

Convex Analysis and Optimization in the Past 50 Years: Some Snapshots

Jean-Baptiste Hiriart-Urruty

Abstract We offer some personal historical snapshots on the evolution of convex analysis and optimization in the past 50 years. We deliberately focus on some specific periods or dates which resonate in our memories.

Keywords Convex analysis • Optimization • Nonsmooth analysis • History of optimization

1 Fifty Years Ago: 1962–1963

We all know that the development of convex analysis during the last 50 years owes much to W. Fenchel (1905–1988), J.-J. Moreau (1923-) and R.T. Rockafellar (1935-). Fenchel was very “geometrical” in his approach; Moreau used to say that he did applied mechanics: he “applied mechanics to mathematics”, while the concept of “dual problem” was a constant leading thread for Rockafellar. The years 1962–1963 can be considered as the date of birth of *modern convex analysis with applications to optimization*. The now familiar appellations like subdifferential, proximal mappings, infimal convolution date back from this period, exactly 50 years ago. In two consecutive notes published by the French Academy of Sciences [16, 17], Moreau introduced the so-called proximal mappings and a way of regularizing a convex function defined on a Hilbert space by performing an inf-convolution with the square of the norm; these preliminary works culminated with the 1965 paper [19], which remains for me the archetype of elegant mathematical paper.

J.-B. Hiriart-Urruty (✉)

Institut de Mathématiques University Paul Sabatier 118, route de Narbonne
31062 Toulouse Cedex 9, France,
e-mail: jbhu@math.univ-toulouse.fr

The short paper by Hörmander [14], on the support functions of sets in a general context of locally convex topological vector spaces, published (in French) some years earlier (1954), was influential in modern developments of convex analysis. These thoughts came to my mind these days since L. Hörmander just passed away (on November 2012); he was very young (less than 23 years old) when he wrote this paper, his Ph.D. thesis on PDE was not yet completed. I remained impressed by the maturity of this mathematician at this early age.

Various names appeared in 1963 to denote a vector s satisfying

$$f(y) \geq f(x) + \langle s, y - x \rangle \text{ for all } y.$$

R.T. Rockafellar in his 1963 Ph.D. thesis [21] called s “a differential of f at x ”; it is J.-J. Moreau who, in a note at the French Academy of Sciences (in 1963) [18], introduced for s the word “sous-gradient” (which became “subgradient” in English). Even the wording “la sous-différentielle” (a feminine word in French, closer to the classical “la différentielle” for differentiable functions) was used in the early days, it became later “le sous-différentiel” (a masculine word in French). As it often happens in research in mathematics, when times are ripe, concepts bloomed in different places of the world at about the same time; in the former USSR, for example, institutes or departments in Moscow and Kiev were on the front; just to give a name, N.Z. Shor’s thesis in Kiev is dated 1964. A little bit earlier, in 1962, N.Z. Shor published a first instance of use of a subgradient method for minimizing a nonsmooth convex function (a piecewise linear one actually).

One of the most specific constructions in convex or nonsmooth analysis is certainly taking the supremum of a (possibly infinite) collection of functions. In the years 1965–1970, various calculus rules concerning the subdifferential of sup-functions started to emerge; working in that direction and using various assumptions, several authors contributed to this calculus rule: B.N. Pshenichnyi, A.D. Ioffe, V.L. Levin, R.T. Rockafellar, A. Sotskov, etc.; however, the most elaborated results of that time were due to M. Valadier (1969); he made use of ε -active indices in taking the supremum of the collection of functions.

The transformation $f \mapsto f^*$ has its origins in a publication of A. Legendre (1752–1833), dated from 1787. Since then, this transformation has received a number of names in the literature: conjugate, polar, maximum transformation, etc. However, it is now generally agreed that an appropriate terminology is *Legendre-Fenchel transform*. In preparing the books with Lemaréchal [12], I remember to have asked by letter L. Hörmander whether the appellation should be *Fenchel transform* or *Legendre-Fenchel transform*; he answered that the name of Legendre should be added to that of Fenchel, which we adopted subsequently. In a letter to C. Kiselman (a colleague from the University of Uppsala, Sweden), dated 1977, W. Fenchel wrote: “*I do not want to add a new name, but if I had to propose one now, I would let myself be guided by analogy and the relation with polarity between convex sets (in dual spaces) and I would call it for example parabolic polarity*”. Fenchel was influenced by his geometric (projective) approach and also by the fact that the “parabolic” function $f(x) = 1/2 \|\cdot\|^2$ is the only one satisfying the relation $f = f^*$.

We have intended to mark this 50th birthday of modern convex analysis by editing a special issue in the Mathematical Programming series B [5].

2 Forty Years Ago: 1971–1973

My first contact with the name of J.-J. Moreau was via his mimeographed lecture notes [20] in the academic year 1971–1972. I was beginning my doctoral studies at the University of Bordeaux, and J.-L. Joly presented his course to us by saying: “These are the notes corresponding to my lectures”, and he gave to each of us a copy. I still have this copy, typed on an old typewriter, by some secretary at Collège de France in Paris (I suppose), comprising my own handwritten annotations. I remember that, with another student next to me, we were impressed by the long list of references authored by J.-J. Moreau and R.T. Rockafellar and posted at the end of lecture notes ([18, 22] references respectively). As beginners in research, we did not know that one of the objectives of researchers in mathematics (the only one?) is to publish as many papers as possible. However, I do not think that the way of publishing at that time was (what is sometimes called) “salami publishing” like it is nowadays. These lecture notes were widely spread in France and elsewhere but never published by an editing house; only in 2003 they were published by a group in Italy (University “Tor Vergata” in Roma). J.-L. Joly was a young professor, just settled in Bordeaux, coming from the University of Grenoble (like others, B. Martinet, A. Auslender, C. Carasso, P.-J. Laurent, C.F. Ducateau, etc.). After some time devoted to convex analysis, he moved to the PDE area. He did some works on convex analysis with P.-J. Laurent; they were presented (some of them exclusively there) in the book entitled “Approximation et Optimisation”, authored by P.-J. Laurent and published in 1972 [15]. I remember exactly when and where (in a bookstore in Bordeaux) I bought this book (students at that time used to buy books, not just photocopy them. . .). I still have this personal copy; the chapters VI (on convex functionals) and VII (on stability and duality in convex optimization) are the most worn ones. This book has been translated into Russian, never into English, I believe. The exam session of June 1972 (a 4-h long written examination) on the lectures in Joly’s course consisted into two parts: the first one was devoted to the construction of some geometrical mean of two convex functions; the second one had for objective to explore the link between “local uniform convexity” of a convex function and the Fréchet-differentiability of its Legendre-Fenchel conjugate. . . A tough exam indeed. I discovered some time later that the matter of the exam was directly taken from a paper by E. Asplund. . . My Master’s thesis was presented in 1972–1973, my first readings of works of mathematical research were those of R.J. Aumann (“Integration of set-valued mappings”) and Z. Artstein (“Set-valued measures”, 1972). I was to cross paths with Z. Artstein several times in my career.

The long papers by A. Ioffe-V. Tikhomirov (Russian Math. Surveys, 1968) and A. Ioffe-V. Levin (Trans. Moscow Math. Soc., 1972), the classical ones by V. F. Demyanov and A. M. Rubinov (1967–1968), were also at our disposal.

R.T. Rockafellar’s book, entitled “Convex Analysis”, was published in 1970. It was quickly spread among interested mathematicians in France. Interestingly enough, this book remains one of the most sold ones in mathematics.

A bit later, in 1974, convex analysis and duality in variational problems were presented in the book by Ekeland and Temam [6], two influential mathematicians from J.-L. Lions’ research group in Paris. The book has been translated into English and Russian.

In those years, techniques and results from convex analysis illuminated several other areas of mathematics: that of monotone operators and PDE (with students and collaborators of H. Brezis), stochastic control theory (Bismut [4] for example), etc.

3 Thirty Years Ago: 1981–1983

I always have been a fan of Russian mathematics. At the end of the 1970s years, I began exchanging letters with B.M. Mordukhovich (Belarus State University in Minsk), colleagues in Kiev (B.N. Pshenichnyi, Yu. Ermoliev, E. Nurminski), and elsewhere. In February 1980, a meeting entitled “Convex Analysis and Optimization” was organized in London in honour of A. Ioffe (Moscow) (see [3]); I presented there (and published in [3]) a survey paper on ε -subdifferential calculus. I like to write survey papers from time to time. Only some years later I had the opportunity to meet (for the first time) B.N. Pshenichnyi, V.F. Demyanov and A. Ioffe; it was at the occasion of these charming meetings organized from time to time in Erice (Sicily).

After doctoral studies under the supervision of A. Auslender and some additional years in Clermont-Ferrand (1973–1981), I moved to Toulouse in September 1981. I left the city of B. Pascal (Clermont-Ferrand) for that of P. Fermat (Toulouse); after all, both lived in the same century, the seventeenth century, the one where the physical notion of motion (velocity, acceleration) was “made mathematics” (birth of differential calculus, tackling extremum problems). In the meantime, between 1973 and 1980, Clarke’s approach of generalized subdifferentials of nonsmooth nonconvex functions had been introduced and solidified. I delivered my first lectures on that subject at the Master level in Toulouse between 1981 and 1983. I also began supervising Ph.D. theses, as it is the role of university professors. The first one, by R. Ellaia (period 1981–1984), was devoted to the analysis and optimization of differences of convex functions [7], a topic I tried to develop and follow for years.

Some years later, in June 1987 precisely, a large meeting on “Applied nonlinear analysis” was organized in Perpignan (extreme south of France), at the occasion of the retirement of J.-J. Moreau¹. With my Ph.D. student Ph. Plazanet, we presented

¹A meeting celebrating the 80th birthday of J.-J. Moreau has been organized later (in 2003), by colleagues in mechanics this time, in Montpellier (where Moreau had spent the major part of his university life). See [1] for more on that.

there and published in [2] a converse to Moreau's theorem, a *factorization theorem* in a way. Since I like this theorem, I reproduce it here.

Theorem (Hiriart-Urruty and Plazanet). *Let g and h be two convex functions defined on a Hilbert space H , satisfying*

$$g(x) + h(x) = \frac{1}{2} \|x\|^2 \text{ for all } x \in H.$$

There then exists a lower-semicontinuous convex function F such that:

$$g = F \diamond \frac{1}{2} \|\cdot\|^2 \text{ and } h = F^* \diamond \frac{1}{2} \|\cdot\|^2.$$

Here, \diamond stands for the infimal convolution operation, and F^* designates the Legendre-Fenchel conjugate of F .

Moreover, an expression of F can be obtained, via g (or h) and $\frac{1}{2} \|\cdot\|^2$, by performing a “deconvolution” of a function by another.

4 Twenty Years Ago: 1993

In 1993 was published the two-volume book co-authored with Lemaréchal [12] (CL in short), final point of 7 years of wrestling with convex analysis, optimization, computers and editing difficulties. We used to call “the HULL” this book (from the initials of our names). So, here is the occasion of some reminiscences of relationships with CL during years. I already told these stories and anecdotes at the occasion of the “CL festchrisft” which took place in Les Houches (Alps region in France) in January 2010 (see [8] for a follow-up as a special issue in the Mathematical Programming series B).

I met CL for the first time in a meeting in the Alps region, during the “Convex analysis days” which took place in January 1974. J.-P. Aubin and P.-J. Laurent were the organizers of this meeting². This was my first international meeting... I remember well that it took place in a charming village called St Pierre-de-Chartreuse and the talks were delivered in a movie theatre or village hall. For me, it was the first time that I saw mathematicians I knew the names (or mathematical results) of: among the 70 participants [confirmed by Laurent (Personal communication, 2010)] were R.T. Rockafellar (who was on sabbatical leave in Grenoble); students or collaborators of H. Brezis (H. Attouch, Ph. Bénéilan, A. Damlamian, etc.), E. Zarantonello, J.-P. Penot, J.-J. Moreau, J.-Ch. Pomerol, M. Valadier, J. Cea, L. Tartar, etc. I remember that I. Ekeland had a pertinent question at each delivered talk.

²Another meeting, with the same title, took place in Murat-Le-Quaire (centre of France) in March 1976; it was organized by A. Auslender. Soon after, in May 1977, a meeting of the same kind, but devoted to nonconvex analysis, was organized in Pau (southwest of France) by J.-P. Penot.

At breakfast, J.-P. Aubin was drinking all the left coffees. I ventured into discussing a bit with M. Valadier (and his inevitable anorak jacket) on the “continuous infimal convolution”. CL was there, a young researcher (just 29 years old) at the research institute called IRIA close to Versailles. The talk by him (in French) was on some “steepest ascent method on the dual function”, the matter of which was written in a research report of IRIA (with a red and white cover). I remember the following anecdote. At the end of the talk, a colleague, the kind of mathematician “who-has-understood-everything-better-and-before-everyone” (you see what I mean), asked CL the following question: “Why do you call that a “steepest ascent”?... I understand that wording only for “steepest descent”... CL answered straight out: “Well... take for example “a deep sky”...” (in French, it is even more striking “méthode de plus profonde montée”, “un ciel profond”... The one who posed the question (I won’t reveal the name) remained speechless... During the lunch, I heard a colleague pursing his lips: “Yeah, we know that some people look for “descent directions”...”. That anecdote leads me to a first theorem.

Theorem 4. *Beware, in meetings, young students or colleagues may be listening to what you are saying... They might remember what you said.*

As a corollary, aimed at beginners.

Corollary 1. *Do not believe that all your colleagues (mathematicians) are fond of mathematics you are doing or theorems you are proving.*

Some of these colleagues just could say: “What you are doing is just routine work, boring...” or “a trivial matter, I can prove it easily”.

About 10 years later, in May 1985, I organized a 1-week long congress in Toulouse, entitled “Mathematics for Optimization”; the main topics of the meeting were variational problems and optimization. Many colleagues came, among the best known ones: P. Ciarlet, J.-P. Aubin, J. Borwein, A. Bensoussan, I. Ekeland, A.B. Kurzhanski, L.C. Young, J. Warga, F. Clarke, B. Dacorogna, J.-P. Penot, R. Temam, H. Tuy, etc. Some participants were there for one of their first meetings abroad their countries: H. Frankowska (Poland and University of Paris IX), M. Lopez and M. Goberna (Valencia, Spain), J.E. Martinez-Legaz (Barcelona), etc. CL was also there, as well as some of our collaborators and colleagues from Chile (R. Correa, A. Jofre). I here would like just to recall the atmosphere during this period, concerning the relationships with other countries, especially with Soviet Union (including Ukraine at that time). Some colleagues from Soviet Union were officially invited: V. Tikhomirov, B.N. Pshenichnyi, V.F. Demyanov... None of them could come, the access to visas was denied. It was typical of what used to happen during those years: you invite (officially) colleagues A, B or C, you get an acknowledgement and answer letter from D, and finally E offers to come... This happened to me several times, especially with Kiev, despite the fact that Kiev and Toulouse are twin cities. I also remember that, during this meeting in Toulouse, a telephone call was organized from J.-P. Aubin, J. Warga, F. Clarke to A. Ioffe (Moscow). All these stories or details are hard to believe nowadays, and yet they took place less than 30 years ago.

CL continued also relationships with colleagues from Soviet Union via IIASA, a research institute close to Vienna in Austria; several meetings on nonsmooth optimization were organized there, with R.T. Rockafellar, R. Wets, B.T. Polyak, Yu. Ermoliev, B.N. Pshenichnyi, V.F. Demyanov, etc.

Another snapshot I would like to offer concerns the two-volume book we wrote together with CL. The initial project was just a 150–200-page book presenting the basics of nonsmooth convex optimization (fundamentals and algorithms); it finally ended up with two volumes, more than 800 pages altogether.

Theorem 5. *If you have a project of writing a book, do not believe it will be finished on time and that its length will be the one you had in mind.*

For the project, I used to go to INRIA (close to Versailles), about one week per month for some years. The INRIA barracks (formerly the NATO headquarters in France) were located in Rocquencourt. The Rocquencourt appellation was known to me because, every morning on the radio news, were evocated the traffic jams at “the triangle of Rocquencourt”. The whole country of France was supposed to be informed of the traffic around this “triangle of Rocquencourt”. So, for me, this triangle was as familiar as the “Bermuda triangle” or the “equilateral triangle”. In CL’s office at INRIA, a large sheet of handwritten paper was posted on the wall, with the list of chapters we had to write for the book project. In front of this office, the one of C. Sagastizabal, doing mathematics on the screen of her computer but also permanently listening at music with her ear flaps. The manuscript of the projected book was written (*he* typed everything) on an Apple Mac⁺ (the screen was just like a stamp!) using Microsoft Word3 and CricketDraw for the pictures. It was then converted to TeX with the help of some home-made code. This took place only about 20 years ago! Here is an excerpt from a letter we exchanged, as the project proceeded: “Like a horse, I feel the smell of the stable, even though the rate of efficiency decreases as and when the tiredness increases”. There is a possible advantage when you write a paper or a book with a co-author (I feel it is difficult to write a book with more than two co-authors, complexity increases a lot, at least that’s my experience), this is what I call the “max rule”: when you are inactive on the project, you may think that your co-author is active. . . so the max of the activities is continuously non-zero.

During that period (around the 1990s), faxes arrived at CL’s office: A. Nemirovski and Yu. Nesterov were organizing their first trips to France (and the West).

A revised printing of the book was published in 1996, but that was not a new edition. Actually, my experience is that a new edition of a book is always. . . augmented, never reduced; C. Byrne from Springer certainly could confirm this statement. J. Dennis commented this statement at the Les Houches meeting in January 2010: “A new edition of a book is always augmented. . . and sometimes worse!”. Later, Springer asked us to write an abridged version of our book, a student version. The project was finalized during a skiing holiday in 2000 in a family house of CL in the Alps. The booklet that we used to call “the soft HULL” was published in 2001. It contained exercises. . . a couple of them are wrong [13].

Despite our numerous exchanges on optimization during years, CL and I never wrote a specific research paper together, except [11]... which remained unpublished.

Here is a further statement that I excerpt from one letter from CL: “I’m sorry but that’s my way of doing research and one has to get used to that. . . I give punches in all directions to find the hole; many may go on the wall but may also be for my unlucky partner”. A final point: CL liked to stud his letters with metaphors or Latin sentences, here is one, a French pun actually, written after some extensive search on properties of the epsilon subdifferential: “Caecum saxa fini (= At this point, that won’t get me anywhere)”.

5 As an Epilogue

I cannot report on all the books (research books or textbooks) written on convex and/or nonsmooth analysis and optimization. The theory and practice are now well established, even if the fields are relatively young if you compare with other fields in applied and/or fundamental mathematics. By experience, I can say that (advanced) students like the geometry and elegance of topics such as Moreau’s decomposition in Hilbert spaces (a typical illustration of techniques in convex analysis). Tools from nonsmooth analysis are now used to handle *nonconvex* variational problems; they can be considered as “basics” when beginning to study variational analysis and optimization. This was precisely the aim of my latest (published) lecture notes on the subject [10].

References

- [1] Alart, P., Maisonneuve, O., Rockafellar, R.T. (eds.): *Nonsmooth Mechanics and Analysis: Theoretical and Numerical Advances*. Springer, New York (2006)
- [2] Attouch, H., Aubin, J.-P., Clarke, F., Ekeland, I. (eds.): *Analyse Non linéaire*. CRM université de Montréal et Gauthier-Villars, Paris (1989)
- [3] Aubin, J.-P., Vinter, R. (eds.): *Convex Analysis and Optimization*. Pitman publishers, London (1982)
- [4] Bismut, J.-M.: Conjugate functions in optimal stochastic control. *J. Math. Anal. Appl.* **44**, 384–404 (1973)
- [5] Combettes, P.L., Hiriart-Urruty, J.-B., Théra, M. (eds.): *Modern convex analysis and its applications*. Special issue *Math. Program. ser. B* (2014, to appear)
- [6] Ekeland, I., Temam, R.: *Analyse convexe et Problèmes variationnels*. Dunod et Gauthier-Villars, Paris (1974)
- [7] Ellaia, R.: *Contributions à l’analyse et l’optimisation de différences de fonctions convexes*. Ph.D. thesis, Paul Sabatier University Toulouse (1984)
- [8] Frangioni, A., Overton, M.L., Sagastizabal, C. (eds.): *Convex analysis, optimization and applications*. Special issue *Math. Program. ser. B* **140**(1), (2013)
- [9] Hiriart-Urruty, J.-B.: Reminiscences... and a little more. *J. Set-Valued Var. Anal.* **18**(3–4), 237–249 (2010)

- [10] Hiriart-Urruty, J.-B.: Bases, Outils et Principes pour l'analyse Variationnelle. SMAI Series. Springer, New York (2012)
- [11] Hiriart-Urruty, J.-B., Lemaréchal, C.: Testing necessary and sufficient conditions for global optimality in the problem of maximizing a convex quadratic function over a convex polyhedron. Seminar of Numerical Analysis. University Paul Sabatier, Toulouse (1990)
- [12] Hiriart-Urruty, J.-B., Lemaréchal, C.: Convex Analysis and Minimization Algorithms, vol. I and II. Springer, Berlin (1993). New printing in 1996
- [13] Hiriart-Urruty, J.-B., Lemaréchal, C.: Fundamentals of Convex Analysis, Text editions. Springer, New York (2001)
- [14] Hörmander, L.: Sur la fonction d'appui des ensembles convexes dans un espace localement convexe. Arkiv för matematik **3**, 181–186 (1954)
- [15] Laurent, P.-J.: Approximation et Optimisation. Hermann, Paris (1972)
- [16] Moreau, J.-J.: Fonctions convexes duales et points proximaux dans un espace hilbertien. Note aux C.R. Acad. Sci. Paris **255**, 2897–2899 (1962)
- [17] Moreau, J.-J.: Propriétés des applications “Prox”. Note aux C.R. Acad. Sci. Paris **256**, 1069–1071 (1963)
- [18] Moreau, J.-J.: Fonctions sous-différentiables. Note aux C.R. Acad. Sci. Paris **257**, 4117–4119 (1963)
- [19] Moreau, J.-J.: Proximité et dualité dans un espace hilbertien. Bull. Soc. Math. France **93**, 273–299 (1965)
- [20] Moreau, J.-J.: Fonctionnelles convexes. Séminaire sur les équations aux dérivées Partielles, Collège de France, 1966–1967. Published in 2003 by the university “Tor Vergata” in Roma
- [21] Rockafellar, R.T.: Convex functions and dual extremum problems. Ph.D. thesis, University Paul Sabatier, Harvard (1963)