leq \alpha] := \{x \in X : f(x) \leq \alpha\}$$
$$[f < \alpha] := \{x \in X : f(x) < \alpha\}$$

to denote the sublevel set and strict sublevel set of f at value α. The sets $[f = \alpha]$, $[f \geq \alpha]$ and $[f > \alpha]$ are defined analogously.

We shall often equip the set X with a topology, denoted by τ. In this case, we denote by $\mathcal{B}(X)$ the σ-algebra of the Borel subsets of the topological space (X, τ).

We say that a function $f : X \to \mathbb{R} \cup \{+\infty\}$ is τ-*lower semicontinuous* if all sublevel sets $[f \leq a]$, $a \in \mathbb{R}$, are τ-closed subsets of X. The function f is called τ-*coercive* if

$$\text{for every } \alpha \in (-\infty, \sup f), \quad \text{the sublevel set } [f \leq \alpha] \text{ is } \tau\text{-compact.} \tag{4}$$

We simply call f lower semicontinuous (respectively, coercive), when no ambiguity on the topology occurs. Notice that the above definition of coercivity encompasses in particular all constant functions.

We further denote by

$$\mathcal{C}(X) \text{ the space of continuous functions from } X \text{ to } \mathbb{R} \tag{5}$$

and we define the subclass of coercive continuous functions by

$$\mathcal{K}(X) := \{f \in \mathcal{C}(X) : f \text{ is } \tau\text{-coercive}\}. \tag{6}$$

If (X, τ) is compact, then every lower semicontinuous functional is coercive and $\mathcal{K}(X) = \mathcal{C}(X)$.

Let $\mathcal{L}_n$ stand for the usual Lebesgue measure over $\mathbb{R}^n$ and let $B_n(x, r)$ (respectively, $\overline{B}_n(x, r)$) be the open (respectively, closed) ball centered at $x \in \mathbb{R}^n$ of radius $r > 0$. We also denote by $\mathbb{B}_n$ (respectively, $\mathbb{S}_n$) the closed unit ball (respectively, unit sphere) of $\mathbb{R}^n$. If there is no ambiguity, we omit the subindex n for each of the elements above. It is well known (see, e.g., [22]) that the n-dimensional volume of the ball of radius $r > 0$ in $\mathbb{R}^n$ is given by

$$\mathcal{L}_n(B(0, r)) = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2} + 1\right)} r^n, \tag{7}$$

3

where Γ stands for the gamma function. In particular, the volume of the n-dimensional ball $B(0,r)$ is proportional to r^n. For any (affine) subspace W of $\mathbb{R}^n$, we denote by $\dim(W)$ its (affine) dimension.

We say that a family $\mathcal{F}$ of real-valued functions is a *cone* if for every $f \in \mathcal{F}$ and $r \geq 0$ we have $rf \in \mathcal{F}$. In addition, we say that $\mathcal{F}$ is a *translation cone* if it is closed by translations (that is, for every $f \in \mathcal{F}$ and every constant $c \in \mathbb{R}$, we have that $f + c \in \mathcal{F}$). Clearly, the set $\mathcal{K}(X)$ of coercive continuous functions is a translation cone in $\mathcal{C}(X)$.

For an operator $T : \mathcal{F} \to (\overline{\mathbb{R}}_+)^X$, we define its *domain*

$$\mathrm{dom}(T) := \{f \in \mathcal{F} : \quad T[f](x) < +\infty, \text{ for all } x \in X\}.$$

If (X, d) is a metric space, we define the metric slope $|\nabla f|(x)$ of an extended real-valued function $f : X \to \mathbb{R} \cup \{+\infty\}$ at the point $x \in \mathrm{dom}(f)$ as follows:

$$|\nabla f|(x) := \begin{cases} \limsup\limits_{y \to x} \dfrac{\{f(x) - f(y)\}_+}{d(y, x)}, & \text{if } x \text{ is not isolated,} \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

In the same setting, the global slope $\mathscr{G}[f](x)$ is defined as follows:

$$\mathscr{G}[f](x) := \sup_{y \in X} \frac{\{f(x) - f(y)\}_+}{d(y, x)}.$$

Notice that $\mathscr{G}[f](x) = 0$ if and only if $x \in \mathrm{argmin}\, f$ (*i.e.* x is global minimum of f), while $|\nabla f|(x) = 0$ whenever x is a local minimum of f. The notions of metric slope (also known as strong slope) and global slope are well known in the literature (see, e.g., [1] and the references therein).

Let us now assume that X is a Banach space with dual X^*. It is well-known that if $f : X \to \mathbb{R}$ is a smooth function, then

$$|\nabla f|(x) = \|\nabla f(x)\|.$$

In the nonsmooth setting, if $f : X \to \mathbb{R} \cup \{+\infty\}$ a lower semicontinuous *convex* function, its (Fenchel-Moreau) subdifferential $\partial f(x)$ at $x \in \mathrm{dom}(f)$ is defined as follows:

$$\partial f(x) = \{p \in X^* : \forall y \in X, f(y) \geq f(x) + \langle p, y - x \rangle\}.$$

It is well-known that $\partial f(x)$ is a closed convex set and it is nonempty if x is a point of continuity of f (see, e.g., [21]). Moreover, it is known (see, e.g., [1, Proposition 1.4.4]) that for any lower semicontinuous convex function over a Banach space, one has

$$|\nabla f|(x) = \mathscr{G}[f](x) = d(0, \partial f(x)). \tag{9}$$

2.1 State-of-the-art

The derivative of a smooth function recovers, up to an additive constant, the function through integration. In the nonsmooth case, Rockafellar [20] showed that every lower semicontinuous convex function can be represented through its subdifferential by means of a *nonsmooth integration*. This result has been refined in [4] for Banach spaces with the Radon-Nikodym property,

provided the function satisfies a mild coercivity property (namely, the asymptotic cone of its epigraph is epi-pointed). In this latter case, a partial knowledge of the subdifferential mapping ∂f is sufficient to recover the function up to a constant.

Historically, this integration result was first stated as a determination result: For every two proper convex lower semicontinuous functions $f, g : X \to \mathbb{R}$ over a Banach space X, one has that

$$(\partial f(x) = \partial g(x), \forall x \in X) \implies f = g + c, \text{ for some } c \in \mathbb{R}. \tag{10}$$

This result was first obtained in Hilbert spaces by Moreau [17], and generalized to Banach spaces one year later by Rockafellar [19]. A more general result was established by Brezis for monotone operators [7], where the same determination result can be obtained in Hilbert spaces only in terms of the element of minimal norm of the subdifferential, that is,

$$(\text{proj}(0, \partial f(x)) = \text{proj}(0, \partial g(x)), \forall x \in \mathcal{H}) \implies f = g + c, \text{ for some } c \in \mathbb{R}, \tag{11}$$

where $\text{proj}(x, A)$ denotes the metric projection of x onto the set A. Notice that knowledge of a full gradient $\nabla f(x)$ (respectively, subdifferential $\partial f(x) \subset X^*$, or $\text{proj}(0, \partial f(x)) \in \mathcal{H}$) at many (all) $x \in X$ is already a rich information: at every such point x we need to know a vector (respectively, a set of vectors). Nonewithstanding, it has recently become clear that much less information (namely, a scalar) is often sufficient if our objective is only to determine functions (rather than recovering them via an explicit formula). This is resumed below:

2.1.1 Determination of convex functions

Let $\mathcal{H}$ be a Hilbert space and $f : \mathcal{H} \to \mathbb{R}$ be a $\mathcal{C}^2$-smooth convex and bounded from below function. Set $V(x) = \frac{1}{2}\|\nabla f(x)\|^2$ and consider the second-order dynamical system

$$\ddot{x}(t) = \nabla V(x(t)), \tag{12}$$

with initial condition $x(0) = x_0 \in \mathcal{H}$. It has been shown in [6] that every evanescent solution of (12) (that is, every solution satisfying $\|\dot{x}(t)\| \to 0$ and $\|\nabla V(x(t))\| \to 0$, as $t \to \infty$) is solution of the first order gradient system:

$$\begin{cases} \dot{x}(t) = -\nabla f(x(t)) \\ x(0) = x_0. \end{cases} \tag{13}$$

On the other hand, $\mathcal{C}^2$-smoothness assumptions guarantees that (13) has unique solution. The fact that f is bounded from below ensures that this solution is evanescent. By straightforward differentiation we deduce that this solution is also solution of the second-order system (12), therefore (12) and (13) have the same solutions. Since (12) depends only on $\|\nabla f\|$ (rather than on ∇f), the following conclusion has been obtained:

- If $f, g : \mathcal{H} \to \mathbb{R}$ are two $\mathcal{C}^2$-smooth convex bounded from below functions, then

$$\|\nabla f\| = \|\nabla g\| \iff \nabla f = \nabla g \iff f = g + c, \text{ for some } c \in \mathbb{R}. \tag{14}$$

Notice that $\mathcal{C}^2$-smoothness was required in order to define property (12). However, this assumption can be relaxed to $\mathcal{C}^1$-smoothness, assuming existence of minimizers [2]. This is based on the remark that

$$\|\nabla f\| = \|\nabla g\| \iff \langle \nabla(f + g), \nabla(f - g)\rangle \tag{15}$$

5

which ensures in turn that the function $f - g$ is constant among the gradient orbits of the (convex) function $f + g$. Since each such orbit *lands* on the (common) set S of minimizers of the functions f, g and $f + g$, and since $f - g$ is constant there (with value $\min f - \min g$), the result follows.

A generalization of (14) has been carried out in [18], where smoothness assumption has been replaced by continuity.

- If $f, g : \mathcal{H} \to \mathbb{R}$ are two convex continuous and bounded from below functions, then

$$
\begin{aligned}
d(0, \partial f(x)) = d(0, \partial g(x)), \text{ for all } x \in \mathcal{H} &\iff \partial f = \partial g \\
&\iff f = g + c, \text{ for some } c \in \mathbb{R}.
\end{aligned}
\tag{16}
$$

To achieve the above result, the authors studied the subgradient system $\dot{x}(t) \in -\partial f(x(t))$ and showed that in this case the assumption $d(0, \partial f(\cdot)) \geq d(0, \partial g(\cdot))$ yields $f \geq g + c$. This is proven based on two key observations: First, the solution $x(t)$ is not only a minimizing curve for f (i.e., $f(x(t)) \to \inf f$ as $t \to +\infty$), but it is also a minimizing curve for g. Second, the chain rule of the convex subdifferential entails that $(f - g)$ is nonincreasing along $x(t)$. Thus, one can consider $c = \inf f - \inf g$. After proving this comparison principle, (16) follows by symmetry.

2.1.2 Determination in metric spaces

Convexity assumption was important for the proofs of (14) and (16). However, the outlined proof via (15) was using convexity for two factors: to conclude that every steepest descent curve lands on a critical point (i.e. has an accumulation point in the set of critical points), and that every critical point is global minimizer, that is,

$$
\text{Crit } f = \{x \in X : \nabla f(x) = 0\} = \arg\min f.
\tag{17}
$$

Assuming (17) and some coercivity condition (instead of convexity), the same argument leads to the following result:

- If $f, g : \mathcal{H} \to \mathbb{R}$ are two $\mathcal{C}^1$-smooth coercive functions, then:

$$
\left.
\begin{aligned}
\|\nabla f\| = \|\nabla g\| \quad \text{and} \\
\text{Crit } f = \arg\min f = \arg\min g \neq \emptyset
\end{aligned}
\right\} \implies f = g + c, \text{ for some } c \in \mathbb{R}.
$$

On the other hand, all results mentioned in the previous subsection are strongly based on (sub)gradient dynamical systems and the Hilbertian structure of the space. Quite surprisingly, it turns out that this structure is eventually not required. Indeed, in the recent work [24], the result (16) has been extended to arbitrary Banach spaces, through a completely different approach, which was based on the notion of *countable orderable sets* introduced in [15]. In that work the authors establish that two continuous and bounded from below functions $f, g : X \to \mathbb{R}$ defined on a metric space (X, d) and with finite global slopes are equal up to a constant, provided they have the same global slopes at every point. In other words:

$$
\mathscr{G}[f] = \mathscr{G}[g] < +\infty \implies f = g + c, \text{ for some } c \in \mathbb{R}.
\tag{18}
$$

The key technique to achieve such a result is the construction of a minimizing sequence by means of the global slope. The construction is based on the following result (proved in [24]): for every sequence $\{x_i\}_i$ of the metric space (X, d) and for every proper extended real-valued function $f : X \to \mathbb{R} \sqcup \{\infty\}$ it holds:

$$\left(\lim_{i \to \infty} \mathscr{G}[f](x_i) = 0 \text{ and } \sum_{i=1}^{\infty} \mathscr{G}[f](x_i) d(x_i, x_{i+1}) < \infty \right) \implies \liminf_{i \to \infty} f(x_i) = \inf_X f. \qquad (19)$$

Although the setting is quite general (metric spaces), the notion of global slope is rather restrictive, since it does not coincide with the modulus of the derivative in the smooth case. But this notion is a very good fit for convex functions defined on a general Banach space X. In this case, combining (9) with (18) yields a generalization of (16).

In an independent work [9], the authors considered the local notion of metric slope and established the following comparison result for the class of continuous coercive functions. In what follows we denote by

$$\mathrm{Crit}\, f = \{x \in X : \ |\nabla f|(x) = 0\}$$

the set of (metrically) critical points.

Proposition 2.1 (slope comparison). *Let (X, d) be a metric space and $f, g : X \to \mathbb{R}$ be two continuous coercive functions. Assume that*

(i). $|\nabla f|(x) > |\nabla g|(x)$, for all $x \in X \backslash \mathrm{Crit} f.$; and

(ii). $f(x) > g(x)$, for all $x \in \mathrm{Crit} f.$

Then, $f > g$.

The proof was obtained by reasoning to contradiction and using discrete iterations and transfinite induction. The following determination result was then obtained as consequence of Proposition 2.1.

Theorem 2.2 (Determination in metric spaces). *Let (X, d) be a metric space and $f, g : X \to \mathbb{R}$ be two continuous coercive functions. Assume that*

(i). $|\nabla f|(x) = |\nabla g|(x) < +\infty$, for all $x \in X$; and

(ii). $f(x) = g(x)$, for all $x \in \mathrm{Crit}\ f.$

Then, $f = g$.

The above result asserts that information on the metric slope $|\nabla f|$ and critical values is sufficient to determine every continuous coercive functions with finite slope (therefore, in particular, every Lipschitz continuous coercive functions). Taking into account the pathologies that prevail Lipschitz functions, the above statement appears close to be optimal: In [9] several counterexamples are presented to show the pertinence of the assumptions. This being said, there is still room for improvements: indeed, assuming X is a complete metric space, it seems plausible to relaxe coercivity/compactness assumption (which is required in the current proof), by an alternative assumption ensuring the existence of appropriate descent (generalized) sequences that link any point to the critical set.

2.2 Description of the current work

Revising the arguments employed in [9] for the proof of Proposition 2.1 and Theorem 2.2 we observe that continuity and coercivity are topological notions, while the metric structure of (M, d) is only required in order to define the metric slope, see (8). In particular, one can assume continuity and coercivity with respect to another topology (not related to the given distance d) and the topological part of the proof can be completely decoupled.

In this work we show that a similar result to Theorem 2.2 holds for any topological space equipped with a Borel probability measure μ, if we replace the metric slope $|\nabla f|(x)$ (corresponding to the steepest descent at x) by the *μ-average descent $T_\mu(f)(x)$ at x* given by

$$T_\mu(f)(x) := \int_X \{f(x) - f(y)\}_+ \, d\mu(y) = \int_{[f \leq f(x)]} [f(x) - f(y)] \, d\mu(y).$$

More generally, we introduce an abstract descent operator $T[f]$ (see Definition 3.1) that encompasses both metric and global slopes (in metric spaces) and average descent (in probability spaces) as well as many other instances, see Subsection 3.3 for further examples and stability properties of this operator. We then establish an abstract determination result (Theorem 3.5) revealing that the metric structure is neither minimal nor optimal framework, as hinted by the topological and metric decoupling observed in [9]. In Section 4 we study general stochastic processes in metric spaces and define adequate oriented operators (particular instances of Definition 3.1) that allow to obtain determination results. Finally, in Section 5 we introduce an equivalence relation among descent moduli for functions $f \in \mathbb{R}^{\mathcal{V}}$ defined in finite spaces $\mathcal{V}$ and show that a typical descent modulus is equivalent to a steepest descent with respect to a prescribed active neighborhood system (see Theorem 5.8).

3 Descent modulus: definition, properties and main examples

Let $\mathcal{F}$ be a family of functions from a nonempty set X to $\mathbb{R}$. For an operator $T : \mathcal{F} \to (\overline{\mathbb{R}}_+)^X$, we define its *domain*

$$\operatorname{dom}(T) := \{f \in \mathcal{F} : \quad T[f](x) < +\infty, \text{ for all } x \in X\}. \tag{20}$$

We also define the set $\mathcal{Z}_T(f)$ of *T-critical points* of $f \in \mathcal{F}$ as follows:

$$\mathcal{Z}_T(f) = \{x \in X \; : \; T[f](x) = 0\}. \tag{21}$$

(Note that every T-critical point of f is a global minimizer for the function $T[f]$.)

In this section we give an axiomatic definition of an abstract descent operator, that is, an operator T acting on (a certain class of) functions f from X to $\mathbb{R}$. This operator associates to each point $x \in X$ an extended nonnegative number $T[f](x) \in \mathbb{R} \cup \{+\infty\}$ which corresponds to an abstract measure of descent (henceforth called *descent modulus*) of f at x.

The required properties of this abstract definition will be kept minimal to encompass several instances stemming from classical and variational analysis, metric geometry and stochastic processes: in particular, the metric slope (used in [9]), the global slope (used in [24]) and the notion of average descent (that will be discussed later in this work) are all captured by the proposed abstract scheme.

3.1 Axiomatic definition

Let $\mathcal{F}$ be a translation cone in the space of functions from a nonempty set X to $\mathbb{R}$.

Definition 3.1 (Abstract descent modulus)**.** Let $T : \mathcal{F} \to (\overline{\mathbb{R}}_+)^X$ be a nonlinear operator. We say that:

(D1). *T preserves global minima*, if for every function $f \in \mathcal{F}$ and $x \in X$ we have

$$x \in \operatorname{argmin} f \implies x \in \mathcal{Z}_T(f).$$

(D2). *T is monotone at x*, if for every $f, g \in \mathcal{F}$ we have:

$$\forall z \in X : \ (f(x) - f(z))_+ \geq (g(x) - g(z))_+ \implies T[f](x) \geq T[g](x). \tag{22}$$

(D3). *T is scalar-monotone at x*, if for every function $f \in \mathcal{F}$ and $r > 1$, we have

$$0 < T[f](x) < +\infty \implies T[f](x) < T[rf](x).$$

The operator T is called (scalar) monotone, if it is (scalar) monotone at every $x \in X$.

We say that T is a *descent modulus* for the class $\mathcal{F}$ if properties (D1)–(D3) hold, that is, if T is monotone, scalar-monotone and preserves global minima.

Before we proceed, let us have a brief discussion on the above properties:

Property (D1) states that there is *no descent at global minima*; thus $T[f](x) = 0$ holds at every $x \in \operatorname{argmin} f$.

Property (D2) expresses the fact that the *amount of descent* of f at a point x depends only on the sublevel set $[f \leq f(x)]$ and is captured by the function $z \mapsto (f(x) - f(z))_+$. Accordingly, for a fixed $x \in X$, the relation

$$g \preceq_x f \quad \underset{\text{def}}{\Longleftrightarrow} \quad \forall z \in X : \quad (g(x) - g(z))_+ \leq (f(x) - f(z))_+$$

is a preorder relation on $\mathcal{F}$ which roughly reads as follows: "f has more descent than g at x". Under this terminology, (D2) requires the mapping $\mathcal{F} \ni f \mapsto T[f](x)$ to be monotone with respect to $\preceq_x$.

Notice further that (22) yields the following: If $g \preceq_x f$, then

$$z \notin [f < f(x)] \implies g(z) \geq g(x) \quad \text{(that is, } z \notin [g < g(x)]\text{)}. \tag{23}$$

This means that $g \preceq_x f$ implies that $[g < g(x)] \subset [f < f(x)]$. Finally, scalar-monotonicity in (D3) expresses the fact that the descent of the function $g = rf$ should be larger than the one of f, when $r > 1$.

In conclusion, the above axioms (D1)–(D3) are natural requirements for an abstract notion of descent of a function f at a point x. The following proposition reveals further properties that can be derived from the axioms of Definition 3.1.

Proposition 3.2 (Properties of descent moduli)**.** *Let $\mathcal{F} \subset \mathcal{C}(X)$ (as before) and $T : \mathcal{F} \to (\overline{\mathbb{R}}_+)^X$ be an operator. Then:*

(a). (one-step descent property) *T is monotone if and only if for every* $f, g \in \mathcal{F}$ *and* $x \in X$

$$T[f](x) > T[g](x) \quad \Longrightarrow \quad \exists z \in [f < f(x)]: \ f(x) - f(z) > g(x) - g(z). \tag{24}$$

(b). (translation-invariance) *If T is a descent modulus for* $\mathcal{F}$, *then for every* $c \in \mathbb{R}$ *and* $f \in \mathcal{F}$ *we have:*

$$T[f] = T[f + c].$$

(c). (strict monotonicity) *Let T be monotone at* $x \in X$. *Then the following are equivalent:*

– (c_1). *T is scalar-monotone at* x.

– (c_2). *For every* $f, g \in \mathcal{F}$ *with* $T[f](x) > 0$, $T[g](x) < +\infty$ *and* $[g \le g(x)] \subset [f \le f(x)]$, *the implication*

$$\exists \delta > 0: \ \forall z \in [g \le g(x)] \quad \Longrightarrow \quad f(x) - f(z) \ge (1 + \delta)(g(x) - g(z)),$$

yields

$$T[f](x) > T[g](x).$$

– (c_3). *For every* $f \in \mathcal{F}$, $x \in X$ *and* $r \in (1, +\infty)$ *such that* $0 < T[f](x)$ *and* $T[rf](x) < +\infty$, *the mapping*

$$[0, r - 1] \ni \delta \ \mapsto \ T[(1 + \delta)f](x)$$

is strictly increasing.

Proof. Let us show the above properties separately:

(a). (*sufficiency*) Reasoning by absurd, assume T verifies the one-step property but it is not monotone. Then, there exist $f, g \in \mathcal{F}$ and $x \in X$ with $(f(x) - f(\cdot))_+ \ge (g(x) - g(\cdot))_+$ but with $T[f](x) < T[g](x)$. By the one-step descent property (24), there exists $z \in [g < g(x)]$ such that

$$g(x) - g(z) > f(x) - f(z).$$

However, the inequality $(f(x) - f(\cdot))_+ \ge (g(x) - g(\cdot))_+$ yields that $[g < g(x)] \subset [f \le f(x)]$, and so the above inequality is a direct contradiction.

(*necessity*) Assume that T is monotone but the one-step descent property does not hold. Then, there exist $f, g \in \mathcal{F}$ and $x \in X$ with $T[f](x) > T[g](x)$ such that for all $z \in X$ we either have $f(x) \le f(z)$ or $f(x) - f(z) \le g(x) - g(z)$. It is not hard to see that for every $z \in X$ one has that

$$(f(x) - f(z))_+ = \begin{cases} f(x) - f(z) \le g(x) - g(z), & \text{if } f(x) > f(z) \\ 0, & \text{otherwise.} \end{cases}$$

Thus, in any case, we get that $(f(x) - f(z))_+ \le (g(x) - g(z))_+$. Then, monotonicity yields that $T[f](x) \le T[g](x)$, which is a contradiction.

(b). We show that for every $f \in \mathcal{F}$ and $c \in \mathbb{R}$, we have $T[f] = T[f + c]$.

Notice that $[(f(x) + c) - (f(z) + c)]_+ \ge [f(x) - f(z)]_+$ holds trivially for all $x, z \in X$. By monotonicity we deduce that $T[f] \le T[f + c]$. Replacing now f by $f' = f + c$ and respectively, c by $c' = -c$, we obtain equality.

(c). Let us show first that $(c_1) \Rightarrow (c_2)$:

Reasoning by absurd, assume that there exist $f, g \in \mathcal{F}$ and $\delta > 0$ satisfying the hypotheses of the statement and $x \in X$ with $0 < T[f](x) \le T[g](x) < +\infty$. Then for all $z \in X$ it holds:

$$(f(x) - f(z))_+ \ge ((1+\delta)g(x) - (1+\delta)g(z))_+ .$$

By monotonicity, we deduce that $T[(1+\delta)g] \le T[f]$. Further, using scalar-monotonicity, we get

$$0 < T[g](x) < +\infty \implies T[g](x) < T[(1+\delta)g](x) \le T[f](x) \le T[g](x),$$

which is a contradiction.

Let us now show that $(c_2) \Rightarrow (c_3)$:

Let $f \in \mathcal{F}$, $x \in X$ and $r > 1$ such that $0 < T[f](x), T[rf](x) < +\infty$. Fix $\delta_1, \delta_2 \in [0, r-1]$ with $\delta_1 < \delta_2$. We set $g = (1+\delta_1)f$ and $h = (1+\delta_2)f$. Monotonicity yields that $0 \le T[g](x)$ and $T[h](x) < +\infty$. Then, setting

$$\delta = \frac{1+\delta_2}{1+\delta_1} - 1$$

we have that for all $y \in X$, $[g \le g(y)] = [h \le h(y)]$, and for all $y, z \in X$:

$$h(y) - h(z) = (1+\delta)(g(y) - g(z)).$$

Thus, (c2) yields that $T[h](x) > T[g](x)$.

Let us finally establish that $(c_3) \Rightarrow (c_1)$.

To this end, let $f \in \mathcal{F}$, $r > 1$ and $x \in X$ such that $0 < T[f](x) < +\infty$. We need to show that $T[f](x) < T[rf](x)$. This holds trivially if $T[rf](x) = +\infty$, therefore we can assume that $T[rf](x) < +\infty$. Since T is monotone, we have that $T[rf](x) \ge T[f](x)$ already, and in particular, $T[rf](x) > 0$. Then, by hypothesis, we have that the mapping

$$[0, r-1] \ni \delta \mapsto T[(1+\delta)f](x)$$

is strictly increasing, which leads us directly to the desired inequality. $\qquad \square$

3.2 Determination in topological spaces

Let (X, τ) be a topological space and let T be a descent modulus for $\mathcal{K}(X)$. Let us define the following equivalent relation on the class $\mathcal{K}(X)$ of continuous coercive functions: we say that the functions $f, g \in \mathcal{K}(X)$ are equivalent (and we denote $f \sim g$) if they have the same T-critical set and they are equal there.

In other words:

$$f \sim g \quad \Longleftrightarrow \quad \mathcal{Z}_T(f) = \mathcal{Z}_T(g) = S \quad \text{and} \quad f|_S = g|_S.$$

In this section, borrowing from techniques developed in [9], we show that properties (D1)–(D3) of the descent modulus (*cf.* Definition 3.1) are sufficient to guarantee that the mapping $f \mapsto T[f]$ is injective on $\mathcal{K}(X)$, modulo the above equivalent relation. Therefore, according to our terminology, the descent modulus *determines* the class $\mathcal{K}(X)$. At this stage, let us also

11

outline the topological nature of this result: no linear or metric structure is required.

The results of this section will be stated in a slightly more general framework. We assume, similarly to the previous section, that $\mathcal{F} \subset \mathcal{K}(X)$ is a translation cone.

We start with the following lemma.

Lemma 3.3 (strict domination of descent modulus). *Let T be a descent modulus for the class $\mathcal{F}$. Let $f, g \in \mathrm{dom}(T)$ such that*

$$\forall x \in X \setminus \mathcal{Z}_T(f), \quad T[f](x) > T[g](x).$$

Then, for all $x \in X$, we have that

$$f(x) \geq g(x) + \mu(x),$$

where

$$\mu(x) := \inf\{(f-g)(z) \; : \; z \in [f \leq f(x)] \cap \mathcal{Z}_T(f)\} \in \mathbb{R} \cup \{-\infty\}.$$

Proof. Since the set of global minimizers of every function $f \in \mathcal{F}$ is nonempty and the abstract descent modulus T preserves global minima, we deduce that $\mathcal{Z}_T(f) \neq \emptyset$ and consequently, $\mu(x) < +\infty$. Let us assume, towards a contradiction, that there exists $x \in X$ such that $f(x) < g(x) + \mu(x)$. Then, clearly $\mu(x) > -\infty$ which readily yields $x \in X \setminus \mathcal{Z}_T(f)$. Therefore, by assumption, $T[f](x) > T[g](x)$. Applying the one-step descent property (24) of T, we infer that there exists $z_0 \in X$ such that

$$f(z_0) < f(x) \quad \text{and} \quad (g-f)(z_0) = c > (g-f)(x) > -\mu(x).$$

Since z_0 is not a T-critical point, we can repeat the above argument to obtain $z_1 \notin \mathcal{Z}_T$ such that $f(z_1) < f(z_0)$ and $(g-f)(z_1) > c = (g-f)(z_0)$. Following the strategy of [9, Proposition 2.2], we construct (by means of a transfinite induction over the ordinals) a generalized sequence $\{z_\alpha\}_\alpha \subset [f \leq f(z_0)]$ such that $\{f(z_\alpha)\}_a$ is decreasing and $\{(g-f)(z_\alpha)\}_\alpha$ is increasing:

– If $\alpha = \beta + 1$ is a successor ordinal then, since $z_\beta \notin \mathcal{Z}_T(f)$ and $g(z_\beta) \geq f(z_\beta) + c$, the one-step descent property (24) yields $z_{\beta+1}$ such that

$$f(z_{\beta+1}) < f(z_\beta) \leq f(z_0) \quad \text{and} \quad (g-f)(z_{\beta+1}) > (g-f)(z_\beta) \geq c.$$

– If α is a limite-ordinal and $\{z_\beta\}_{\beta<\alpha} \subset [f \leq f(z_0)]$ is defined accordingly, then since the sublevel set $[f \leq f(z_0)]$ is compact, the ω-limit set

$$A = \bigcap_{\beta<\alpha} \overline{\{z_\eta \; : \; \beta \leq \eta < \alpha\}},$$

is nonempty. Pick any $z_\alpha \in A$. Clearly, $z_\alpha \in [f \leq f(z_0)]$, $f(z_\alpha) \leq f(z_\beta)$ for all $\beta \leq \alpha$, and by continuity

$$(g-f)(z_\beta) = \inf\{(g-f)(z_\eta) \; : \; \beta \leq \eta < \alpha\} \leq (g-f)(z_\alpha).$$

Notice that the above construction never meets a T-critical point of f. Indeed, if $z_\alpha \in \mathcal{Z}_T(f)$ for some ordinal α, then since $f(z_\alpha) < f(x)$ we would have that

$$\begin{aligned}
-\mu(x) &= \sup\{(g-f)(z) \; : \; z \in [f \leq f(x)] \cap \mathcal{Z}_T(f)\} \\
&\geq (g-f)(z_\alpha) \geq c > -\mu(x),
\end{aligned}$$

12

which is a contradiction. Due to a cardinality obstruction, we necessarily deduce that $z_\alpha = z_\beta$ for some ordinals α, β with $\alpha > \beta$. This yields

$$(g - f)(z_{\beta+1}) > (g - f)(z_\beta) = (g - f)(z_\alpha) \geq (g - f)(z_{\beta+1}),$$

which is clearly a contradiction. The conclusion follows. □

Theorem 3.4 (Comparison principle)**.** *Let T be a descent modulus for $\mathcal{F}$ and let $f, g \in \mathrm{dom}(T)$ and $c \in \mathbb{R}$ such that*

(i). $T[f](x) \geq T[g](x)$, for all $x \in X$; and

(ii). $f(\bar{x}) \geq g(\bar{x}) + c$, for all $\bar{x} \in \mathcal{Z}_T(f)$.

Then, $f \geq g + c$.

Proof. Let $x \in X \setminus \mathcal{Z}_T(f)$ be arbitrarily chosen. Fix $\varepsilon > 0$, set $f_\varepsilon = (1 + \varepsilon)f$ and notice that monotonicity of T yields that $\mathcal{Z}_T(f_\varepsilon) \subset \mathcal{Z}_T(f)$. Let $z \in X \setminus \mathcal{Z}_T(f_\varepsilon)$. We have two cases:

Case 1: $z \in X \setminus \mathcal{Z}_T(f)$. Then scalar-monotonicity of T yields

$$T[f_\varepsilon](z) = T[(1 + \varepsilon)f](z) > T[f](z) \geq T[g](z).$$

Case 2: $z \in \mathcal{Z}_T(f) \setminus \mathcal{Z}_T(f_\varepsilon)$. Then $T[f_\varepsilon](z) > 0 = T[f](z) \geq T[g](z)$.

In both cases, $T[f_\varepsilon](z) > T[g](z)$. Thus, by Lemma 3.3, we have that

$$
\begin{aligned}
f_\varepsilon(x) &\geq g(x) + \inf\{(f_\varepsilon - g)(z) \ : \ z \in \mathcal{Z}_T(f_\varepsilon) \cap [f_\varepsilon \leq f_\varepsilon(x)]\} \\
&\geq g(x) + \inf\{(f_\varepsilon - g)(z) \ : \ z \in \mathcal{Z}_T(f) \cap [f \leq f(x)]\} \\
&\geq g(x) + c + \varepsilon \inf\{f(z) : \ z \in \mathcal{Z}_T(f) \cap [f \leq f(x)]\} \\
&\geq g(x) + c + \varepsilon \min f.
\end{aligned}
$$

Finally, by taking $\varepsilon \to 0$, we deduce that $f(x) \geq g(x) + c$. The proof is complete. □

Applying twice Theorem 3.4, we deduce easily the following determination result.

Theorem 3.5 (Determination of continuous coercive functions)**.** *Let T be a descent modulus for a translation cone $\mathcal{F}$ of $\mathcal{K}(X)$. Let $f, g \in \mathrm{dom}(T)$ and $c \in \mathbb{R}$ be such that*

(i). $T[f](x) = T[g](x)$ for all $x \in X$ (whence $\mathcal{Z}_T(f) = \mathcal{Z}_T(g)$); and

(ii). $f(x) = g(x) + c$, for all $x \in \mathcal{Z}_T(f)$.

Then, $f = g + c$.

Remark 3.6. A descent modulus T for a class $\mathcal{F}$ is meant to assign a quantified measure of descent at every point of $f \in \mathcal{F}$. This quantity is also allowed to be infinite at some points of some functions and whenever this happens the determination result cannot apply. Therefore, T does not determine the whole class $\mathcal{F}$, but instead only functions in $\mathrm{dom}(T) \subset \mathcal{F}$.

3.3 Stability properties of descent moduli and examples

The metric slope (used in [1], [10] *e.g.*) is a natural instance of abstract descent modulus and the results of the previous section can be seen as a minimal axiomatic presentation of the slope determination result given in [9]. In this section, we show that the axiomatic descent modules also captures the notion of global slope (used in [24]) as well as several natural adaptations of the notion of slope to topological spaces, emancipating from the metric framework.

Throughout this section, $\mathcal{F}$ will denote a translation cone of $\mathcal{C}(X)$.

Proposition 3.7 (*m*-slope). *Let $m : X \times X \to \mathbb{R}_+$ be a mapping satisfying:*

$$m(x, y) = 0 \quad \Longleftrightarrow \quad x = y \qquad \text{(separation axiom).}$$

Let further $\mathcal{D} = \{\mathcal{D}_x\}_{x \in X}$ be a family of subsets of X satisfying $x \in \mathcal{D}_x$ for every $x \in X$. Then, the m-slope

$$s_f(x) := \begin{cases} \limsup_{y \to x} \Delta_f^+(x, y) & \text{if } x \text{ is not isolated,} \\ 0 & \text{otherwise,} \end{cases}$$

and the semiglobal $(\mathcal{D}, m)$-slope

$$\mathscr{G}_{\mathcal{D}}[f](x) = \sup_{y \in \mathcal{D}_x} \Delta_f^+(x, y) \tag{25}$$

are moduli of descent for the class $\mathcal{F}$, where

$$\Delta_f^+(x, y) = \begin{cases} \dfrac{(f(x) - f(y))_+}{m(x, y)} & \text{if } y \neq x, \\ 0 & \text{if } y = x. \end{cases} \tag{26}$$

Proof. Let us show that the above operators of (local) m-slope and (semiglobal) $(\mathcal{D}, m)$-slope satisfy axioms (D1)–(D3) of Definition 3.1. It is straightforward to see that (D1) (preservation of global minima) is fulfilled. Axiom (D3) (scalar monotonicity) is also fulfilled, since for every $f \in \mathcal{F}$ and $r > 0$ we have

$$s_{rf}(x) = r s_f(x) \quad \text{and} \quad \mathscr{G}_{\mathcal{D}}[rf](x) = r\mathscr{G}_{\mathcal{D}}[f](x).$$

It remains to show that both operators also satisfy axiom (D2) (Monotonicity). To this end, let $f, g \in \mathcal{F}$ such that

$$(f(x) - f(z))_+ \geq (g(x) - g(z))_+.$$

Then for every $y \in X$ we have $\Delta_f^+(x, y) \geq \Delta_g^+(x, y)$, which readily yields that $s_f(x) \geq s_g(x)$ and $\mathscr{G}_{\mathcal{D}}[f](x) \geq \mathscr{G}_{\mathcal{D}}[g](x)$. The proof is complete. $\square$

Remark 3.8. When (X, τ) is a metric space and m is the distance function, the m-slope $s_f(x)$ coincides with the usual metric slope $|\nabla f|(x)$ and the main result of [9] follows directly from Theorem 3.5. Taking now $\mathcal{D}_x = X$ for all $x \in X$, the semiglobal $(\mathcal{D}, m)$-slope $\mathscr{G}_{\mathcal{D}}[f](x)$ coincides with the global slope $\mathscr{G}[f](x)$ (see, e.g., [1, Definition 1.2.4]) which was used in [24].

Notice that the semiglobal slope $\mathscr{G}_{\mathcal{D}}[f]$ is intrinsically different from the metric slope (or the norm of the gradient $\|\nabla f\|$ in the differentiable case), which already reveals that Definition 3.1 represents a much more general setting. The next proposition shows that we can go even further.

Proposition 3.9 (Constructing descent moduli). (i). *Let T_1, T_2 be descent moduli for the class $\mathcal{F}$. Then $T_1 + T_2$ is also a descent modulus for $\mathcal{F}$, where*

$$(T_1 + T_2)[f](x) := T_1[f](x) + T_2[f](x), \quad \text{for all } f \in \mathcal{F} \text{ and } x \in \operatorname{dom} f.$$

(ii). *Let T be a descent modulus for $\mathcal{F}$ and let $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ be a strictly increasing function with $\phi(0) = 0$ and $\limsup_{t \to +\infty} \phi(t) = +\infty$. Then*

$$(\phi T)[f](x) := (\phi \circ T[f])(x), \quad \text{for all } f \in \mathcal{F} \text{ and } x \in \operatorname{dom} f,$$

is also a descent modulus for $\mathcal{F}$, under the convention $\phi(+\infty) = \limsup_{t \to +\infty} \phi(t) = +\infty$.

In particular, rT, $r \geq 0$ is a descent modulus for $\mathcal{F}$, where

$$(rT)[f](x) := r \cdot T[f](x),$$

under the convention $r \cdot (+\infty) = +\infty$ for $r > 0$, and $0 \cdot (+\infty) = 0$.

Proof. Let T_1, T_2, T and ϕ as in the statements (i) and (ii). We show that axioms (D1)–(D3) of Definition 3.1 are fulfilled:

– (D1) (*Preservation of global minima*) Let $f \in \mathcal{F}$ and $x \in \operatorname{argmin} f$. Then

$$T_1[f](x) = T_2[f](x) = T[f](x) = 0$$

and consequently $(T_1 + T_2)[f](x) = 0$ and $\phi(T[f](x)) = \phi(0) = 0$. Therefore, $T_1 + T_2$ and ϕT preserve global minima.

– (D2) (*Monotonicity*) Let $f, g \in \mathcal{F}$ and $x \in X$ such that

$$(f(x) - f(z))_+ \geq (g(x) - g(z))_+, \quad \forall x \in X.$$

Then, since T_1 and T_2 are monotone, we have that

$$(T_1 + T_2)[f](x) = T_1[f](x) + T_2[f](x) \geq T_1[g](x) + T_2[g](x) = (T_1 + T_2)[g](x).$$

Similarly, since T is monotone and ϕ is non-decreasing, we get that

$$(\phi T)[f](x) = \phi(T[f](x)) \geq \phi(T[g](x)) = (\phi T)[g](x).$$

Thus, $T_1 + T_2$ and ϕT are monotone.

– (D3) (*Scalar monotonicity*) Let $f \in \mathcal{F}$, $x \in X$ and $r > 1$ and assume $0 < (T_1 + T_2)[f](x) < +\infty$. Up to a mutual change of T_1 and T_2, we many assume $0 < T_1[f](x) < +\infty$. Then, using the scalar monotonicity of T_1 and the monotonicity of T_2, we deduce

$$\begin{aligned}
(T_1 + T_2)[rf](x) = T_1[rf](x) + T_2[rf](x) &> T_1[f](x) + T_2[rf](x) \\
&\geq T_1[f](x) + T_2[f](x) = (T_1 + T_2)[f](x).
\end{aligned}$$

Thus, $(T_1 + T_2)$ is scalar-monotone.

Let us now assume $0 < (\phi T)[f](x) < +\infty$. Since $\phi(0) = 0$ and $\phi(+\infty) = +\infty$, we obtain again $0 < T[f](x) < +\infty$. Thus, $T[rf](x) > T[f](x)$ and

$$(\phi T)[rf](x) = \phi(T[rf](x)) > \phi(T[f](x)) = (\phi T)[rf](x),$$

yielding that (ϕT) is scalar-monotone. We conclude that both $(T_1 + T_2)$ and (ϕT) are descent moduli for $\mathcal{F}$. $\square$

Notice that the family of descent moduli for the class $\mathcal{F}$ has the structure of a convex cone (i.e., it is a cone closed for the sum), with the sum and the scalar multiplication being defined as in Proposition 3.9.

The following proposition provides other types of operations, based on *truncations*, that preserve descent moduli.

Proposition 3.10 (Truncated descents). *Let T be a descent modulus for the class $\mathcal{F}$. Then:*

(i). For every $\varepsilon > 0$, the operator T_ε given by

$$T_\varepsilon[f](x) = \begin{cases} T[f](x), & \text{if } f(x) > \inf f + \varepsilon \\ 0, & \text{otherwise,} \end{cases}$$

is a descent modulus for $\mathcal{F}$.

(ii). For every $K \subset X$, the operator $T\big|_K$ given by

$$T\big|_K[f](x) = \begin{cases} T[f](x), & \text{if } x \in K \\ 0, & \text{otherwise,} \end{cases}$$

is a descent modulus for $\mathcal{F}$.

Proof. Let T, $\varepsilon > 0$ and $K \subset X$ as in the statement of the proposition. We will show that the operators T_ε and $T\big|_K$ satisfy properties (D1)–(D3) of Definition 3.1. Notice that for every $f \in \mathcal{F}$ and $x \in X$ we have $T[f](x) \geq T_\varepsilon[f](x)$ and $T[f](x) \geq T\big|_K[f](x)$. Therefore, if $T[f](x) = 0$, the above readily yields $T_\varepsilon[f](x) = \big|_K(x) = 0$, and (D1) holds trivially.

Let us now prove (D2). To this end, Let $f, g \in \mathcal{F}$ and $x \in X$ such that

$$(f(x) - f(z))_+ \geq (g(x) - g(z))_+, \quad \forall x \in X.$$

Let us first deal with T_ε: if $f(x) > \inf f + \varepsilon$, then $T_\varepsilon[f](x) = T[f](x) \geq T[g](x) \geq T_\varepsilon[g](x)$. On the other hand, if $f(x) \leq \inf f + \varepsilon$, then $[f(x) - f(z)]_+ \leq \varepsilon$ for all $z \in X$, whence $g(x) \leq \inf g + \varepsilon$ and $T_\varepsilon[f](x) = T_\varepsilon[g](x) = 0$ (by definition of T_ε). We conclude that $T_\varepsilon[f](x) \geq T_\varepsilon[g](x)$.

Let us now deal with $T\big|_K$: If $x \in K$, then $T\big|_K[f](x) = T[f](x) \geq T[g](x) = T\big|_K[g](x)$, while if $x \in X \setminus K$, then $T\big|_K[f](x) = T\big|_K[g](x) = 0$. In both cases $T\big|_K[f](x) \geq T\big|_K[g](x)$.

It remains to prove (D3). Let $f \in \mathcal{F}$, $x \in X$ and $r > 1$. If $\inf f = -\infty$, then $T_\varepsilon[f] = T[f]$ and the result follows. Therefore, we may assume $\inf f > -\infty$ and $0 < T_\varepsilon[f](x) < +\infty$. This yields $f(x) > \inf f + \varepsilon$ and consequently, $T_\varepsilon[f](x) = T[f](x)$. Noting that

$$r f(x) > r(\inf f + \varepsilon) > \inf r f + \varepsilon,$$

we conclude that $T_\varepsilon[rf](x) = T[rf](x)$, as well. Then, since $T_\varepsilon[f](x) = T[f](x) < T[rf](x) = T_\varepsilon[rf](x)$, we conclude that T_ε is scalar-monotone.

Let us now assume $0 < T\big|_K[f](x) < +\infty$. This yields in particular that $x \in K$ and so $T\big|_K[f](x) = T[f](x)$ and $T\big|_K[rf](x) = T[rf](x)$. Then, since $T\big|_K[f](x) = T[f](x) < T[rf](x) = T\big|_K[rf](x)$, we conclude that $T\big|_K$ is scalar-monotone. $\square$

The last stability property that we study is the pointwise limit. In general, this operation does not preserve moduli of descent, since scalar-monotonicity can be lost in the limit process, as the following example reveals.

16

Example 3.11 (Axiom (D3) is not preserved under pointwise limits). Let $X = \mathbb{R}^n$ and consider the class $\mathcal{F} = \mathcal{C}^1(\mathbb{R}^n)$ of $\mathcal{C}^1$-smooth functions. Let us further consider the sequence of descent moduli

$$T_n[f](x) = \sqrt[n]{\|\nabla f(x)\|}, \quad n \in \mathbb{N},$$

and its pointwise limit operator:

$$T[f](x) = \lim_{n \to \infty} T_n[f](x) = \begin{cases} 0, & \text{if } \nabla f(x) = 0, \\ 1, & \text{otherwise.} \end{cases}$$

The operator T preserves global minima and is monotone. However, it is not scalar-monotone (and it clearly fails to determine functions in the sense of Theorem 3.5.) ◇

The following definition introduces a large subclass of abstract descent moduli which provides a remedy to the above situation.

Definition 3.12 (Homogeneous descent moduli). Let $\mathcal{F} \subset \mathcal{C}(X)$ be a translation cone, and let $p \in (0, +\infty)$. An operator $T : \mathcal{F} \to (\overline{\mathbb{R}}_+)^X$ is said to be

(i). *p–homogeneous* if $T[rf](x) = r^p\, T[f](x)$, for every $f \in \mathcal{F}$ and $r > 0$.

(ii). *p–superhomogeneous* if $T[rf](x) \geq r^p\, T[f](x)$, for every $f \in \mathcal{F}$ and $r > 0$.

Clearly all p–homogeneous and all p–superhomogeneous operator are also scalar-monotone. The interest of this class is that every operator T which is defined as a pointwise limit of a sequence of p-(super)homogeneous descent moduli $\{T_n\}_{n \in \mathbb{N}}$, that is,

$$T[f](x) = \lim_{n \to +\infty} T_n[f](x), \qquad \text{for all } f \in \mathcal{F} \text{ and } x \in \text{dom}(f),$$

is itself a p–(super)homogeneous descent modulus. In other words, axiom (D3) (scalar-monotonicity) is preserved in this context. One can also observe that up to a composition with the strictly increasing function $\varphi(t) := t^{1/p}$, p–(super)homogenicity reduces to 1 –(super)homogenicity.

Proposition 3.13. *Let $(\Lambda, \preccurlyeq)$ be a directed set, $p \in (0, +\infty)$ and $(T_\alpha)_{\alpha \in \Lambda}$ be a generalized sequence of p –(super)homogeneous descent moduli for the class $\mathcal{F}$. Then the following operators, defined for every $f \in \mathcal{F}$ and $x \in \text{dom}(f)$, are descent moduli for the class $\mathcal{F}$:*

(i). $\left(\limsup\limits_{\alpha \in \Lambda} T_\alpha \right) [f](x) := \limsup\limits_{\alpha \in \Lambda} T_\alpha[f](x);$

(ii). $\left(\sup\limits_{\alpha \in \Lambda} T_\alpha \right) [f](x) := \sup\limits_{\alpha \in \Lambda} T_\alpha[f](x);$

(iii). $\left(\liminf\limits_{\alpha \in \Lambda} T_\alpha \right) [f](x) := \liminf\limits_{\alpha \in \Lambda} T_\alpha[f](x)$;

(iv). $\left(\inf\limits_{\alpha \in \Lambda} T_\alpha \right) [f](x) := \inf\limits_{\alpha \in \Lambda} T_\alpha[f](x).$

17

Proof. Let us verify that $T := \limsup_\alpha T_\alpha$ satisfies axioms (D1)–(D3) of Definition 3.1. (A similar reasoning will apply to the other three operators.)

– (D1) (*Preservation of global minima*) Choose $f \in \mathcal{F}$ and $x \in \arg\min f$. Then, $T_\alpha[f](x) = 0$ for all $\alpha \in \Lambda$ and so $T[f](x) = 0$. Thus, T preserves global minima.

– (D2) (*Monotonicity*). Let $f, g \in \mathcal{F}$ and $x \in X$ be such that $(f(x) - f(z))_+ \geq (g(x) - g(z))_+$, for all $z \in X$. Then, $T_\alpha[f](x) \geq T_\alpha[g](x)$ for each $\alpha \in \Lambda$. Thus, $T[f](x) \geq T[g](x)$ as well, showing that T is monotone

– (D3) (*Scalar-monotonicity*): Let $f \in \mathcal{F}$ and $x \in X$. We readily deduce from p-superhomogeneity that $T[rf](x) = \limsup_\alpha T_\alpha[rf](x) \geq r^p \limsup_\alpha T_\alpha[f](x) = r^p T[f](x)$. It follows that T is also p–superhomogeneous, therefore, in particular, scalar-monotone.

The proof is complete. $\qquad\square$

3.4 Slope-like operators that are not descent moduli

We finish this section by discussing two examples in the literature that have being introduced as "slope operators" on a metric space (X, d), but fail to verify Definition 3.1 of descent modulus. The first concept is the so-called *weak slope*, introduced in [11, 8]. For a continuous function $f : X \to \mathbb{R}$, the weak slope at a point $x \in X$, denoted by $|df|(x)$, is defined as the supremum of $\sigma \in \mathbb{R}_+$ such that there exist $\delta > 0$ and a continuous map $\mathcal{H} : [0, \delta] \times B(x, \delta) \to X$ such that

$$\forall s \in [0, \delta], \; \forall y \in B(x, \delta), \quad d(\mathcal{H}(s, y), y) \leq s \; \text{ and } \; f(\mathcal{H}(s, y)) \leq f(y) - \sigma s. \tag{27}$$

Notice that $|df(x)| \geq \sigma$ whenever it is possible to find a continuous deformation $\mathcal{H}$ over a neighborhood of x, such that the descent of f through that deformation is at least σ for every point y over which $\mathcal{H}$ is acting. Thus, one might interpret the weak-slope as the slowest descent around x. This concept has been largely studied in the setting of nonsmooth variational analysis and critical point theory.

The second concept is the *limiting slope* (see, e.g. [16, Definition 8.4]), which is defined as the lower semicontinuous envelope (or closure) of the strong slope $|\nabla f|$. That is, for a lower semicontinuous function $f : X \to \mathbb{R}$ and a point $x \in X$, the limiting slope of f at x is given by

$$\overline{|\nabla f|}(x) := \lim_{\varepsilon \to 0} \inf \left\{ |\nabla f|(y) \; : \; d(x, y) \leq \varepsilon, \text{ and } f(y) \leq f(x) + \varepsilon \right\}. \tag{28}$$

Since the slope that can be very ill-behaved, the limiting slope provides a regularized alternative. It is worth to mention that using this notion, Drusvyatskiy, Ioffe and Lewis were able to deal with the long-standing problem of existence of steepest descent curves [12].

The following example shows that the weak slope and the limiting slope are not descent moduli for $\mathcal{K}(X)$, since they fail to determine coercive continuous functions even in the interval $[0, 1]$.

Example 3.14. Let $\mathfrak{c} : [0, 1] \to [0, 1]$ be the well-known Cantor Staircase and let us consider the function $f : [0, 1] \to \mathbb{R}$ given by $f(t) = \mathfrak{c}(t) + t$. By construction, it is not hard to see that $|\nabla f|(t) \in \{1, +\infty\}$ for every $t \in (0, 1]$, that $|\nabla f|(0) = 0$ (since $0 \in \arg\min f$), and that the slope is $+\infty$ only in a subset of the Cantor set. Thus,

$$\overline{|\nabla f|}(t) = \mathbb{1}_{(0,1]}(t) := \begin{cases} 1, & \text{if } t \in (0, 1] \\ 0, & \text{if } t = 0. \end{cases}$$

18

Similarly, we claim that $|df|(t)$ takes the same values as $\overline{|\nabla f|}(t)$. Clearly $|df|(0) = 0$ and $|df|(t) \geq 1$ for all $t \in (0, 1]$. Now, fix $\bar{t} \in (0, 1]$ and take any $\sigma > 0$, $\delta > 0$ and $\mathcal{H}$ satisfying (27). Since f is strictly increasing, $\mathcal{H}(s, t) < t$ for every $t \in B(\bar{t}, \delta)$ and every $s \in [0, \delta]$. In particular, $0 < t - \mathcal{H}(s, t) = d(\mathcal{H}(s, t), t) \leq s$. Whence $t - s \leq \mathcal{H}(s, t)$ and consequently, $f(t-s) \leq f(\mathcal{H}(s, t))$. Since the Cantor set is totally disconnected, there exists $t \in (\bar{t} - \delta, \bar{t})$ such that $|\nabla f|(t) = 1$. Then,

$$|\nabla f|(t) \geq \limsup_{s \to 0^+} \frac{f(t) - f(t-s)}{s} \geq \limsup_{s \to 0^+} \frac{f(t) - f(\mathcal{H}(s, t))}{s} \geq \sigma.$$

Thus, $\sigma \leq 1$, which proves that $|df(\bar{t})| = 1$. This proves the claim. By taking $g : [0, 1] \to \mathbb{R}$ given by $g(t) = t$, we get that $\overline{|\nabla g|}(t) = |dg|(t) = \mathbb{1}_{(0,1]}(t)$, and so, the conclusion of Theorem 3.5 fails to hold for both the weak and the limiting slope. Since clearly both operators preserve global minima and are scalar-monotone (by homogeneity), we conclude that both operators fail to be monotone in the sense of Definition 3.1. ◇

4 The paradigm of averaged descent

It was shown in [6, Theorem 3.8] that two $\mathcal{C}^2$-smooth convex and bounded from below functions f, g defined on a Hilbert space $\mathcal{H}$ are equal up to a constant, provided $\|\nabla f(x)\| = \|\nabla g(x)\|$, for all $x \in \mathcal{H}$. In other words, the operator:

$$f \mapsto \Gamma[f] := \|\nabla f\|^2 \tag{29}$$

is injective, modulo the constant functions, on the class of $\mathcal{C}^2$-smooth convex and bounded from below functions. Notice that the Γ-operator defined in (29) (also known as *carré-du-champ* operator) is strongly related to the Wiener diffusion process, generated by the Laplacian operator. This hints towards a new important instance of descent modulus, namely the average descent, giving rise to a determination result of probabilistic nature. This will be developed in this section, in full generality.

4.1 Extension of dispersion measures

We first recall that for a $\mathcal{C}^1$-smooth function $f : \mathbb{R}^n \to \mathbb{R}$ the following formula holds:

$$\|\nabla f(x)\|^2 = \lim_{\varepsilon \to 0} \frac{n}{\mathcal{L}_n(B_n(x, \varepsilon))} \int_{B_n(x, \varepsilon)} \left[\frac{f(x) - f(y)}{\|x - y\|} \right]^2 \mathcal{L}_n(dy), \tag{30}$$

where, as mentioned in Section 2, $\mathcal{L}_n$ stands for the usual Lebesgue measure on $\mathbb{R}^n$. The above formula is well-known and can be deduced from the following (also well-known) lemma, for which we provide a simple proof for completeness.

Lemma 4.1. *For any $k \geq 1$, any $r > 0$ and $V \in \mathbb{R}^k$ it holds:*

$$\|V\|^2 = \frac{k}{\mathcal{L}_k(B_k(0, r))} \int_{B_k(0, r)} \left\langle V, \frac{u}{\|u\|} \right\rangle^2 du. \tag{31}$$

Proof. The proof is a consequence of the invariance by rotations of the ball. Consider $(e_i)_{i=1}^k$ the usual orthonormal basis of $\mathbb{R}^k$. By symmetry, we can restrict to the case where $V = \|V\| \cdot e_1$,

so that

$$\int_{B_k(0,r)} \left\langle V, \frac{u}{\|u\|} \right\rangle^2 du = \|V\|^2 \int_{B_k(0,r)} \left\langle e_1, \frac{u}{\|u\|} \right\rangle^2 du$$

$$= \|V\|^2 \int_{B_k(0,r)} \left\langle e_i, \frac{u}{\|u\|} \right\rangle^2 du$$

for any $i \in \{1, ..., n\}$. We deduce

$$\int_{B_k(0,r)} \left\langle V, \frac{u}{\|u\|} \right\rangle^2 du = \frac{\|V\|^2}{k} \int_{B_k(0,r)} \sum_{i=1}^{n} \left\langle e_i, \frac{u}{\|u\|} \right\rangle^2 du$$

$$= \frac{\|V\|^2}{k} \int_{B_k(0,r)} \left\| \frac{u}{\|u\|} \right\|^2 du$$

$$= \frac{\|V\|^2}{k} \int_{B_k(0,r)} 1 \, du$$

$$= \frac{\|V\|^2 \mathcal{L}_k(B_k(0,r))}{k}$$

leading to the desired equality.

$\square$

Based on equation (30), we propose an extension of the Γ-operator (29), that we call *dispersion operator*, for functions defined on a topological space (X, τ).

To this end, we consider the family $\beta = \{\beta_x\}_{x \in X}$ of neighborhood bases: β_x is a neighborhood base at x of the topology τ. We further denote by

$$\mu : X \times \mathcal{B}(X) \to \mathbb{R}_+$$

a mapping that associates to every $x \in X$, a locally finite measure $\mu(x, \cdot) \equiv \mu_x$ (that is, for every $y \in X$, μ_x is finite on a neighborhood V_y of y), with positive measure at every element of β_x. Let further $m : X \times X \to \mathbb{R}_+$ be as in Proposition 3.7, that is,

$$m(x, y) = 0 \iff x = y.$$

Finally, let us consider the *local dimension* mapping $n : X \to \mathbb{R}_+$, where we interpret $n(x)$ to be the local dimension of X at x. (Obviously, if $X = \mathbb{R}^n$ or if X is a manifold of dimension n, then $n(x) \equiv n$, for all $x \in X$.)

We are now ready to give the following definition:

Definition 4.2 (Dispersion operator). Let $p \in (0, +\infty)$. We define the *p-dispersion operator* T_μ (depending also on β and $n : X \to \mathbb{R}_+$) as follows:

$$T_\mu[f](x) := \limsup_{B \in \beta_x} \frac{n(x)}{\mu_x(B)} \int_B |\Delta_f(x, y)|^p \, \mu_x(dy) \tag{32}$$

where the limit-superior is taken over the inductive set β_x endowed with the partial order of the reverse inclusion and

$$\Delta_f(x, y) := \begin{cases} \frac{f(x) - f(y)}{m(x, y)}, & \text{if } y \neq x, \\ \\ 0, & \text{if } y = x. \end{cases} \tag{33}$$

20

Remark 4.3. (i). We kept the notation simple and denoted the above dispersion operator by T_μ (rather than $T_{\mu,\beta,m,n,p}$) in order to emphasize that T_μ is the limit-superior of integral operators. The action at x in these operators is integrated by the measure μ_x.

(ii). Definition 4.2 is inspired by a construction used in [23] to extend diffusion processes to metric spaces. The "lim sup" ensures that T_μ is always well-defined, with possibly $+\infty$–values. When X is a metric space and m is the distance function, the domain $\mathrm{dom}(T_\mu)$ contains at least all (locally) Lipschitz functions. This makes the dispersion operator to be a nontrivial extension of (29) beyond the differentiable setting.

(iii). The family β in Definition 4.2 encompasses several natural choices when the structure of the space (X, τ) is known. For example, if (X, τ) is a (pseudo)metric space, then we can take the set of corresponding balls $\beta_x = \{B(x,r)\}_{r>0}$, for all $x \in X$. More generally, if the topological space (X, τ) is first-countable, then a natural choice is $\beta_x = \{\mathcal{V}_n\}_{n\in\mathbb{N}}$, where $\{\mathcal{V}_n\}_{n\in\mathbb{N}}$ is any countable basis of the neighborhoods of x.

If $X = \mathbb{R}^n$, then our default choice will be $\beta_x := \{B(x,r)\}_{r>0}$.

We denote by $\mathcal{S}_n^+$ the set of $(n \times n)$–positive semidefinite matrices, and let us consider a map $R : \mathbb{R}^n \to \mathcal{S}_n^+$. The following proposition shows that the operators of the form

$$\Gamma_R[f](x) = \|R(x)\nabla f(x)\|^2, \ x \in \mathbb{R}^n$$

can be obtained as particular cases of (32), under suitable choices of the parameter $p > 0$, the separation map m, the measure map $\mu : \mathbb{R}^n \times \mathcal{B}(\mathbb{R}^n) \to \mathbb{R}_+$ and a local dimension map $x \mapsto n(x)$.

In what follows $\mathrm{supp}\,(\mu_x)$ stands for the support of the measure $\mu_x := \mu(x, \cdot)$. We say that a measure μ is absolutely continuous with respect to ν (and denote $\mu << \nu$) if both measures are defined on the same measurable space $(X, \mathcal{B})$ and it holds:

$$\nu(A) = 0 \implies \mu(A) = 0, \quad \text{ for all } A \in \mathcal{B}.$$

We are now ready to state and prove the following result:

Proposition 4.4. *Let $R : \mathbb{R}^n \to \mathcal{S}_n^+$ and set $W_x := x + \mathrm{Ker}(R(x))^\perp$, for each $x \in \mathbb{R}^n$. Then for $m(x,y) := \|x - y\|$, and $p = 2$, there exist a measure map $\mu : \mathbb{R}^n \times \mathcal{B}(\mathbb{R}^n) \to \mathbb{R}_+$ and a dimension map $n : X \to \mathbb{R}_+$ such that $\mathrm{supp}\,(\mu_x) \subset W_x$ for all $x \in \mathbb{R}^n$ and*

$$T_\mu[f](x) = \|R(x)\nabla f(x)\|^2, \quad \text{for every } f \in \mathcal{C}^1(\mathbb{R}^n).$$

Proof. Let us fix $x \in \mathbb{R}^n$. We are going to define a positive real value $n(x)$ and a measure μ_x whose support is contained in W_x, in a way that:

$$T_\mu[f](x) = \limsup_{r>0} \frac{n(x)}{\mu_x(B(x,r))} \int_{B(x,r)} \Delta_f(x,y)^2 \mu_x(dy) = \|R(x)\nabla f(x)\|^2. \tag{34}$$

Set $k = \dim (\mathrm{Ker}\, R(x))^\perp$, $0 \le k \le n$.

If $k = 0$, then $\mathrm{Ker}\, R(x) = \mathbb{R}^n$, $W_x \equiv x$ and $R(x)\nabla f(x) = 0$. Then (34) holds trivially by setting $\mu_x = \delta_x$ (the Dirac measure at x) and using the fact that $\Delta_f(x,x) = 0$ (*cf.* (33)).

Let us now assume that $1 \le k \le n$. Let $\{e_j\}_{j=1}^n$ be an orthonormal base of $\mathbb{R}^n$ such that

$$(\mathrm{Ker}\, R(x))^\perp = \mathrm{span}\,(e_j)_{j=1}^k = \mathbb{R}^k$$

21

and

$$\text{Ker } R(x) \equiv \mathbb{R}^{n-k} = \begin{cases} \text{span } (e_j)_{j=k+1}^n, & \text{for } k < n \\ \{0\} & \text{for } k = n. \end{cases}$$

Then there exists $R \in \mathcal{S}_k^+$ (the trace of $R(x) \in \mathcal{S}_n^+$ on the subspace $\mathbb{R}^k \times \{0\}^{n-k}$ of $\mathbb{R}^n$) such that decomposing $z \in \mathbb{R}^n$ as $z = (v, w) \in \mathbb{R}^k \times \mathbb{R}^{n-k}$, it holds $R(x)z = Rv$.

Let $\Psi : \mathbb{R}^k \to \mathbb{R}^k$ be given by

$$\Psi(u) = \begin{cases} \frac{\|u\|}{\|Ru\|} Ru & \text{if } u \neq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{35}$$

Clearly Ψ is an isometric automorphism of $\mathbb{R}^k$ (depending on x, which is fixed) with inverse:

$$\Psi^{-1}(v) = \begin{cases} \frac{\|v\|}{\|R^{-1}v\|} R^{-1}v & \text{if } v \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

In particular $\Psi(B_k(0, r)) = B_k(0, r)$, for every $r > 0$. Furthermore, Ψ is a $\mathcal{C}^1$-diffeomorphism of $\mathbb{R}^k \setminus \{0\}$. Following the notation of [14, Chapter 3], let us define the Jacobian operator as

$$J\Psi(u) = |\det(D\Psi(u))|, \tag{36}$$

where $D\Psi$ is the derivative of Ψ. We define $h_k : \mathbb{R}^k \to \mathbb{R}$ (depending on Ψ, therefore on x) such that

$$h_k(v) = \frac{\|R\Psi^{-1}(v)\|^2}{\|\Psi^{-1}(v)\|^2} \left[J\Psi(\Psi^{-1}(v)) \right]^{-1}, \quad v \in \mathbb{R}^k.$$

Notice that for $v = \Psi(u)$ the above yields:

$$h_k(\Psi(u)) = \frac{\|Ru\|^2}{\|u\|^2} [J\Psi(u)]^{-1}, \quad u \in \mathbb{R}^k. \tag{37}$$

We set

$$\begin{cases} h : \mathbb{R}^n \to \mathbb{R} \\ h(z) := h_k(v), \quad \text{for } z = (v, w) \in \mathbb{R}^n \end{cases} \tag{38}$$

and consider the measure $\lambda : \mathcal{B}(\mathbb{R}^n) \to \mathbb{R}_+$ (depending on $k = \dim(W_x)$) given by the formula

$$\lambda(A) = \mathcal{L}_k \left(A \cap \left(\mathbb{R}^k \times \{0\}^{n-k} \right) \right), \quad \text{for all } A \in \mathcal{B}(\mathbb{R}^n). \tag{39}$$

Let π_k denote the projection of $\mathbb{R}^n$ to the first k-coordinates. Notice that λ is the trivial extension to $\mathcal{B}(\mathbb{R}^n)$ of the Lebesgue measure $\mathcal{L}_k$ on $\mathcal{B}(\mathbb{R}^k)$. Let us define

$$\kappa := \kappa(x) = \mathcal{L}_k(B_k(0, 1))^{-1} \left(\int_{B_k(0,1)} \frac{\|Ru\|^2}{\|u\|^2} \mathcal{L}_k(du) \right). \tag{40}$$

where $B_k(0, r) = \pi_k \left(B(0, r) \cap \left(\mathbb{R}^k \times \{0\}^{n-k} \right) \right)$. We finally set $n(x) := \kappa(x) \dim(W_x)$ and define the measure $\mu_x : \mathcal{B}(\mathbb{R}^n) \to \mathbb{R}_+$ as follows:

$$\mu_x(A) := \int_{A-x} h(z)\lambda(dz) \equiv \int_{\pi_k\left((A-x)\cap\left(\mathbb{R}^k\times\{0\}^{n-k}\right)\right)} h_k(v)\, \mathcal{L}_k(dv), \quad \text{for all } A \in \mathcal{B}(\mathbb{R}^n). \tag{41}$$

22

This operation eliminates the last $n - k$ coordinates (which are equal to 0 for all elements of $(A - x) \cap (\mathbb{R}^k \times \{0\}^{n-k})$), adjusting vectors to the right dimension for integration. By means of a change of variables induced by Ψ (see, e.g., [14, Theorem 3.9]), we deduce

$$\mu_x(B(x,r)) = \int_{B_k(0,r)} h_k(v)\mathcal{L}_k(dv) = \int_{B_k(0,r)} h_k(\Psi(u))J\Psi(u)\mathcal{L}_k(du)$$

$$= r^k \int_{B(0,1)} \frac{\|Ru\|^2}{\|u\|^2} \mathcal{L}_k(du) = \kappa\mathcal{L}_k(B_k(0,1))r^k = \kappa\mathcal{L}_k(B_k(0,r))$$

Now, using the first-order Taylor approximation of f at x, we deduce from (33) that

$$\Delta_f(x,y)^2 = \left[\left\langle \nabla f(x), \frac{y - x}{\|y - x\|} \right\rangle + \varepsilon(\|y - x\|) \right]^2, \qquad \text{where } \lim_{r \to 0} \varepsilon(r) = 0.$$

For any $r > 0$ we deduce from (41) that:

$$\int_{B(x,r)} \Delta_f(x,y)^2 \mu_x(dy) = \int_{B(0,r)} \left[\langle \nabla f(x), \frac{z}{\|z\|} \rangle + \varepsilon(\|z\|) \right]^2 h(z)\lambda(dz)$$

$$= \int_{B(0,r)} \langle \nabla f(x), \frac{z}{\|z\|} \rangle^2 h(z)\,\lambda(dz) +$$

$$+ \int_{B(0,r)} 2 \langle \nabla f(x), \frac{z}{\|z\|} \rangle \varepsilon(\|z\|)\,h(z)\,\lambda(dz) + \int_{B(0,r)} \varepsilon(\|z\|)^2\,h(z)\,\lambda(dz).$$

Let $M > 0$ be an upper bound of the function h on $B(0,1)$. Since $\mu_x(B(x,r)) = \kappa\mathcal{L}_k(B_k(0,r))$ and $n(x) = k \cdot \kappa$, it follows that

$$\frac{2n(x)}{\mu_x(B(x,r))} \int_{B(0,r)} \left\langle \nabla f(x), \frac{z}{\|z\|} \right\rangle \varepsilon(\|z\|)h(z)\lambda(dz) \leq 2\,k\,M\,\|\nabla f(x)\|\,\varepsilon(r) \longrightarrow 0 \quad \text{and}$$

$$\frac{n(x)}{\mu_x(B(x,r))} \cdot \int_{B(0,r)} \varepsilon(\|z\|)^2 h(z)\lambda(dz) \leq k\,M\,\varepsilon(r)^2 \longrightarrow 0 \quad \text{(as } r \to 0\text{)}.$$

Denoting by $[\nabla f(x)]_k \in \mathbb{R}^k$ the vector consisting of the first k-coordinates of $\nabla f(x)$ and recalling the decomposition $z = (v,w) \in \mathbb{R}^k \times \mathbb{R}^{n-k}$ we deduce from (38):

$$\int_{B(0,r)} \left\langle \nabla f(x), \frac{z}{\|z\|} \right\rangle^2 h(z)\lambda(dz) = \int_{B(0,r)} \left\langle \nabla f(x), \frac{(v,w)}{\|(v,w)\|} \right\rangle h_k(v)\mathcal{L}_k(dv).$$

Using the change of variables $v = \Psi(u)$ (recall that $\Psi(B(0,r)) = B(0,r)$ for every $r > 0$) we obtain from (35) and (37)

$$\int_{B(0,r)} \left\langle \nabla f(x), \frac{z}{\|z\|} \right\rangle^2 h(z)\lambda(dz)$$

$$= \int_{B_k(0,r)} \left\langle [\nabla f(x)]_k, \frac{\Psi(u)}{\|\Psi(u)\|} \right\rangle^2 \frac{\|Ru\|^2}{\|u\|^2} [J\Psi(u)]^{-1}J\Psi(u)du$$

$$= \int_{B_k(0,r)} \left\langle [\nabla f(x)]_k, \frac{Ru}{\|u\|} \right\rangle^2 du = \int_{B_k(0,r)} \left\langle R[\nabla f(x)]_k, \frac{u}{\|u\|} \right\rangle^2 du.$$

Therefore we deduce from (31) and from the definitions of $n(x)$ and μ_x:

$$
\begin{aligned}
T_\mu[f](x) &= \limsup_{r>0} \frac{n(x)}{\mu_x(B(x,r))} \int_{B_k(0,r)} \left\langle R[\nabla f(x)]_k, \frac{u}{\|u\|} \right\rangle^2 du \\
&= \limsup_{r>0} \frac{k}{\mathcal{L}_k(B_k(0,r))} \int_{B_k(0,r)} \left\langle R[\nabla f(x)]_k, \frac{u}{\|u\|} \right\rangle^2 du \\
&= \|R[\nabla f(x)]_k\|^2 \equiv \|R(x)\nabla f(x)\|^2 .
\end{aligned}
$$

The proof is complete. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

4.2 Oriented dispersion operators

The operator T_μ defined in (32) fails to determine continuous coercive functions, and consequently is not a descent modulus outside the differentiable setting. The reason for this failure will be illustrated in the following example.

Example 4.5. Let $X = [-1, 1]$ and let m be its usual metric. For each $x \in [-1, 1]$, let $\mu(x, \cdot)$ be the usual Lebesgue measure over $[-1, 1]$ and $n(x) = 1$. Set $p = 2$.

$$
f(x) = x^2 \qquad \text{and} \qquad g(x) = -x^2 .
$$

By (30), we have that

$$
\forall x \in (-1, 1), \quad T_\mu[g](x) = T_\mu[f](x) = |2x|^2 .
$$

Furthermore, it is not hard to see that at $x = \pm 1$, we have that

$$
T_\mu[g](x) = T_\mu[f](x) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_{1-\varepsilon}^1 \left(\frac{1 - t^2}{1 - t} \right)^2 dt = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_{1-\varepsilon}^1 (1 + t)^2 dt = 4
$$

Since the only T-critical point of g is 0, we deduce that T does not preserve global minima and so it is not a descent modulus. Furthermore, since the only T- critical point of f is 0 as well, we have constructed two different functions with $T_\mu[f] = T_\mu[g]$ and that coincide over $\mathcal{Z}_T(f)$. In conclusion, T fails to determine continuous coercive functions in general metric spaces, in the sense of Theorem 3.5. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\Diamond$

In the above example, the points $x = -1$ and $x = 1$ should have been critical for the function $g(x) = -x^2$, since they are global minimizers. However, this fails to be the case because the operator T_μ is not oriented. This leads to the following definition, which induces asymmetry between descent and ascent directions (by penalizing the latter). As we shall see, this is particularly relevant in nonsmooth settings.

Definition 4.6 (Oriented dispersion operator)**.** Let μ, β, m, n and p be as in Definition 4.2. We define the oriented dispersion operator, denoted by T_μ^+, as

$$
\begin{aligned}
T_\mu^+[f](x) &:= \limsup_{B \in \beta_x} \frac{n(x)}{\mu_x(B)} \int_{B \cap [f \le f(x)]} [\Delta_f(x, y)]^p \, \mu(x, dy) \\
&= \limsup_{B \in \beta_x} \frac{n(x)}{\mu_x(B)} \int_B \left[\Delta_f^+(x, y) \right]^p \mu(x, dy),
\end{aligned}
$$

24

where

$$\Delta_f^+(x,y) := \begin{cases} \frac{(f(x)-f(y))_+}{m(x,y)} & \text{if } x \neq y \\ \\ 0 & \text{if } x = y. \end{cases} \tag{42}$$

The value $T_\mu^+[f](x)$ corresponds to the dispersion of f at x which is exclusively due to the directions of descent. In the smooth case, the value of the oriented dispersion $T_\mu^+[f](x)$ is the half of the value of the non-oriented dispersion $T_\mu[f](x)$, as expected by symmetry. This is the content of the following proposition.

Proposition 4.7. *Let $X = \mathbb{R}^n$, $n(x) \equiv n$, $\beta_x = \{B(x,\varepsilon) \; : \; \varepsilon > 0\}$ and μ_x be the n-dimensional Lebesgue measure for every $x \in \mathbb{R}^n$. Take $p = 2$ and $m(x,y) = \|x - y\|$. Then*

$$T_\mu^+[f](x) = \frac{1}{2}\|\nabla f(x)\|^2, \quad \text{for every } f \in \mathcal{C}^1(\mathbb{R}^n).$$

Proof. If $\|\nabla f(x)\| = 0$, then $0 \leq T_\mu^+[f](x) \leq T_\mu[f](x) = \|\nabla f(x)\|^2 = 0$ and the conclusion follows trivially.

Let us now consider the case $\|\nabla f(x)\| \neq 0$. By a change of coordinates, we may assume that $x = 0$, $f(0) = 0$ and $\nabla f(0) = r\,e_n$, where $r = \|\nabla f(0)\| > 0$ and e_n be the n-th vector of an orthonormal base of $\mathbb{R}^n$. In this setting, we denote

$$S := [f \leq f(0)] \quad \text{and} \quad \mathbb{R}^{n-1} := \text{span}\{e_j\}_{j=1}^{n-1} \equiv \{x \in \mathbb{R}^n : \langle e_n, z \rangle = 0\}.$$

Following a similar development as in the proof of Proposition 4.4, for the particular case $R(x) = \mathbb{I}_n$ (the identity map on $\mathbb{R}^n$), we deduce

$$T_\mu^+[f](x) = \limsup_{r > 0} \frac{n}{\mathcal{L}_n(B(0,r))} \int_{B(0,r) \cap S} \left\langle \nabla f(0), \frac{u}{\|u\|} \right\rangle^2 du.$$

Consider the semispace $H = \{x \in \mathbb{R}^n : \langle e_n, v \rangle \leq 0\}$. Then we have:

$$\int_{B(0,r) \cap S} \left\langle \nabla f(0), \frac{u}{\|u\|} \right\rangle^2 du = \int_{B(0,r) \cap H} \left\langle \nabla f(0), \frac{u}{\|u\|} \right\rangle^2 du + \int_{B(0,r) \cap (S \triangle H)} \left\langle \nabla f(0), \frac{u}{\|u\|} \right\rangle^2 du.$$

In what follows we show that $\mathcal{L}_n(B(0,r) \cap (S \triangle H))$ is small, where $S \triangle H$ denotes the symmetric difference between S and H. To this end, it is easy to see that

$$B(0,r) \cap (S \triangle H) \subset B_{n-1}(0,r) \times [-d(r), d(r)],$$

where $d(r)$ stands for the maximal distance between the subspace $\mathbb{R}^{n-1} \times \{0\}$ and the elements of the following set (see Figure 4.2)

$$D(r) = \{(y,z) \in B(0,r) : \; y \in B_{n-1}(0,r) \text{ and } f(y,z) = 0\} \bigcap B(0,r).$$

Using the Implicit Function Theorem, we deduce the existence of an open subset $\mathcal{U} \subset \mathbb{R}^{n-1}$ containing 0, an open set $\mathcal{V} \subset \mathbb{R}^n$ containing 0 and a function $\varphi : \mathcal{U} \to \mathbb{R}$ of class $\mathcal{C}^1$ such that its graph coincides with $[f = 0] \cap \mathcal{V}$ and

$$\nabla \varphi(0) = -(\partial_n f(0))^{-1} \begin{pmatrix} \partial_1 f(0) \\ \vdots \\ \partial_{n-1} f(0) \end{pmatrix} = \mathbf{0}_{n-1}.$$
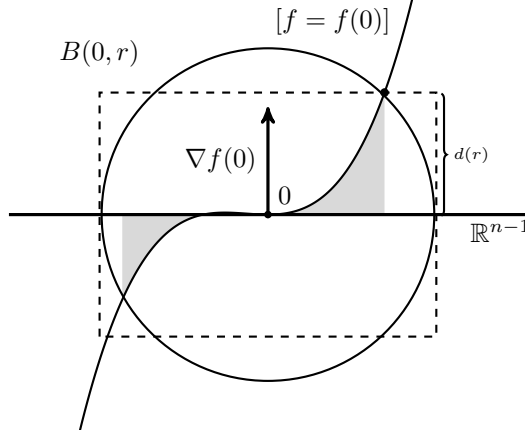
Figure 1: The gray area corresponds to the asymmetric difference $S \triangle H$. The dashed line outlines the set $B_{n-1}(0,r) \times [-d(r), d(r)]$, where the dot depicts the farthest point of the set $D(r)$ to the linear subspace $\mathbb{R}^{n-1}$.

Therefore, for $r > 0$ sufficiently small, we have $B_{n-1}(0,r) \subset \mathcal{U}$ and $B(0,r) \subset \mathcal{V}$, which yields

$$\{(y,z) \in B_{n-1}(0,r) \times \mathbb{R} : f(y,z) = 0\} = \{(y, \varphi(y)) : y \in B(0,r)\}.$$

Therefore $d(r) = \sup\{|\varphi(y)| : y \in B_{n-1}(0,r)\}$. Evoking the mean value theorem we deduce

$$\sup\{|\varphi(y)| : y \in B_{n-1}(0,r)\} \leq r \cdot \sup\{\|\nabla\varphi(y)\| : y \in B_{n-1}(0,r)\}.$$

By continuity of $\nabla\varphi$ and recalling that $\nabla\varphi(0) = 0$ we deduce that $d(r) = o(r)$. Recalling formula (7) for the volume of the $(n-1)$-dimensional ball $B_{n-1}(0,r)$, we set

$$K = \frac{2\pi^{(n-1)/2}}{\Gamma\left(\frac{n-1}{2} + 1\right)}$$

and we obtain:

$$\mathcal{L}_n(B(0,r) \cap (S \triangle H)) \leq \mathcal{L}_{n-1}(B_{n-1}(0,r)) \cdot 2d(r) = \left(K \cdot r^{n-1}\right) o(r) \equiv o(r^n)$$

Therefore,

$$\frac{n}{\mathcal{L}_n(B(0,r))} \int_{B(0,r) \cap (S \triangle H)} \left\langle \nabla f(0), \frac{u}{\|u\|} \right\rangle^2 du \leq n \|\nabla f(0)\|^2 \frac{\mathcal{L}_n(B(0,r) \cap (S \triangle H))}{\mathcal{L}_n(B(0,r))}$$

$$= \frac{n\|\nabla f(0)\|^2}{K_n} \frac{o(r^n)}{r^n} \xrightarrow{r \to 0} 0.$$

Since $B(0,r) \cap H$ is the south–half of the ball $B(0,r)$, a symmetry argument ensures

$$T_\mu^+[f](0) = \limsup_{r \to 0} \frac{n}{\mathcal{L}_n(B(0,r))} \int_{B(0,r) \cap H} \left\langle \nabla f(0), \frac{u}{\|u\|} \right\rangle^2 du$$

$$= \frac{1}{2} \limsup_{r \to 0} \frac{n}{\mathcal{L}_n(B(0,r))} \int_{B(0,r)} \left\langle \nabla f(0), \frac{u}{\|u\|} \right\rangle^2 du = \frac{1}{2}\|\nabla f(0)\|^2.$$

The proof is complete. $\qquad\square$

Remark 4.8. The above arguments can be easily adapted to show that when μ_x and $n(x)$ are as in Proposition 4.4, then the oriented dispersion operator T_μ^+ (*cf.* Definition 4.6) satisfies:

$$T_\mu^+[f](x) = \frac{1}{2}\|R(x)\nabla f(x)\|^2, \qquad \text{for all } f \in \mathcal{C}^1.$$

The following proposition justifies the introduction of the oriented dispersion in a nonsmooth setting. Given a metric space (X, d) we denote by $\mathrm{Lip}(X)$ the class of real-valued Lipschitz continuous functions on X.

Theorem 4.9. *Let (X, d) be a metric space and $\mu : X \times \mathcal{B}(X) \to \mathbb{R}_+$ a measure mapping such that for every $x \in X$, μ_x is a locally finite measure with positive measure on open sets and $\beta = \{\beta_x\}_{x \in X}$ be any family of neighborhood bases. Then, the oriented dispersion operator T_μ^+ is a descent modulus for $\mathcal{K}(X)$ and verifies that $\mathcal{K}(X) \cap \mathrm{Lip}(X) \subset \mathrm{dom}(T_\mu^+)$.*

Proof. The conditions over μ ensure that $\mathcal{K}(X) \cap \mathrm{Lip}(X) \subset \mathrm{dom}(T_\mu^+)$. Clearly the operator T_μ^+ preserves global minima and is scalar-monotone. Let us now show that it is monotone. Let $f, g \in \mathcal{K}(X)$ and let $x \in X$ such that

$$(f(x) - f(z))_+ \geq (g(x) - g(z))_+.$$

Then, $\Delta_f^+(x, z) \geq \Delta_g^+(x, z)$ for all $z \in X$ and the conclusion follows. $\square$

Remark 4.10. The measure map $\mu : X \times \mathcal{B}(X) \to \mathbb{R}_+$ is assumed to be locally finite, which yields in particular that each measure μ_x is finite on the compact sets of (X, τ). Apart from this assumption and the existence of a neighborhood system $\{\beta_x\}_{x \in X}$ where μ_x takes nonzero values, no other property is required. In this setting, the superior limit in Definition 4.2 and Definition 4.6 are well-defined, yielding that the dispersion operators are descent moduli. Even less will be required to define nonlocal operators (see next section), namely, μ_x to be finite on compact sets.

4.3 Oriented nonlocal operators

Apart from the diffusion operators, which are of local nature, one can also consider nonlocal operators. These latter serve to model jump dynamics, see *e.g.* [13]. We shall now define dispersion measures for these processes.

Definition 4.11 (Nonlocal dispersion operators)**.** Let $\mu : X \times \mathcal{B}(X) \to \mathbb{R}_+$ be a measure mapping such that for every $x \in X$, μ_x is finite on all compact sets, and let $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ be a strictly increasing function with $\phi(0) = 0$. We define the nonlocal dispersion operator induced by ϕ and μ as

$$T_{\phi,\mu}[f](x) = \int_X \phi(|f(x) - f(y)|)\mu(x, dy),$$

By construction, $T_{\phi,\mu}$ is finite for every measurable bounded function with compact support. When $X = V$ is a finite space, the nonlocal operators are particularly relevant, due to the fact that all points are isolated and so diffusion is not possible. In this setting, the measure map μ can be represented by a matrix $L : V \times V \to \mathbb{R}_+$, in the form of

$$T_{\phi,\mu}[f](x) = \sum_{y \in V} L(x, y)\, \phi(|f(x) - f(y)|)$$

Remark 4.12. In the context of Markov generators, the nonlocal operators are of the form

$$L[f](x) = \int_X (f(y) - f(x))\mu_x(dy),$$

where μ is assumed to be regular in the sense that $x \mapsto \mu_x(A)$ is measurable for every $A \in \mathcal{B}(X)$. When $\phi(t) = t^2$ we are working with the dispersion operator $T_{\phi,\mu}[f] := \Gamma[f]$, where Γ is the carré-du-champ operator associated to L, which in all generally is defined by the identity $\Gamma[f] = L[f^2] - 2fL[f]$ (as soon as $f, f^2 \in \mathrm{dom}(L)$, see e.g. [3]).

In general, a nonlocal operator $T_{\phi,\mu}$ might fail to be a descent modulus and to determine functions in the sense of Theorem 3.5.

Example 4.13. Let $\phi(t) := t^2$. Fix $N \in \mathbb{N}$ even, and set $\mathcal{V} := \mathbb{Z}_N \cup \{\bar{0}\}$, where $\bar{0} \notin \mathbb{Z}_N$ and $\mathbb{Z}_N = \{0, 1, \ldots, N-1\}$ stands for the usual cyclic additive group modulo N. We define a nonlocal operator L as follows:

$$L(x,y) = \begin{cases} 1/2, & \text{if } x \in \mathbb{Z}_N \setminus \{0\} \text{ and } y = x \pm 1, \\ 1/3, & \text{if } x = 0 \text{ and } y \in \{\bar{0}, 1, -1\}, \\ 1, & \text{if } x = \bar{0} \text{ and } y = 0, \\ 0, & \text{otherwise.} \end{cases}$$

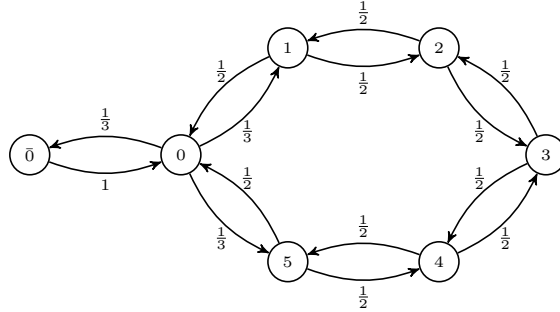Now, choose two functions $f_1, f_2 \in \mathbb{R}^{\mathcal{V}}$ satisfying that

Figure 2: Case $N = 6$.

$$f_i(\bar{0}) = f_i(0) = 0 \quad \text{and} \quad |f(x \pm 1) - f(x)| = 1, \forall x \in \mathbb{Z}_N.$$

There is at least $\binom{N}{N/2} > 1$ functions verifying the above requirements, so we can take $f_1 \neq f_2$. However, it is not hard to see that for the measure map μ associated with L, the nonlocal operator T_μ verifies that

$$T_\mu[f_i](x) = \begin{cases} 0 & \text{if } x = \bar{0}, \\ 2/3 & \text{if } x = 0, \\ 1 & \text{otherwise.} \end{cases}$$

for $i = 1, 2$. Thus, T_μ does not preserve the global minima since either $\mathrm{argmin}\, f_i \supseteq \{0, \bar{0}\}$ or $\mathrm{argmin}\, f_i \subset \mathcal{V} \setminus \{0, \bar{0}\}$. Furthermore, T_μ fails the determination theorem even for functions with $\mathcal{Z}_{T_\mu}(f) \neq \emptyset$. $\Diamond$

Example 4.13 is very illustrative as concerns the following: when $\phi(t) = t^2$, nonlocal operators do not preserve global minima in general. Indeed, if $\mathcal{V}$ is a finite state space, $T_\mu[f](x)$ measures the dispersion around point $x \in \mathcal{V}$, when $L(x, y)$ represents the probability to jump from the point x to the point y. Therefore it is natural for $T_\mu[f](x)$ to be strictly positive. However, by imposing $f(0) = f(\bar{0})$ in Example 4.13 we are forcing a point with no dispersion: starting from $x = \bar{0}$, the only possibility is to jump to 0.

Definition 4.14 (Oriented nonlocal operators)**.** Let ϕ and μ be as in Definition 4.11. We define the oriented nonlocal operator induced by ϕ and μ as

$$T_{\phi,\mu}^+[f](x) = \int_{[f \leq f(x)]} \phi(f(x) - f(y))\mu_x(dy) = \int_X \phi((f(x) - f(y))_+)\mu_x(dy) \qquad (43)$$

Similarly to the (local) oriented dispersion operator, the above operator is a descent modulus for $\mathcal{K}(X)$, and always determines a suitable subclass of continuous coercive functions $\mathcal{K}(X)$. In the nonlocal case, we do not need to assume Lipschitz continuity, and consequently, X can be a mere topological space. The class is given by the *strictly coercive* functions, which is given by

$$\mathcal{K}_s(X) = \{f : X \to \mathbb{R} \ : \ \forall x \in X, \ [f \leq f(x)] \text{ is compact}\}. \qquad (44)$$

The main difference between $\mathcal{K}_s(X)$ and $\mathcal{K}(X)$ is that the latter class admits functions attaining their maximum value since the set $[f \leq \max_X f]$ does not have to be compact. If X is compact, then the classes $\mathcal{K}(X)$ and $\mathcal{K}_s(X)$ coincide, however, if X is noncompact, then functions in $\mathcal{K}_s(X)$ cannot attain their supremum.

Theorem 4.15. *Let $\mu : X \times \mathcal{B}(X) \to \mathbb{R}_+$ be a measure mapping such that μ_x is finite on all compact sets, for every $x \in X$. Then the oriented nonlocal operator $T_{\phi,\mu}^+$ is a descent modulus for $\mathcal{K}(X)$ and verifies that $\mathcal{K}_s(X) \subset \text{dom}(T_{\phi,\mu}^+)$.*

Proof. Since μ_x is finite on all compact sets, for each $x \in X$, we deduce that $\mathcal{K}_s(X) \subset \text{dom}(T_{\phi,\mu}^+)$. Furthermore, since $\phi(0) = 0$, it is clear that $T_{\phi,\mu}^+$ preserves global minima.

Let us show now that $T_{\phi,\mu}^+$ is monotone: let $f, g \in \mathcal{K}(X)$ and $x \in X$ such that

$$(f(x) - f(z))_+ \geq (g(x) - g(z))_+, \quad \forall z \in X.$$

Then, since ϕ is non-decreasing, we have that

$$T_{\phi,\mu}^+[f](x) = \int_X \phi((f(x) - f(z))_+)\mu_x(dz) \geq \int_X \phi((g(x) - g(z))_+)\mu_x(dz) = T_{\phi,\mu}^+[g](x).$$

We conclude that $T_{\phi,\mu}^+$ is monotone.

Finally, let us show that $T_{\phi,\mu}^+$ is scalar-monotone. Let $f \in \mathcal{K}(X)$, let $r > 1$ and let $x \in X$ such that $0 < T[f](x) < +\infty$. By monotonicity, we have that $T_{\phi,\mu}^+[rf](x) \geq T_{\phi,\mu}^+[f](x) > 0$. Let us now define the sets

$$A_n = \left\{z \in X : f(x) - f(z) \geq \frac{1}{n}\right\} \bigcap \left\{z \in X \ : \ \phi(r(f(x) - f(z))) - \phi(f(x) - f(z)) \geq \frac{1}{n}\right\}.$$

Clearly $\{A_n\}_n$ is an increasing sequence of μ_x-measurable sets satisfying:

$$\bigcup_{n \geq 1} A_n = [f < f(x)].$$

Thus, by monotone convergence theorem, we have that

$$\lim_n \int_{A_n} \phi(f(x) - f(z))\mu_x(dz) = \int_{[f<f(x)]} \phi(f(x) - f(z))\mu_x(dz) = T_{\phi,\mu}^+[f](x) > 0.$$

Choose then $n \in \mathbb{N}$ such that $\int_{A_n} \phi(f(x) - f(z))\mu_x(dz) > 0$. Then, $\mu(x, A_n) > 0$ and

$$\begin{aligned}
T_{\phi,\mu}^+[rf](x) &= \int_X \phi(r(f(x) - f(z))_+)\mu_x(dz) \\
&= \int_{A_n} \phi(r(f(x) - f(z)))\mu_x(dz) + \int_{X \setminus A_n} \phi(r(f(x) - f(z))_+)\mu_x(dz) \\
&\geq \int_{A_n} \phi((f(x) - f(z))) + \frac{1}{n}\mu_x(dz) + \int_{X \setminus A_n} \phi((f(x) - f(z))_+)\mu(x, dz) \\
&= \int_X \phi((f(x) - f(z))_+)\mu(x, dz) + \frac{1}{n}\mu(A_n) > T_{\phi,\mu}^+[f](x).
\end{aligned}$$

All three properties of Definition 3.1 are satisfied and the proof is complete. $\qquad\square$

Remark 4.16. Any Γ-operator (carré-du-champ operator) coming from a regular Markov generator in $\mathbb{R}^n$ (with the euclidean distance) has the form

$$\Gamma[f](x) = \limsup_{r \to 0} \frac{n(x)}{\mu_{1,x}(B(x,r))} \int_{B(x,r)} [\Delta_f(x,y)]^2 \, \mu_{1,x}(dy) + \int_X (f(x) - f(y))^2 \mu_{2,x}(dy).$$

The above operator measures the dispersion of the function f around a point x, when the point evolves following a local diffusion process linked to $(\mu_{1,x})_x$ and a nonlocal jump process given by $(\mu_{2,x})_x$. The oriented dispersion is only taking into account the descent directions and has the form

$$\Gamma^+[f](x) = \limsup_{\varepsilon \to 0} \frac{n(x)}{\mu_{1,x}(B(x,\varepsilon))} \int_{B(x,\varepsilon)} \left[\Delta_f^+(x,y)\right]^2 \mu_{1,x}(dy) + \int_X [(f(x) - f(y))_+]^2 \mu_{2,x}(dy)$$

The above oriented operator is a descent modulus for $\mathcal{K}(X)$. Thus, if we know the oriented dispersion of a continuous coercive function f (with finite oriented dispersion), and we know its values on the critical points (that is, points with zero oriented dispersion), we completely determine the function f, in the spirit of Theorem 3.5.

5 Descent moduli over finite sets

Finite state spaces provide a simple and experimental framework to investigate further properties of moduli of descent. We shall use the terminology "*finite descent modulus*" to refer to a descent modulus over a finite set. In this section we study two particular features of finite descent moduli:

- an alternative proof, based on a probabilistic approach, of (an enhanced version of) the determination theorem for descent moduli mimicking Markov generators; and

- a characterization of *homogeneous* finite descent moduli, up to a natural equivalence relation based on the corresponding critical map.

We have already encountered a finite descent moduli in Example 4.13. Let us present a general procedure generating finite descent moduli: on a finite state space $\mathcal{V}$ (neither empty nor a singleton), consider a Markov generator $L := (L(x,y))_{x,y \in \mathcal{V}}$, namely a matrix satisfying

$$\begin{cases} \forall x, y \in \mathcal{V} : & x \neq y \implies L(x,y) \ \geq \ 0 \\ \forall x \in \mathcal{V} : & \sum_{y \in \mathcal{V}} L(x,y) \ = \ 0 \end{cases}$$

Such a generator acts linearly on any function $f \in \mathbb{R}^{\mathcal{V}}$ (which coincides with $\mathcal{K}(\mathcal{V})$) via

$$\forall\, x \in \mathcal{V}, \qquad L[f](x) = \sum_{y \in \mathcal{V}} L(x,y)(f(y) - f(x)) \tag{45}$$

By analogy to Definition 4.6 and Definition 4.14, we consider the non-linear operator T_L acting on any function $f \in \mathbb{R}^{\mathcal{V}}$ via

$$\forall\, x \in \mathcal{V}, \qquad T_L[f](x) = \sum_{y \in \mathcal{V}} L(x,y)(f(x) - f(y))_+ \tag{46}$$

From Theorem 4.15, T_L is a descent modulus. In Subsection 5.1, we will recover the determination theorem for this kind of descent modulus via a probabilistic approach.

More generally, for any $m > 0$, one can consider $T_{L,m}$ given by

$$\forall\, x \in \mathcal{V}, \qquad T_{L,m}[f](x) = \left(\sum_{y \in \mathcal{V}} L(x,y)((f(x) - f(y))_+)^m \right)^{1/m} \tag{47}$$

as well as its limit $T_{L,\infty}$ as m goes to infinity:

$$\forall\, x \in \mathcal{V}, \qquad T_{L,\infty}[f](x) = \max\{(f(x) - f(y))_+ \,:\, y \in D_x\} \tag{48}$$

where for every $x \in \mathcal{V}$ we set:

$$D_x := \{x\} \sqcup \{y \in \mathcal{V} \,:\, L(x,y) > 0\} \tag{49}$$

Let us recall (see Definition 3.12 for $p = 1$) that a descent modulus T is said to be *homogeneous* (or 1–*homogeneous*) if for all $r \geq 0$ and $f \in \mathbb{R}^{\mathcal{V}}$ we have: $T[rf] = rT[f]$.

All the above operators $T_{L,m}$, $m \in (0, +\infty]$, are homogeneous descent moduli. It should be noticed that there are many more homogeneous descent moduli: for instance in (47) we can allow the exponent m to depend on $x \in \mathcal{V}$. Moreover, given n homogeneous descent moduli $T_1, ..., T_n$, and positive numbers $a_1, ..., a_n > 0$, the weighted sum $a_1 T_1 + \cdots + a_n T_n$ is again a homogeneous descent moduli. Even fancier constructions are possible. This being said, there exist non-homogeneous descent moduli. Indeed, for any non-decreasing mapping $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ with $\phi(0) = 0$, the descent modulus T_ϕ defined by

$$\forall\, x \in \mathcal{V}, \qquad T_L[f](x) = \sum_{y \in \mathcal{V}} L(x,y)\phi((f(x) - f(y))_+) $$

is homogeneous if and only if ϕ is linear, as long as $L \neq 0$.

Given a descent modulus T we recall from (21) the *critical map* $\mathcal{Z}_T$, which associates to every function $f \in \mathbb{R}^{\mathcal{V}}$ its set of critical points $\mathcal{Z}_T(f) = (T[f])^{-1}(0)$. Notice that the critical maps $\mathcal{Z}_{T_{L,m}}$ related to the moduli $T_{L,m}$ in (47)–(48) are all the same as m varies in $(0, +\infty]$.

In Subsection 5.2 we introduce an equivalence relation among homogeneous descent moduli, using the critical maps. Under this relation, all moduli $T_{L,m}$ in (47) turn out to be equivalent to each other (for different values of $m \in \mathbb{N}$) and also equivalent to $T_{L,\infty}$. The main result of this section is to show that every homogeneous descent modulus on a general finite set $\mathcal{V}$ (without generator L) is still of the form (48) for some family $\mathcal{D} = \{\mathcal{D}_x\}$ which is naturally associated to T, provided it satisfies a (necessary and sufficient) mild condition.

5.1 A probabilistic approach

Let $L := (L(x,y))_{x,y \in \mathcal{V}}$ be a Markov generator on the finite set $\mathcal{V}$.

For every $f \in \mathbb{R}^{\mathcal{V}}$ the associated f-oriented Markov generator $L^f := (L^f(x,y))_{x,y \in \mathcal{V}}$ is defined for $x, y \in \mathcal{V}$ with $x \neq y$ as follows:

$$
L^f(x,y) := \begin{cases} L(x,y), & \text{if } f(y) \leq f(x) \\ 0, & \text{otherwise.} \end{cases}
$$

The values $L^f(x,x)$ on the diagonal are determined by the fact that the sum of the rows $\sum_{x \in \mathcal{V}} L(x,y)$ should vanish.

Let $T : \mathbb{R}^{\mathcal{V}} \to \mathbb{R}^{\mathcal{V}}$ be defined for every $f \in \mathbb{R}^{\mathcal{V}}$ and $x \in \mathcal{V}$ as follows:

$$
T[f](x) := -L^f[f](x) = -\sum_{y \in \mathcal{V}} L^f(x,y)(f(y) - f(x)) = \sum_{y \in \mathcal{V}} L(x,y)(f(x) - f(y))_+.
$$

This non-linear operator T coincides with T_L defined in (46) and is a descent modulus. For every $f \in \mathbb{R}^{\mathcal{V}}$ the set of T-critical points is given by the formula

$$
\mathcal{Z}_T(f) := \{x \in \mathcal{V} : T[f](x) = 0\}. \tag{50}
$$

Given $x, y \in \mathcal{V}$, an L-path from x to y is a finite sequence $\{x_k\}_{0 \leq k \leq N}$ with $N \geq 0$, $x_0 = x$, $x_N = y$ and such that for all $0 \leq k < N$, $L(x_k, x_{k+1}) > 0$. This path is called an L^f-path from x to y if in addition $\{f(x_k)\}_{0 \leq k \leq N}$ is a non-increasing finite sequence. We write $x \xrightarrow{f} y$ to indicate that there exists a L^f-path from x to y. We set:

$$
x \succeq_f y \iff x \xrightarrow{f} y \qquad \text{and} \qquad x \approx_f y \iff \begin{cases} x \xrightarrow{f} y \\ y \xrightarrow{f} x \end{cases}
$$

It is straighforward to check that $\succeq_f$ is an order relation on $\mathcal{V}$ and $\approx_f$ is its corresponding equivalence relation ($x \approx_f y$ if and only if $x \succeq_f y$ and $y \succeq_f x$). The set of minima of $\succeq_f$ is defined as follows:

$$
M(f) := \{\bar{x} \in \mathcal{V} : \ \forall x \in \mathcal{V}, \ (\bar{x} \succeq_f x \Rightarrow \bar{x} \approx_f x)\}. \tag{51}
$$

Notice that $\bar{x} \in M(f)$ if and only if for any $y \in \mathcal{V}$ with $f(y) < f(\bar{x})$ and any L-path $\{x_k\}_{0 \leq k \leq N}$ from x to y, we have $\max_{0 \leq k \leq N} f(x_k) > f(\bar{x})$. Moreover, we always have $M(f) \subset \mathcal{Z}_T(f)$ and the inclusion may be strict.

Example 5.1. Let $\mathcal{V} = \mathbb{Z}_9$, and set L such that

$$L(x, y) > 0 \iff y = x \pm 1.$$

Consider $f = (1, 0, 0, 1, 2, 1, 1, 2, 1)$. The set $\mathcal{V}$, its connections through L and the level sets of f are depicted in Figure 3.
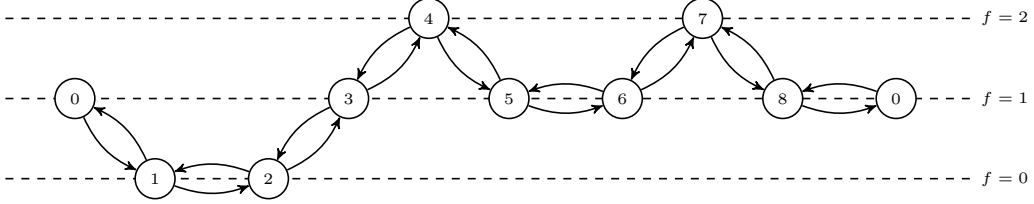

Figure 3: Node 0 has been replicated at the beginning and end of the representation. Only connections (x, y) with $L(x, y) > 0$ have been drawn.

Here, $M(f) = \{1, 2, 5, 6\}$ and $\mathcal{Z}_T(f) = \{1, 2, 5, 6, 8\}$. The node 8 is critical since L does not allow to jump to any node with smaller value in one step. However, the path $8 \to 0 \to 1$ is an L^f-path leading to a point with smaller f-value. Note that 5 and 6 are in $M(f)$ since there is no L^f-path emanating from any of them and landing at a different node with smaller f-value.$\diamond$

For $x \in \mathcal{V}$, let $X_x^f := (X_x^f(t))_{t \geq 0}$ stand for a Markov process starting from x and whose generator is L^f. For such a process, the function

$$\mathbb{R}_+ \ni t \mapsto f(X_x^f(t))$$

is almost surely non-increasing and bounded, thus converging. Furthermore, the finite Markov process $X_x^f(t)$ is converging in law for large $t \geq 0$ toward a distribution which may depend on the initial point x and whose support is included into the set $M(f)$.

Fix $f, g \in \mathbb{R}^{\mathcal{V}}$. Since $\mathcal{V}$ is finite, the functions f, g are trivially continuous and coercive. Therefore, Theorem 3.5 directly yields:

$$\left. \begin{array}{r} T[f] = T[g] \\ f = g \text{ on } \mathcal{Z}_T(f) \end{array} \right\} \implies f = g. \tag{52}$$

In what follows, we obtain (52) via a probabilistic approach, in a slightly enhanced version, namely replacing the set $\mathcal{Z}_T(f) = \mathcal{Z}_T(g)$ (where f and g are assumed to be equal) by the (potentially smaller) set $M(f) \cup M(g)$. The technical ingredient of the proof is contained in the following lemma.

Lemma 5.2. *For any $f, g \in \mathbb{R}^{\mathcal{V}}$ with $T[f] \geq T[g]$, we have $L^f[g] \geq L^f[f]$.*

Proof. Indeed, for any $x \in \mathcal{V}$, we have

$$
\begin{aligned}
-L^f[g](x) &= \sum_{y\,:\,f(y)\leq f(x)} L(x,y)(g(x)-g(y)) = \\
&= \sum_{y\,:\,f(y)\leq f(x),\,g(y)\leq g(x)} L(x,y)(g(x)-g(y)) + \sum_{y\,:\,f(y)\leq f(x),\,g(y)>g(x)} L(x,y)(g(x)-g(y)) \\
&\leq \sum_{y\,:\,f(y)\leq f(x),\,g(y)\leq g(x)} L(x,y)(g(x)-g(y)) \leq \sum_{y\,:\,g(y)\leq g(x)} L(x,y)(g(x)-g(y)) \\
&= T[g](x) \leq T[f](x) = -L^f[f](x).
\end{aligned}
$$

$\square$

We are now ready to give a probabilistic proof of the following comparison result. (Recall from (51) the definition of $M(f)$.)

Proposition 5.3. *Let $f, g \in \mathbb{R}^{\mathcal{V}}$ be two functions satisfying:*

(i). $T[f](x) \geq T[g](x)$, *for all $x \in \mathcal{V}$; and*

(ii). $f(x) \geq g(x)$, *for all $x \in M(f)$.*

Then $f \geq g$.

Proof. Due to the martingale problem characterization of X_x^f (see [13] e.g.), there exists a martingale $\{M_g^f(t)\}_{t\geq 0}$ starting from 0 such that

$$
g(X_x^f(t)) = g(x) + \int_0^t L^f[g](X_x^f(s))\,ds + M_g^f(t), \qquad \text{for all } t \geq 0.
$$

Taking expectations we get

$$
\mathbb{E}[g(X_x^f(t))] = g(x) + \int_0^t \mathbb{E}[L^f[g](X_x^f(s))]\,ds, \qquad \text{for all } t \geq 0. \tag{53}
$$

Denote by π^f the limit law of the distributions of $X_x^f(t)$ for large $t \geq 0$. Then π^f is supported on $M(f)$ and

$$
\lim_{t\to+\infty} \mathbb{E}[g(X_x^f(t))] = \pi^f[g].
$$

In particular the integral of the right hand side of (53) converges for large $t \geq 0$ and it holds:

$$
\pi^f[g] = g(x) + \int_0^{+\infty} \mathbb{E}[L^f[g](X_x^f(s))]\,ds. \tag{54}
$$

Applying the above arguments with g replaced by f, we also obtain

$$
\pi^f[f] = f(x) + \int_0^{+\infty} \mathbb{E}[L^f[f](X_x^f(s))]\,ds. \tag{55}
$$

The assumption (ii) yields $\pi^f[f] \geq \pi^f[g]$. On the other hand, from Lemma 5.2, we have

$$
\mathbb{E}[L^f[g](X_x^f(s))] \geq \mathbb{E}[L^f[f](X_x^f(s))], \qquad \text{for all } s \geq 0.
$$

Combining the above with (54) we deduce

$$\pi^f[f] \geq \pi^f[g] = g(x) + \int_0^{+\infty} \mathbb{E}[L^f[g](X_x^f(s))] \geq g(x) + \int_0^{+\infty} \mathbb{E}[L^f[f](X_x^f(s))] \, ds.$$

Comparing the above inequality with (55) yields $f(x) \geq g(x)$ and the result follows. $\qquad\square$

By symmetry we obtain the following corollary:

Corollary 5.4. *Let $f, g \in \mathbb{R}^{\mathcal{V}}$ be such that*
(i). $T[f](x) = T[g](x)$, for all $x \in \mathcal{V}$; and
(ii). $f(x) = g(x)$, for all $x \in M(f) \cup M(g)$.
Then $f = g$.

5.2 Classification of descent moduli on $\mathbb{R}^{\mathcal{V}}$

Denote $\mathcal{P}(\mathcal{V})^*$ the family of nonempty subsets of $\mathcal{V}$. Given a descent modulus T on $\mathcal{V}$ we recall from (50) the critical map

$$\mathcal{Z}_T \, : \, \mathbb{R}^{\mathcal{V}} \to \mathcal{P}(\mathcal{V})^*.$$

Definition 5.5 (equivalence of descent moduli). Let T, S be two descent moduli on $\mathcal{V}$. We say that the moduli T and S are *equivalent* (and denote $T \sim S$) if $\mathcal{Z}_T = \mathcal{Z}_S$.

Notice that if $T \sim S$, then T and S determine the same functions via Theorem 3.5.

A family $\mathcal{D} := \{\mathcal{D}_x\}_{x \in \mathcal{V}}$ is called an *active neighborhood system*, provided $x \in \mathcal{D}_x \subset \mathcal{V}$ for every $x \in \mathcal{V}$. An example of such system has been defined in (49) in the particular case where the set $\mathcal{V}$ is equipped with a generator L.

We henceforth denote by $\mathcal{E}(\mathcal{V})$ the set of active neihborhood systems on $\mathcal{V}$. Then for any such system $\mathcal{D} \in \mathcal{E}(\mathcal{V})$ we associate a descent modulus $T_{\mathcal{D}}$ defined for $f \in \mathbb{R}^{\mathcal{V}}$ and $x \in \mathcal{V}$ as follows (compare with (25) in Proposition 3.7):

$$T_{\mathcal{D}}[f](x) := \max_{y \in \mathcal{D}_x} (f(x) - f(y))_+ \tag{56}$$

Conversely, given any (abstract) descent modulus T we set:

$$\begin{cases} \mathcal{K}_x(T) & := & \{K \subset \mathcal{V} \, : \, x \in K \cap \mathcal{Z}_T(\mathbb{1}_K)\} \\ D_x(T) & := & \bigcap_{K \in \mathcal{K}_x(T)} K \end{cases} \tag{57}$$

where $\mathbb{1}_K$ denotes the characteristic function of the set K, that is:

$$\mathbb{1}_K(x) = \begin{cases} 1, & \text{if } x \in K \\ 0, & \text{if } x \notin K. \end{cases}$$

The interest of these notions is illustrated by the following result:

Theorem 5.6 (classification of moduli). *If a homogeneous descent modulus T satisfies*

$$\forall \, x \in \mathcal{V} : \quad D_x(T) \in \mathcal{K}_x(T) \tag{$\mathcal{H}$}$$

then there exists a family $\mathcal{D} \in \mathcal{E}(\mathcal{V})$ such that T is equivalent to $T_{\mathcal{D}}$ given in (56).

35

Before we proceed, let us introduce the following definition.

Definition 5.7. For a critical map $\mathcal{Z} : \mathbb{R}^{\mathcal{V}} \to \mathcal{P}(\mathcal{V})^*$ and for each $x \in \mathcal{V}$, we define

$$\begin{cases} \mathcal{K}_x(\mathcal{Z}) & := & \{K \subset \mathcal{V} : x \in K \cap \mathcal{Z}(\mathbb{1}_K)\} \\ \mathcal{D}_x(\mathcal{Z}) & := & \bigcap_{K \in \mathcal{K}_x} K \end{cases} \tag{58}$$

If there is no confusion, we might simply write $\mathcal{K}_x$ and $\mathcal{D}_x$, respectively.

The proof of Theorem 5.6 is based on a characterization of those maps $\mathcal{Z} : \mathbb{R}^{\mathcal{V}} \to \mathcal{P}(\mathcal{V})^*$ for which there exists a descent modulus T such that $\mathcal{Z} = \mathcal{Z}_T$. More precisely, let $\mathcal{Z} : \mathbb{R}^{\mathcal{V}} \to \mathcal{P}(\mathcal{V})^*$ be an *abstract critical map* satisfying the following conditions :

(Z1) for every $f \in \mathbb{R}^{\mathcal{V}}$ and $r \in \mathbb{R}$: $\mathcal{Z}[f + r] = \mathcal{Z}[f]$

(Z2) for every $f \in \mathbb{R}^{\mathcal{V}}$ and $r > 0$: $\mathcal{Z}[rf] = \mathcal{Z}[f]$

(Z3) for every $f \in \mathbb{R}^{\mathcal{V}}$ and $r \in \mathbb{R}$: $\quad \mathcal{Z}[f] = (\mathcal{Z}[\phi_r(f)] \cap [f \leq r]) \sqcup (\mathcal{Z}[\varphi_r(f)] \cap [f > r])$

where

$$\forall\, r \in \mathbb{R}, \forall\, s \in \mathbb{R}, \qquad \begin{cases} \phi_r(s) & := & r \wedge s \\ \varphi_r(s) & := & r \vee s. \end{cases}$$

We also assume

(Z4) for every $K \subset \mathcal{V}$ we have: $K^{\mathrm{c}} \subset \mathcal{Z}(\mathbb{1}_K)$; and

(Z5) for every $x \in \mathcal{V}$:

$$\mathcal{K}_x = \{K \subset \mathcal{V} : \mathcal{D}_x \subset K\}.$$

The announced characterization of critical maps is the following:

Theorem 5.8 (characterization of critical maps)**.** *An abstract critical map $\mathcal{Z} : \mathbb{R}^{\mathcal{V}} \to \mathcal{P}(\mathcal{V})^*$ is associated to some homogeneous descent modulus T (that is, $\mathcal{Z} = \mathcal{Z}_T$) if and only if conditions (Z1)–(Z5) hold.*
In this case $\mathcal{Z} = \mathcal{Z}_{T_{\mathcal{D}}}$, where $T_{\mathcal{D}}$ is defined by (56) for $\mathcal{D} := \{\mathcal{D}_x\}_{x \in \mathcal{V}}$ constructed in (58).

The last assertion of Theorem 5.8 is implicitly assuming that $\mathcal{D} \in \mathcal{E}(\mathcal{V})$. The following lemma confirms that this is indeed the case:

Lemma 5.9. *Let $\mathcal{Z} : \mathbb{R}^{\mathcal{V}} \to \mathcal{P}(\mathcal{V})^*$ be an abstract critical mapping that satisfies conditions (Z1)–(Z4). Let further $\mathcal{D} := \{\mathcal{D}_x\}_{x \in \mathcal{V}}$ be constructed as in (58). Then $\mathcal{Z}[\mathbb{1}_{\mathcal{V}}] = \mathcal{V}$ and $\mathcal{D} \in \mathcal{E}(\mathcal{V})$.*

Proof. Applying (Z4) with $K = \emptyset$, we get $\mathcal{V} \subset \mathcal{Z}(\mathbb{1}_{\emptyset}) = \mathcal{Z}[\mathbf{0}]$, where $\mathbf{0}$ denotes the null function on $\mathcal{V}$. We deduce from (Z1) (for $r = 1$) that $\mathcal{Z}[\mathbb{1}_{\mathcal{V}}] = \mathcal{Z}[\mathbf{0}] = \mathcal{V}$. Recall that the family $\mathcal{D} := (\mathcal{D}_x)_{x \in \mathcal{V}}$ belongs to $\mathcal{E}(\mathcal{V})$ if and only if it satisfies

$$x \in \mathcal{D}_x, \ \forall x \in \mathcal{V}.$$

Fix $x \in \mathcal{V}$. Since $\mathcal{Z}[\mathbb{1}_{\mathcal{V}}] = \mathcal{V}$, we have $x \in \mathcal{Z}[\mathbb{1}_{\mathcal{V}}]$, which in conjunction with $x \in \mathcal{V}$ yields $\mathcal{V} \in \mathcal{K}_x$. It follows that $\mathcal{K}_x \neq \emptyset$. By definition of $\mathcal{K}_x$, for any $K \in \mathcal{K}_x$, we have $x \in K$, so that $x \in \bigcap_{K \in \mathcal{K}_x} K = \mathcal{D}_x$. Therefore $\mathcal{D} \in \mathcal{E}(\mathcal{V})$. $\qquad\square$

Let us postpone for a while the proof of Theorem 5.8 and show instead that Theorem 5.8 implies Theorem 5.6. To this end, let T be a homogeneous descent modulus on $\mathcal{V}$. Then using Theorem 5.8, it is easy to see that T is equivalent to $T_\mathcal{D}$ for some active neighborhood system $\mathcal{D} \in \mathcal{E}(\mathcal{V})$ provided the following proposition is proven:

Proposition 5.10. *Under Assumption ($\mathcal{H}$), the critical map $\mathcal{Z}_T$ satisfies (Z1)–(Z5).*

Proof. We verify successively that conditions (Z1)–(Z5) hold. Indeed, condition (Z1) comes from the translation invariance property of T, see Proposition 3.2, while (Z2) is consequence of the homogeneity assumption for T.

Verifying (Z3) requires some extra work: fix $f \in \mathbb{R}^\mathcal{V}$ and $r \in \mathbb{R}$. Since (Z3) holds trivially for constant functions, we may assume that f takes at least two different values. Then for any $s, s' \in \mathbb{R}$ we have

$$
\begin{aligned}
(\phi_r(s) - \phi_r(s'))_+ &\leq (s - s')_+ \\
(\varphi_r(s) - \varphi_r(s'))_+ &\leq (s - s')_+
\end{aligned}
$$

and monotonicity yields

$$
\begin{aligned}
T[\phi_r(f)] &\leq T[f] \\
T[\varphi_r(f)] &\leq T[f]
\end{aligned}
$$

Consequently:

$$
\mathcal{Z}_T[f] \subset \mathcal{Z}_T[\phi_r(f)] \cap \mathcal{Z}_T[\varphi_r(f)]
$$

so that

$$
\begin{aligned}
\mathcal{Z}_T[f] &= (\mathcal{Z}_T[f] \cap [f \leq r]) \sqcup (\mathcal{Z}_T[f] \cap [f > r]) \\
&\subset (\mathcal{Z}_T[\phi_r(f)] \cap [f \leq r]) \sqcup (\mathcal{Z}_T[\varphi_r(f)] \cap [f > r]).
\end{aligned}
$$

To get the reserved inclusion, consider $x \in \mathcal{V}$ with $f(x) \leq r$, in particular $\phi_r(f(x)) = f(x)$. For any $z \in \mathcal{V}$ with $\phi_r(f(z)) < \phi_r(f(x))$, we have $\phi_r(f(z)) = f(z)$, so that

$$
(\phi_r(f(x)) - \phi_r(f(z)))_+ = (f(x) - f(z))_+
$$

For any $z \in \mathcal{V}$ with $\phi_r(f(z)) \geq \phi_r(f(x))$, we must have $f(z) \geq f(x)$, thus

$$
(\phi_r(f(x)) - \phi_r(f(z)))_+ = 0 = (f(x) - f(z))_+
$$

From the monotonicity property, we deduce $T[\phi_r(f)](x) = T[f](x)$. These considerations show that

$$
\mathcal{Z}_T[\phi_r(f)] \cap [f \leq r] \subset \mathcal{Z}_T[f] \tag{59}
$$

Finally, consider $x \in \mathcal{V}$ with $f(x) > r$, in particular $\varphi_r(f(x)) = f(x)$. Since f is not constant, we can define

$$
a := \frac{f(x) - (r \vee \min f)}{\max f - \min f} > 0.
$$

Now, on the one hand, for any $z \in \mathcal{V}$ with $\varphi_r(f(z)) \geq \varphi_r(f(x))$, we must have $f(z) \geq f(x)$, thus

$$
(\varphi_r(f(x)) - \varphi_r(f(z)))_+ = 0 = (f(x) - f(z))_+ = (af(x) - af(z))_+
$$

On the other hand, for any $z \in \mathcal{V}$ with $\varphi_r(f(z)) < \varphi_r(f(x))$, we have

$$\varphi_r(f(x)) - \varphi_r(f(z)) \geq f(x) - (r \vee \min f) \geq a(f(x) - f(z))$$

We deduce that

$$\forall \, z \in \mathcal{V}, \qquad (\varphi_r(f(x)) - \varphi_r(f(z)))_+ \geq (af(x) - af(z))_+$$

and by monotonicity $T[\varphi_r(f)](x) \geq T[af](x) = aT[f](x)$, by homogeneity. It follows that

$$\mathcal{Z}_T[\varphi_r(f)] \cap [f > r] \subset \mathcal{Z}_T[f].$$

Combining with (59), we get the reverse inclusion

$$\mathcal{Z}_T[f] \supset (\mathcal{Z}_T[\phi_r(f)] \cap [f \leq r]) \sqcup (\mathcal{Z}_T[\varphi_r(f)] \cap [f > r]),$$

therefore (Z3) holds.

Condition (Z4) is a consequence of the preservation of global minima, since the set of global minima of $\mathbb{1}_K$ coincides with K^{c}.

It remains to show (Z5). Set

$$\widetilde{\mathcal{K}_x} = \{K \subset \mathcal{V} : \mathcal{D}_x \subset K\}.$$

Then for every $K \in \mathcal{K}_x$ we have $\mathcal{D}_x \subset K$, that is, $K \in \widetilde{\mathcal{K}_x}$ and $\mathcal{K}_x \subset \widetilde{\mathcal{K}_x}$. To prove the reverse inclusion, consider $K \subset \mathcal{V}$ with $\mathcal{D}_x \subset K$. We need to verify that $K \in \mathcal{K}_x$. Since $x \in \mathcal{D}_x$, we get $x \in K$. Furthermore, for every $z \in \mathcal{V}$ we have:

$$\left(\mathbb{1}_K(x) - \mathbb{1}_K(z)\right)_+ = 1 - \mathbb{1}_K(z) \leq 1 - \mathbb{1}_{\mathcal{D}_x}(z) = \left(\mathbb{1}_{\mathcal{D}_x}(x) - \mathbb{1}_{\mathcal{D}_x}(z)\right)_+$$

and thus by monotonicity, $T[\mathbb{1}_K](x) \leq T[\mathbb{1}_{\mathcal{D}_x}](x) = 0$, where the last equality is obtained via $(\mathcal{H})$. It follows that $x \in \mathcal{Z}_T(K)$ whence $K \in \mathcal{K}_x$.

The proof is complete. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

For $\mathcal{D} \in \mathcal{E}(\mathcal{V})$, denote for simplicity by $\mathcal{Z}_\mathcal{D}$ the critical map $\mathcal{Z}_{T_\mathcal{D}}$ associated to the homogeneous descent modulus $T_\mathcal{D}$. The following lemma shows how to recover the active neighborhood system $\mathcal{D} = \{\mathcal{D}_x\}_x$ from $\mathcal{Z}_\mathcal{D}$:

Lemma 5.11. *For any $x \in \mathcal{V}$, we have*

$$\mathcal{D}_x = \bigcap_{K \in \mathcal{K}_x} K$$

where

$$\mathcal{K}_x := \{K \subset \mathcal{V} : x \in K \cap \mathcal{Z}_\mathcal{D}(\mathbb{1}_K)\}.$$

Proof. For any $f \in \mathbb{R}^\mathcal{V}$, recall that

$$\begin{aligned} \mathcal{Z}_\mathcal{D}[f] &= \{x \in \mathcal{V} : T_\mathcal{D}[f](x) = 0\} = \{x \in \mathcal{V} : \max_{y \in \mathcal{D}_x}(f(x) - f(y))_+ = 0\} \\ &= \{x \in \mathcal{V} : \forall \, y \in \mathcal{D}_x, \, f(y) \geq f(x)\}. \end{aligned}$$

38

In particular taking $f = \mathbb{1}_K$ with $K \in \mathcal{P}(\mathcal{V})^*$, we get

$$\mathcal{Z}_\mathcal{D}(\mathbb{1}_K) = \{x \in \mathcal{V} : \forall\, y \in \mathcal{D}_x,\ \mathbb{1}_K(y) \geq \mathbb{1}_K(x)\} = \{x \in K : \mathcal{D}_x \subset K\} \cup K^c.$$

Fix $x \in \mathcal{V}$ and consider $K \in \mathcal{K}_x$. Since $x \in K$ and $x \in \mathcal{Z}_\mathcal{D}(\mathbb{1}_K)$, we deduce that $\mathcal{D}_x \subset K$. It follows that

$$\mathcal{D}_x \quad \subset \quad \bigcap_{K \in \mathcal{K}_x} K.$$

To get the reverse implication, it is sufficient to check that $\mathcal{D}_x \in \mathcal{K}_x$. Note that

$$\mathcal{Z}_\mathcal{D}[\mathbb{1}_{\mathcal{D}_x}] = \{y \in \mathcal{D}_x : \mathcal{D}_y \subset \mathcal{D}_x\} \cup \mathcal{D}_x^c.$$

Therefore, $x \in \mathcal{Z}_\mathcal{D}[\mathbb{1}_{\mathcal{D}_x}]$. Since we also have $x \in \mathcal{D}_x$, we deduce that $\mathcal{D}_x \in \mathcal{K}_x$. $\qquad\square$

The above lemma justifies the introduction of the objects $\mathcal{K}_x$, $\mathcal{D}_x$, for $x \in \mathcal{V}$, for any mapping $\mathcal{Z} : \mathbb{R}^\mathcal{V} \to \mathcal{P}(\mathcal{V})^*$ in (58) by analogy to (57). Denote by $\hat{\mathcal{Z}}$ the set of mappings $\mathcal{Z} : \mathbb{R}^\mathcal{V} \to \mathcal{P}(\mathcal{V})^*$ satisfying $\mathcal{Z}[\mathbb{1}] = \mathcal{V}$ and by $\hat{\mathcal{Z}}_\mathcal{E}$ the set of critical maps $\mathcal{Z}_\mathcal{D}$ associated to $T_\mathcal{D}$ with $\mathcal{D} \in \mathcal{E}(\mathcal{V})$. Let $\mathcal{Q}$ be the mapping $\hat{\mathcal{Z}} \ni \mathcal{Z} \mapsto \mathcal{Z}_\mathcal{D} \in \hat{\mathcal{Z}}_\mathcal{E}$ considered above. Lemma 5.11 shows that $\mathcal{Q}^2 = \mathcal{Q}$, that is, $\mathcal{Q}$ is a kind of non-linear projection.

Let us show that in Proposition 5.10 we don't need to assume $(\mathcal{H})$ if the critical map is of the form $\mathcal{Z}_\mathcal{D}$:

Lemma 5.12. *For any $\mathcal{D} \in \mathcal{E}(\mathcal{V})$, $\mathcal{Z}_\mathcal{D}$ satisfies $(\mathcal{H})$ and thus $(Z5)$.*

Proof. Thanks to Lemma 5.11, the family $\{\bigcap_{K \in \mathcal{K}_x} K\}_{x \in \mathcal{V}}$ constructed in (58) coincides with the active neighborhood system $\mathcal{D} = \{\mathcal{D}_x\}_{x \in \mathcal{V}}$ in the definition of $\mathcal{Z}_\mathcal{D}$. Thus to check $(\mathcal{H})$, it suffices to show that $x \in \mathcal{Z}_\mathcal{D}[\mathbb{1}_{\mathcal{D}_x}]$, for any $x \in \mathcal{V}$, or equivalently $T_\mathcal{D}[\mathbb{1}_{\mathcal{D}_x}](x) = 0$. A direct computation gives:

$$T_\mathcal{D}[\mathbb{1}_{\mathcal{D}_x}](x) = \max_{z \in \mathcal{D}_x} \mathbb{1}_{\mathcal{D}_x}(x) - \mathbb{1}_{\mathcal{D}_x}(z) = 0$$

Condition $(Z5)$ then follows from Proposition 5.10. $\qquad\square$

Here is the first step towards Theorem 5.8:

Proposition 5.13. *Let $\mathcal{Z} : \mathbb{R}^\mathcal{V} \to \mathcal{P}(\mathcal{V})^*$ satisfying $(Z1)$–$(Z3)$ and $\mathcal{Z}[\mathbb{1}_\mathcal{V}] = \mathcal{V}$. Let $\mathcal{D}$ be constructed as in (58) and define $\mathcal{Z}_\mathcal{D}$ the critical map associated to $T_\mathcal{D}$ given in (56). Assume that*

$$\forall\, K \subset \mathcal{V}, \qquad \mathcal{Z}(\mathbb{1}_K) = \mathcal{Z}_\mathcal{D}(\mathbb{1}_K)$$

Then we have $\mathcal{Z} = \mathcal{Z}_\mathcal{D}$.

Proof. From Proposition 5.10 and Lemma 5.9, $\mathcal{Z}_\mathcal{D}$ also satisfies $(Z1)$–$(Z3)$ and $\mathcal{Z}_\mathcal{D}[\mathbb{1}_\mathcal{V}] = \mathcal{V}$. Let $f \in \mathbb{R}^\mathcal{V}$. We prove that $\mathcal{Z}[f] = \mathcal{Z}_\mathcal{D}[f]$ via induction over the number $n \in \mathbb{N}$ of values taken by f.
● We begin with the case where $n = 1$, that is, f is constant. Denote by $a \in \mathbb{R}$ the value of f. Taking into account condition $(Z1)$ and the fact that $\mathcal{Z}[\mathbb{1}_\mathcal{V}] = \mathcal{V}$, we obtain

$$\mathcal{Z}[f] = \mathcal{Z}[f - a + 1] = \mathcal{Z}[\mathbb{1}_\mathcal{V}] = \mathcal{Z}_\mathcal{D}[\mathbb{1}_\mathcal{V}] = \mathcal{V} = \mathcal{Z}_\mathcal{D}[f]$$

39

• Consider the case where $n = 2$ and let $f(\mathcal{V}) = \{a, b\}$ with $a < b$. Set $K := [f = b]$. Using (Z1) and (Z2), we get

$$\mathcal{Z}[f] = \mathcal{Z}\left[\frac{f - a}{b - a}\right] = \mathcal{Z}(\mathbb{1}_K) = \mathcal{Z}_\mathcal{D}(\mathbb{1}_K) = \mathcal{Z}_\mathcal{D}[f].$$

• Consider the case where $n > 2$, assuming that $\mathcal{Z}[g] = \mathcal{Z}_\mathcal{D}[g]$ for all $g \in \mathbb{R}^\mathcal{D}$ taking at most $n - 1$ values. Write $f_1 < f_2 < \cdots < f_n$ the values taken by f. Take $k = \lfloor\frac{n+1}{2}\rfloor$ (integer part), set $r = f_k$ and

$$
\begin{aligned}
g_- &:= \phi_r \circ f \\
g_+ &:= \varphi_r \circ f
\end{aligned}
$$

By the choice of r, both g_- and g_+ take at most $n - 1$ values.
Condition (Z3) then yields

$$
\begin{aligned}
\mathcal{Z}[f] &= (\mathcal{Z}[g_-] \cap [f \le r]) \sqcup (\mathcal{Z}[g_+] \cap [f > r]) \\
&= (\mathcal{Z}_\mathcal{D}[g_-] \cap [f \le r]) \sqcup (\mathcal{Z}_\mathcal{D}[g_+] \cap [f > r]) = \mathcal{Z}_\mathcal{D}[f].
\end{aligned}
$$

as desired. $\qquad\square$

Having established Proposition 5.13, the following result finishes the proof of Theorem 5.8:

Proposition 5.14. *Let $\mathcal{Z} : \mathbb{R}^\mathcal{V} \to \mathcal{P}(\mathcal{V})^*$ be a mapping satisfying (Z1)–(Z5) and let $\mathcal{D} = (\mathcal{D}_x)_{x \in \mathcal{V}} \in \mathcal{E}(\mathcal{V})$ be as in (58). Then we have*

$$\forall\, K \subset \mathcal{V}: \qquad \mathcal{Z}[\mathbb{1}_K] = \mathcal{Z}_\mathcal{D}[\mathbb{1}_K]$$

Proof. From (Z4), we have

$$K^\mathrm{c} \cap \mathcal{Z}(\mathbb{1}_K) = K^\mathrm{c} = K^\mathrm{c} \cap \mathcal{Z}_\mathcal{D}(\mathbb{1}_K).$$

Now, let $K' = \{x \in K \ : \ \mathcal{D}_x \subset K\}$. For every $x \in K'$, due to (Z5), we have $K \in \mathcal{K}_x$, so $x \in \mathcal{Z}[\mathbb{1}_K]$. Since this is true for any $x \in K'$, we get $K' \subset \mathcal{Z}[\mathbb{1}_K]$.
According to Lemma 5.12, $\mathcal{Z}_\mathcal{D}$ also satisfies (Z5), and it follows as above that $K' \subset \mathcal{Z}_\mathcal{D}[\mathbb{1}_K]$. We deduce

$$K' \cap \mathcal{Z}(\mathbb{1}_K) = K' = K' \cap \mathcal{Z}_\mathcal{D}(\mathbb{1}_K).$$

Finally, let $K \setminus K' = \{x \in K \ : \ \mathcal{D}_x \setminus K \ne \emptyset\}$. For every $x \in K \setminus K'$, since $\mathcal{D}_x \not\subset K$, we have $K \notin \mathcal{K}_x$, due to the definition of $\mathcal{K}_x$ in (58). Since $x \in K$, the only possibility is that $x \notin \mathcal{Z}(\mathbb{1}_K)$. The same reasoning applies to $\mathcal{Z}_D$ (recalling Lemma 5.11) and we get

$$K \setminus K' \cap \mathcal{Z}(\mathbb{1}_K) = \emptyset = K \setminus K' \cap \mathcal{Z}_\mathcal{D}(\mathbb{1}_K)$$

Since $\mathcal{V} = K^\mathrm{c} \sqcup K' \sqcup (K \setminus K')$, we conclude that

$$\mathcal{Z}(\mathbb{1}_K) = K^\mathrm{c} \sqcup K' = \mathcal{Z}_\mathcal{D}(\mathbb{1}_K),$$

finishing the proof. $\qquad\square$

Remark 5.15. Note that (Z5) was only used to prove that $K' \subset \mathcal{Z}(\mathbb{1}_K)$. Thus, when (Z5) is not verified, the constructed modulus of descent $T_\mathcal{D}$ might enlarge the critical map $\mathcal{Z}_\mathcal{D}$ with respect to $\mathcal{Z}$, as illustrated by Example 5.17 below.

The following two examples show that homogeneity and (Z4) are necessary assumptions.

Example 5.16 (A non-homogeneous descent modulus failing (Z2))**.** Let $\varepsilon > 0$ and consider the operator $T_\varepsilon : \mathbb{R}^\mathcal{V} \to \mathbb{R}_+^\mathcal{V}$ given by

$$T_\varepsilon[f](x) := \begin{cases} f(x) - \min f & \text{if } f(x) > \min f + \varepsilon \\ 0 & \text{if } f(x) \le \min f + \varepsilon. \end{cases} = \phi_\varepsilon(f(x) - \min f)$$

where the mapping ϕ_ε is defined for $r \ge 0$ by

$$\phi_\varepsilon(r) := \begin{cases} 0, & \text{if } r \in [0, \varepsilon] \\ r, & \text{if } r > \varepsilon. \end{cases}$$

We claim that T_ε is a descent modulus.

- Let $f \in \mathbb{R}^\mathcal{V}$. For every $x \in \operatorname{argmin} f$, we have that $T_\varepsilon[f](x) = 0$, and so T_ε preserves global minima.

- Let $f, g \in \mathbb{R}^\mathcal{V}$ and $x \in \mathcal{V}$ such that

$$(f(x) - f(z))_+ \ge (g(x) - g(z))_+, \quad \forall z \in \mathcal{V}.$$

  On the one hand, if $f(x) \le \min f + \varepsilon$, we have that

$$\varepsilon \ge (f(x) - f(z))_+ \ge (g(x) - g(z))_+, \quad \forall z \in \mathcal{V},$$

  and so, $g(x) \le \min g + \varepsilon$ as well. Then $T_\varepsilon[f](x) = T_\varepsilon[g](x)$. On the other hand, if $f(x) \ge \min f + \varepsilon$, by taking $z^* \in \operatorname{argmin} g$, we have that

$$\begin{aligned} T_\varepsilon[f](x) = f(x) - \min f &\ge (f(x) - f(z^*))_+ \\ &\ge (g(x) - g(z^*))_+ = g(x) - \min g \ge T_\varepsilon[g](x). \end{aligned}$$

  Thus, T_ε is monotone.

- Let $f \in \mathbb{R}^\mathcal{V}$ and $r > 1$. Then, for every $x \in \mathcal{V}$,

$$T_\varepsilon[f](x) > 0 \implies \varepsilon \le T_\varepsilon[f](x) = f(x) - \min f < r(f(x) - \min f) = rf(x) - \min rf = T_\varepsilon[rf](x).$$

  Thus, T_ε is scalar-monotone.

Then, by definition, T_ε is a descent modulus. However, choose $f \in \mathbb{R}^\mathcal{V}$ such that $\alpha = \max f - \min f > \varepsilon$, and choose $r = \frac{\varepsilon}{\alpha}$. Then, we have that $\mathcal{Z}_T(f) \ne \mathcal{V}$ but

$$\forall x \in \mathcal{V}, \quad rf(x) - \min rf \le r(\max f - \min f) = \varepsilon \implies \mathcal{Z}_T(rf) = \mathcal{V}.$$

$\diamond$

Example 5.17 (A family of homogeneous moduli of descent failing (Z4)). Let $\mathcal{D}' = (\mathcal{D}'_x)_{x \in \mathcal{V}} \in \mathcal{E}(\mathcal{V})$ such that there exists $\bar{x} \in \mathcal{V}$ where $|\mathcal{D}'_{\bar{x}}| \geq 3$, and let $T : \mathbb{R}^{\mathcal{V}} \to \mathbb{R}^{\mathcal{V}}_+$ be the operator given by

$$T[f](x) := \begin{cases} \left(f(x) - \max\limits_{y \in \mathcal{D}'_x \setminus \{x\}} f(y) \right)_+ & \text{if } \mathcal{D}'_x \neq \{x\} \\ 0 & \text{if } \mathcal{D}'_x = \{x\}. \end{cases}$$

Clearly T is homogeneous and preserves global minima. Let us prove that T is monotone: let $f, g \in \mathbb{R}^{\mathcal{V}}$ and $x \in \mathcal{V}$ such that

$$(f(x) - f(z))_+ \geq (g(x) - g(z))_+, \quad \forall z \in \mathcal{V}.$$

If $\mathcal{D}'_x = \{x\}$, then $T[f](x) = 0 \geq 0 = T[g](x)$. If $\mathcal{D}'_x \neq \{x\}$, then there exist $y^* \in \mathcal{D}'_x \setminus \{x\}$ such that $f(y^*) = \max_{y \in \mathcal{D}'_x \setminus \{x\}} f(y)$. Then,

$$T[f](x) = (f(x) - f(y^*))_+ \geq (g(x) - g(y^*))_+ \geq \left(g(x) - \max\limits_{y \in \mathcal{D}'_x \setminus \{x\}} g(y) \right)_+ = T[g](x).$$

Thus, T is monotone and therefore it is a descent modulus. Now, let $(\mathcal{K}_x)_{x \in \mathcal{V}}$ and $\mathcal{D} = (\mathcal{D}_x)_{x \in \mathcal{V}}$ constructed as in (58) for $\mathcal{Z}_T$. Then:

- If $\mathcal{D}'_x = \{x\}$, then $\mathcal{K}_x = \{K \in \mathcal{P}(\mathcal{V})^* : x \in K\}$ and so $\mathcal{D}_x = \mathcal{D}'_x$.

- If $\mathcal{D}'_x = \{x, y\}$ for some $y \neq x$, then $\mathcal{K}_x = \{K \in \mathcal{V} : \{x, y\} \subset K\}$. Indeed, consider $K \in \mathcal{K}_x$, we have $x \in K$ and $x \in \mathcal{Z}_T(K)$. We compute

$$T(\mathbb{1}_K)(x) = (\mathbb{1}_K(x) - \mathbb{1}_K(y))_+ = 1 - \mathbb{1}_K(y)$$

  so for this expression to vanish, we must have $y \in K$. Conversely, if $\{x, y\} \subset K$, then $x \in K$ and

$$T(\mathbb{1}_K)(x) = (\mathbb{1}_K(x) - \mathbb{1}_K(y))_+ = 0$$

  so $K \in \mathcal{K}_x$. We deduce $\mathcal{D}_x = \{x, y\}$ and so $\mathcal{D}_x = \mathcal{D}'_x$.

- If $|\mathcal{D}'_x| \geq 3$, we have that there is $y, z \in \mathcal{D}'_x \setminus \{x\}$ with $y \neq z$ such that $\{x, y\}, \{x, z\} \in \mathcal{K}_x$. Thus, $\mathcal{D}_x \subset \{x, y\} \cap \{x, z\} = \{x\} \neq \mathcal{D}'_x$, it follows that $\mathcal{D}_x = \{x\}$. Furthermore, $\{x\} \notin \mathcal{K}_x$ since $T[\mathbb{1}_{\{x\}}](x) = 1$.

Thus, since $|\mathcal{D}'_{\bar{x}}| \geq 3$, T fails (Z4). $\qquad \diamond$

References

[1] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Mathematics. Birkhäuser, second edition, 2008.

[2] J.-B. Baillon. Personal communication, October 2018.

[3] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators.* Grundlehren der mathematischen Wissenschaften 346. Springer, Cham, 2014.

[4] J. Benoist and A. Daniilidis. Subdifferential representation of convex functions: refinements and applications. *J. Convex Anal.*, 12(2):255–265, 2005.

[5] L. E. Blumenson. Classroom Notes: A Derivation of n-Dimensional Spherical Coordinates. *Amer. Math. Monthly*, 67(1):63–66, 1960.

[6] T. Z. Boulmezaoud, P. Cieutat, and A. Daniilidis. Gradient flows, second-order gradient systems and convexity. *SIAM J. Optim.*, 28(3):2049–2066, 2018.

[7] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert.* North-Holland Publishing Co., 1973.

[8] J. N. Corvellec, M. Degiovanni, and M. Marzocchi. Deformation properties for continuous functionals and critical point theory. *Topol. Methods Nonlinear Anal.*, 1(1):151–171, 1993.

[9] A. Daniilidis and D. Salas. A determination theorem in terms of the metric slope. *Proc. Amer. Math. Soc.*, 150(10):4325–4333, 2022.

[10] E. De Giorgi, A. Marino, and M. Tosques. Problems of evolution in metric spaces and maximal decreasing curve. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat. (8)*, 68(3):180–187, 1980.

[11] M. Degiovanni and M. Marzocchi. A critical point theory for nonsmooth functionals. *Ann. Mat. Pura Appl. (4)*, 167:73–100, 1994.

[12] D. Drusvyatskiy, A. D. Ioffe, and A. S. Lewis. Curves of descent. *SIAM J. Control Optim.*, 53(1):114–138, 2015.

[13] S. N. Ethier and T. G. Kurtz. *Markov processes.* Wiley Series in Probability and Mathematical Statistics. New York, 1986.

[14] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions.* Textbooks in Mathematics. CRC Press, Boca Raton, FL, revised edition, 2015.

[15] L. Gajek and D. Zagrodny. Countably orderable sets and their applications in optimization. *Optimization*, 26(3-4):287–301, 1992.

[16] A. D. Ioffe. *Variational analysis of regular mappings.* Springer Monographs in Mathematics. Springer, Cham, 2017. Theory and applications.

[17] J.-J. Moreau. Proximité et dualité dans un espace hilbertien. *Bull. Soc. Math. France*, 93:273–299, 1965.

[18] P. Pérez-Aros, D. Salas, and E. Vilches. Determination of convex functions via subgradients of minimal norm. *Math. Program.*, 190(1-2, Ser. A):561–583, 2021.

[19] R. T. Rockafellar. Characterization of the subdifferentials of convex functions. *Pacific J. Math.*, 17:497–510, 1966.

[20] R. T. Rockafellar. On the maximal monotonicity of subdifferential mappings. *Pacific J. Math.*, 33:209–216, 1970.

[21] R. T. Rockafellar and Roger J.-B. Wets. *Variational analysis.* Grundlehren der mathematischen Wissenschaften 317. Springer, 1998.

[22] D. J. Smith and M. K. Vamanamurthy. How small is a unit ball? *Math. Mag.*, 62(2):101–107, 1989.

[23] K. T. Sturm. Diffusion processes and heat kernels on metric spaces. *Ann. Probab.*, 26(1):1–55, 1998.

[24] L. Thibault and D. Zagrodny. Determining functions by slopes. *Comm. Contemp. Math.*, 2022. (In press).

Aris DANIILIDIS

Institute of Statistics and Mathematical Methods in Economics, VADOR E105-04
TU Wien, Wiedner Hauptstraße 8, A-1040 Wien

(on leave) DIM-CMM, CNRS IRL 2807
Beauchef 851, FCFM, Universidad de Chile

E-mail: `aris.daniilidis@tuwien.ac.at`
`https://www.arisdaniilidis.at/`

Laurent MICLO

Toulouse School of Economics, University Toulouse 1 (Capitole)
1, Esplanade de l'Université, F-31080 Toulouse Cedex 06, France

E-mail: `laurent.miclo@math.cnrs.fr`
`https://perso.math.univ-toulouse.fr/miclo/`

David SALAS

Instituto de Ciencias de la Ingenieria, Universidad de O'Higgins
Av. Libertador Bernardo O'Higgins 611, Rancagua, Chile

E-mail: `david.salas@uoh.cl`
`http://davidsalasvidela.cl`