Minimizing irregular convex functions: Ulam stability for approximate minima

Michel Théra

Université de Limoges and XLIM (UMR 6172)
michel.thera@unilim.fr

Collaborator: **Emil Ernst**

Advanced methods and perspectives in nonlinear optimization and control

ENSIACET, Toulouse – february 3-5 th, 2010

Outline of the talk

1. Position of the problem
   - Background on Attouch-Wets convergence
   - Existing results
   - A characterization of closed and convex sets with a bounded boundary

2. A characterization of Ulam stability of approximate minima

Outline of the talk

1. Position of the problem
   - Background on Attouch-Wets convergence
   - Existing results
   - A characterization of closed and convex sets with a bounded boundary

2. A characterization of Ulam stability of approximate minima

J.-L. Lions

«Ce que j'aime dans les mathématiques appliquées, c'est qu'elles ont pour ambition de donner du monde des systèmes une représentation qui permette de comprendre et d'agir. Et, de toutes les représentations, la représentation mathématique, lorsqu'elle est possible, est celle qui est la plus souple et la meilleure. Du coup, ce qui m'intéresse, c'est de savoir jusqu'où on peut aller dans ce domaine de la modélisation des systèmes, c'est d'atteindre les limites. »

«What I like about applied mathematics is that it has as its aim giving the world of complex systems a certain representation that allows us to both understand and to act. And of all possible representations, mathematical modeling – when it is possible – is the most flexible and the best. In this direction, what interests me is to know just how far we can go in pushing the limits of modeling such systems »

References

Basic references :

- E. Ernst, M. Théra, *Minimizing irregular convex functions : Ulam stability for approximate minima.* Preprint 2010.
- H. Attouch, M. Théra, *Convergences en analyse multivoque et variationnelle*, MATAPLI, 36(1993), pp. 22-39.
- G. Beer, *Topologies on closed and closed convex sets*, Kluwer Academics, 1993.
- J.J. Moreau, *Fonctionelles convexes*, Séminaire "Equations aux dérivées partielles," Collège de France, 1966 and Edizioni del Dipartimento di Ingegneria Civile dell'Università di Roma Tor Vergata, Roma, 2003.

Let $f \in \Gamma_0(X)$, bounded below i.e., $f \in \Gamma_b(X)$ and $\varepsilon > 0$.

$$\{x \in X : f(x) \leq \inf_X f + \varepsilon\}$$

the set of all the ε-approximate minima of f is a nonempty subset of X.

Consider $\mathfrak{A}_\varepsilon : f \in \Gamma_b(X) \mapsto \{x \in X : f(x) \leq \inf_X f + \varepsilon\}$.

This talk focuses on the study of the upper semi-continuity of $\mathfrak{A}_\varepsilon$.

A function $f \in \Gamma_b(X)$ such that $\mathfrak{A}_\varepsilon$ is upper semi-continuous at f is said to have a *Ulam stable ε-approximate minima.*

- Our main objective in this presentation is to characterize the subclass of $\Gamma_b(X)$ formed of those functions at which $\mathfrak{A}_\varepsilon$ is upper semi-continuous.
- Despite its abstract appearance, this type of stability turns out to be essential in numerical optimization, namely in answering the natural question of defining the largest class of functionals $f \in \Gamma_b(X)$ for which minimization algorithms exist.

Ulam stability of approximate minima is achieved for all coercive functions $f \in \Gamma_0(X)$.

Several convincing examples prove that Ulam stability may fail outside the coercive setting.

Suggest that coercivity and Ulam stability are closely connected notions.

Unexpected (and, to our best knowledge, new) the fact that Ulam stability holds also for some non-coercive functions $f \in \Gamma_b(X)$, since the mapping $\mathfrak{A}_\varepsilon$ is upper semi-continuous at any constant function, and also is upper semi-continous everywhere on $\Gamma_b(\mathbb{R})$.

Thus, Ulam stability is a property shared by three apparently dissimilar type of $\Gamma_b(X)$-functionals : the class of all $f \in \Gamma_0(X)$ which are **coercive**, or **constant** or when $X = \mathbb{R}$.

The objective of this presentation is to provide an explanation to this bizarre situation.

To this respect,

- we characterize the mappings $f \in \Gamma_b(X)$ with Ulam-stable approximate minima as being those functions $f \in \Gamma_b(X)$ for which the boundary of every sub-level set is bounded,
- and we also remark that a $\Gamma_0(X)$-functional has all its sub-level sets with a bounded boundary if and only if it is coercive, constant, or defined on the real axis.

Accordingly, the key concept for Ulam stability is the boundedness of the boundary of any sub-level set of the function, and not, as the emphasis classically put on coerciveness might suggest, the boundedness of the sub-level sets themselves.

The interpretation of Ulam stability as ensuring the existence of a minimizing algorithm allows us now to state that $f \in \Gamma_b(X)$ may be numerically minimized as long as the boundary of each of its sub-level sets is bounded.

Ulam stability is crucial when applied to the minimization of a function $f \in \Gamma_b(X)$. Indeed, in order to correctly work, most (if not all) of the numerical minimizing algorithms require enough regularity on the function to be minimized. This regularity in general fails in real world applications.

To overcome this difficulty and be able to minimize also less regular functions in the class $\Gamma_b(X)$, it is customary to use one of the available regularizing techniques

- the rolling ball technique (Noll)
- the Lipschitzian regularization (Hiriart-Urruty, Azé, Attouch & Wets, Fitzpatrick & Phelps, $\cdots$)
- the Moreau-Yosida regularization (Attouch & Théra, Moreau, Yosida, $\cdots$,
- the robust regularization (Ben Tal & Nemirovski, Seeger)

and construct in this way a sequence $(f_n)_{n\in\mathbb{N}}$ in $\Gamma_b(X)$ "very regular" (typically coercive and two times continuously differentiable) converging to f; $(f_n)_{n\in\mathbb{N}}$ is called a regularizing sequence for f.

For every function f_n and error bound $\varepsilon > 0$, it is now possible to use a minimizing algorithm and compute an ε-approximate minimizer of f_n, that is a vector $x_{n,\varepsilon}$ such that

$$f_n\left(x_{n,\varepsilon}\right) - \min_X f_n \leq \varepsilon; \tag{1}$$

$x_{n,\varepsilon}$ is called a ε-asymptotically minimizing sequence (Zolezzi).

Of course, this technique is of some use only if the ε-asymptotically minimizing sequence actually approaches the set of ε-approximate minima of f, that is if

$$\lim_{n\to\infty} \left(\operatorname{dist} x_{n,\varepsilon}, \ \{x \in X : \ f(x) - \inf_X f \leq \varepsilon\}\right) = 0. \tag{2}$$

This talk is concerned with the class of those mappings in $\Gamma_b(X)$ for which the above-described technique works for one of the regularizing techniques ordinarily in use.

When does relation (2) hold for all the sequences $(f_n)_{n\in\mathbb{N}}$, $f_n \in \Gamma_b(X)$ Attouch-Wets converging to f ?

as the Attouch-Wets convergence appears as the main common feature of all the above-mentionned regularizing sequences.

We note $(C(X), \tau_H)$ the topology induced on $C(X)$ by the uniform Hausdorff metric, and $(\Gamma_b(X), \tau_{AW})$, the subspace of those functions in $\Gamma_0(X)$ that are bounded below endowed with the Attouch-Wets topology (nothing but the topology of the uniform convergence on bounded sets applied to the distance functionals to the sets from $C(X)$).

More precisely, we seek for all $f \in \Gamma_b(X)$ such that, for every sequence $(f_n)_{n\in\mathbb{N}}$, $f_n \in \Gamma_b(X)$ Attouch-Wets converging to f, and every error bound $\varepsilon > 0$, it holds that

$$\varepsilon - \operatorname{argmin} f_n \subset \ \varepsilon - \operatorname{argmin} f + \delta_n B_X \qquad (3)$$

for some sequence $\delta_n > 0$ converging to zero (as usually, B_X stands for the closed unit ball of X).

Outline

1. Position of the problem
   - Background on Attouch-Wets convergence
   - Existing results
   - A characterization of closed and convex sets with a bounded boundary
2. A characterization of Ulam stability of approximate minima

As the Attouch-Wets topology is metrizable, relation (3) says in fact that the application

$$\mathfrak{A}_\varepsilon : (\Gamma_b(X), \tau_{AW}) \to (C(X), \tau_H),$$

defined as

$$\Gamma_b(X) \ni g \to \mathfrak{A}_\varepsilon(g) = \varepsilon - \operatorname{argmin} g \in C(X), \tag{4}$$

is upper semi-continuous at f.

Our interest for the upper semi-continuity of $\mathfrak{A}_\varepsilon$ is thus motivated by the fact that, although potentially irregular, any $\Gamma_b(X)$- functional f at which the mapping $\mathfrak{A}_\varepsilon$ is upper semi-continuous may still be minimized.

Indeed, it is enough to apply a minimizing algorithm to any regularizing sequence of f, as the ε-asymptotically minimizing sequence thus obtained must approach the ε-approximate minima of the function f to be minimized.

Outline

1. Position of the problem
   - Background on Attouch-Wets convergence
   - Existing results
   - A characterization of closed and convex sets with a bounded boundary

2. A characterization of Ulam stability of approximate minima

Existing results

Moreau (the upper semi-continuity of the approximate subdifferential on the interior of the effective domain of a $\Gamma_0(X)$-functional) :

means that the restriction of the mapping $\mathfrak{A}_\varepsilon$ to the subset

$$\{f - y^\star \in \Gamma_b(X) : y^\star \in X^\star\}$$

of $\Gamma_b(X)$ is upper semi-continuous at f provided that f is coercive.

Originally : Prove the upper semi-continuity of the approximate subdifferential on the interior of the effective domain of a $\Gamma_0(X)$-functional (see also Asplund & Rockafellar), as a partial answer to the Ulam-stability question.

Beer & Luchetti :

any coercive $\Gamma_0(X)$ functional admits a Ulam-stable ε-approximate minima for any positive value $\varepsilon > 0$.

Attouch-Wets topology

$C(X)$ is the class of all the closed and nonempty convex subsets of a normed space $(X, \|\cdot\|)$ The Hausdorff distance between two sets :

$$\mathrm{d}_{\mathrm{H}}(C_1, C_2) = \max\left(e(C_1, C_2),\ e(C_2, C_1)\right) \tag{5}$$

where

$$e(C_1, C_2) = \sup_{x_1 \in C_1} \inf_{x_2 \in C_2} \|x_1 - x_2\|,$$

is the excess of C_1 over C_2.
Endowed with this metric, $C(X)$ becomes a metric space, denoted hereafter by $(C(X), \tau_H)$.

This topology turns out to be too fine for its applications to variational problems, in the sense that many sequences of sets which "intuitively" must converge, are not Hausdorff converging.

Hence, several coarser topologies have been introduced. Among them, a prominent place is held by the Attouch-Wets topology (sometimes called the bounded Hausdorff topology) :

a net $(C_i)_{i\in I} \subset C(X)$ Attouch-Wets converges to $C \in C(X)$ if and only if it converges with respect to any of the distances d_ρ, $\rho > 0$, where

$$\mathrm{d}_\rho(C_1, C_2) = \max\left(e(C_1 \cap \rho B_X, C_2), e(C_2 \cap \rho B_X, C_1)\right).$$

We shall denote this topology by $(C(X), \tau_{AW})$; we also use notation $(\Gamma_0(X), \tau_{AW})$ for the topology induced on $\Gamma_0(X)$ by the one-to-one mapping

$$\mathrm{epi} : \Gamma_0(X) \to (C(X \times \mathbb{R}), \tau_H).$$

where $\mathrm{epi}\, f$ is the epigraph of the $\Gamma_0(X)$-functional f.
the Attouch-Wets topology τ_{AW} on $C(X)$ is nothing but the topology of the uniform convergence on bounded sets applied to the distance functionals to the sets from $C(X)$.

This topology is metrizable

$$\mathrm{d}_{AW}(C_1,\ C_2) = \Sigma_{i=1}^{+\infty} \frac{\mathrm{d}_i\,(C_1,\ C_2)}{2^i\,(1+\mathrm{d}_i\,(C_1,\ C_2))};$$

A sequence $(f_n)_{n\in\mathbb{N}}$ in $\Gamma_0(X)$, Attouch-Wets converges to f provided the sequence $(\mathrm{epi}\, f_n)_{n\in\mathbb{N}}$ of the epigraphs of the f_n is convergent in the Attouch-Wets topology to $\mathrm{epi}\, f$.

We note $(C(X),\ \tau_H)$ the topology induced on $C(X)$ by the uniform Hausdorff metric, and $(\Gamma_b(X),\ \tau_{AW})$, the subspace of those functions in $\Gamma_0(X)$ that are bounded below endowed with the Attouch-Wets topology.

The identity mapping $\iota : (C(X), \tau_{AW}) \to (C(X), \tau_H)$

Lower semicontinuity of the identity mapping

it can be observed that the three following statements are equivalent :
i) the mapping ι is continuous at $C \in C(X)$;
ii) the mapping ι is lower semi-continuous at $C \in C(X)$;
iii) the set C is bounded.

Lemma (☼)

Let C be a closed convex subset of a normed space $(X, \|\cdot\|)$. The two following sentences are equivalent :
i) *the boundary of C is unbounded ;*
ii) *there is an unbounded sequence $(x_n)_{n\in\mathbb{N}} \subset X$ such that*

$$a \leq \mathrm{dist}(x_n, C) \leq b \quad \forall n \in \mathbb{N} \tag{6}$$

for some values a and b such that $0 < a < b$.

Lemma (☼☼)

Let X be a topological vector space and C be a closed subset of X of non-void interior. Any line segment of ends $x \in \overset{\circ}{C}$ and $y \in X \setminus C$ meets the boundary of C.

As a consequence of Lemma (☼☼), let us mention :

Lemma (☼☼☼)

Let X be a normed vector space and C be a closed subset of X. Then, for any $x \in X \setminus C$ it holds that

$$\operatorname{dist}(x, C) = \operatorname{dist}(x, \partial C),$$

where by ∂S we mean the topological boundary of the set S.

Remark

The convexity assumption on C is essential in establishing that the implication $i) \Rightarrow ii)$ holds true. Indeed, the closed subset

$$C = \left\{(x, y) : |x\,y| \leq \frac{1}{2}\right\}$$

of the normed space $\left(\mathbb{R}^2, \|(x, y)\| = \sqrt{x^2 + y^2}\right)$ *possesses an unbounded boundary. Straightforward calculations prove that*

$$\operatorname{dist}((x, y),\ C) \leq \frac{1}{2\sqrt{x^2 + y^2}} \quad \forall (x, y) \in \mathbb{R}^2,\ (x, y) \neq (0, 0).$$

Hence, no sequence fulfilling relation (6) may exists.

Proposition (♠)

The identity mapping

$$\iota : (C(X), \tau_{AW}) \to (C(X), \tau_H)$$

is upper semi-continuous at C provided that the boundary of C is bounded.

Proof by contradiction + use of Lemma (☼)

Prove the existence of a sequence $(C_n)_{n\in\mathbb{N}} \subset C(X)$ Attouch-Wets converging to C, and of a value $\gamma > 0$ such that

$$\mathrm{e}\,(C_n,\, C) \geq \gamma \qquad \forall n \in \mathbb{N}. \tag{7}$$

Then construct an unbounded sequence $(x_n)_{n\in\mathbb{N}} \subset X$ such that

$$\gamma \leq \mathrm{dist}(x_n, C) \leq 2\,\gamma \tag{8}$$

Converse of Proposition (♠) : a first step

The following result provides a sistematic manner to construct sequences of closed and convex sets Attouch-Wets converging to a given set.

Lemma (♋)

Let C be a closed and convex subset of a normed space $(Y\|\cdot\|)$ and set $C_n := C \cap n\, B_Y$, the intersection of C with the closed ball of center θ_Y and radius n.

For any sequence $(y_n)_{n\in\mathbb{N}} \subset Y$, denote by $A_n = \mathrm{co}(y_n, C_n)$.

Then the sequence $(A_n)_{n\in\mathbb{N}} \subset C(Y)$ Attouch-Wets converges to C provided that

$$\|y_n\| \leq n^2 \quad \text{and} \quad \mathrm{dist}\,(y_n, C) \leq k \quad \forall\, n \in \mathbb{N}, \tag{9}$$

where $k > 0$ is a postive real number.

Proposition (◑)

If the identity mapping $\iota : (C(X), \tau_{AW}) \to (C(X), \tau_H)$ *is upper semi-continuous at* C*, then the boundary of* C *is bounded.*

Let $D \in C(X)$ be a set at which the mapping ι is u.s.c, and assume that ∂D is unbounded. Apply implication $i) \Rightarrow ii)$ from Lemma (☼) to D : $\exists\, (x_n)_{n\in\mathbb{N}} \subset X$ unbounded, $\exists$ two values a and b such that $a \leq \mathrm{dist}(x_n, D) \leq b, \;\; \forall n \in \mathbb{N}$.
As the sequence $(x_n)_{n\in\mathbb{N}}$ is unbounded, take a subsequence, still denoted by $(x_n)_n$ such that $\|x_n\| \geq n^2$. Define $D_n = \mathrm{co}\,(x_n, D \cap \, n\, B_X)\,$;

Apply Lemma (☺) for $X = Y$, $D = C$, $x_n{=}y_n$ and $b = k$ to deduce that the sequence $(D_n)_{n\in\mathbb{N}}$ Attouch-Wets converges to D.
On the other hand,

$$\mathrm{e}(D_n, D) \geq \mathrm{dist}(x_n, D) \geq a \qquad \forall\, n \in \mathbb{N}.$$

Hence, the mapping i is not upper semi-continuous at D, a contradiction.

Outline

1. Position of the problem
   - Background on Attouch-Wets convergence
   - Existing results
   - A characterization of closed and convex sets with a bounded boundary
2. A characterization of Ulam stability of approximate minima

A characterization of closed and convex sets with a bounded boundary

Propositions (♠) and (◑) motivate our interest for the study of closed and convex sets with a bounded boundary, as their class turns out to be a key notion when studying the interplay between the topologies $(C(X), \tau_H)$ and $(C(X), \tau_{AW})$.

The following result allows us to completely characterize the family of closed and convex sets with a bounded boundary, as well as $\Gamma_0(X)$-functionals with sublevel sets possessing this property.

Proposition (✠)

Let X be a normed vector space of dimension greater than one, and let C be an unbounded and proper closed and convex subset of X. Then, the boundary of C is unbounded.

As the boundary of X is empty, the boundary of a bounded set is bounded (is a part of the set), and a boundary of a convex subset of the real axis is composed of at most two points, it follows that the boundaries of all the three above-mentionned classes of sets are bounded. We have thus proved the following characterization of the subfamily of $C(X)$ consisting of sets with a bounded boundary.

Proposition (Π)

The boundary of a closed and convex subset C of a normed space X is bounded if and only if at least one of the following three statements holds true :

s_1) $X = \mathbb{R}$;
s_2) C *is bounded* ;
s_3) $C = X$.

Continued

Accordingly, all the sub-level sets of a $\Gamma_0(X)$- functional have bounded boundaries, if and only if at least one of the following three statements holds true :

f_1) $X = \mathbb{R}$;

f_2) the mapping f is coercive ;

f_3) the mapping f is constant.

A sequence of Attouch-Wets converging functions

A standard way to associate an extended real-valued function to a subset S of the product space $(X \times \mathbb{R}, \|\cdot,\cdot\|_B)$ (here $\|\cdot,\cdot\|_B$ is the standard box norm on $X \times \mathbb{R}$, $\|x, s\|_B = \max(\|x\|, |s|)$) is to consider k_S, the lower-boundary function to S. This application is defined (Rockafellar, Seeger), as the unique extended-real-valued function $k_S : X \to \mathbb{R} \cup \{-\infty, +\infty\}$ such that

$$\mathrm{epi}(k_S) = S + \{0\} \times \mathbb{R}_+;$$

more precisely,

$$X \ni u \to k_S(u) = \inf\{s \in \mathbb{R} : (x, s) \in S\} \in X \cup \{-\infty, +\infty\}.$$

When S is a convex subset of $X \times \mathbb{R}$, then k_S is also a convex function. Assuming that S is closed and convex does not automatically implies that $k_S \in \Gamma_0(X)$. if and only if k_S is proper (in other words, if the vector $(0, -1)$ does not belong to the recession cone of S

Let us now consider a sequence $(S_n)_{n\in\mathbb{N}} \subset X \times \mathbb{R}$ which Attouch-Wets converges ; even if k_{A_n} exists for every n, it may happen that the sequence $(k_{A_n})_{n\in\mathbb{N}}$ is not Attouch-Wets convergent in $\Gamma_0(X)$ (take for instance $X = \mathbb{R}$ and $A_n = \{(x, \, n\, x) : \, x \in \mathbb{R}\}$).
The following result provides a very general frame in which the Attouch-Wets convergence of a sequence of sets $(S_n)_{n\in\mathbb{N}}$ implies the Attouch-Wets convergence of the sequence $(k_{S_n})_{n\in\mathbb{N}}$.

Lemma

Let $(S_n)_{n\in\mathbb{N}} \subset X \times \mathbb{R}$ be a sequence of sets Attouch-Wets converging to the epigraph of a functional $f \in \Gamma_0(X)$. Then :
i) for n large enough, the functions k_{S_n} belong to $\Gamma_0(X)$;
ii) the sequence $(k_{S_n})_{n\in\mathbb{N}}$ Attouch-Wets converges to f.

A technical result

Proposition (KK)

Let $f \in \Gamma_b(X)$ and $\eta > 0$. The application $\mathfrak{A}_\varepsilon$ is not upper semi-continuous at f for any $\varepsilon < \eta$, provided that the sub-level set $L_\eta = \{x \in X : f(x) \leq \inf_X f + \eta\}$ of f possesses an unbounded boundary.

The main result.

Main result

Complete characterization of the subclass of $\Gamma_b(X)$ of those functions which have a Ulam-stable ε-approximate minima.

Theorem

Let $(X, \|\cdot\|)$ be a normed space. The application $\mathfrak{A}_\varepsilon$ is upper semi-continuous at f for any $\varepsilon > 0$ if and only if the boundary of every sub-level set of f is bounded.

Thank you for your attention