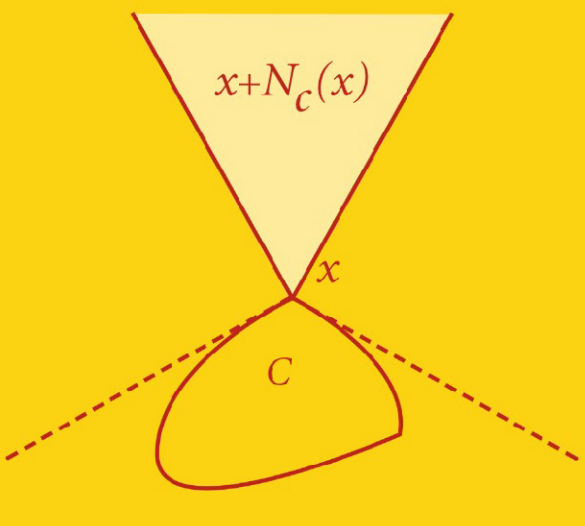
CMS Books in Mathematics

Roberto Lucchetti

Convexity and Well-Posed Problems



Canadian Mathematical Society
Société mathématique du Canada

Roberto Lucchetti

Convexity and Well-Posed Problems

With 46 Figures

Springer

Roberto Lucchetti
Dipto. Matematica
Politecnico di Milano
Milano, 20133
Italy
rel@como.polimi.it

Editors-in-Chief
Rédacteurs-en-chef
Jonathan Borwein
Karl Dilcher
Department of Mathematics and Statistics
Dalhousie University
Halifax, Nova Scotia B3H 3J5
Canada
cbs-editors@cms.math.ca

Mathematics Subject Classification (2000): 49-01, 46N10, 26B25, 90-01

Library of Congress Control Number: 2005932085

ISBN 978-0387-28719-5

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

springeronline.com

Dedicated to my family, pets included.

Contents

Preface

This book deals mainly with the study of convex functions and their behavior from the point of view of stability with respect to perturbations. We shall consider convex functions from the most modern point of view: a function is defined to be *convex* whenever its epigraph, the set of the points lying above the graph, is a convex set. Thus many of its properties can be seen also as properties of a certain convex set related to it. Moreover, we shall consider extended real valued functions, i.e., functions taking possibly the values $-\infty$ and $+\infty$. The reason for considering the value $+\infty$ is the powerful device of including the constraint set of a constrained minimum problem into the objective function itself (by redefining it as $+\infty$ outside the constraint set). Except for trivial cases, the minimum value must be taken at a point where the function is not $+\infty$, hence at a point in the constraint set. And the value $-\infty$ is allowed because useful operations, such as the inf-convolution, can give rise to functions valued $-\infty$ even when the primitive objects are real valued.

Observe that defining the objective function to be $+\infty$ outside the closed constraint set preserves lower semicontinuity, which is the pivotal and minimal continuity assumption one needs when dealing with minimum problems. Variational calculus is usually based on derivatives. In the convex case too, of course, the study of the derivative is of the utmost importance in the analysis of the problems. But another concept naturally arises, which is a very important tool for the analysis. This is the *subdifferential* of a function at a given point x, which, as opposed to the derivative, does not require the function to be finite on a whole ball around x. It also exists when the graph of the function has angles, and preserves many important properties of the derivatives. Thus a chapter is dedicated to the study of some properties of the subdifferential: its connections with the directional derivatives and the Gâteaux and Fréchet differentials whenever they exist, and its behavior as a multifunction. The following chapter, after introducing the most fundamental existence theorem in minimum problems, the Weierstrass theorem, is dedicated to the Ekeland variational principle which, among other things, establishes, for a very general class $\mathcal{F}$ of functions (lower semicontinuous, lower bounded) defined on a complete metric space X, an existence theorem on a dense (for a natural topology on $\mathcal{F}$) set. This gives a way around lack of a topology on X, and allows for application of the Weierstrass theorem. We also analyze in some detail some of the very interesting consequences of the principle, mainly in the convex setting.

Next, we introduce the fundamental operation of Fenchel conjugation. This is the basis of all the duality theory which we develop, essentially following the approach of Ekeland–Temam (see [ET]). We then give a representative number of examples of its applications, including zero sum games, including the beautiful proof of the famous von Neumann theorem on the existence of an equilibrium in mixed strategies for finite games. This also allows us to get interesting results for linear programming. I want to stress at this point that, notwithstanding that the minimization of a scalar convex function is the primary subject of study of this book, the basic underlying concept that motivated me to write it is "optimization". For this reason, I include in it some game theory, one of the most modern and challenging aspects of optimization, with a glance as well to vector optimization. My hope is that readers will be stimulated and encouraged to bring the ideas, developed here for the convex, extended real valued functions, (mainly stability and well-posedness) to these domains too. To this end I must however say that some research is already in progress in this direction, although it is not so well established as to have a place in this book.

Coming back to the content of the book, I have to mention that my primary goal is to illustrate the ideas of stability and well-posedness, mainly in the convex case. Stability means that the basic parameters of a minimum problem, the infimal value and the set of the minimizers, do not vary much if we slightly change the initial data, the objective function and the constraint set. On the other hand, well-posedness means that points with values close to the value of the problem must be close to actual solutions. In studying this, one is naturally led to consider perturbations of functions and of sets. But it turns out that neither traditional convergences of functions, pointwise convergence, compact-open topology, nor classical convergence of sets, Hausdorff and Vietoris, are well suited to our setting. The stability issue explains why scholars of optimization have devoted so much time to defining and studying various convergence structures on the space of closed subsets of a metric space. Moreover, this approach perfectly fits with the idea of regarding functions as sets. Thus beginning with Chapter 8, the second part of the book starts with an introduction to the basic material concerning convergence of the closed subsets of a metric space X, and the topological nature of these convergences. These topologies are usually called hypertopologies, in the sense that the space X can be embedded in the hyperspace (whose points are closed sets), and the topology in the hyperspace respects the topology of X. A sequence $\{x_n\}$ in X converges in X if and only if the sequence of sets $\{\{x_n\}\}$ converges in the hyperspace. Since this topic appears to be interesting in itself, Appendix B is dedicated to exploring in more detail some basic ideas underlying the construction and study of these topologies/convergences, but it is not necessary to the comprehension of the rest of the book.

Using these topologies requires also knowing the continuity of basic operations involving them. For instance, when identifying functions with sets, it is not clear (nor even true) whether the sum of two convergent (in some particular sense) sequences converges to the sum of the limits. Yet having this property is very fundamental, for instance to ensure a good Lagrange multipliers rule in constrained problems. Thus, Chapter 9 is dedicated to this issue.

We then turn our attention to the study of well-posed problems, and the connection between stability and well-posedness. In doing this, we give some emphasis to a very recent and fruitful new well-posedness concept, which in some sense contains at the same time the two classical notions of stability and Tykhonov well-posedness.

Since there are many important classes of minimization problems for which existence cannot be guaranteed universally for all elements of the class, it is interesting to know "how many" of these problems will have solutions and also enjoy the property of being well-posed. This is the subject of Chapter 11. We consider here the idea of "many" from the point of view of the Baire category, and in the sense of σ-porosity, a recent and interesting notion which provides more refined results than the Baire approach. This part contains the most recent results in the book, and is mainly based on some papers by Ioffe, Revalski and myself.

The book ends with some appendices, entitled "Functional analysis" (a quick review of the Hahn–Banach theorem and the Banach–Dieudonné–Krein–Smulian theorem), "Topology" (the theorem of Baire, and a deeper insight to hypertopologies) and "More game theory".

A few words on the structure of the book. The part on convexity is standard, and much of the inspiration is taken from the classical and beautiful books cited in the References, such as those by Ekeland–Temam, Rockafellar, Phelps, and Lemaréchal–Hiriart-Urruty. I also quote more recent and equally interesting books, such as those of Borwein–Lewis and of Zalinescu. The study of hypertopologies is instead a less classical issue, the only book available is the one by G. Beer [Be]. However my point of view here is different from his and I hope that, though very condensed, this section will help people unfamiliar with hypertopologies to learn how to use them in the context of optimization problems. Finally, the sections related to stability have roots in the book by Dontchev–Zolezzi, but here we focus mainly on convexity.

About the (short) bibliography, I should emphasize that, as far as the first part is concerned, I do not quote references to original papers, since most of the results which are presented are now classical; thus I only mention the most important books in the area, and I refer the reader to them for a more complete bibliography. The references for hypertopologies and classical notions of well-posedness are the books by [Be],[DZ] respectively. When dealing with more recent results, which are not yet available in a book, I quote the original papers. Finally, the section concerning game theory developed in the duality chapter is inspired by [Ow].

The book contains more than 120 exercises, and some 45 figures. The exercises, which are an essential part of this work, are not all of the same level of difficulty. Some are suitable for students, while others are statements one can find in recent papers. This does not mean that I consider these results to be straightforward. I have merely used the exercise form to establish some interesting facts worth mentioning but whose proof was inessential to a reading of the book. I have chosen to start each chapter with one of my favorite quotations, with no attempt to tie the quote directly to the chapter.

Since this is my first and last book of this type, I would like to make several acknowledgements. First of all, I want to thank all my coauthors. I have learned much from all of them, in particular, A. Ioffe and J. Revalski. Most of the material concerning the genericity results is taken from some of their most recent papers with me. More importantly, I am very happy to share with them a friendship going far beyond the pleasure of writing papers together. For several years these notes were used to teach a class at the Department of Mathematics and Physics at the Catholic University of Brescia, and a graduate class held at the Faculty of Economics at the University of Pavia. I would like to thank my colleagues M. Degiovanni and A. Guerraggio for inviting me to teach these classes, and all students (in particular I want to mention Alessandro Giacomini) who patiently helped me in greatly improving the material, and correcting misprints. I also wish to thank some colleagues whom I asked to comment on parts of the book, in particular G. Beer, who provided me with some excellent remarks on the chapters dedicated to hypertopologies. Also, comments by the series editors J. Borwein and K. Dilcher to improve the final version of the book were greatly appreciated. I owe thanks to Mary Peverelli and Elisa Zanellati for undertaking the big task of outlining figures copied from my horrible and incomprehensible drawings. Last but not least, I would like to express my appreciation for an invitation from CNRS to spend three months at the University of Limoges, attached to LACO. The nice, quiet and friendly atmosphere of the department allowed me to complete the revision of all material. In particular, I thank my host M. Théra, and the director of the LACO, A. Movahhedi.

While going over the book for the last time, I learned of the passing away of my friend and colleague Jan Pelant. A great man and a great mathematician, his loss hurts me and all who had the good fortune to meet and know him.

1

Convex sets and convex functions: the fundamentals

Nobody realizes that some people expend
a tremendous amount of energy
merely to be normal.
(A. Camus)

In this first chapter we introduce the basic objects of this book: convex sets and convex functions. For sets, we provide the notions of convex set, convex cone, the convex, conic and affine hulls of a set, and the recession cone. All these objects are very useful in highlighting interesting properties of convex sets. For instance, we see that a closed convex set, in finite dimensions, is the closure of its relative interior, and we provide a sufficient condition in order that the sum of two closed convex sets be closed, without using any compactness assumption. To conclude the introduction of these basic geometric objects of the convex analysis, we take a look at the important theorems by Carathéodory, Radon and Helly.

We then introduce the idea of extended real valued convex function, mainly from a geometric point of view. We provide several important examples of convex functions and see what type of operations between functions preserve convexity. We also introduce the very important operation of inf-convolution.

In this introductory chapter we mainly focus on the geometry of convexity, while in the second chapter we shall begin to consider the continuity properties of the extended real valued convex functions.

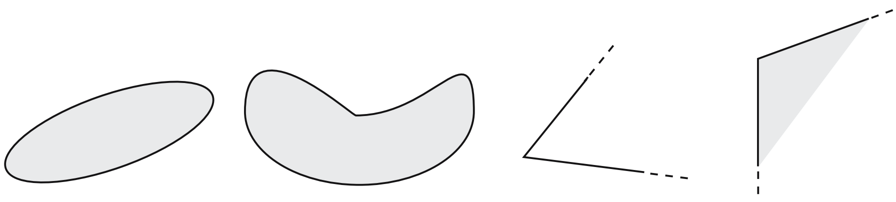
1.1 Convex sets: basic definitions and properties

Let X be a linear space and C a subset of X.

Definition 1.1.1 C is said to be *convex* provided

$$x, y \in C, \lambda \in (0,1) \text{ imply } \lambda x + (1-\lambda) y \in C.$$

The empty set is assumed to be convex by definition. C is a *cone* if $x \in C$, $\lambda \geq 0$ imply $\lambda x \in C$.



Convex set. Nonconvex set. Cone. Convex cone.

Figure 1.1.

Exercise 1.1.2 A cone is convex if and only if $x, y \in C$ implies $x + y \in C$.

For sets A, C and for $t \in \mathbb{R}$, we set

$$A + C := \{a + c : a \in A, c \in C\}, \quad tA := \{ta : a \in A, t \in \mathbb{R}\}.$$

Exercise 1.1.3 Let A, C be convex (cones). Then $A + C$ and tA are convex (cones). Also, if C_α is an arbitrary family of convex sets (convex cones), then $\bigcap_\alpha C_\alpha$ is a convex set (convex cone). If X, Y are linear spaces, $L\colon X \to Y$ a linear operator, and C is a convex set (cone), then $L(C)$ is a convex set (cone). The same holds for inverse images.

Definition 1.1.4 We shall call a *convex combination* of elements $x_1, \dots, x_n$ any vector x of the form

$$x = \lambda_1 x_1 + \cdots + \lambda_n x_n,$$

with $\lambda_1 \geq 0, \dots, \lambda_n \geq 0$ and $\sum_{i=1}^n \lambda_i = 1$.

We now see that a set C is convex if and only if it contains any convex combination of elements belonging to it.

Proposition 1.1.5 *A set C is convex if and only if for every $\lambda_1 \geq 0, \dots, \lambda_n \geq 0$ such that $\sum_{i=1}^n \lambda_i = 1$, for every $c_1, \dots, c_n \in C$, for all n, then $\sum_{i=1}^n \lambda_i c_i \in C$.*

Proof. Let

$$A = \left\{ \sum_{i=1}^n \lambda_i c_i \ : \ \lambda_i \geq 0, \sum_i \lambda_i = 1, c_i \in C \ \forall i, n \in \mathbb{R} \right\}.$$

We must prove that $A = C$ if and only if C is convex. Observe that A contains C. Next, A is convex. This is very easy to see, and tedious to write, and so we omit it. Thus the proof will be concluded once we show that $A \subset C$ provided C is convex. Take an element $x \in A$. Then

$$x = \sum_{i=1}^{n} \lambda_i c_i,$$

with $\lambda_i \geq 0, \sum_i \lambda_i = 1, c_i \in C$. If $n = 2$, then $x \in C$ just by definition of convexity. Suppose now $n > 2$ and that the statement is true for any convex combination of (at most) $n-1$ elements. Then

$$x = \lambda_1 c_1 + \cdots + \lambda_n c_n = \lambda_1 c_1 + (1 - \lambda_1) y,$$

where

$$y = \frac{\lambda_2}{1-\lambda_1} c_2 + \cdots + \frac{\lambda_n}{1-\lambda_1} c_n.$$

Now observe that y is a convex combination of $n-1$ elements of C and thus, by inductive assumption, it belongs to C. Then $x \in C$ as it is a convex combination of two elements. $\square$

If C is not convex, then there is a smallest convex set (convex cone) containing C: it is the intersection of all convex sets (convex cones) containing C.

Definition 1.1.6 The *convex hull* of a set C, denoted by $\operatorname{co} C$, is defined as

$$\operatorname{co} C := \bigcap \{A : C \subset A,\ A \text{ is convex}\}.$$

The *conic hull* denoted by $\operatorname{cone} C$, is

$$\operatorname{cone} C := \bigcap \{A : C \subset A,\ A \text{ is a convex cone}\}.$$

Proposition 1.1.7 *Given a set C,*

$$\operatorname{co} C = \left\{ \sum_{i=1}^{n} \lambda_i c_i : \ \lambda_i \geq 0, \sum_{i=1}^{n} \lambda_i = 1, c_i \in C \ \forall i, n \in \mathbb{R} \right\}.$$

Proof. It easily follows from Proposition 1.1.5. $\square$

Definition 1.1.8 Let A be a convex set. A point $x \in A$ is said to be an *extreme point* of A if it is not the middle point of a segment contained in A. A *simplex* S is the convex hull of a finite number of points $x_1, \dots, x_k$.

Exercise 1.1.9 Given a simplex S as in the above definition, show that the extreme points of S are a subset of $\{x_1, \dots, x_k\}$.