

Introduction

The primary aim of this book is to present the conjugate and subdifferential calculus using the method of perturbation functions in order to obtain the most general results in this field. The secondary aim is to give important applications of this calculus and of the properties of convex functions. Such applications are: the study of well-conditioned convex functions, uniformly convex and uniformly smooth convex functions, best approximation problems, characterizations of convexity, the study of the sets of weak sharp minima, well-behaved functions and the existence of global error bounds for convex inequalities, as well as the study of monotone multifunctions by using convex functions.

The method of perturbation functions is based on the “fundamental duality theorem” which says that under certain conditions one has

$$\inf_{x \in X} \Phi(x, 0) = \max_{y^* \in Y^*} (-\Phi^*(0, y^*)). \quad (\text{FDF})$$

For many problems in convex optimization one can associate a useful perturbation function. We give here four examples; see [Rockafellar (1974)] for other interesting ones.

Example 1 (Convex programming; see Section 2.9) Let $f, g_1, \dots, g_n : X \rightarrow \overline{\mathbb{R}}$ be proper convex functions with $\text{dom } f \cap \bigcap_{i=1}^n \text{dom } g_i \neq \emptyset$. The problem of minimizing $f(x)$ over the set of those $x \in X$ satisfying $g_i(x) \leq 0$ for all $i = 1, \dots, n$ is equivalent to the minimization of $\Phi(x, 0)$ for $x \in X$, where

$$\Phi : X \times Y \rightarrow \overline{\mathbb{R}}, \quad \Phi(x, y) := \begin{cases} f(x) & \text{if } g_i(x) \leq y_i \ \forall 1 \leq i \leq n, \\ +\infty & \text{otherwise,} \end{cases}$$

and $Y := \mathbb{R}^n$; the element y^* obtained from the right-hand side of (FDF) will furnish the Lagrange multipliers.

Example 2 (Control problems) Let $F : X \times Y \rightarrow \overline{\mathbb{R}}$ be a proper convex function and $A : X \rightarrow Y$ a linear operator. A control problem (in its abstract form) is to minimize $F(x, y)$ for $x \in X$ and $y = Ax + y_0$. The perturbation function to be considered is $\Phi : X \times Y \rightarrow \overline{\mathbb{R}}$ defined by $\Phi(x, y) := F(x, Ax + y_0 + y)$.

Example 3 (Semi-infinite programming) We are as in Example 1 but $\{1, \dots, n\}$ is replaced by a general nonempty set I . In this case $Y = \mathbb{R}^I$ and $\Phi(x, y) := f(x)$ if $g_i(x) \leq y_i$ for all $i \in I$, $\Phi(x, y) := \infty$ otherwise.

Formula (FDF), or more precisely the Fenchel–Rockafellar duality formula, can also be used for deriving results similar to that in the next example.

Example 4 ([Simons (1998b)]; see Exercise 2.37) Let X be a linear space, $(Y, \|\cdot\|)$ be a normed linear space, $A : X \rightarrow Y$ be a linear operator, $y_0 \in Y$ and $f : X \rightarrow \overline{\mathbb{R}}$ be a proper convex function. Then $f(x) + \|Ax + y_0\|^2 \geq 0$ for all $x \in X$ if and only if there exists $y^* \in Y^*$ such that $f(x) - 2\langle Tx + y_0, y^* \rangle - \|y^*\|^2 \geq 0$ for all $x \in X$.

It is worth mentioning that adequate perturbation functions can be used for deriving formulas for the conjugate and ε -subdifferential for many types of convex functions; this method is used by Rockafellar (1974) for $f \circ A$ and $f_1 + \dots + f_n$, but we use it for almost all the operations which preserve convexity (see Section 2.8).

The formula (FDF) is automatically valid when $\inf_{x \in X} \Phi(x, 0) = -\infty$ and is equivalent to the subdifferentiability at $0 \in Y$ of the marginal function $h : Y \rightarrow \overline{\mathbb{R}}$, $h(y) := \inf_{x \in X} \Phi(x, y)$, when $\inf_{x \in X} \Phi(x, 0) \in \mathbb{R}$. A sufficient condition for this is the continuity of the restriction of h to the affine hull of its domain at 0; note that 0 is in the relative algebraic interior of the domain of h in this case (without this condition one can give simple examples in which the subdifferential of h at 0 is empty).

Considering the multifunction $\mathcal{R} : X \times \mathbb{R} \rightrightarrows Y$ whose graph is the set $\text{gr } \mathcal{R} = \{(x, t, y) \mid (x, y, t) \in \text{epi } \Phi\}$, the continuity of $h|_{\text{aff}(\text{dom } h)}$ at 0 is ensured if \mathcal{R} is relatively open at some (x_0, t_0) with $(x_0, t_0, 0) \in \text{gr } \mathcal{R}$. This fact was observed for the first time by Robinson (1976). This remark shows the importance of open mapping theorems for convex multifunctions in convex analysis. In Banach spaces such a result is the well-known Robinson–

Ursescu theorem. The preceding examples show that the consideration of more general spaces is natural: In Example 3 Y is a locally convex space while in Example 4 X can be endowed with the topology $\sigma(X, X')$. The original result of Ursescu (1975) is stated in very general topological vector spaces. The inconvenient of Ursescu's theorem is that one asks the multifunction to be closed, condition which is quite strong in certain situations. For example, when calculating the conjugate or subdifferential of $\max(f, g)$ with f, g proper lower semicontinuous convex functions one has to evaluate conjugate or the subdifferential of $0 \cdot f + 1 \cdot g$ which is not lower semicontinuous convex. Fortunately we dispose of another open mapping theorem in which the closedness condition is replaced by a weaker one, but one must pay for this by asking (slightly) more on the spaces involved.

As said above, the second aim of the book is to give some interesting applications of conjugate and subdifferential calculus, less treated in other books.

In many algorithms for the minimization problem (P) $\min f(x)$, s.t. $x \in X$, one obtains a sequence (x_n) which is minimizing (*i.e.* $(f(x_n)) \rightarrow \inf f$) or stationary (*i.e.* $(d_{\partial f(x_n)}(0)) \rightarrow 0$). It is important to know if such a sequence converges to a solution of (P). Assuming that $S := \operatorname{argmin} f := \{x \mid f(x) = \inf f\} \neq \emptyset$, one says that f is well-conditioned if $(d_S(x_n)) \rightarrow 0$ whenever (x_n) is a minimizing sequence, and f is well-behaved (asymptotically) if (x_n) is minimizing whenever (x_n) is a stationary sequence; when S is a singleton well-conditioning reduces to the well-known notion of well-posedness in the sense of Tikhonov. If f is well-conditioned with linear rate the set $\operatorname{argmin} f$ is a set of weak sharp minima. When f is convex, we establish several characterizations of well-conditioning using the conjugate or the subdifferential of f . When S is a singleton one of the characterizations is close to uniform convexity of f at a point.

One says that the proper function $f : (X, \|\cdot\|) \rightarrow \overline{\mathbb{R}}$ is strongly convex if

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) - \frac{c}{2}\lambda(1 - \lambda)\|x - y\|^2$$

for some $c > 0$ and for all $x, y \in \operatorname{dom} f$, $\lambda \in [0, 1]$. This notion is not very adequate for non-Hilbert spaces; for general normed spaces, one says that f is uniformly convex if there exists $\rho : \mathbb{R}_+ \rightarrow \overline{\mathbb{R}}_+$ with $\rho(0) = 0$ such that

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) - \lambda(1 - \lambda)\rho(\|x - y\|)$$

for all $x, y \in \operatorname{dom} f$ and all $\lambda \in [0, 1]$. From a numerical point of view

the class of uniformly convex functions is important because on a Banach space every uniformly convex and lower semicontinuous proper function has a unique minimum point and the corresponding minimization problem is well-conditioned. It turns out that uniformly convex functions have very nice characterizations using their conjugates and subdifferentials. The dual notion for uniformly convex function is that of uniformly smooth convex function. An important fact is that f is uniformly convex (uniformly smooth) if and only if f^* is uniformly smooth (uniformly convex).

Another interesting application of convex analysis is in the study of monotone operators. This became possible by using a convex function associated to a multifunction introduced by M. Coodey and S. Simons. So, one obtains quite easily characterizations of maximal monotone operators, local boundedness of monotone operators and maximal monotonicity of the sum of two maximal monotone operators using continuity properties of convex functions, the formula for the subdifferential of a sum of convex functions and a minimax theorem (whose proof is also included).

A more detailed presentation of the book follows.

The book is divided into three chapters, every chapter ending with exercises and bibliographical notes; there are more than 80 exercises. It also includes the complete solutions of the exercises, the bibliography, the list of notations and the index of terms.

No prior knowledge of convex analysis is assumed, but basic knowledge of topology, linear spaces, topological (locally convex) linear spaces and normed spaces is needed.

In Chapter 1, as a preliminary, we introduce the notions and results of functional analysis we need in the rest of the book. For easy reference, in Section 1.1 we recall several notions, notations and results (without proofs) which can be found in almost all books on functional analysis; let us mention four separation theorems for convex sets, the Dieudonné and Alaoglu–Bourbaki theorems, as well as the bipolar theorem.

In Section 1.2 we introduce cs-closed, cs-complete, lcs-closed (*i.e.* lower cs-closed) and ideally convex, bcs-complete, li-convex (*i.e.* lower ideally convex) sets and prove several results concerning them. We point out the good stability properties of li-convex and lcs-closed sets. We also introduce two conditions denoted (Hx) and (Hwx) which refer to sets in product spaces that are stronger than the cs-closedness and ideal convexity, but weaker than the cs-completeness and bcs-completeness of the sets, respectively. Then, besides the classical algebraic interior A^i and relative algebraic in-

terior iA of a subset A of a linear space X , we introduce, when X is a topological vector space, the sets ${}^{ic}A$ and ${}^{ib}A$, which reduce to iA when the affine hull $\text{aff } A$ of A is closed or barreled, respectively, and are the empty set otherwise. The quasi interior of a set and united sets are also studied.

In Section 1.3 we state and prove the famous Ursescu's theorem as well as a slight amelioration of Simons' open mapping theorem. As application of these results one reobtain the Banach–Steinhaus theorem and the closed graph theorem as well as two results of O. Cârjă which are useful in controllability problems. Because the notions (with the exception of *cs*-closed and ideally convex sets) and results from Sections 1.2 and 1.3 are not treated in many books (to our knowledge only [Kusraev and Kutateladze (1995)] contains some similar material), we give complete proofs of the results.

The chapter ends with Section 1.4 in which we state and prove the Ekeland's variational principle, the smooth variational principle of Borwein and Preiss, as well as two (dual) results of Ursescu which generalize Baire's theorem.

Chapter 2 is dedicated, mainly, to conjugate and ε -subdifferential calculus. Because no prior knowledge of convex analysis is assumed, we introduce in Section 2.1 convex functions, give several characterizations using the epigraph, or the gradients in case of differentiability, point out the operations which preserve convexity and study the important class of convex functions of one variable; the existence of the (ε -)directional derivative and some of its properties are also studied. We close this section with a characterization of convex functions using the upper Dini directional derivative.

Section 2.2 is dedicated to the study of continuity properties of convex functions. To the classes of sets introduced in Section 1.2 correspond *cs*-closed, *cs*-complete, *lcs*-closed, ideally convex, *bcs*-complete and *li*-convex functions. We mention the fact that almost all operations which preserve convexity also preserve the *lcs*-closedness and the *li*-convexity of functions as seen in Proposition 2.2.19. The most part of the results of this section are not present in other books; among them we mention the result on the convexity of a quasiconvex positively homogeneous function and the results on *cs*-closed, *cs*-complete, *cs*-convex, *lcs*-closed, ideally convex, *bcs*-complete and *li*-convex functions.

Section 2.3 concerns conjugate functions; all the results are classical.

Section 2.4 is dedicated to the introduction and study of direct properties of the subdifferential. Using such properties one obtains easily the for-

mulas for the subdifferentials of Af and $f_1 \square \cdots \square f_n$ which are valid without additional hypothesis. The classical theorem which states that the ε -subdifferential of a proper convex function is nonempty and w^* -compact at a continuity point of its domain, as well as the formula for the ε -directional derivative as the support function of the ε -subdifferential is also established. The less classical result which states that the same formula holds for $\varepsilon > 0$ when the function is not necessarily continuous (but is lower semi-continuous) is established, too. We mention also Theorem 2.4.14 related to the subdifferential of sublinear functions; some of its statements are not very spread. Other interesting results are introduced for completeness or further use.

In Section 2.5 we introduce the general problem of convex programming and establish sufficient conditions for the existence and uniqueness of solutions, respectively. We mention especially Theorems 2.5.2 and 2.5.5; Theorem 2.5.2 ameliorates a result of Polyak (1966), which shows that the reflexivity of the space, needed in proving the existence of solutions, is almost necessary, while Theorem 2.5.5 shows that the coercivity condition is essential for the existence of solutions.

Section 2.6 is dedicated to perturbed functions. One introduces primal and dual problems, the marginal function, and give some direct properties of them. Then one obtains the formula for the ε -subdifferential of the marginal function using the $(\varepsilon + \eta)$ -subdifferentials (with $\eta > 0$) of the perturbed function. Applying this result one obtains formulas for the ε -subdifferential of several types of convex functions.

In the main result of Section 2.7 we provide nine (non-independent) sufficient conditions which ensure the validity of the fundamental duality formula (FDF). The most known of them is that $(x_0, 0) \in \text{dom } \Phi$ and $\Phi(x_0, \cdot)$ is continuous at 0 for some $x_0 \in X$. For the proof of the sufficiency of some conditions one uses the open mapping theorems established in Section 1.3. A related result involves also a convex multifunction; this will be useful for obtaining the formulas for the conjugate and the ε -subdifferential of a function of the forms $g \circ A$ with A a densely defined and closed linear operator and of $g \circ H$ with g being increasing and H convex.

Section 2.8 is dedicated entirely to conjugate and ε -subdifferential calculus for convex functions. The considered functions φ have the form: $\varphi(x) = F(x, A(x))$ and $\varphi = f + g \circ A$ with A a continuous linear operator, $\varphi = f + g$, $\varphi(x) = \inf\{g(y) \mid y \in \mathcal{C}(x)\}$ with \mathcal{C} a convex process, $\varphi = g \circ H$ with H a convex operator and g an increasing convex func-

tion, $\varphi = \max\{f_1, \dots, f_n\}$ and $\varphi = f_1 \diamond f_2$. Besides classical conditions one points out very recent ones. For the proof one constructs an adequate perturbation function and uses the fundamental duality theorem.

In Section 2.9 we apply the fundamental duality theorem for obtaining necessary and sufficient optimality conditions in convex optimization problems with constraints. These conditions involve the subdifferentials of the functions considered or the corresponding Lagrangian. The results are well-known. However we mention the formula for the normal cone to a level set stated in Corollary 2.9.5 for not necessarily finite-valued functions which is quite new.

The minimax theorem presented in Section 2.10 will be used in the section dedicated to monotone multifunctions.

Throughout Chapter 3 the involved spaces are normed spaces. In Section 3.1 besides the classical theorems of Borwein, Brøndsted–Rockafellar, Bishop–Phelps and Rockafellar (on the maximal monotonicity of the subdifferential of a convex function) we present a recent theorem of Simons and use it for a very short proof of Rockafellar’s theorem (mentioned before). As a consequence of the Brøndsted–Rockafellar theorem we obtain other three conditions for the validity of the formulas for the conjugate and subdifferential of the function $F(\cdot, A(\cdot))$ (and therefore for the functions $f + g \circ A$ and $f + g$).

The aim of Section 3.2 is to characterize the convex functions using other types of subdifferentials. In fact we use an abstract subdifferential. An example of such subdifferential is Clarke’s one for which we establish several properties. The main tool for such characterizations is the well-known Zagrodny’s approximate mean value theorem; the version we present subsumes several results met in the literature. We present also an integration theorem of Thibault and Zagrodny which yields the fact that two lower semicontinuous convex functions on a Banach space which has the same Fenchel subdifferential coincide up to an additive constant.

In Section 3.3 we introduce the class \mathcal{A} of functions $\varphi : \mathbb{R}_+ \rightarrow \overline{\mathbb{R}}_+$ with $\varphi(0) = 0$ and several useful subclasses. To any $\varphi \in \mathcal{A}$ we associate $\varphi^\# \in \mathcal{A}$ defined by $\varphi^\#(t) = \sup\{ts - \varphi(s) \mid s \geq 0\}$. These classes of functions turn out to be useful in studying well-conditioned convex functions, uniformly convex and uniformly smooth convex functions, as well as in the study of the geometry of normed spaces. As an illustration of the use of these classes of functions we study the differentiability of convex functions with

respect to arbitrary bornologies. Using one of the characterizations and the Brøndsted–Rockafellar theorem one obtains the following interesting result of Asplund and Rockafellar: Let X be a Banach space and $f : X \rightarrow \overline{\mathbb{R}}$ a proper lower semicontinuous convex function; if f^* is Fréchet differentiable at $\bar{x}^* \in \text{int}(\text{dom } f^*)$ then $\nabla f^*(\bar{x}^*) \in X$.

In Section 3.4 we introduce the well-conditioned convex functions and give several characterizations of this notion using the conjugate and the subdifferential of the function. An important special case is that of well-conditioning with linear rate. This situation is studied in Section 3.10.

In Section 3.5 we study uniformly convex and uniformly smooth convex functions, respectively. To any convex function f one associates the gages ρ_f and σ_f of uniform convexity and uniform smoothness, respectively. The gage ρ_f has an important property: the mapping $0 < t \mapsto t^{-2}\rho_f(t)$ is nondecreasing. Because $\sigma_{f^*} = (\rho_f)^\#$ and $\sigma_f = (\rho_{f^*})^\#$ for any proper lower semicontinuous convex function f , the mapping $0 < t \mapsto t^{-2}\sigma_f(t)$ is non-increasing for such a function; moreover, one obtains that for such an f , f is uniformly convex if and only if f^* is uniformly smooth and f is uniformly smooth if and only if f^* is uniformly convex. Then one establishes many characterizations of uniformly convex functions and of uniformly smooth convex functions. In these characterizations appear functions (gages or moduli) belonging to different subclasses of \mathcal{A} introduced in Section 3.3. These gages and moduli are sharp enough in order to obtain that f is c -strongly convex if and only if f^* is Fréchet differentiable on X^* and ∇f^* is c^{-1} -Lipschitz. Even if the results are established in general Banach spaces the natural framework for uniformly convex and uniformly smooth convex function is that of reflexive Banach spaces. This is due to the fact that when there exists a proper lower semicontinuous and uniformly convex function on a Banach space whose domain has nonempty interior, the space is necessarily reflexive.

Section 3.6 is dedicated to the study of those convex functions which are uniformly convex on bounded sets and uniformly smooth on bounded sets, respectively. Under strong coercivity of the function one shows that these notions are dual.

In Section 3.7 we study the function $f_\varphi : X \rightarrow \mathbb{R}$, $f_\varphi(x) = \int_0^{\|x\|} \varphi(t) dt$, where φ is a weight function, in connection with the geometric properties of the norm. So, one establishes characterizations of the strict convexity, the smoothness and the reflexivity of X by the strict convexity, the Gâteaux differentiability of f_φ and the surjectivity of ∂f_φ , respectively.

One obtains also characterizations of (local) uniform convexity and (local) uniform smoothness of X with the help of the properties of f_φ . For example one obtains: X is uniformly convex $\Leftrightarrow X^*$ is uniformly smooth $\Leftrightarrow f_\varphi$ is uniformly convex on bounded sets $\Leftrightarrow (f_\varphi)^*$ is uniformly smooth on bounded sets $\Leftrightarrow (f_\varphi)^*$ is Fréchet differentiable and $\nabla(f_\varphi)^*$ is uniformly continuous on bounded sets. Note that a part of the results of this section can be found in the book [Cioranescu (1990)], but the proofs are different; note also that some notions are introduced differently in Cioranescu's book.

Another application of convex analysis is emphasized in Section 3.8; here we apply the results on the existence, the uniqueness and the characterizations of optimal solutions of convex programs to the problem of the best approximation with elements of a convex subset of a normed space.

In Section 3.9 it is shown that there exists a strong relationship between the well-posedness of the minimization problem $\min f(x)$ s.t. $x \in X$, and the differentiability at 0 of the conjugate f^* of f ; when f is convex these properties are equivalent. Using this result we establish a very interesting characterization of Chebyshev sets in Hilbert spaces and show that the class of weakly closed Chebyshev sets coincides with the class of closed convex sets in Hilbert spaces.

Section 3.10 deals with sets of weak sharp minima, well-behaved convex functions and the study of the existence of global error bounds for convex inequalities. These notions were studied separately for a time, but they are intimately related. As noted above, $\operatorname{argmin} f$ is a set of weak sharp minima for f exactly when f is well-conditioned with linear rate. But the inequality $f(x) \leq 0$ has a global error bound exactly when $\operatorname{argmin}[f]_+$ is a set of weak sharp minima for $[f]_+ := \max(f, 0)$. We give several characterizations of the fact that $\operatorname{argmin} f$ is a set of weak sharp minima for f , one of them being the fact that up to a constant, the conjugate f^* is sublinear on a neighborhood of the origin. Several numbers associated to a convex function are introduced which are related to the conditioning number from numerical analysis. Although the most part of the results from this section are stated in the literature in finite dimensional spaces, we present them in infinite dimensions.

The last section of this book, Section 3.11, is dedicated to the study of monotone multifunctions on Banach spaces. We use in the presentation two recent articles of Simons. The proofs are quite technical and use the lower semicontinuous convex function χ_M associated to the multifunction M :

$X \rightrightarrows X^*$, the minimax theorem and a few results of convex analysis. One obtains: two characterizations of maximal monotone multifunctions; the fact that the condition $0 \in \text{int}(\text{dom } T_1 - \text{dom } T_2)$ is equivalent to other three conditions involving $\text{dom } T_i$ and $\text{dom } \chi_{T_i}$, and is sufficient for the maximal monotonicity of $T_1 + T_2$; $\overline{\text{dom } T}$ and $\overline{\text{Im } T}$ are convex if X is reflexive and T is maximal monotone; $\overline{\text{dom } T}$ is convex if $\text{int}(\text{dom } T) \neq \emptyset$ and T is maximal monotone; T is locally bounded at $x_0 \in (\text{co}(\text{dom } T))^i$ if T is a monotone multifunction; Rockafellar's theorem on the local boundedness of maximal monotone multifunctions. The result stating that for a maximal monotone multifunction T on the Banach space X for which $\overline{\text{dom } T}$ is convex the local boundedness of T at $\bar{x} \in \overline{\text{dom } T}$ implies that $\bar{x} \in \text{int}(\overline{\text{dom } T})$ seems to be new. When applied to the subdifferential of a proper lower semicontinuous convex function f on the Banach space X , this result gives (for example): f is continuous $\Leftrightarrow \text{dom } f$ is open $\Leftrightarrow \partial f$ is locally bounded at any $x \in \text{dom } f$.

The exercises are intended to exemplify the topics treated in the book. Many exercises are auxiliary results spread in recent articles, although some of them are extracted from other books. Some exercises are important results which could be parts of textbooks, but which do not fit very well with the aim of the present book. Among them we mention Exercise 3.11 on the Moreau regularization.