A continuous gradient-like dynamical approach to Pareto-optimization in Hilbert spaces

Hedy Attouch, Xavier Goudou

A continuous gradient-like dynamical approach to Pareto-optimization in Hilbert spaces *

Hédy Attouch and X. Goudou,

Université Montpellier II
Institut de Mathématiques et de Modélisation de Montpellier – UMR CNRS 5149
34095 Montpellier, France

March 23, 2013

Abstract

In a Hilbert space setting, we consider new continuous gradient-like dynamical systems for constrained multiobjective optimization. This type of dynamics was first investigated by Cl. Henry, and B. Cornet, as a model of allocation of resources in economics. Based on the Yosida regularization of the discontinuous part of the vector field which governs the system, we obtain the existence of strong global trajectories. We prove a descent property for each objective function, and in the quasi-convex case, convergence of the trajectories to Pareto critical points. We give an interpretation of the dynamic in terms of Pareto equilibration for cooperative games. By time discretization, we make a link to recent studies of Svaiter et al. on the algorithm of steepest descent for multiobjective optimization.

Keywords Multiobjective optimization · Continuous gradient systems · Asymptotic behavior · Pareto critical · Pareto optimization · Cooperative games · Steepest descent.

Mathematics Subject Classifications (2010) 34E10 · 37L05 · 37L65 · 90B50 · 90C29 · 90C31 · 91A12 · 91A35 · 91B06 · 91B55

Introduction

In this paper, we study a continuous gradient-like dynamical system which enjoys remarkable properties with respect to constrained Pareto optimization. It was first investigated by Henry [30] and Cornet [21], [22], [23] in the seventies, as a dynamical model of allocation of resources in economics. We revisit it from a modern perspective. In a general Hilbert space setting, we provide a new constructive proof of the existence of strong global trajectories, and show their convergence in the case of quasi-convex functions. We establish a link between this dynamical system and the steepest descent algorithms for multiobjective optimization which have been recently introduced by Svaiter et al. in [25], [26]. As a general rule, making a link between optimization

*Contact author: H. Attouch, hedy.attouch@univ-montp2.fr, phone: +33 04 67 14 35 72, fax: +33 04 67 14 35 58.

algorithms and continuous dynamics has proved to be fruitful. For further research, it may serve as a mathematical basis to study complex interactions in game theory (for example by coupling cooperative dynamic and noncooperative dynamic). As well, it may be adapted in order to fit with topics of current interest in multiobjective optimization like, control, inertia, multiscale aspects, Newton methods.

Let us now describe the system. Throughout the paper, $\mathcal{H}$ is a real Hilbert space with scalar product and norm denoted by $\langle \cdot, \cdot \rangle$ and $\| \cdot \| = \sqrt{\langle \cdot, \cdot \rangle}$ respectively. We make the following standing assumptions on the multiple objective functions $(f_i)_{i=1,2,..,q}$, $(q \in \mathbb{N})$, and constraint K:

H0) $K \subset \mathcal{H}$ is a closed convex nonempty set.

For each $i = 1, 2, ..., q$, $f_i : \mathcal{H} \to \mathbb{R}$ is a real-valued function which satisfies:

H1) f_i is differentiable, its gradient ∇f_i is Lipschitz continuous on bounded sets;

H2) $f_i : \mathcal{H} \to \mathbb{R}$ is a quasi-convex function;

H3) f_i is bounded from below on K.

The description of our dynamic involves the following convex sets in $\mathcal{H}$: for $u \in K$,

$$N_K(u) = \{z \in \mathcal{H} : \ \langle z, v - u \rangle \leq 0 \quad \text{for any } v \in K\}$$

is the normal cone to K at u. It is a closed convex cone which models the *contact forces* which are attached to the constraint K. For $u \in \mathcal{H}$,

$$\text{Conv}\,\{\nabla f_i(u);\ i = 1, 2, ..., q\} = \{\sum_{i=1}^{q} \lambda_i \nabla f_i(u) : \quad 0 \leq \lambda_i \leq 1,\ \sum_{i=1}^{q} \lambda_i = 1\}$$

is the closed convex hull of the vectors $\{\nabla f_i(u);\ i = 1, 2, ..., q\}$. It models the *driving forces* which governs our system.

For any closed convex set $C \subset \mathcal{H}$ we denote by

$$C^0 = \text{proj}_C 0$$

the projection of the origin onto C, which is the element of minimal norm of C.

For any $u \in K$ the set

$$N_K(u) + \text{Conv}\,\{\nabla f_i(u);\ i = 1, 2, ..., q\}$$

is a closed convex set, as being equal to the vectorial sum of two closed convex sets, one of them being compact. Thus, it has a unique element of minimal norm, which allows us to define the differential equation

$$\text{(MCSD)} \quad \dot{u}(t) + \Big(N_K(u(t)) + \text{Conv}\,\{\nabla f_i(u(t));\ i = 1, 2, ..., q\} \Big)^0 = 0, \qquad (1)$$

called the Multiobjective Continuous Steepest Descent, (MCSD) for short. From the point of view of modeling, it has the following properties:

a) It is a descent method, i.e., for *each* $i = 1, 2, ..., q$, $t \mapsto f_i(u(t))$ is nonincreasing.

b) Its trajectories converge to Pareto critical points.

c) The scalarization of the multiobjective optimization problem is done endogeneously. At time t, the vector field which governs the system involves a convex combination $\sum_{i=1}^{q} \lambda_i(t) \nabla f_i(\cdot)$ of the gradients $\nabla f_i(\cdot)$, with scalars $\lambda_i(t)$ which are not fixed in advance. They are part of the process, whence the decentralized features of this dynamic. Indeed, this system and its companions provide a rich dynamical model of the interaction between cooperating agents.

The mathematical analysis of (MCSD) is quite difficult, because the dynamic is governed by a discontinuous vector field. Indeed, the multivalued operator $v \mapsto N_K(v)$ is discontinuous at the boundary of K. Trajectories of (MCSD) are *lazy* solutions (also called slow solutions, see [9, Ch. 6, section 8]) of the differential inclusion

$$\dot{u}(t) + N_K(u(t)) + \text{Conv}\,\{\nabla f_i(u(t));\ i = 1, 2, ..., q\} \ni 0, \tag{2}$$

which is governed by the sum of the maximal monotone operator N_K and the multivalued continuous operator $v \mapsto \text{Conv}\,\{\nabla f_i(v);\ i = 1, 2, ..., q\}$. Lazy solution means that the trajectory chooses a velocity which has minimal norm among all the possible directions offered by the differential inclusion. This type of differential inclusion occurs in various domains (mechanics, economics, control...), and has subsequently be the object of active research, see for example [5], [8], [14], [20].

In this paper we provide several new results concerning (MCSD):

i) In a Hilbert space setting, in Theorem 1.5 we prove convergence of the trajectories of (MCSD) to Pareto critical solutions of the multiobjective constrained problem

$$\min\{F(v) :\ v \in K\}$$

where $F : \mathcal{H} \to \mathbb{R}^q,\ F(v) = (f_i(v))_{i=1,2,...,q}$. We obtain this result under the assumption of quasi-convexity of the f_i and convexity of K. Our proof is in the line of the proof of convergence of the steepest descent by Goudou-Munier [29] in the case of a single objective function (scalar case); it makes use of Lyapunov analysis and Opial's lemma.

ii) In Theorem 2.5, we prove the existence of strong global solutions of the more general evolution equation

$$\dot{u}(t) + \Big(\partial\Phi(u(t)) + \text{Conv}\,\{\nabla f_i(u(t));\ i = 1, 2, ..., q\}\Big)^0 = 0, \tag{3}$$

where $\partial\Phi$ is the subdifferential of a closed convex proper function $\Phi : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$. (MCSD) corresponds to $\Phi = \delta_K$, the indicator function of K. We provide a new constructive proof of the existence of strong (i.e., absolutely continuous on each bounded time interval) global solutions. It is based on the Yosida regularization of $\partial\Phi$, and Peano existence theorem for differential equations. The difficult point is to pass to the limit on the regularized differential equations, as the regularization parameter goes to zero, because the vector field which governs our dynamic is not continuous, nor monotone. This new approach is flexible: it can be easily adapted to the case of non-autonomous equations with constraint and/or criteria which vary (changing environment, uncertainty). From the modeling and numerical point of view, another interesting aspect is the ability to handle the constraint by various ways, such as exterior penalty, or interior point methods.

iii) In Theorem 3.5 we provide an equivalent formulation of the steepest descent direction for multiobjective constrained optimization, and its interpretation for cooperative games. We briefly analyze the Pareto equilibration process associated to (MCSD). Then, by time discretization of (MCSD), we establish the link with some recent numerical algorithms for multiobjective optimization: the steepest descent of Fliege and Svaiter [25], Grana Drummond and Svaiter [26], and the proximal methods in vector optimization of Bonnel, Iusem and Svaiter [16].

The paper is organized following the three above items: in section 1, we show the convergence to Pareto critical points of the trajectories of (MCSD), as t goes to infinity. Then, in section 2, we prove the existence of strong global solutions to (MCSD). Finally, in sections 4 and 5 we successively examine some natural connections between (MCSD), cooperative games, and numerical algorithms for multiobjective optimization.

1 Asymptotic convergence to Pareto critical points

In this section, we study the asymptotic behavior ($t \to +\infty$) of the trajectories of (MCSD). We take for granted the existence of strong global solutions of (MCSD). This question will be examined into detail in the next section. Such solution $u : [0, +\infty) \to \mathcal{H}$ is continuous, and absolutely continuous on each interval $[0, T]$, $0 < T < +\infty$, (see Definition 2.4). In order to prove the weak convergence of the trajectories of (MCSD) we use the classical Opial's lemma [35]. We recall in its continuous form, and give a short proof of it:

Lemma 1.1. *Let S be a non empty subset of $\mathcal{H}$ and $u : [0, +\infty) \to \mathcal{H}$ a map. Assume that*

$$(i) \qquad \text{for every } z \in S,\ \lim_{t\to+\infty} \|u(t) - z\| \text{ exists};$$

$$(ii) \qquad \text{every weak limit point of the map } u \text{ belongs to } S.$$

Then

$$w - \lim_{t\to+\infty} u(t) = u_\infty \quad \text{exists, for some element } u_\infty \in S.$$

Proof. By (i) and $S \neq \emptyset$, the trajectory u is bounded in $\mathcal{H}$. In order to obtain its weak convergence, we just need to prove that the trajectory has a unique weak limit point. Let $u(t_n^1) \rightharpoonup z^1$ and $u(t_n^2) \rightharpoonup z^2$, with $t_n^1 \to +\infty$, and $t_n^2 \to +\infty$. By (ii), $z^1 \in S$, and $z^2 \in S$. By (i), it follows that $\lim_{t\to+\infty} \|u(t) - z^1\|$ and $\lim_{t\to+\infty} \|u(t) - z^2\|$ exist. Hence, $\lim_{t\to+\infty}(\|u(t) - z^1\|^2 - \|u(t) - z^2\|^2)$ exists. Developing and simplifying this last expression, we deduce that

$$\lim_{t\to+\infty} \left\langle u(t), z^2 - z^1 \right\rangle \quad \text{exists.}$$

Hence

$$\lim_{n\to+\infty} \left\langle u(t_n^1), z^2 - z^1 \right\rangle = \lim_{n\to+\infty} \left\langle u(t_n^2), z^2 - z^1 \right\rangle,$$

which gives $\|z^2 - z^1\|^2 = 0$, and hence $z^2 = z^1$. □

The stationary points of our dynamic are Pareto critical points, as defined below.

Definition 1.2. *We say that $\bar{u}$ is a Pareto critical point of the multiobjective optimization problem $\{(f_i)_{i=1,...,q};\ K\}$ if there exists $(\lambda_i)_{i=1,2,...,q}$ such that*

$$0 \leq \lambda_i \leq 1,\ \sum_{i=1}^{q} \lambda_i = 1 \quad \sum_{i=1}^{q} \lambda_i \nabla f_i(\bar{u}) \ + N_K(\bar{u}) \ni 0. \tag{4}$$

This notion has been considered by Smale in [38], Cornet in [21], see [11], [26], [40] for recent account of this notion, and various extensions of it. Note that equivalent formulations of this notion can be given, thanks to the positive homogeneity property of the formula: the condition $\sum_{i=1}^{q} \lambda_i = 1$ can be dropped, just assuming the λ_i to be nonnegative, and at least one of them positive. This notion is naturally linked to the notion of Pareto optimality. Set $F : \mathcal{H} \to \mathbb{R}^q,\ F(v) = (f_i(v))_{i=1,2,...,q}$. The order in $\mathbb{R}^q$ is defined by $y \preceq z \Leftrightarrow y_i \leq z_i$ for all $i = 1, 2, ..., q$. We also consider the strict order relation $y \prec z \Leftrightarrow y_i < z_i$ for all $i = 1, 2, ..., q$.

Definition 1.3. *i) An element $u \in K$ is called Pareto optimal if there does not exist $v \in K$ such that $F(v) \preceq F(u)$ and $F(v) \neq F(u)$. Equivalently, there there does not exist $v \in K$ such that $f_i(v) \leq f_i(u)$ for all $i = 1, 2, ..., q$, and $f_j(v) < f_j(u)$ for one $j \in 1, 2, ..., q$.*
ii) An element $u \in K$ is called weak Pareto optimal if there does not exist $v \in K$ such that $F(v) \prec F(u)$. Equivalently, there does not exist $v \in K$ such that $f_i(v) < f_i(u)$ for all $i = 1, 2, ..., q$.

Let us respectively denote by $\mathcal{P}$, $\mathcal{P}_w$, and $\mathcal{P}_c$ the set of Pareto optima, weak Pareto optima, and Pareto critical points. Indeed, Pareto criticality is a first-order necessary optimality condition for (local) vectorial optimization. It is a multiobjective extension of the Fermat rule. Let us state it in a precise way.

Proposition 1.4. *i) The following inclusions hold:*

$$\mathcal{P} \subset \mathcal{P}_w \subset \mathcal{P}_c.$$

ii) The equality $\mathcal{P}_w = \mathcal{P}_c$ is true when all functions $f_i,\ i = 1, 2, ..., q$ are convex.

Proof. The inclusion $\mathcal{P}_w \subset \mathcal{P}_c$ in item i) is obtained by a direct application of the Hahn-Banach separation theorem (see for example [21, Proposition 1.1], [40], [11]).
Item ii) treats the case where all functions $f_i,\ i = 1, 2, ..., q$ are convex. Let $u \in \mathcal{P}_c$. Then, u is a (global) solution of the convex minimization problem

$$\min \left\{ \quad \sum_{i=1}^{q} \lambda_i f_i(v) :\ v \in K \right\} \tag{5}$$

with λ_i which are all nonnegative, and at least one of them positive. Indeed, (5) forces u to be a weak Pareto minima. Otherwise, there would exist some $v \in K$ such that $f_i(v) < f_i(u)$ for all $i = 1, 2, ..., q$, which would imply (one uses the fact at least one of the λ_i is positive) $\sum_{i=1}^{q} \lambda_i f_i(v) < \sum_{i=1}^{q} \lambda_i f_i(u)$, a clear contradiction. □

We can now state our main convergence result.

Theorem 1.5. *Let us make assumptions* H0), H1), H2), H3). *Then for any strong global solution* $t \in [0, +\infty[\mapsto u(t) \in \mathcal{H}$ *of* (MCSD)*, the following properties hold:*
i) Descent property: for each $i = 1, 2, ..., q$, $t \mapsto f_i(u(t))$ *is nonincreasing, and for almost all* $t > 0$

$$\|\dot{u}(t)\|^2 + \frac{d}{dt} f_i(u(t)) \leq 0. \tag{6}$$

ii) Finite energy property:

$$\int_0^{+\infty} \|\dot{u}(t)\|^2 dt < +\infty. \tag{7}$$

iii) Weak convergence: Assume that the trajectory $t \in [0, +\infty[\mapsto u(t) \in \mathcal{H}$ *is bounded in* $\mathcal{H}$*. Then* $u(t)$ *converges weakly in* $\mathcal{H}$ *as* $t \to +\infty$*.*
iv) Pareto optimization: Assume moreover that the trajectory $t \in [0, +\infty[\mapsto u(t) \in \mathcal{H}$ *is relatively compact in* $\mathcal{H}$*. Then* $u(t)$ *converges strongly in* $\mathcal{H}$ *as* $t \to +\infty$ *to a Pareto critical point* u_∞*.*

Proof. By definition of (MCSD), $-\dot{u}(t)$ is the projection of the origin onto the closed convex set $C(t) := N_K(u(t)) + \text{Conv}\{\nabla f_i(u(t))\}$. Hence, for any $\xi \in C(t)$

$$\langle 0 - (-\dot{u}(t)), \xi - (-\dot{u}(t)) \rangle \leq 0, \tag{8}$$

that is

$$\langle \dot{u}(t), \xi + \dot{u}(t) \rangle \leq 0. \tag{9}$$

Noticing that $0 \in N_K(u(t))$ and $\nabla f_i(u(t)) \in \text{Conv}\{\nabla f_i(u(t))\}$, we can take $\xi = \nabla f_i(u(t))$ in (9), which yields

$$\langle \dot{u}(t), \nabla f_i(u(t)) + \dot{u}(t) \rangle \leq 0. \tag{10}$$

The derivation chain rule is valid in our situation, see for example [19, Corollaire VIII.10]. Hence

$$\|\dot{u}(t)\|^2 + \frac{d}{dt} f_i(u(t)) \leq 0. \tag{11}$$

As a consequence, $\frac{d}{dt} f_i(u(t)) \leq 0$, and for each $i = 1, 2, ..., q$ the function $t \mapsto f_i(u(t))$ is nonincreasing. Moreover by integrating (11) and using that f_i is bounded from below on K we obtain

$$\int_0^{+\infty} \|\dot{u}(t)\|^2 dt \leq f_i(u(0)) - \inf_K f_i. \tag{12}$$

This proves points *i)* and *ii)*.

Let us now prove the weak convergence of any bounded trajectory u of (MCSD). To that end we use Opial's Lemma 1.1 with

$$S = \left\{ v \in K : \ \forall i = 1, 2, ..., q \ \ f_i(v) \leq \inf_{t \geq 0} f_i(u(t)) \right\}. \tag{13}$$

Functions f_i are continuous, quasi-convex, and hence lower semicontinuous for the weak topology of $\mathcal{H}$. As well, the closed convex set K is closed for the weak topology of $\mathcal{H}$. The trajectory $t \in [0, +\infty[\mapsto u(t) \in \mathcal{H}$ has been assumed to be bounded in $\mathcal{H}$. Hence,

any weak limit point of the trajectory belongs to S, which is a closed convex non empty subset of $\mathcal{H}$.

i) Take $z \in S$ and set, for any $t \geq 0$

$$h_z(t) = \frac{1}{2}\|u(t) - z\|^2. \tag{14}$$

We have

$$\dot{h}_z(t) = \langle u(t) - z, \dot{u}(t) \rangle . \tag{15}$$

Since u is a solution of (MCSD) there exists

$$\eta(t) \in N_K(u(t)) \text{ and } 0 \leq \lambda_i(t) \leq 1, \ \sum_{i=1}^{q} \lambda_i(t) = 1 \tag{16}$$

such that, for almost all $t > 0$

$$\dot{u}(t) + \sum_{i=1}^{q} \lambda_i(t) \nabla f_i(u(t)) + \eta(t) = 0. \tag{17}$$

By combining (15) and (17) we obtain

$$\dot{h}_z(t) + \sum_{i=1}^{q} \lambda_i(t) \langle u(t) - z, \nabla f_i(u(t)) \rangle + \langle u(t) - z, \eta(t) \rangle = 0. \tag{18}$$

On the one hand, since $\eta(t) \in N_K(u(t))$ and $z \in K$

$$\langle \eta(t), u(t) - z \rangle \geq 0. \tag{19}$$

On the other hand, by quasi-convexity of f_i, the set $\{v \in \mathcal{H} : \ f_i(v) \leq f_i(u(t))\}$ is convex. Since $u(t)$ and z belong to this set, we have that, for any $0 \leq \theta \leq 1$, $\theta z + (1-\theta)u(t)$ remains in this set. Hence

$$f_i(u(t) + \theta(z - u(t))) - f_i(u(t)) \leq 0. \tag{20}$$

Dividing by $\theta > 0$, and letting θ go to zero gives

$$\langle \nabla f_i(u(t)), z - u(t) \rangle \leq 0. \tag{21}$$

As a consequence

$$\sum_{i=1}^{q} \lambda_i(t) \langle \nabla f_i(u(t)), u(t) - z \rangle \geq 0. \tag{22}$$

Combining (18) with (19) and (22) we obtain

$$\dot{h}_z(t) \leq 0. \tag{23}$$

Hence, h_z is a decreasing function, which proves item *i*) of Lemma 1.1.

Let us verify item *ii*) of Lemma 1.1. Let $w - \lim u(t_n) = z$ for some sequence $t_n \to +\infty$. Since $u(t_n) \in K$ and K is a closed convex subset of $\mathcal{H}$, we have $z \in K$. Moreover

$$\inf_{t\geq 0} f_i(u(t)) = \lim_{t\to+\infty} f_i(u(t)) \tag{24}$$

$$= \lim_{n\to+\infty} f_i(u(t_n)) \tag{25}$$

$$\geq f_i(z) \tag{26}$$

where the last inequality follows from the fact that f_i is continuous and quasi-convex, and hence lower semicontinuous for the weak topology of $\mathcal{H}$. This being true for each $i = 1, 2, ..., q$ we conclude that $z \in S$. The two conditions of the Opial Lemma 1.1 are satisfied, which gives the weak convergence of all bounded trajectory u of (MCSD).

Let us now assume that the trajectory $t \mapsto u(t)$ is relatively compact. As a consequence, it is bounded, and by *iii*) it converges weakly, and hence strongly in $\mathcal{H}$ as $t \to +\infty$ (on relatively compact sets weak and strong convergence are equivalent notions). Set

$$u(t) \to u_\infty \quad \text{strongly in } \mathcal{H}, \tag{27}$$

and show that u_∞ is a Pareto critical point. Clearly, the finite energy property (7)

$$\int_0^{+\infty} \|\dot{u}(t)\|^2 dt < +\infty$$

implies

$$\text{liminfess}_{t\to+\infty} \|\dot{u}(t)\| = 0.$$

Equivalently, there exists a sequence $t_n \to +\infty$ such that

$$\dot{u}(t_n) \to 0. \tag{28}$$

By (16) and (17)

$$\dot{u}(t_n) + \sum_{i=1}^{q} \lambda_i(t_n) \nabla f_i(u(t_n)) + \eta(t_n) = 0, \tag{29}$$

with

$$\eta(t_n) \in N_K(u(t_n)), \text{ and } 0 \leq \lambda_i(t_n) \leq 1, \ \sum_{i=1}^{q} \lambda_i(t_n) = 1. \tag{30}$$

By compactness of the unit simplex in $\mathbb{R}^q$, we can extract a further subsequence (still noted t_n to simplify the notation) such that, for each $i = 1, 2, ..., q$

$$\lambda_i(t_n) \to \lambda_{i,\infty}, \tag{31}$$

with

$$0 \leq \lambda_{i,\infty} \leq 1, \ \sum_{i=1}^{q} \lambda_{i,\infty} = 1. \tag{32}$$

By (27), $u(t_n) \to u_\infty$ strongly in $\mathcal{H}$. By continuity of $\nabla f_i : \mathcal{H} \to \mathcal{H}$,

$$\nabla f_i(u(t_n)) \to \nabla f_i(u_\infty). \tag{33}$$

We can now pass to the limit on (29). By (27), (28), (31), (33)

$$\eta(t_n) \in N_K(u(t_n))$$
$$u(t_n) \to u_\infty \text{ strongly in } \mathcal{H}$$
$$\eta(t_n) = -\dot{u}(t_n) - \sum_{i=1}^{q} \lambda_i(t_n) \nabla f_i(u(t_n)) \to -\sum_{i=1}^{q} \lambda_{i,\infty} \nabla f_i(u_\infty).$$

By closedness of the graph of the set-valued mapping $v \mapsto N_K(v)$ (it is a maximal monotone operator), we finally obtain

$$\sum_{i=1}^{q} \lambda_{i,\infty} \nabla f_i(u_\infty) + N_K(u_\infty) \ni 0 \tag{34}$$

which, with (32), expresses that u_∞ is a Pareto critical point. □

When the objective functions f_i are convex, one can get rid of the compactness assumption on the trajectory, and, at the limit, still obtain Pareto optimality.

Proposition 1.6. *Assume that all the functions f_i, $i = 1, 2, ..., q$ are convex. Let $t \in [0, +\infty[\mapsto u(t) \in \mathcal{H}$ be a strong global solution of* (MCSD)*. Let us assume that u is bounded. Then $u(t)$ converges weakly in $\mathcal{H}$ as $t \to +\infty$ to a weak Pareto minimum.*

Proof. By Theorem 1.5 iii), we know that $u(t)$ converges weakly in $\mathcal{H}$ as $t \to +\infty$ to some u_∞. We want to prove that u_∞ is a Pareto critical point. We just need to adapt the end of the proof of Theorem 1.5 iii). Let us show how, thanks to convexity, we can get rid of the strong compactness assumption on the trajectory, and instead, use a weak compactness argument. We have

$$\dot{u}(t_n) \to 0 \text{ strongly in } \mathcal{H}; \tag{35}$$
$$u(t_n) \to u_\infty \text{ weakly in } \mathcal{H}; \tag{36}$$
$$-\dot{u}(t_n) \in \sum_{i=1}^{q} \lambda_i(t_n) \nabla f_i(u(t_n)) + N_K(u(t_n)). \tag{37}$$

We can rewrite the above inclusion as

$$-\dot{u}(t_n) \in \partial G_n(u(t_n)) \tag{38}$$

with $G_n : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ defined by

$$G_n(v) = \sum_{i=1}^{q} \lambda_i(t_n) f_i(v) + \delta_K(v)$$

(δ_K is the indicator function of K). By using (31), it can be easily verified that the sequence of closed convex functions $(G_n)_n$ Mosco-epiconverges to $G : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$

$$G(v) = \sum_{i=1}^{q} \lambda_{i,\infty} f_i(v) + \delta_K(v).$$

We now use the equivalence between Mosco-epiconvergence of a sequence of closed convex functions, and the graph-convergence of their subdifferentials, [3, Theorem 3.66, Proposition 3.59], see also [2], [13, Theorem 8.3.9]. We obtain the graph-convergence of the sequence of maximal monotone operators ∂G_n to ∂G, and by [3, Proposition 3.59], their convergence in the weak $-\mathcal{H}\times$ strong $-\mathcal{H}$ topology. By applying this result to (38), and taking account of (35), we obtain $0\in\partial G(u_\infty)$. Equivalently,

$$\sum_{i=1}^{q}\lambda_{i,\infty}\nabla f_i(u_\infty)+N_K(u_\infty)\ni 0,$$

which, with (32), expresses that u_∞ is a Pareto critical point. By Proposition 1.4, and the convexity of the f_i, it follows that u_∞ is a weak Pareto minimum. □

Remark 1.7. a) In the unconstrained single criteria case, it is proven in [29] that the trajectory converges towards a critical (Pareto) point as soon as the set of optimal (Pareto) points is non empty. This is an open question in our multiobjective framework.

b) Since each function $t\mapsto f_i(u(t))$ is nonincreasing, a natural condition insuring that the trajectory remains bounded (respectively relatively compact) is that one of the function f_i has bounded (respectively relatively compact) sublevel sets.

2 Existence of strong global solutions

In this section, we consider the abstract evolution equation

$$\dot{u}(t)+\Big(\partial\Phi(u(t))-B(u(t))\Big)^0=0. \tag{39}$$

As a particular case, (MCSD) is obtained by taking $\Phi=\delta_K$ the indicator function of K, and $B(v)=-\mathrm{Conv}\,\{\nabla f_i(v);\ i=1,2,...,q\}$ (the minus sign before B is introduced for convenience of mathematical analysis).

2.1 Statement of the results

We make the standing assumptions:

- $\mathcal{A}1$) $\mathcal{H}=\mathbb{R}^d$ is a finite dimensional Euclidian space.
- $\mathcal{A}2$) $\Phi:\mathcal{H}\to\mathbb{R}\cup\{+\infty\}$ is a closed convex proper function, and $\partial\Phi:\mathrm{dom}\partial\Phi\subset\mathcal{H}\to 2^{\mathcal{H}}$ is the subdifferential of Φ.
- $\mathcal{A}3$) $B:\mathcal{H}\to 2^{\mathcal{H}}$ is a continuous set-valued mapping with convex compact values, which satisfies the growth condition: there exists some $c>0$ such that

$$\text{for all } x\in\mathcal{H},\ \text{ and all } y\in B(x),\ \ \|y\|\le c(1+\|x\|). \tag{40}$$

The continuity of the set-valued mapping $x\mapsto B(x)$ is taken in the Hausdorff sense. It means that, whenever $x_n\to x$ then $\mathrm{haus}(B(x_n),B(x))\to 0$. Recall the definition of the Hausdorff distance.

Definition 2.1. *The Hausdorff distance between two closed convex subsets C and D of $\mathcal{H}$ is defined by*

$$\text{haus}(C,D) = \max\{e(C,D);\ e(D,C)\} \tag{41}$$

where $e(C,D) = \sup_{x\in C} d(x,D)$ is the excess of C on D, and $e(D,C) = \sup_{x\in D} d(x,C)$ is the excess of C on D. Equivalently (see [13, Lemma 1.5.1; Definition 3.2.1.])

$$\text{haus}(C,D) = \sup_{x\in\mathcal{H}} |d(x,C) - d(x,D|. \tag{42}$$

Assumptions $\mathcal{A}2$) and $\mathcal{A}3$) are satisfied in the case of the (MCSD) dynamic. The continuity of the set-valued mapping $B : v \in \mathcal{H} \mapsto B(v) = -\text{Conv}\{\nabla f_i(v);\ i = 1,2,...,q\} \in 2^{\mathcal{H}}$ is a consequence of the following

Lemma 2.2. *Set $L(R) = \max\{L_i(R);\ i = 1,2,...,q\}$, where $L_i(R)$ is the Lipschitz constant of ∇f_i on the ball $B(0,R)$. Then for any $u, v \in B(0,R)$*

$$\text{haus}(B(u),B(v)) \le L(R)\|u - v\|.$$

Proof. Let $\xi \in B(v)$, i.e., $\xi = -\sum_{i=1}^q \lambda_i \nabla f_i(v)$, with $\sum_{i=1}^q \lambda_i = 1,\ \lambda_i \ge 0$. Since $-\sum_{i=1}^q \lambda_i \nabla f_i(u) \in B(u)$ we infer

$$\begin{aligned} d(\xi, B(u)) &\le \|\xi + \sum_{i=1}^q \lambda_i \nabla f_i(u)\| \\ &\le \|\sum_{i=1}^q \lambda_i \nabla f_i(v) - \sum_{i=1}^q \lambda_i \nabla f_i(u)\| \\ &\le \sum_{i=1}^q \lambda_i \|\nabla f_i(v) - \nabla f_i(u)\| \\ &\le L(R)\|u - v\|. \end{aligned}$$

This being true for any $\xi \in B(v)$, it follows that $e(B(v),B(u)) \le L(R)\|u - v\|$, and from the symmetry of the role played by u and v, we obtain the result. □

Our proof of the existence of solutions to (39) relies on Peano, and not on Cauchy-Lipschitz theorem. This explains assumption $\mathcal{A}1$), and why in this section we restrict our analysis to the finite dimensional case.

In order to define the notion of strong solution, let us first recall some notions concerning vector-valued functions of real variables (see [18, Appendix] for more details).

Definition 2.3. *Given $T \in \mathbb{R}^+$, a function $f : [0,T] \longrightarrow \mathcal{H}$ is said to be absolutely continuous if one of the following equivalent properties holds:*

i) there exists an integrable function $g : [0,T] \longrightarrow \mathcal{H}$ such that

$$f(t) = f(0) + \int_0^t g(s)\,ds \qquad \forall t \in [0,T];$$

ii) f is continuous and its distributional derivative belongs to the Lebesgue space $L^1([0,T];\mathcal{H})$;

iii) for every $\epsilon > 0$, there exists some $\eta > 0$ such that for any finite family of intervals $I_k = (a_k, b_k)$

$I_k \cap I_j = \emptyset$ for $i \ne j$ and $\sum_k |b_k - a_k| \le \eta \Longrightarrow \sum_k \|f(b_k) - f(a_k)\| \le \epsilon$.

Moreover, an absolutely continuous function is differentiable almost everywhere, its derivative coincide with its distributional derivative almost everywhere, and one can recover the function from its derivative $f' = g$ using the integration formula (i).

Definition 2.4. *We say that $u(\cdot)$ is a strong global solution of (39) if the following properties are satisfied:*

(i) $u : [0, +\infty) \to \mathcal{H}$ is continuous, and absolutely continuous on each interval $[0, T]$, $0 < T < +\infty$;

(ii) there exists $v : [0, +\infty) \to \mathcal{H}$ and $w : [0, +\infty) \to \mathcal{H}$ which satisfy

$$v \in L^2(0, T; \mathcal{H}), \quad w \in L^2(0, T; \mathcal{H}) \quad \text{for all } T > 0; \tag{43}$$

$$v(t) \in \partial\Phi(u(t)), \ w(t) \in B(u(t)) \quad \text{for almost all } t > 0; \tag{44}$$

$$v(t) - w(t) = \Big(\partial\Phi(u(t)) - B(u(t))\Big)^0 \quad \text{for almost all } t > 0; \tag{45}$$

$$\dot{u}(t) + v(t) - w(t) = 0 \quad \text{for almost all } t > 0. \tag{46}$$

Theorem 2.5. *Let us make assumptions $\mathcal{A}1)$, $\mathcal{A}2)$, $\mathcal{A}3)$. Then, for any $u_0 \in dom\partial\Phi$ there exists a strong global solution $u : [0, +\infty) \to \mathcal{H}$ of the Cauchy problem*

$$\begin{cases} \dot{u}(t) + \Big(\partial\Phi(u(t)) - B(u(t))\Big)^0 = 0, \\ u(0) = u_0. \end{cases} \tag{47}$$

In the above theorem, we only claim existence. Without any further assumptions, uniqueness is not guarantied. As we already claimed, the following proof of existence relies on Peano, not on Cauchy-Lipschitz theorem.

2.2 Approximate equations

The main difficulty comes from the discontinuity of the multivalued operator $\partial\Phi$, and hence of the vector field which governs the differential equation (39). As a main ingredient of our approach, we use the Yosida approximation of the maximal monotone operator $\partial\Phi$. This regularization method is widely used in nonsmooth convex analysis, see [3], [9], [12], [18], [41] for a detailed presentation. Its main properties are summarized in the following

Proposition 2.6. *For any $\lambda > 0$, the Moreau-Yosida approximation of index λ of Φ is the function $\Phi_\lambda : \mathcal{H} \to \mathbb{R}$ which is defined for all $x \in \mathcal{H}$ by*

$$\Phi_\lambda(x) = \inf\left\{\Phi(\xi) + \frac{1}{2\lambda}\|x - \xi\|^2 : \quad \xi \in \mathcal{H}\right\}. \tag{48}$$

1. *The infimum in (48) is attained at a unique point $J_\lambda x \in \mathcal{H}$, which satisfies*

$$\Phi_\lambda(x) = \Phi(J_\lambda x) + \frac{1}{2\lambda}\|x - J_\lambda x\|^2; \tag{49}$$

$$J_\lambda x + \lambda\partial\Phi(J_\lambda x) \ni x. \tag{50}$$

$J_\lambda = (I + \lambda\partial\Phi)^{-1} : \mathcal{H} \to \mathcal{H}$ is everywhere defined and nonexpansive. It is is called the resolvent of index λ of $A = \partial\Phi$.

2. *Φ_λ is convex, and continuously differentiable. Its gradient at $x \in \mathcal{H}$ is equal to*

$$\nabla\Phi_\lambda(x) = \frac{1}{\lambda}(x - J_\lambda x). \tag{51}$$

3. *The operator $A_\lambda = \nabla\Phi_\lambda = \frac{1}{\lambda}(I - J_\lambda)$ is called the Yosida approximation of index λ of the maximal monotone operator $A = \partial\Phi$. It is Lipschitz continuous with Lipschitz constant $\frac{1}{\lambda}$.*

We are going to adapt to our situation the classical proof of the existence of strong solutions to evolution equations governed by maximal monotone operators, see [18]. For each $\lambda > 0$, we consider the Cauchy problem which is obtained by replacing in (47) $\partial\Phi$ by its Yosida approximation $\nabla\Phi_\lambda$. Firstly, we are going to show the existence of a global classical solution $u_\lambda : [0, +\infty) \to \mathcal{H}$ of the Cauchy problem

$$\begin{cases} \dot{u}_\lambda(t) + \Big(\nabla\Phi_\lambda(u_\lambda(t)) - B(u_\lambda(t))\Big)^0 = 0, \\ u_\lambda(0) = u_0. \end{cases} \tag{52}$$

We now use the following elementary result

Lemma 2.7. *Let C be a closed convex set in $\mathcal{H}$. Then for any $z \in \mathcal{H}$*

$$\Big(z - C\Big)^0 = z - \mathrm{proj}_C z.$$

Proof. We have

$$\Big(z - C\Big)^0 = \{z - x : \quad x \in C,\ \|z - x\| \le \|z - \xi\| \text{ for all } \xi \in C\}.$$

This uniquely defines $x = \mathrm{proj}_C z$, which gives the result. □

Taking $z = \nabla\Phi_\lambda(u_\lambda(t))$, and $C = B(u_\lambda(t))$ in Lemma 2.7, we can equivalently rewrite (52) as

$$\begin{cases} \dot{u}_\lambda(t) + \nabla\Phi_\lambda(u_\lambda(t)) - \mathrm{proj}_{B(u_\lambda(t))}\nabla\Phi_\lambda(u_\lambda(t)) = 0, \\ u_\lambda(0) = u_0. \end{cases} \tag{53}$$

To study (53), we need the following result on the Hölder-continuity of the mapping $C \mapsto \mathrm{proj}_C(x)$. The family of the closed convex bounded subsets C of $\mathcal{H}$ is equipped with the Hausdorff metric. The following classical result can be found in [1, Theorem 2], [7, Proposition 5.1], [36, Theorem 3.1].

Proposition 2.8. *Let C and D be two closed convex subsets of $\mathcal{H}$. Then, for any $x \in \mathcal{H}$ the mapping $C \mapsto \mathrm{proj}_C(x)$ is Hölder-continuous. More precisely, for any two closed convex subsets C and D of $\mathcal{H}$*

$$\|\mathrm{proj}_C(x) - \mathrm{proj}_D(x)\| \le \rho(\|x\|)\mathrm{haus}(C, D)^{\frac{1}{2}} \tag{54}$$

where $\rho(\|x\|) = (\|x\| + d(x, C) + d(x, D))^{\frac{1}{2}}$.

In the above result, the Hölder exponent $\frac{1}{2}$ is sharp, see [7, Proposition 5.1] and the accompanying example. We easily derive from these results the continuity of the vector field

$$F_\lambda : \ x \in \mathcal{H} \mapsto F_\lambda(x) = \nabla\Phi_\lambda(x) - \mathrm{proj}_{B(x)}\nabla\Phi_\lambda(x) \tag{55}$$

which governs (53). More precisely,

Lemma 2.9. *For any $x, y \in \mathcal{H}$, for any $\lambda > 0$,*

$$\|F_\lambda(x) - F_\lambda(y)\| \leq \frac{2}{\lambda}\|x - y\| + \rho(\|\nabla\Phi_\lambda(y)\|)\mathrm{haus}(B(x), B(y))^{\frac{1}{2}}. \tag{56}$$

Proof.

$$F_\lambda(x) - F_\lambda(y) = \nabla\Phi_\lambda(x) - \nabla\Phi_\lambda(y) - \Big(\mathrm{proj}_{B(x)}\nabla\Phi_\lambda(x) - \mathrm{proj}_{B(y)}\nabla\Phi_\lambda(y)\Big). \tag{57}$$

By the Lipschitz continuity property of $\nabla\Phi_\lambda$ (with Lipschitz constant $\frac{1}{\lambda}$) we obtain

$$\|F_\lambda(x) - F_\lambda(y)\| \leq \frac{1}{\lambda}\|x - y\| + \|\mathrm{proj}_{B(x)}\nabla\Phi_\lambda(x) - \mathrm{proj}_{B(y)}\nabla\Phi_\lambda(y)\|. \tag{58}$$

Let us set

$$B^\lambda(x) = \mathrm{proj}_{B(x)}\nabla\Phi_\lambda(x). \tag{59}$$

We have

$$\begin{aligned}
\|B^\lambda(x) - B^\lambda(y)\| &\leq \|\mathrm{proj}_{B(x)}\nabla\Phi_\lambda(x) - \mathrm{proj}_{B(x)}\nabla\Phi_\lambda(y)\| \\
&\quad + \|\mathrm{proj}_{B(x)}\nabla\Phi_\lambda(y) - \mathrm{proj}_{B(y)}\nabla\Phi_\lambda(y)\| \\
&\leq \|\nabla\Phi_\lambda(x) - \nabla\Phi_\lambda(y)\| + \|\mathrm{proj}_{B(x)}\nabla\Phi_\lambda(y) - \mathrm{proj}_{B(y)}\nabla\Phi_\lambda(y)\| \\
&\leq \frac{1}{\lambda}\|x - y\| + \rho(\|\nabla\Phi_\lambda(y)\|)\mathrm{haus}(B(x), B(y))^{\frac{1}{2}},
\end{aligned}$$

where we successively use the triangle inequality, the nonexpansive property of the projection mapping $x \mapsto \mathrm{proj}_C(x)$, and the Hölder-continuity of the mapping $C \mapsto \mathrm{proj}_C(x)$. Combining (58) and the above inequality we obtain (56). □

Even if the mapping $x \mapsto B(x)$ is Lipschitz continuous with respect to the Hausdorff metric, Lemma 2.9 tells us that the vector field F_λ which governs (53) is only locally Hölder continuous with exponent $\frac{1}{2}$. Thus we cannot apply Cauchy-Lipschitz theorem. Instead, the classical assumptions of Peano theorem are satisfied:

i) By Lemma 2.9 and assumption $\mathcal{A}3$), the vector field F_λ is continuous.

ii) The space $\mathcal{H}$ has been assumed to be finite dimensional.

Hence, there exists of a local solution u_λ to the Cauchy problem (53) (it is not necessarily unique). In order to prove global existence, we first establish energy estimates: Let us write the differential equation (53) as

$$\dot{u}_\lambda(t) + \nabla\Phi_\lambda(u_\lambda(t)) - B^\lambda(u_\lambda(t)) = 0. \tag{60}$$

By definition (59) of B^λ and assumption $\mathcal{A}3$), for any $x \in \mathcal{H}$

$$\|B^\lambda(x)\| \leq \sup\{\|y\| : \quad y \in B(x)\} \tag{61}$$

$$\leq c(1 + \|x\|). \tag{62}$$

Lemma 2.10. *Let u_λ be a solution of the Cauchy problem* (53) *on some interval* $[0,T]$. *Let x_0 be a given element in dom$\partial\Phi$. Then, for any $t \in [0,T]$*

$$\|u_\lambda(t) - x_0\| \leq e^{\left(\frac{1}{2}+c\right)t}\left(\|u_0 - x_0\| + \|\partial\Phi^0(x_0)\| + c\|x_0\| + c\right); \tag{63}$$

$$\frac{1}{2}\int_0^T \|\dot{u}_\lambda(\tau)\|^2 d\tau + \Phi_\lambda(u_\lambda(T)) \leq \Phi(u_0) + \frac{c^2}{2}\int_0^T (1 + \|u_\lambda(\tau)\|^2) d\tau. \tag{64}$$

Proof. By taking the scalar product with $u_\lambda(t) - x_0$ in (60) we obtain

$$\frac{1}{2}\frac{d}{dt}\|u_\lambda(t) - x_0\|^2 + \langle \nabla\Phi_\lambda(u_\lambda(t)), u_\lambda(t) - x_0\rangle - \left\langle B^\lambda(u_\lambda(t)), u_\lambda(t) - x_0 \right\rangle = 0. \tag{65}$$

By using the monotonicity property of $\nabla\Phi_\lambda$

$$\langle \nabla\Phi_\lambda(u_\lambda(t)) - \nabla\Phi_\lambda(x_0), u_\lambda(t) - x_0\rangle \geq 0, \tag{66}$$

and Cauchy-Schwarz inequality in (65), we deduce that

$$\frac{1}{2}\frac{d}{dt}\|u_\lambda(t) - x_0\|^2 \leq \|\nabla\Phi_\lambda(x_0)\|\|u_\lambda(t) - x_0\| + \|B^\lambda(u_\lambda(t))\|\|u_\lambda(t) - x_0\|. \tag{67}$$

By using the majorization $\|\nabla\Phi_\lambda(x_0)\| \leq \|\partial\Phi^0(x_0)\|$ ([18, Proposition 2.6]), where $\partial\Phi^0(x_0)$ denotes the element of minimal norm of the nonempty closed convex set $\partial\Phi(x_0)$, and the growth property (61) on B^λ, we obtain

$$\frac{1}{2}\frac{d}{dt}\|u_\lambda(t) - x_0\|^2 \leq \left(\|\partial\Phi^0(x_0)\| + c\|x_0\| + c\right)\|u_\lambda(t) - x_0\| + c\|u_\lambda(t) - x_0\|^2. \tag{68}$$

By integration of the above inequality (Gronwall type argument) and elementary majorization we obtain

$$\|u_\lambda(t) - x_0\| \leq e^{\left(\frac{1}{2}+c\right)t}\left(\|u_0 - x_0\| + \|\partial\Phi^0(x_0)\| + c\|x_0\| + c\right).$$

Let us now estimate $\dot{u}_\lambda$. By taking the scalar product with $\dot{u}_\lambda$ in (60) we obtain

$$\|\dot{u}_\lambda(t)\|^2 + \frac{d}{dt}\Phi_\lambda(u_\lambda(t)) = \left\langle B^\lambda(u_\lambda(t)), \dot{u}_\lambda(t)\right\rangle. \tag{69}$$

After elementary majorization, and by integration on $[0,T]$ we obtain

$$\frac{1}{2}\int_0^T \|\dot{u}_\lambda(\tau)\|^2 d\tau + \Phi_\lambda(u_\lambda(T)) \leq \Phi(u_0) + \frac{c^2}{2}\int_0^T (1 + \|u_\lambda(\tau)\|^2) d\tau,$$

where we have used that $\Phi_\lambda(u_0) \leq \Phi(u_0)$ ([18, Proposition 2.11]). □

We have all the ingredients for the global existence of solutions for approximated equations (53).

Proposition 2.11. *For any $\lambda > 0$, and any $u_0 \in$ dom$\partial\Phi$ there exists a classical global solution $u_\lambda : [0,+\infty) \to \mathcal{H}$ of the Cauchy problem*

$$\begin{cases} \dot{u}_\lambda(t) + \nabla\Phi_\lambda(u_\lambda(t)) - \mathrm{proj}_{B(u_\lambda(t))}\nabla\Phi_\lambda(u_\lambda(t)) = 0, \\ u_\lambda(0) = u_0. \end{cases} \tag{70}$$

Proof. From estimations (63) and (64), by a standard argument, let us show the existence of a global solution u_λ to the Cauchy problem (53), which is defined on the whole interval $[0,+\infty[$. By Zorn's lemma, let u_λ be a maximal solution, and suppose that it is defined on a bounded interval $[0,T_{max}[$ with $T_{max}<+\infty$. Since $T_{max}<+\infty$, by (63) we have that the trajectory remains bounded on $[0,T_{max}[$. Using this estimation in (64) we deduce that $\int_0^{T_{max}}\|\dot{u}_\lambda(t)\|^2dt<+\infty$. Note that $\Phi_\lambda(u_\lambda(t))$ remains bounded, because Φ_λ is continuous and $u_\lambda(t)$ remains bounded. Since $\int_0^{T_{max}}\|\dot{u}_\lambda(t)\|^2dt<+\infty$ and $T_{max}<+\infty$ we easily obtain, by Cauchy-Schwarz inequality, that $\lim_{t\to T_{max}}u_\lambda(t):=z$ exists. By applying again the local existence result with z as a new Cauchy data at time T_{max}, we obtain the existence of a solution which is defined on an interval which is strictly larger than T_{max}, a clear contradiction. □

2.3 Estimations on the sequence (u_λ)

Let us prove that the preceding estimations on the sequence $(u_\lambda)_\lambda$ are uniform with respect to λ. By (63), for any $0<T<\infty$

$$\sup_{\lambda>0}\|u_\lambda\|_{L^\infty(0,T;\mathcal{H})}\leq\|x_0\|+e^{\left(\frac{1}{2}+c\right)T}\left(\|u_0-x_0\|+\|\partial\Phi^0(x_0)\|+c\|x_0\|+c\right). \tag{71}$$

Hence for any $0<T<\infty$

$$\sup_{\lambda>0}\|u_\lambda\|_{L^\infty(0,T;\mathcal{H})}<+\infty. \tag{72}$$

On the other hand, by (64)

$$\frac{1}{2}\int_0^T\|\dot{u}_\lambda(t)\|^2dt+\Phi_\lambda(u_\lambda(T))\leq\Phi(u_0)+\frac{Tc^2}{2}(1+\|u_\lambda\|^2_{L^\infty(0,T;\mathcal{H})}). \tag{73}$$

Let us minorize $\Phi_\lambda(u_\lambda(t))$. Let take some $y_0\in\partial\Phi(x_0)$. We have $x_0=J_\lambda(x_0+\lambda y_0)$. By (49), which is a direct consequence of the definition of Φ_λ, and the convex subdifferential inequality at x_0

$$\begin{aligned}\Phi_\lambda(u_\lambda(t))&\geq\Phi(J_\lambda(u_\lambda(t)),&(74)\\&\geq\Phi(x_0)+\langle y_0,J_\lambda(u_\lambda(t))-x_0\rangle&(75)\\&\geq\Phi(x_0)-\|y_0\|\|J_\lambda(u_\lambda(t))-x_0\|.&(76)\end{aligned}$$

Since J_λ is nonexpansive, and $x_0=J_\lambda(x_0+\lambda y_0)$, we have

$$\begin{aligned}\|J_\lambda(u_\lambda(t))-x_0\|&=\|J_\lambda(u_\lambda(t))-J_\lambda(x_0+\lambda y_0)\|&(77)\\&\leq\|u_\lambda(t)-(x_0+\lambda y_0)\|.&(78)\end{aligned}$$

Combining (74) and (77) we obtain

$$\Phi_\lambda(u_\lambda(t))\geq\Phi(x_0)-\|y_0\|\left(\|u_\lambda(t)-(x_0+\lambda y_0)\|\right). \tag{79}$$

Combining (73) and (79) we obtain

$$\frac{1}{2}\int_0^T\|\dot{u}_\lambda(t)\|^2dt\leq\Phi(u_0)-\Phi(x_0)+\frac{Tc^2}{2}(1+\|u_\lambda\|^2_{L^\infty(0,T)})+\|y_0\|\left(\|u_\lambda\|_{L^\infty(0,T)}+\|x_0+\lambda y_0\|\right). \tag{80}$$

By (72) we obtain

$$\sup_{\lambda>0} \|\dot{u}_\lambda\|_{L^2(0,T;\mathcal{H})} < +\infty. \tag{81}$$

Returning to (73) we obtain

$$\sup_{\lambda>0,\ 0<t<T} \Phi_\lambda(u_\lambda(t)) < +\infty. \tag{82}$$

Finally, by (60) and the growth condition (61)

$$\|\nabla\Phi_\lambda(u_\lambda(t))\| \le \|B^\lambda(u_\lambda(t))\| + \|\dot{u}_\lambda(t)\| \tag{83}$$

$$\le \|\dot{u}_\lambda(t)\| + c(1 + \|u_\lambda(t)\|), \tag{84}$$

which, by (72)) and (81) gives

$$\sup_{\lambda>0} \|\nabla\Phi_\lambda(u_\lambda)\|_{L^2(0,T;\mathcal{H})} < +\infty. \tag{85}$$

Of course, in (72), (81), (82) and (85) the bounds depend on T.

2.4 Passing to the limit, $\lambda \to 0$

As we have already pointed out, the difficulty comes from the discontinuous nature of the multivalued operator $\partial\Phi$, and hence of the vector field which governs the differential equation (39). Indeed, we are going to play on the monotonicity property of some components of this vector field, and hence of the closedness property of their graphs for the strong×weak topology (demiclosedness property).

By estimations (72),(81), and the fact that $\mathcal{H}$ is finite dimensional, we deduce from Ascoli's theorem that, for any $0 < T < +\infty$, the generalized sequence (u_λ) is relatively compact for the uniform convergence topology on $[0,T]$. Thus, by a diagonal argument (we keep the notation (u_λ) for simplicity) we get the existence of $u \in \mathcal{C}(0,+\infty;\mathcal{H})$, and $v, w \in L^2_{loc}(0,+\infty;\mathcal{H})$ such that, for any $0 < T < +\infty$,

$$u_\lambda \to u \qquad \text{strong} - \mathcal{C}(0,T;\mathcal{H}) \tag{86}$$

$$\dot{u}_\lambda \rightharpoonup \dot{u} \qquad \text{weak} - L^2(0,T;\mathcal{H}) \tag{87}$$

$$\nabla\Phi_\lambda(u_\lambda) \rightharpoonup v \qquad \text{weak} - L^2(0,T;\mathcal{H}) \tag{88}$$

$$B^\lambda(u_\lambda) \rightharpoonup w \qquad \text{weak} - L^2(0,T;\mathcal{H}). \tag{89}$$

By passing to the limit on (60) we obtain

$$\dot{u} + v - w = 0. \tag{90}$$

Let us identify v and w. From now on we argue on $[0,T]$ with an arbitrary fixed $0 < T < +\infty$.

a) By (51), we have $\nabla\Phi_\lambda(x) = \frac{1}{\lambda}(x - J_\lambda x)$. Hence

$$\|u_\lambda - J_\lambda(u_\lambda)\|_{L^2(0,T;\mathcal{H})} \le \lambda \|\nabla\Phi_\lambda(u_\lambda)\|_{L^2(0,T;\mathcal{H})}. \tag{91}$$

By (85), the generalized sequence $(\nabla\Phi_\lambda(u_\lambda))$ is bounded in $L^2(0,T;\mathcal{H})$. Hence

$$\lim_{\lambda\to 0}(u_\lambda - J_\lambda(u_\lambda)) = 0 \ \text{ in } L^2(0,T;\mathcal{H}). \tag{92}$$

By (86), $u_\lambda \to u$ strongly in $\mathcal{C}(0,T;\mathcal{H})$, and hence strongly in $L^2(0,T;\mathcal{H})$. Hence

$$J_\lambda(u_\lambda) \to u \ \text{ strongly in } L^2(0,T;\mathcal{H}). \tag{93}$$

Let us summarize the previous results:

$$\nabla\Phi_\lambda(u_\lambda) \in \partial\Phi(J_\lambda(u_\lambda)), \tag{94}$$

$$J_\lambda(u_\lambda) \to u \ \text{ strong} - L^2(0,T;\mathcal{H}), \tag{95}$$

$$\nabla\Phi_\lambda(u_\lambda) \rightharpoonup v \ \text{ weak} - L^2(0,T;\mathcal{H}). \tag{96}$$

Let us consider the canonical extension to $L^2(0,T;\mathcal{H})$ of the maximal monotone operator $\partial\Phi$:

$$\mathbb{A} = \big\{(\xi,\eta) \in L^2(0,T;\mathcal{H}) \times L^2(0,T;\mathcal{H}) : \quad \eta(t) \in \partial\Phi(\xi(t)) \ \ a.e.\ t \in (0,T).\big\} \tag{97}$$

Classically, $\mathbb{A}$ is maximal monotone in $L^2(0,T;\mathcal{H})$. Indeed, it is the subdifferential of the canonical extension of Φ to $L^2(0,T;\mathcal{H})$, see [18, Proposition 2.16]. As a consequence, it is closed for the strong-$L^2(0,T;\mathcal{H})\times$weak-$L^2(0,T;\mathcal{H})$ topology. From (94)-(95)-(96) we deduce that

$$v(t) \in \partial\Phi(u(t) \ \ a.e.\ t \in (0,T), \tag{98}$$

which, combined with (90) gives

$$\dot{u}(t) + \partial\Phi(u(t)) - w(t) \ni 0 \ \ a.e.\ t \in (0,T). \tag{99}$$

Since Φ is a closed convex function, its subdifferential is maximal monotone. As a consequence, given $w \in L^2(0,T;\mathcal{H})$, the unique solution of the above differential inclusion is the lazy solution, see [9, Theorem 1, ch.6, sec.8], [18, Remark 3.9], which means

$$\dot{u}(t) + \Big(\partial\Phi(u(t)) - w(t)\Big)^0 = 0 \ \ a.e.\ t \in (0,T). \tag{100}$$

Equivalently

$$\dot{u}(t) + \mathrm{proj}_{\partial\Phi(u(t))}w(t) - w(t) = 0 \ \ a.e.\ t \in (0,T). \tag{101}$$

Comparing (101) with (90) we obtain

$$v(t) = \mathrm{proj}_{\partial\Phi(u(t))}w(t) \ \ a.e.\ t \in (0,T). \tag{102}$$

b) Let us now introduce the operator $\mathcal{B} : L^2(0,T;\mathcal{H}) \to L^2(0,T;\mathcal{H})$ which is the canonical extension to $L^2(0,T;\mathcal{H})$ of the multivalued operator B:

$$\mathcal{B} = \big\{(\xi,\eta) \in L^2(0,T;\mathcal{H}) \times L^2(0,T;\mathcal{H}) : \quad \eta(t) \in B(\xi(t)) \ \ a.e.\ t \in (0,T).\big\} \tag{103}$$

The operator $\mathcal{B}$ is demiclosed in $L^2(0,T;\mathcal{H}) \times L^2(0,T;\mathcal{H})$, (see [5, Proposition 3.4]). Since

$$\begin{aligned} &B^\lambda(u_\lambda) \in \mathcal{B}(u_\lambda) \\ &u_\lambda \to u \ \text{ strong} - L^2(0,T;\mathcal{H}) \\ &B^\lambda(u_\lambda) \rightharpoonup w \ \text{ weak} - L^2(0,T;\mathcal{H}) \end{aligned} \qquad .$$

we deduce that

$$w(t) \in B(u(t)) \quad a.e.\ t \in (0,T). \tag{104}$$

c) Suppose a moment that we have proved

$$w(t) = \mathrm{proj}_{B(u(t))} v(t) \quad a.e.\ t \in (0,T). \tag{105}$$

Putting together (102) and (105), we have

$$v(t) = \mathrm{proj}_{\partial\Phi(u(t))} w(t) \quad a.e.\ t \in (0,T), \tag{106}$$

$$w(t) = \mathrm{proj}_{B(u(t))} v(t) \quad a.e.\ t \in (0,T), \tag{107}$$

and the proof can be easily completed thanks to the following lemma.

Lemma 2.12. *Let C and D be closed convex sets in $\mathcal{H}$. Suppose that one of the two sets is bounded. Then their vectorial difference $C-D$ is a closed convex set. Its unique element of minimal norm $(C-D)^0$ is characterized by*

$$(C-D)^0 = v - w : \quad v = \mathrm{proj}_C w \quad \textit{and} \quad w = \mathrm{proj}_D v. \tag{108}$$

Proof. The element of minimal norm of $C-D$ is equal to $v-w$, with (v,w) solution of the convex minimization problem

$$\min \left\{ \|v-w\|^2 + \delta_C(v) + \delta_D(w) : \ v \in \mathcal{H},\ w \in \mathcal{H} \right\}. \tag{109}$$

Writing down the necessary and sufficient optimality conditions of this convex minimization problem gives the result. □

Note that, in the above writing, v and w are not necessarily unique. The important point is that the property $(v = \mathrm{proj}_C w,\ w = \mathrm{proj}_D v)$ implies that $(C-D)^0 = v - w$. By taking $v = v(t)$, $w = w(t)$, $C = \partial\Phi(u(t))$, and $D = B(u(t))$, (106)-(107) tells us that we are in the situation described in Lemma 2.12. Hence,

$$v(t) - w(t) = (\partial\Phi(u(t)) - B(u(t)))^0 \quad a.e.\ t \in (0,T) \tag{110}$$

which combined with (90) gives

$$\dot{u}(t) + (\partial\Phi(u(t)) - B(u(t))^0 = 0 \quad a.e.\ t \in (0,T) \tag{111}$$

the desired result.

d) Thus, the last point we have to prove in order to complete the proof of Theorem 2.5 is

$$w(t) = \mathrm{proj}_{B(u(t))} v(t) \quad a.e.\ t \in (0,T). \tag{112}$$

To that end, we use the following demiclosedness property (see [12, Theorem 4.17]).

Lemma 2.13. *Let C be a closed convex set in $\mathcal{H}$. Then, the operator $I - \mathrm{proj}_C$ is firmly nonexpansive. In particular, it is monotone, Lipschitz continuous, hence maximal monotone. As a consequence, it is demiclosed, i.e., its graph is closed for the strong-$L^2(0,T;\mathcal{H})$×weak-$L^2(0,T;\mathcal{H})$ topology.*

By Proposition 2.8

$$\|\mathrm{proj}_{B(u_\lambda(t))}\nabla\Phi_\lambda(u_\lambda(t))-\mathrm{proj}_{B(u(t))}\nabla\Phi_\lambda(u_\lambda(t))\| \tag{113}$$

$$\leq \rho(\|\nabla\Phi_\lambda(u_\lambda(t))\|)\mathrm{haus}(B(u_\lambda(t)),B(u(t)))^{\frac{1}{2}} \tag{114}$$

where $\rho(\|x\|) = (\|x\| + d(x,B(u_\lambda(t))) + d(x,B(u(t))))^{\frac{1}{2}}$. Since $u_\lambda \to u$ $\mathcal{C}(0,T;\mathcal{H})$, and B is continuous for the Hausdorff metric, it follows from (113) and the Lebesgue dominated convergence theorem that

$$\mathrm{proj}_{B(u_\lambda(t))}\nabla\Phi_\lambda(u_\lambda(t)) - \mathrm{proj}_{B(u(t))}\nabla\Phi_\lambda(u_\lambda(t)) \to 0 \quad L^2(0,T;\mathcal{H}). \tag{115}$$

Thus we can rewrite (53) as

$$\dot{u}_\lambda(t) + \nabla\Phi_\lambda(u_\lambda(t)) - \mathrm{proj}_{B(u(t))}\nabla\Phi_\lambda(u_\lambda(t)) = \epsilon_\lambda(t), \tag{116}$$

with $\epsilon_\lambda \to 0$ strongly in $L^2(0,T;\mathcal{H})$ as $\lambda \to 0$.
For each $t > 0$, let us introduce the operator $M(t) : \mathcal{H} \to \mathcal{H}$

$$M(t) = I - \mathrm{proj}_{B(u(t))} \tag{117}$$

i.e., for every $x \in \mathcal{H}$, $M(t)(x) = x - \mathrm{proj}_{B(u(t))}(x)$. By Lemma 2.13, it is maximal monotone. One can easily verify that its canonical extension to $L^2(0,T;\mathcal{H})$

$$\mathcal{M} = \big\{(u,v) \in L^2(0,T;\mathcal{H}) \times L^2(0,T;\mathcal{H}) : \quad v(t) = M(t)(u(t)) \ \ a.e.\ t \in (0,T)\big\} \tag{118}$$

is still maximal monotone. Let us reformulate our results with the help of $\mathcal{M}$.

$$\epsilon_\lambda - \dot{u}_\lambda = \mathcal{M}(\nabla\Phi_\lambda(u_\lambda)) \tag{119}$$

$$\epsilon_\lambda - \dot{u}_\lambda \to -\dot{u} \ \ \mathrm{weak} - L^2(0,T;\mathcal{H}), \tag{120}$$

$$\nabla\Phi_\lambda(u_\lambda) \rightharpoonup v \ \ \mathrm{weak} - L^2(0,T;\mathcal{H}). \tag{121}$$

Suppose for a moment that we have been able to prove that $\dot{u}_\lambda \to \dot{u}$ strongly in $L^2(0,T;\mathcal{H})$. Then $\epsilon_\lambda - \dot{u}_\lambda \to -\dot{u}$ strong $- L^2(0,T;\mathcal{H})$ and, by the demiclosedness property of $\mathcal{M}$ we obtain

$$\dot{u}(t) + M(t)(v(t)) = 0 \ \ a.e.\ t \in (0,T) \tag{122}$$

that is

$$\dot{u}(t) + v(t) - \mathrm{proj}_{B(u(t))}(v(t)) = 0 \ \ a.e.\ t \in (0,T). \tag{123}$$

Comparing with (90), we obtain $w(t) = \mathrm{proj}_{B(u(t))}(v(t))$, which is the desired result (recall (112)). Thus, in order to complete the proof, we need to show that the filtered sequence of derivatives $(\dot{u}_\lambda)$ converges strongly in $L^2(0,T;\mathcal{H})$. This is the last step of the proof, this result is not only technical, it has its own interest.

2.5 Strong L^2 convergence of the sequence $(\dot{u}_\lambda)$

The idea is to prove that

$$\int_0^T \|\dot{u}_\lambda(t)\|^2 dt \to \int_0^T \|\dot{u}(t)\|^2 dt. \tag{124}$$

This property and the weak-L^2 convergence of the sequence $(\dot{u}_\lambda)$ will clearly imply the strong convergence of the sequence. To this end, we follow an argument which is in line with that developed in [2, Theorem 2.1] and [10]. We use the following elementary argument.

Lemma 2.14. *Let $(a_{n,1})_{n\in\mathbb{N}}, (a_{n,2})_{n\in\mathbb{N}}, ..., (a_{n,l})_{n\in\mathbb{N}}$ be a finite family of sequences of real numbers which satisfy:*

$$\begin{aligned}
&\sum_{k=1}^{l} a_{n,k} \leq 0 \quad \textit{for each } n \in \mathbb{N};\\
&a_k \leq \liminf_n a_{n,k} \quad \textit{for each } k = 1, 2, ..., l;\\
&\sum_{k=1}^{l} a_k = 0.
\end{aligned}$$

Then, $a_{n,k} \to a_k$ for each $k = 1, 2, ..., l$.

By (104) we have

$$w(t) \in B(u(t)) \quad a.e.\ t \in (0,T). \tag{125}$$

Since $u_\lambda \to u$ strongly in L^2, and B is a continuous multivalued mapping, by taking

$$z_\lambda(t) = \mathrm{proj}_{B(u_\lambda(t))} w(t) \quad a.e.\ t \in (0,T). \tag{126}$$

we obtain

$$z_\lambda \to \mathrm{proj}_{B(u)} w = w \quad \text{strongly in} \quad L^2(0,T;\mathcal{H}). \tag{127}$$

By definition (59) of $B^\lambda(u_\lambda(t)) = \mathrm{proj}_{B(u_\lambda(t))} \nabla\Phi_\lambda(u_\lambda(t))$, $z_\lambda(t) \in B(u_\lambda(t))$ and the obtuse-angle property

$$\Big\langle \nabla\Phi_\lambda(u_\lambda(t)) - B^\lambda(u_\lambda(t)), z_\lambda(t) - B^\lambda(u_\lambda(t)) \Big\rangle \leq 0. \tag{128}$$

By (60)

$$\nabla\Phi_\lambda(u_\lambda(t)) - B^\lambda(u_\lambda(t)) = -\dot{u}_\lambda(t). \tag{129}$$

Hence

$$\Big\langle -\dot{u}_\lambda(t), z_\lambda(t) - B^\lambda(u_\lambda(t)) \Big\rangle \leq 0. \tag{130}$$

On the other hand, after scalar multiplication of (60) by $\dot{u}_\lambda(t)$, and integration on $(0,T)$, we obtain

$$\int_0^T \|\dot{u}_\lambda(t)\|^2 dt + \Phi_\lambda(u_\lambda(T)) - \Phi_\lambda(u_0) - \int_0^T \Big\langle \dot{u}_\lambda(t), B^\lambda(u_\lambda(t)) \Big\rangle = 0. \tag{131}$$

By combining (130) and (131) we obtain

$$\int_0^T \|\dot{u}_\lambda(t)\|^2 dt + \Phi_\lambda(u_\lambda(T)) - \Phi_\lambda(u_0) - \int_0^T \langle \dot{u}_\lambda(t), z_\lambda(t)\rangle \leq 0. \tag{132}$$

We are now in position to apply Lemma 2.14. By using the lower semicontinuity for the weak topology of the square of the norm, and the respective weak/strong convergence properties of $\dot{u}_\lambda$ and z_λ, we have

$$\begin{cases} \int_0^T \|\dot{u}(t)\|^2 dt \leq \liminf \int_0^T \|\dot{u}_\lambda(t)\|^2 dt; \\ \Phi(u(T)) - \Phi(u_0) \leq \liminf \left(\Phi_\lambda(u_\lambda(T)) - \Phi_\lambda(u_0)\right); \\ -\int_0^T \langle \dot{u}(t), w(t)\rangle = \lim \left(-\int_0^T \langle \dot{u}_\lambda(t), z_\lambda(t)\rangle\right). \end{cases} \tag{133}$$

On the other hand, by (99)

$$\dot{u}(t) + v(t) - w(t) = 0 \quad a.e.\ t \in (0,T), \tag{134}$$

with $v(t) \in \partial\Phi(u(t))$ $a.e.$ $t \in (0,T)$ and $v \in L^2(0,T;\mathcal{H})$.
After scalar multiplication of (134) by $\dot{u}(t)$, and integration on $(0,T)$, and by using the generalized derivation chain rule (see [18, Lemma 3.3]) we have

$$\int_0^T \|\dot{u}(t)\|^2 dt + \Phi(u(T)) - \Phi(u_0) - \int_0^T \langle \dot{u}(t), w(t)\rangle \leq 0. \tag{135}$$

By combining (133) with (134), and applying Lemma 2.14 we finally obtain

$$\int_0^T \|\dot{u}_\lambda(t)\|^2 dt \to \int_0^T \|\dot{u}(t)\|^2 dt.$$

which ends the proof of Theorem 2.5.

3 Multiobjective gradient processes and steepest descent

3.1 Multiobjective gradient processes

Following Smale [38], let us define a notion of gradient process for a multiobjective optimization problem

$$(P) \quad \min\{F(v) :\ v \in K\} \tag{136}$$

with $F : v \in \mathcal{H} \to (f_i(v))_{i=1,2,\dots,q} \in \mathbb{R}^q$, $K \subset \mathcal{H}$. Let us recall (see Definition 1.2) that

$$\mathcal{P}_c = \left\{ v \in \mathcal{H} :\ \text{there exist } 0 \leq \lambda_i \leq 1,\ \sum_{i=1}^q \lambda_i = 1 \quad \sum_{i=1}^q \lambda_i \nabla f_i(v) \ + N_K(v) \ni 0 \right\}$$

is the set of Pareto critical points. For all $u \in K$, set

$$C(u) = \begin{cases} \{v \in T_K(u) :\ \langle \nabla f_i(u), v\rangle < 0 \text{ for all } i = 1,2,...,q\} & \text{if } u \notin \mathcal{P}_c \\ \{0\} \quad \text{if } u \in \mathcal{P}_c \end{cases} \tag{137}$$

where $T_K(u)$ is the closed convex tangent cone to K at $u \in K$.

Definition 3.1. *A gradient process for the multiobjective optimization (P) is a differential equation*

$$\dot{u}(t) = \phi(u(t)) \tag{138}$$

where $\phi : K \to \mathcal{H}$ is a mapping such that $\phi(u) \in C(u)$ for all $u \in K$.

In [38], the vector field ϕ is continuous, and K is a manifold in $\mathbb{R}^q$. In the above definition, these assumptions have been extended in order to cover the (MCSD) dynamic, which involves a discontinuous vector field on a general Hilbert space.
In what follows we will make frequent use of the Moreau decomposition theorem [34]:

Theorem 3.2. *(Moreau) Let T be a closed convex cone of a real Hilbert space $\mathcal{H}$, and N be the polar cone, i.e., $N = \{v \in \mathcal{H} : \langle v, \xi\rangle \leq 0$ for all $\xi \in T\}$. Then, for all $v \in \mathcal{H}$ there exists a unique decomposition*

$$\begin{aligned} &v = v_T + v_N, \quad v_T \in T,\ v_N \in N; \\ &\langle v_T, v_N\rangle = 0. \end{aligned}$$

Moreover, $v_T = \mathrm{proj}_T(v)$, and $v_N = \mathrm{proj}_N(v)$.

Proposition 3.3. *The (MCSD) dynamic is a multiobjective gradient process. It is governed by the vector field $u \mapsto s(u)$, where*

$$s(u) = \Big(-N_K(u) - \mathrm{Conv}\,\{\nabla f_i(u);\ i = 1, 2, ..., q\} \Big)^0$$

is called the multiobjective steepest descent direction at $u \in K$.

Proof. We follow an argument which is similar to the proof of the descent property in Theorem 1.5. By definition of $s(u)$, $s(u) = \Big(z - N_K(u)\Big)^0$ for some $z \in -\mathrm{Conv}\,\{\nabla f_i(u);\ i = 1, 2, ..., q\}$. By Moreau decomposition theorem,

$$\Big(z - N_K(u)\Big)^0 = z - \mathrm{proj}_{N_K(u)} z \tag{139}$$

$$= \mathrm{proj}_{T_K(u)} z \tag{140}$$

and hence $s(u) \in T_K(u)$.
Let us now show that $v = s(u)$ is a descent direction, i.e.,

$$\langle \nabla f_i(u), s(u)\rangle < 0 \text{ for all } i = 1, 2, ..., q \quad \text{ if } u \notin \mathcal{P}_c. \tag{141}$$

By definition, $-s(u)$ is the projection of the origin onto the closed convex set $C := N_K(u) + \mathrm{Conv}\,\{\nabla f_i(u)\}$. Hence, for any $\xi \in C$

$$\langle 0 - (-s(u)), \xi - (-s(u))\rangle \leq 0, \tag{142}$$

that is

$$\langle s(u), \xi + s(u)\rangle \leq 0. \tag{143}$$

Noticing that $0 \in N_K(u(t))$ and $\nabla f_i(u) \in \mathrm{Conv}\{\nabla f_i(u)\}$, we can take $\xi = \nabla f_i(u)$ in (143), which yields

$$\|s(u)\|^2 + \langle s(u), \nabla f_i(u)\rangle \leq 0. \tag{144}$$

Hence as long as $s(u) \neq 0$ we have, for all $i = 1, 2, ..., q$

$$\langle s(u), \nabla f_i(u)\rangle < 0.$$

Let us now notice that $s(u) \neq 0$ is equivalent to say that $0 \notin N_K(u) + \mathrm{Conv}\{\nabla f_i(u))\}$, which is equivalent to say that u is not a critical point. This proves (141), which completes the proof. □

Remark 3.4. The above result has a simple geometrical interpretation. Take for simplicity the unconstrained problem, i.e., $K = \mathcal{H}$ and two criteria f_1, f_2. Then $-s(u)$ is the orthogonal projection of the origin on the vectorial segment $[\nabla f_1(u), \nabla f_2(u)]$. By the classical result on the sum of the angles of a triangle, this forces the angles bewteen $-s(u)$ and $\nabla f_i(u)$, $i = 1, 2$ to be accute and hence $\langle s(u), \nabla f_i(u)\rangle \leq 0$.

3.2 The multiobjective steepest descent direction

In the case of a single criteria f, and a constraint K, the steepest descent direction $s(u)$ at $u \in K$ is given by

$$s(u) = \Big(- N_K(u) - \nabla f(u)) \Big)^0, \tag{145}$$

$$= \mathrm{proj}_{T_K(u)}(-\nabla f(u)). \tag{146}$$

It is associated to the lazy solution property (see [18, Theorem 3.2]) of the trajectories of the differential inclusion

$$\dot{u}(t) + N_K(u(t)) + \nabla f(u(t)) \ni 0.$$

Equivalently $\psi(u) = \frac{s(u)}{\|s(u)\|}$ is the solution of the minimization problem:

$$\min\left\{\langle \nabla f(u), v\rangle : \quad v \in T_K(u),\ \|v\| = 1\right\}.$$

The equivalence between these two formulations of the direction of steepest descent can be generalized to the case of multiple objectives as follows:

Theorem 3.5. *For any $u \in K$, $u \notin \mathcal{P}_c$, the minimization problem*

$$\min_{v \in T_K(u),\ \|v\|=1} \ \max_{i=1,2,...,q} \langle \nabla f_i(u), v\rangle \tag{147}$$

admits a unique solution $\psi(u)$ which is related to the multiobjective steepest descent direction $s(u)$ by

$$\psi(u) = \frac{s(u)}{\|s(u)\|}.$$

Proof. Our proof is an extension to the case of a Hilbert space of the proof of [21, Proposition 3.1]. In all that follows, u is a fixed element of K, with $u \notin \mathcal{P}_c$. By a positive homogeneity argument, (147) can be reformulated as

$$\min_{v \in \mathcal{K}} \max_{i=1,2,\dots,q} \langle \nabla f_i(u), v \rangle \,, \tag{148}$$

where

$$\mathcal{K} = \{ v \in T_K(u),\ \|v\| \leq 1 \} \,. \tag{149}$$

Since $\mathcal{K}$ is a closed convex bounded set in $\mathcal{H}$, and $v \mapsto \max_{i=1,2,\dots,q} \langle \nabla f_i(u), v \rangle$ is a convex continuous function, (148) admits at least a solution $\bar{v}$. Let us show that this solution is unique, and equal to $\psi(u) = \frac{s(u)}{\|s(u)\|}$. We use a duality argument, and reformulate (148) as a convex-concave saddle value problem. We first notice that

$$\max_{i=1,2,\dots,q} \langle \nabla f_i(u), v \rangle = \max_{\Lambda=(\lambda_i) \in S} \left\langle \sum_i \lambda_i \nabla f_i(u), v \right\rangle , \tag{150}$$

where $S = \{ \Lambda = (\lambda_i) \in \mathbb{R}^q : \quad 0 \leq \lambda_i \leq 1,\ \sum_{i=1}^q \lambda_i = 1 \}$ is the unit simplex in $\mathbb{R}^q$. Thus, problem (148) is equivalent to the saddle value problem on $\mathcal{K} \times S$

$$\min_{v \in \mathcal{K}} \max_{\Lambda=(\lambda_i) \in S} \left\langle \sum_i \lambda_i \nabla f_i(u), v \right\rangle , \tag{151}$$

associated to the bilinear Lagrangian function

$$L(v, \Lambda) = \left\langle \sum_i \lambda_i \nabla f_i(u), v \right\rangle .$$

Note that $\mathcal{K}$ and S are two closed convex bounded sets in reflexive Banach spaces. Hence, by the Von Neumann's minimax theorem, see [4, Theorem 9.7.1], there exists $\bar{\Lambda} = (\bar{\lambda}_i) \in S$ such that $(\bar{v}, \bar{\Lambda})$ is a saddle point of (151). Set

$$\bar{\psi}(u) = \mathrm{proj}_{T_K(u)} \left(- \sum_i \bar{\lambda}_i \nabla f_i(u) \right) . \tag{152}$$

a) Let us first prove that the solution $\bar{v}$ of (148) is unique, and equal to

$$\bar{v} = \frac{\bar{\psi}(u)}{\|\bar{\psi}(u)\|} . \tag{153}$$

To that end, we use one of the two conditions of the saddle value property (151),

(namely $L(\bar v,\bar\Lambda)=\inf_{v\in\mathcal K}L(v,\bar\Lambda)$), and Moreau theorem, to obtain

$$\begin{aligned}
\langle\bar\psi(u),\bar v\rangle &= -\left\langle\sum_i\bar\lambda_i\nabla f_i(u),\bar v\right\rangle-\left\langle\mathrm{proj}_{N_K(u)}\left(-\sum_i\bar\lambda_i\nabla f_i(u)\right),\bar v\right\rangle\\
&\geq -\left\langle\sum_i\bar\lambda_i\nabla f_i(u),\bar v\right\rangle\quad\text{because }\bar v\in T_K(u)\\
&\geq -\left\langle\sum_i\bar\lambda_i\nabla f_i(u),\frac{\bar\psi(u)}{\|\bar\psi(u)\|}\right\rangle\quad\text{because }\frac{\bar\psi(u)}{\|\bar\psi(u)\|}\in\mathcal K\\
&\geq\left\langle\mathrm{proj}_{T_K(u)}\left(-\sum_i\bar\lambda_i\nabla f_i(u)\right),\frac{\bar\psi(u)}{\|\bar\psi(u)\|}\right\rangle+\\
&\frac{1}{\|\bar\psi(u)\|}\quad\left\langle\mathrm{proj}_{N_K(u)}\left(-\sum_i\bar\lambda_i\nabla f_i(u)\right),\mathrm{proj}_{T_K(u)}\left(-\sum_i\bar\lambda_i\nabla f_i(u)\right)\right\rangle\\
&\geq\left\langle\bar\psi(u),\frac{\bar\psi(u)}{\|\bar\psi(u)\|}\right\rangle=\|\bar\psi(u)\|.
\end{aligned}$$

Hence, $\left\langle\frac{\bar\psi(u)}{\|\bar\psi(u)\|},\bar v\right\rangle\geq 1$, which immediately gives

$$\begin{aligned}
\|\bar v-\frac{\bar\psi(u)}{\|\bar\psi(u)\|}\|^2&=\|\bar v\|^2+\|\frac{\bar\psi(u)}{\|\bar\psi(u)\|}\|^2-2\left\langle\frac{\bar\psi(u)}{\|\bar\psi(u)\|},\bar v\right\rangle\\
&\leq 1+1-2\leq 0.
\end{aligned}$$

Hence $\bar v=\frac{\bar\psi(u)}{\|\bar\psi(u)\|}\|$, that's precisely (153).

b) In order to complete the proof we are going to show that

$$\bar\psi(u)=s(u),$$

which combined with (153), will clearly imply $\bar v=\frac{s(u)}{\|s(u)\|}$.
Thus, we have to prove that $\bar\psi(u)$ is the element of minimal norm of the convex set

$$C:=\left\{\left(-\sum_i\lambda_i\nabla f_i(u)\right)-\eta:\quad\Lambda=(\lambda_i)\in S,\ \text{and}\ \ \eta\in N_K(u)\right\}.$$

By definition (152) and Moreau theorem, we have $\bar\psi(u)\in C$. Moreover, by (153) and Moreau theorem

$$\begin{aligned}
\|\bar\psi(u)\|&=\left\langle\bar\psi(u),\frac{\bar\psi(u)}{\|\bar\psi(u)\|}\right\rangle\\
&=\langle\bar\psi(u),\bar v\rangle\\
&=\left\langle\left(-\sum_i\bar\lambda_i\nabla f_i(u)\right),\bar v\right\rangle.
\end{aligned}$$

As a consequence, by using the other condition of the saddle value property (151), (namely $L(\bar{v},\bar{\Lambda}) = \sup_{\Lambda\in S} L(\bar{v},\Lambda)$), and Moreau theorem, we obtain

$$\begin{aligned}
\|\bar{\psi}(u)\| &\leq -\left\langle \sum_i \lambda_i \nabla f_i(u), \bar{v} \right\rangle \quad \text{for all } \Lambda \in S \\
&\leq \left\langle \text{proj}_{T_K(u)} \left(-\sum_i \lambda_i \nabla f_i(u) \right), \bar{v} \right\rangle \quad \text{for all } \Lambda \in S \\
&\leq \|\text{proj}_{T_K(u)} \left(-\sum_i \lambda_i \nabla f_i(u) \right)\| \quad \text{for all } \Lambda \in S \\
&\leq \| \left(-\sum_i \lambda_i \nabla f_i(u) \right) - \eta\| \quad \text{for all } \Lambda \in S, \text{ and } \eta \in N_K(u).
\end{aligned}$$

Hence, if u is not a Pareto critical point, $\bar{\psi}(u)$ is the element of minimal norm of the convex set C that is, $\bar{\psi}(u) = s(u)$. Combining with (153) we obtain $\bar{v} = \frac{s(u)}{\|s(u)\|}$. The argument being valid for any solution $\bar{v}$ of (148), we conclude that (148) admits a unique solution $\psi(u)$ which satisfies $\psi(u) = \frac{s(u)}{\|s(u)\|}$. □

4 Links with cooperative game theory

Let us consider q agents (consumers, social actors, deciders,...). The agent i acts on a decision space $\mathcal{H}_i$, and takes decision $v_i \in \mathcal{H}_i$, $i = 1, 2, ..., q$. Let $\mathcal{K}$ be a given closed subset of $\mathcal{H} = \mathcal{H}_1 \times \mathcal{H}_2 \times ... \times \mathcal{H}_q$, which reflects the limitation of ressources, and/or various constraints. Feasible decisions $v \in \mathcal{H}$ satisfy

$$v = (v_1, v_2, ..., v_q) \in \mathcal{K}.$$

Each agent i has a disutility (loss) function $f_i : \mathcal{H} \to \mathbb{R}$ which associates to each feasible decision $v \in \mathcal{K}$ the scalar $f_i(v)$. The game in normal form is given by the triplet $(\mathcal{H}, \mathcal{K}, (f_i)_{i=1,...q})$. The (MCSD) dynamic has been designed in order to satisfy some desirable properties with respect to Pareto equilibration: each trajectory $t \mapsto u(t)$ of (MCSD) satisfies

i) for each $i = 1, 2, ..., q$, $t \mapsto f_i(u(t))$ is nonincreasing (Theorem 1.5, item i));

ii) $u(t)$ converges to a Pareto critical point as $t \to +\infty$ (Theorem 1.5, item iii));

iii) at each time $t > 0$, the multiobjective steepest descent direction followed by u is a Pareto optimal solution of the linearized tangent game.

Item iii) is proved below. It is related to the following notion of tangent game, which extends the classical linearization idea to the multiobjective constrained case. To our knowledge, it has been first considered in [21, chaper 3, section 4].

Definition 4.1. *Let us give a game $\mathcal{G}$ in normal form $(\mathcal{H}, \mathcal{K}, (f_i)_{i=1,...q})$. The tangent game of $\mathcal{G}$ at $u \in \mathcal{K}$ is the game $(\mathcal{H}, \mathcal{K}(u), (\langle \nabla f_i(u), \cdot \rangle)_{i=1,...q})$ with*

- *$K(u) = \{v \in T_{\mathcal{K}}(u) : \ \|v\| \leq 1\}$;*
- *$v \mapsto \langle \nabla f_i(u), v \rangle$ is the directional derivative of f_i at u in the direction v.*

Proposition 4.2. *The multiobjective steepest descent direction at $u \in \mathcal{K}$ is a Pareto optimal solution of the tangent game of $\mathcal{G}$ at $u \in \mathcal{K}$.*

Proof. In the proof of Theorem 3.5, we showed that the steepest descent direction is solution of a minimization problem

$$\min \left\{ \left\langle \sum_i \lambda_i \nabla f_i(u), v \right\rangle : \quad v \in K(u) \right\}$$

for some $\lambda = (\lambda_i)_i$ belonging to the unit simplex of $\mathbb{R}^q$. This implies Pareto criticality with respect to the tangent game of $\mathcal{G}$ at u. The tangent game of $\mathcal{G}$ at $u \in \mathcal{K}$ is associated to convex objective functions $(\langle \nabla f_i(u), \cdot \rangle)_{i=1,\dots q})$. Hence, by Proposition 1.4 we obtain weak Pareto optimality. Indeed, following [21], one can show the stronger property of Pareto optimality. □

Remark 4.3. As we already stressed, in the (MCSD) dynamic we do not scalarize the original vector optimization problem. Neither ordering information nor weighting factors for the different objective functions are assumed to be known. That's what we call the endogenous property of the system. There is much to say about the interpretation of the steepest descent in decision sciences. Taking the worst directional derivative (indeed, in view of minimization, it is the largest one), confers to (MCSD) robustness (minimization in the worst case), and nice convergence properties.

5 Numerical descent methods for multiobjective optimization

In [25], Fliege and Svaiter analyze the convergence properties of various descent methods for multiobjective optimization in the unconstrained, finite dimensional setting. Let us recall the formulation of the multiobjective steepest descent direction, which has been given in Theorem 3.5. We have that $\psi(u) = \frac{s(u)}{\|s(u)\|}$ is the solution of the convex minimization problem

$$\min_{v \in T_K(u),\ \|v\|=1} \max_{i=1,2,\dots,q} \langle \nabla f_i(u), v \rangle . \tag{154}$$

This formulation is close to the notion of multiobjective steepest descent direction, in the unconstrained case, which has been introduced in [25]. It is the solution of the strongly convex problem

$$\min_{v \in \mathcal{H}} \left\{ \frac{1}{2}\|v\|^2 + \max_{i=1,2,\dots,q} \langle \nabla f_i(u), v \rangle \right\}.$$

Indeed, in the same paper, the authors mention the possibility to replace $\frac{1}{2}\|v\|^2$ by other functions, including our choice which is the indicator function of the unit ball in $\mathcal{H}$. They indicate that a quite similar analysis can be developed in these various situations.

Thus, the algorithms developed in [25], [26], and [28] are close to the explicit time discretization of our dynamic. As well, the proximal algorithm, as developed in [16], is

close to its implicit discretization. In turn, these algorithmic results suggest that there is a large class of continuous dynamics which contains (MCSD), and having similary properties with respect to Pareto equilibration. Similarly, the recent Newton method for vector optimization which has been developed in [27] suggests the existence of corresponding continuous dynamics (see for example [6] in the case of a single objective). Enriching the class of dynamics can be a useful approach for understanding the complex interactions in Pareto equilibration (coalitions, negociating, bargaining, dealing with uncertainty, changing environment, psychological aspects). These are interesting subjects for further research.

References

[1] Y. I. Alber and A. I. Notik , *On some estimates for projection operators in Banach spaces*, Comm. Appl. Nonlinear Anal., **2** (1995), no. 1, pp. 47–55.

[2] H. Attouch, *Convergences de fonctionnelles convexes*, Journées d'Analyse Non Linéaire, Besançon, Lecture Notes in Mathematics, Springer-Verlag **665**, (1977), pp. 1-40.

[3] H. Attouch, *Variational convergence for functions and operators*, Pitman Advanced Publishing Program, Applicable Mathematics Series, 1984.

[4] H. Attouch, G. Buttazzo, and G. Michaille, Variational analysis in Sobolev and BV spaces. Applications to PDE's and optimization, *MPS/SIAM Series on Optimization*, 6, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006, 634 pages.

[5] H. Attouch and A. Damlamian, *On multivalued evolution equations in Hilbert spaces*, Israël J. Math., **12** (1972), pp. 373–390.

[6] H. Attouch and B. F. Svaiter, *A continuous dynamical Newton-Like approach to solving monotone inclusions*, SIAM J. Control Optim., **49**(2) (2011), pp. 574-598.

[7] H. Attouch and R. Wets, *Quantitative stability of variational systems: II, A framework for nonlinear conditioning*, SIAM J. Optim., **3** (1993), pp. 359–381.

[8] J.P. Aubin and A. Cellina, *Differential inclusions*, Springer, Berlin, 1984.

[9] J.P. Aubin and I. Ekeland, *Applied nonlinear analysis*, Wiley, 1984.

[10] J.B. Baillon and A. Haraux, *Comportement à l'infini pour les équations d'évolution avec forcing périodique*, Arch. Rat. Mech. Anal, **67**, 1 (1977), pp. 101–109.

[11] T. Q. Bao, and B. Mordukhovich, *Sufficient conditions for global weak Pareto solutions in multiobjective optimization*, Positivity, **16** (2012), pp. 579-602.

[12] H. Bauschke and P. Combettes, *Convex analysis and monotone operator theory*, CMS books in Mathematics, Springer, 2011.

[13] G. Beer, *Topologies on closed and closed convex sets*, Mathematics and its Applications, Kluwer, **268**, 1993.

[14] H. Benabdellah, C. Castaing, and A. Salvadori, *Compactness and discretization methods for differential inclusions and evolution problems*, Atti. Sem. Mat. Univ. Modena, XLV (1997), pp. 9–51.

[15] G. Bento, O. Ferreira, and P. Oliveira, *Unconstrained Steepest descent method for multicriteria optimization on riemannian manifolds*, J. Optim. Theory Appl., **154**, (2012), pp. 88107.

[16] H. Bonnel, A. N. Iusem, and B. F. Svaiter, *Proximal methods in vector optimization*, SIAM Journal on Optimization, **15** (2005), pp. 953–970.

[17] P.A.N. Bosman, *On gradients and hybrid evolutionary algorithms for real-valued multi-objective optimization*, IEEE Transactions on Evolutionary Computation **16**, (1), (2012), pp. 51–69.

[18] H. Brézis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland/Elsevier, New-York, 1973.

[19] H. Brézis, *Analyse fonctionnelle*, Masson, 1983.

[20] A. Cellina and V. Staicu, *On evolution equations having monotonicities of opposite sign*, Journal Diff. Equations, **90** (1991), pp. 71–80.

[21] B. Cornet, *Contributions to the mathematical theory of dynamical mechanisms of resource allocation*, Thèse de doctorat d'état, Université Paris IX Dauphine (1981).

[22] B. Cornet, *Monotone planning procedures and accessibility of Pareto optima*, in New Trends in Dynamic System Theory and Economics, Aoki et Marzollo (eds.), Academic Press (1979), pp. 337–349.

[23] B. Cornet, *Existence of slow solutions for a class of differential inclusions*, Journal of Mathematical Analysis and Applications, **96** (1983), pp. 130–147.

[24] G. P. Crespi and M. Rocca, *Minty variational inequalities and monotone trajectories of differential inclusions*, Journal of Inequalities in Pure and Applied Mathematics, **5**, Issue 2, (2004)

[25] J. Fliege and B. F. Svaiter, *Steepest descent methods for multicriteria optimization*, Mathematical Methods of Operations Research, **51**(3) (2000), pp. 479–494.

[26] L. M. Graña Drummond and B. F. Svaiter, *A steepest descent method for vector optimization*, Journal of Computational and Applied Mathematics, **175** (2005), pp. 395–414.

[27] L. M. Graña Drummond, F.M.P. Raupp and B. F. Svaiter, *A quadratically convergent Newton method for vector optimization*, Optimization, 2012.

[28] L. M. Graña Drummond and L.M. Iusem, *A projected gradient method for vector optimization problems*, Comput. Optim. Appl., **28**(1), (2004), pp. 529.

[29] X. Goudou and J. Munier, *The gradient and heavy ball with friction dynamical systems: the quasiconvex case*, Math. Program., **116**(1-2) (2009), pp. 173-191.

[30] Cl. Henry, *An existence theorem for a class of differential equations with multivalued right-hand side*, Journal of Mathematical Analysis and Applications, **41** (1973), pp. 168–179.

[31] K.C. Kiwiel, and K. Murty, *Convergence of the steepest descent method for minimizing quasiconvex functions*, J. Optim. Theory Appl., **89**(1), (1996), pp. 221–226.

[32] C. Malivert, *A descent method for Pareto optimization*, Journal of Mathematical Analysis and Applications, **88**, No. 2, (1982), pp. 610–631.

[33] K.M. Miettinen, *Nonlinear multiobjective optimization*, Kluwer Academic Publishers, Norwell, 1999.

[34] J.J. Moreau, *Décomposition orthogonale d'un espace hilbertien selon deux cônes mutuellement polaires*, Comptes Rendus de l'Académie des Sciences Paris, Série A, **225**, (1962), pp. 238–240.

[35] Z. Opial, *Weak convergence of the sequence of successive approximations for nonexpansive mappings*, Bull. Amer. Math. Soc. **73** (1967), pp. 591–597.

[36] J.P. Penot, *Continuity properties of projection operators*, Journal of Inequalities and Applications, Springer, open access, 2003.

[37] Y. Sawaragi, H. Nakyama, and T. Tanino, *Theory of multiobjective optimization*, Academic Press, Orlando, Florida, 1985.

[38] S. Smale, *Global analysis and economics I: Pareto optimum and a generalization of Morse theory*, Dynamical Systems, Academic Press, 1973.

[39] S. Smale, *Exchange process with price adjustment*, Journal of Mathematical Economics, **3**, 1976, pp. 211–216.

[40] D. The Luc, *Theory of vector optimization*, Springer, Berlin, 1989.

[41] E. Zeidler, *Nonlinear functional analysis and its applications, Part II: Monotone operators*, Springer-Verlag, New-York, 1990.