<table>
<thead>
<tr>
<th>No.</th>
<th>Author(s)</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HERMAN/KUČERA/ŠIMŠA</td>
<td>Equations and Inequalities</td>
</tr>
<tr>
<td>2</td>
<td>ARNOLD</td>
<td>Abelian Groups and Representations of Finite Partially Ordered Sets</td>
</tr>
<tr>
<td>3</td>
<td>BORWEIN/LEWIS</td>
<td>Convex Analysis and Nonlinear Optimization, 2nd Ed.</td>
</tr>
<tr>
<td>4</td>
<td>LEVIN/LUBINSKY</td>
<td>Orthogonal Polynomials for Exponential Weights</td>
</tr>
<tr>
<td>5</td>
<td>KANE</td>
<td>Reflection Groups and Invariant Theory</td>
</tr>
<tr>
<td>6</td>
<td>PHILLIPS</td>
<td>Two Millennia of Mathematics</td>
</tr>
<tr>
<td>7</td>
<td>DEUTSCH</td>
<td>Best Approximation in Inner Product Spaces</td>
</tr>
<tr>
<td>8</td>
<td>FABIAN ET AL.</td>
<td>Functional Analysis and Infinite-Dimensional Geometry</td>
</tr>
<tr>
<td>9</td>
<td>Křížek/LUCA/SOMER</td>
<td>17 Lectures on Fermat Numbers</td>
</tr>
<tr>
<td>10</td>
<td>BORWEIN</td>
<td>Computational Excursions in Analysis and Number Theory</td>
</tr>
<tr>
<td>11</td>
<td>REED/SALES (Editors)</td>
<td>Recent Advances in Algorithms and Combinatorics</td>
</tr>
<tr>
<td>12</td>
<td>HERMAN/KUČERA/ŠIMŠA</td>
<td>Counting and Configurations</td>
</tr>
<tr>
<td>13</td>
<td>NAZARETH</td>
<td>Differentiable Optimization and Equation Solving</td>
</tr>
<tr>
<td>14</td>
<td>PHILLIPS</td>
<td>Interpolation and Approximation by Polynomials</td>
</tr>
<tr>
<td>15</td>
<td>BEN-ISRAEL/GREVILLE</td>
<td>Generalized Inverses, 2nd Ed.</td>
</tr>
<tr>
<td>16</td>
<td>ZHAO</td>
<td>Dynamical Systems in Population Biology</td>
</tr>
<tr>
<td>17</td>
<td>GÖPFERT ET AL.</td>
<td>Variational Methods in Partially Ordered Spaces</td>
</tr>
<tr>
<td>18</td>
<td>AKIVIS/GOLDBERG</td>
<td>Differential Geometry of Varieties with Degenerate Gauss Maps</td>
</tr>
<tr>
<td>19</td>
<td>MIKHALEV/SHPILRAY/YU</td>
<td>Combinatorial Methods</td>
</tr>
<tr>
<td>20</td>
<td>BORWEIN/ZHU</td>
<td>Techniques of Variational Analysis</td>
</tr>
<tr>
<td>21</td>
<td>VAN BRUMMELEN/KINYON</td>
<td>Mathematics and the Historian’s Craft</td>
</tr>
<tr>
<td>22</td>
<td>LUCCHETTI</td>
<td>Convexity and Well-Posed Problems</td>
</tr>
<tr>
<td>23</td>
<td>NICULESCU/PERSSON</td>
<td>Convex Functions and Their Applications</td>
</tr>
<tr>
<td>24</td>
<td>SINGER</td>
<td>Duality for Nonconvex Approximation and Optimization</td>
</tr>
<tr>
<td>25</td>
<td>HIGGINSON/PIMM/SINCLAIR</td>
<td>Mathematics and the Aesthetic</td>
</tr>
</tbody>
</table>
Ivan Singer

Duality for Nonconvex Approximation and Optimization

With 17 Figures
To the memory of my wonderful wife, Crina
## Contents

### List of Figures

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>xi</td>
</tr>
</tbody>
</table>

### Preface

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>xiii</td>
</tr>
</tbody>
</table>

### 1 Preliminaries

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>27</td>
</tr>
<tr>
<td>1.3</td>
<td>39</td>
</tr>
<tr>
<td>1.4</td>
<td>46</td>
</tr>
<tr>
<td>1.4.1</td>
<td>47</td>
</tr>
<tr>
<td>1.4.2</td>
<td>71</td>
</tr>
</tbody>
</table>

### 2 Worst Approximation

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>86</td>
</tr>
<tr>
<td>2.2</td>
<td>93</td>
</tr>
</tbody>
</table>

### 3 Duality for Quasi-convex Supremization

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>103</td>
</tr>
<tr>
<td>3.2</td>
<td>108</td>
</tr>
<tr>
<td>3.3</td>
<td>121</td>
</tr>
<tr>
<td>3.4</td>
<td>127</td>
</tr>
<tr>
<td>3.4.1</td>
<td>127</td>
</tr>
<tr>
<td>3.4.2</td>
<td>129</td>
</tr>
</tbody>
</table>
3.5 Duality for quasi-convex supremization over structured primal constraint sets ........................................ 131

4 Optimal Solutions for Quasi-convex Maximization .................. 137
4.1 Maximum points of quasi-convex functions .......................... 137
4.2 Maximum points of continuous convex functions ................... 144
4.3 Some basic subdifferential characterizations of maximum points .... 149

5 Reverse Convex Best Approximation ................................. 153
5.1 The distance to the complement of a convex set .................... 154
5.2 Characterizations and existence of elements of best approximation in complements of convex sets .................... 161

6 Unperturbational Duality for Reverse Convex Infimization ........ 169
6.1 Some hyperplane theorems of surrogate duality ..................... 171
6.2 Unconstrained surrogate dual problems for reverse convex infimization ........................................... 175
6.3 Constrained surrogate dual problems for reverse convex infimization ........................................... 184
6.4 Unperturbational Lagrangian duality for reverse convex infimization ........................................... 189
6.5 Duality for infimization over structured primal reverse convex constraint sets ........................................ 190
6.5.1 Systems ....................................................... 190
6.5.2 Inequality constraints ......................................... 198

7 Optimal Solutions for Reverse Convex Infimization ................ 203
7.1 Minimum points of functions on reverse convex subsets of locally convex spaces .......................... 203
7.2 Subdifferential characterizations of minimum points of functions on reverse convex sets ............... 209

8 Duality for D.C. Optimization Problems .............................. 213
8.1 Unperturbational duality for unconstrained d.c. infimization .... 213
8.2 Minimum points of d.c. functions .................................. 221
8.3 Duality for d.c. infimization with a d.c. inequality constraint .... 225
8.4 Duality for d.c. infimization with finitely many d.c. inequality constraints ........................................ 232
8.5 Perturbational theory ............................................... 244
8.6 Duality for optimization problems involving maximum operators . 247
8.6.1 Duality via conjugations of type Lau .......................... 248
8.6.2 Duality via Fenchel conjugations ................................ 252

9 Duality for Optimization in the Framework of Abstract Convexity 259
9.1 Additional preliminaries from abstract convex analysis ............ 259
9.2 Surrogate duality for abstract quasi-convex supremization, using polarities $\Delta_G : 2^X \rightarrow 2^W$ and $\Delta_G : 2^X \rightarrow 2^{W \times R}$ .......................... 267
9.3 Constrained surrogate duality for abstract quasi-convex supremization, using families of subsets of $X$ .......................... 270
9.4 Surrogate duality for abstract reverse convex infimization, using polarities $\Delta_G : 2^X \to 2^W$ and $\Delta_G : 2^X \to 2^{W \times R}$ ................................. 271
9.5 Constrained surrogate duality for abstract reverse convex infimization, using families of subsets of $X$ .......................... 273
9.6 Duality for unconstrained abstract d.c. infimization ................................. 275

10 Notes and Remarks .......................................................... 279

References ................................................................. 329

Index ................................................................. 347
List of Figures

1.1 ............................................................. 3
1.2 ............................................................. 12
1.3 ............................................................. 40
1.4 ............................................................. 44
1.5 ............................................................. 49

2.1 ............................................................. 86
2.2 ............................................................. 88
2.3 ............................................................. 90
2.4 ............................................................. 91
2.5 ............................................................. 92
2.6 ............................................................. 95
2.7 ............................................................. 96

5.1 ............................................................. 153
5.2 ............................................................. 156
5.3 ............................................................. 157
5.4 ............................................................. 159
5.5 ............................................................. 163
In this monograph we present some approaches to duality in nonconvex approximation in normed linear spaces and to duality in nonconvex global optimization in locally convex spaces.

At the first stage of development of approximation theory in normed linear spaces, the “best approximation” of an element by linear subspaces, and more generally, by convex sets (i.e., the minimization of the distance of an element to a convex set) was studied. Later, the following two main classes of nonconvex approximation problems were considered: “worst approximation,” i.e., the maximization of the distance of an element to an arbitrary set, and “reverse convex best approximation,” i.e., the minimization of the distance of an element to the complement of a convex set. These may be called “anticonvex” problems (following Penot [175], who has used this term in the more general context of optimization theory). The first results on duality for these problems were obtained in the papers [73], [74].

In optimization theory in locally convex spaces, first linear optimization problems, and more generally, convex optimization problems, i.e., the minimization of a convex function on a convex set (clearly, best approximation of an element by the elements of a convex set belongs to this class of problems) were studied. Later, the duality results obtained in this direction were extended to duality results for nonconvex problems, based on generalizations of convexity and of the methods of convex analysis by Elster and Nehse [60], Balder [13], Lindberg [134], Dolecki and Kurcyusz [48], Dolecki [47], and others.

Independently, some classes of nonconvex optimization problems of a different type were studied, which Hiriart-Urruty [102] called “convex-anticonvex” problems (and we shall also adopt this terminology), since they have the following specific structure. They are minimization problems in which convexity is present twice,
in the constraint set and/or in the objective function, but once in the reverse way; namely, these are “convex maximization,” i.e., maximization of a convex function on a convex set (or equivalently, minimization of a concave function on a convex set), “reverse convex minimization,” i.e., minimization of a convex function on the complement of a convex set, and “d.c. optimization,” i.e., optimization problems involving differences of convex functions. Of course, the latter also encompasses convex optimization problems as a particular case. The first results on duality for these problems were obtained in the papers [215], [218], [220], [217], and the paper [280] of Toland.

For some time, approximation theory and optimization theory have developed independently, in parallel. In the 1960s it was observed that optimization, i.e., the minimization or maximization of a function, contains approximation as a particular case. Indeed, approximation is the minimization or maximization of a particular function on a normed linear space \( X \), namely, the function

\[
    f(y) = \|x_0 - y\| \quad (y \in X).
\]

Thus, in the 1970s there appeared naturally the idea of studying them together in this spirit, as reflected for example by the titles of the monographs of Laurent, *Approximation et optimisation* (1970) [129], Holmes, *A Course on Optimization and Best Approximation* (1972) [106], Krabs, *Optimierung und Approximation* (1975) [122], Hettich and Zencke, *Numerische Methoden der Approximation und semi-infiniter Optimierung* (1982) [96], Glashoff and Gustaffson, *Linear Approximation and Optimization* (1983) [84], and Jongen, Jonker, and Twilt, *Nonlinear Optimization in \( \mathbb{R}^n \). I. Morse Theory, Chebyshev Approximation* (1986) [114]. The same point of view also appeared in parts of other monographs on optimization theory. On the other hand, going in the opposite direction, Cheney and Goldstein [32] have extended a result on the existence of best approximations to a result on the existence of optimal solutions of minimization problems. Starting with [212], [213], there was suggested and systematically carried out a program of work in this direction, namely, to show that many methods and results of approximation theory are so strong that they can be generalized to yield new methods and results in optimization theory. Subsequently, others also adopted this latter point of view (e.g., Wriedt [295], Berdyshev [17]). In the present monograph we shall study these two theories and their interactions, going from approximation to optimization and vice versa.

It has long been known that duality is a powerful tool in the study of approximation and optimization problems. For problems of approximation in a normed linear space \( X \), namely, of minimization or maximization of the distance to a given subset of \( X \), “duality” means simply their study with the aid of the elements of the conjugate space \( X^* \). In a general setting, “duality theory” in optimization means the simultaneous study of a pair of optimization problems, related in some way, namely, the initial problem, called the “primal problem,” of minimization or maximization of a function on a subset of a locally convex space \( X \), and the “dual problem” of minimization or maximization of a function on a subset of a locally convex space \( W \), with the aim of obtaining more information on the primal problem (on its “optimal value,” on its “optimal solutions,” etc.). In general (with the exceptions of Sec-
tions 9.3 and 9.5), \( W \) is a set of functions on \( X \), or alternatively, \( W \) is an arbitrary set, but paired with \( X \) with the aid of a function on the Cartesian product \( X \times W \) called a “coupling function.” In fact, although the latter is apparently more general, it turns out that these two methods are equivalent. We shall avoid the use of the term “duality” in other senses (so instead of “dual space” we shall use “conjugate” space; instead of “duality” between families of subsets we shall use “polarity”; etc.).

The monographs devoted to approximation theory in normed linear spaces ([210], [211]) and those containing some chapters or sections on approximation in such spaces (e.g., Akhiezer [1], Cheney [31], Tikhomirov [277]) treat duality mainly for the case of best approximation by convex sets or special classes of convex sets (linear subspaces, cones) or do not consider duality at all (Deutsch [41], devoted to best approximation in inner product spaces; Braess [25]). Also, the monographs on approximation and optimization, mentioned above, of Laurent, Krabs, and others consider duality mainly for convex sets and functions, or like the one of Jongen, Jonker, and Twilt, do not consider duality at all.

Furthermore, most of the existing monographs on optimization theory or convex analysis and optimization treat duality mainly for the convex and quasi-convex cases (e.g., Stoer and Witzgall [262], Auslender [11], Ioffe and Tikhomirov [111], Ekeland and Temam [54], Elster, Reinhardt, Schäuble, and Donath [61], Barbu and Precupanu [14], Pshenichnyi [182], Ponstein [180], Hettich and Zenke [96], Glashoff and Gustaffson [84], Ekeland and Turnbull [55], Hiriart-Urruty and Lemaréchal [104], Gol'shtein and Tretyakov [91], Borwein and Lewis [21]) or include some brief parts on non-convex duality, especially on d.c. duality (e.g., Konno, Thach, and Tuy [120], Strékalovsky [267], Rubinov [193], Pallabchya and Rolewicz [169], Rockafellar and Wets [187], Tuy [284]). A section of the recent monograph of Rubinov and Yang, *Lagrange-Type Functions in Constrained Nonconvex Optimization* [201], presents the general theory of Lagrange-type functions and duality, developed mainly by the authors [201, Ch. 3, Section 3.2].

The monographs devoted especially to duality in optimization theory, by Gol'shtein, *The Theory of Duality in Mathematical Programming and Its Applications* (in Russian, 1971) [90], Rockafellar, *Conjugate Duality and Optimization* (1974) [185], and Walk, *Theory of Duality in Mathematical Programming* (1989) [293], treat only duality for convex optimization and some nonconvex generalizations of it. The monograph of Gao, *Duality Principles in Nonconvex Systems: Theory, Methods and Applications* (2000) [76], addressed to those working in applied mathematics, physics, mechanics, and engineering, presents a brief combination of Rockafellar’s perturbational duality theory for convex problems and Auchmuty’s [10] extended Lagrange duality theory as part of Gao’s larger original theory of duality, which aims to encompass “duality in natural phenomena.” Finally, the recent monograph of Goh and Yang, *Duality in Optimization and Variational Inequalities* (2001) [89], contains a short chapter on a nonconvex duality theory due to the authors (Goh and Yang [88]) for the classical mathematical programming problem in \( \mathbb{R}^n \). However, there is no monograph devoted to duality for nonconvex approximation and optimization problems.
There are detailed surveys on some of the approaches to nonconvex duality, described above. Thus, for the nonconvex duality results based on generalizations of convexity and generalizations of the methods of convex analysis see Martínez-Legaz ([143], [140]), and for the nonconvex duality results based on various Lagrange-type functions see the respective chapters of the monographs of Goh and Yang [89] and Rubinov and Yang [201]. Therefore, these approaches will be presented here more briefly, mainly in Chapters 1 and 10. The present monograph is devoted to the study of duality for the anticonvex approximation and convex-anticonvex optimization problems, in the above-mentioned senses. Note that these include a very broad class of nonconvex problems. For example, as we shall see in Chapter 8, the infimization of a lower semicontinuous function over a closed subset of a Hilbert space can be easily reformulated as the problem of infimization of a continuous linear function subject to a d.c. constraint, or alternatively, of a convex function subject to a reverse convex constraint.

We shall concentrate here only on duality, so we shall not consider, for example, characterizations of primal optimal solutions involving only the primal constraint set and the primal objective function (with the exception of those, such as Remarks 4.1 and 7.1, that are used to prove duality results). We shall study duality only for global approximation and optimization, but some results for the local case will be also mentioned briefly in the Notes and Remarks. We shall not consider here duality for multiobjective optimization. In order to limit the size of this monograph, quadratic optimization and differentiable optimization will not be considered here; also, algorithms are not given here (for the latter, see for example the survey article Tuy [283] and the monographs Konno, Thach, and Tuy [120] and Tuy [284]).

Being the first of this kind in the literature, the present monograph is based entirely on articles in mathematical journals. Some unpublished results and some new proofs are also given.

Let us describe, briefly, the contents of the chapters of the book.

In Chapter 1, after some preliminaries from convex analysis and abstract convex analysis, we give some results on duality for best approximation by elements of convex sets in normed linear spaces, and on duality for the infimization of convex and quasi-convex functions on convex sets in locally convex spaces. These will serve as a basis of comparison with the nonconvex duality results of the subsequent chapters and with the methods of obtaining them.

In Chapter 2 we consider the deviation $\delta(G, x_0)$ of a set $G$ from an element $x_0$ in a normed linear space $X$, i.e., the supremum of the distances $\|g - x_0\| = \text{dist}(g, x_0)$, over all $g \in G$. We give duality formulas for $\delta(G, x_0)$ and characterizations of the elements $x_0 \in G$ for which the above supremum is attained, i.e., of the so-called elements of worst approximation (or farthest points).

Chapter 3 is devoted to the more general problem of quasi-convex supremization $\sup f(G)$, where $G$ is a set in a locally convex space $X$ and $f : X \to \mathbb{R}$ is a quasi-convex function. We introduce and study both unconstrained and constrained surrogate dual problems, as well as unperturbational and perturbational Lagrangian dual problems for quasi-convex supremization. Also, we consider surrogate duality for the case that the primal constraint set $G$ is expressed with the aid of a "system."
In Chapter 4 we present various characterizations of the optimal solutions for quasi-convex supremization problems \( \sup f(G) \), i.e., of the elements \( g_0 \in G \) such that \( f(g_0) = \max f(G) \).

In Chapter 5 we study the best approximation \( \text{dist}(x_0, \mathbb{C}G) \) by the complement \( \mathbb{C}G = X \setminus G \) of a convex set \( G \) in a normed linear space \( X \), i.e., the infimum of the distances \( \|x_0 - z\| = \text{dist}(x_0, z) \), over all \( z \in \mathbb{C}G \). We give duality formulas for \( \text{dist}(x_0, \mathbb{C}G) \), and characterizations of the elements \( z_0 \in \mathbb{C}G \) for which the above infimum is attained.

Chapter 6 is devoted to the more general problem of reverse convex infimization \( \inf f(\mathbb{C}G) \), where \( G \) is a convex set in a locally convex space \( X \) and \( f : X \to \mathbb{R} \) is a function. We introduce and study both unconstrained and constrained surrogate dual problems, and unperturbational Lagrangian dual problems for reverse convex infimization. Also, we consider surrogate duality for the case that the primal constraint set \( G \) is expressed with the aid of a system or with the aid of inequalities.

In Chapter 7 we present various characterizations of the optimal solutions for reverse convex infimization problems \( \inf f(\mathbb{C}G) \) (where \( G \) is a convex subset of a locally convex space \( X \)), i.e., of the elements \( z_0 \in \mathbb{C}G \) such that \( f(z_0) = \min f(\mathbb{C}G) \).

Chapter 8 is devoted to "d.c. optimization," i.e., to optimization problems involving differences of convex functions. We first give duality results for the unconstrained infimization of the difference \( f - h \) of two functions on a locally convex space \( X \), the first of them being arbitrary and the second one convex and lower semicontinuous. Next we give some characterizations for optimal solutions of such problems. We also study duality for the infimization of the difference \( f - h \), where \( f, h \) are convex functions, on a constraint set defined by an inequality \( l(x) - k(x) \leq 0 \) or \( l(x) - k(x) < 0 \), where \( l, k \) are convex functions, or on a constraint set defined by finitely many such inequalities. Furthermore, we present some results of perturbational Lagrangian duality for d.c. infimization. Finally, we present some duality results for the unconstrained problem of infimization of the pointwise maximum of two functions \( f \) and \( -h \) on a locally convex space \( X \), the first of them being arbitrary and the second one quasi-convex (or more particularly, convex) and lower semicontinuous (it turns out that this is, essentially, a d.c. problem).

The framework of abstract convexity, which encompasses various generalizations of convex sets and convex functions, permits us to study optimization of more general functions on more general sets. In Chapter 9 we present briefly some duality results for such optimization problems.

The concluding Chapter 10 contains some comments, bibliographical references, and additional results for each of the preceding chapters.

We hope that this book will interest a large circle of readers, including those who want to use it for research or as a reference book, or for a graduate course, or for independent study (to this end, we have given detailed proofs of the results and several illustrations).

I would like to express my profound gratitude to my long-time friend J.-E. Martínez-Legaz for his support of the project of this book and his generous help in its materialization. He has patiently and carefully read several versions of the whole manuscript, making valuable suggestions for corrections, improvements and
additions. Furthermore, I thank A.M. Rubinov for his stimulating interest and encouragement and for helpful comments on some parts. Also, I thank C. Zălinescu for prompt answers to some questions. I am grateful for the excellent working conditions ensured by the Simion Stoilow Institute of Mathematics of the Romanian Academy during the writing and preparation for print of the manuscript. Finally, I wish to thank J.M. Borwein for accepting to publish this book in his prestigious series and for his invaluable strong support in various stages of its production. Last, but not least, my thanks are due to Springer and the Canadian Mathematical Society, for their efforts and care in the production process.

Bucharest, Romania

Ivan Singer

October 2005
Duality for Nonconvex Approximation and Optimization
1
Preliminaries

1.1 Some preliminaries from convex analysis

In this section we recall some basic definitions and results about convex analysis in
the framework of normed linear spaces, in which we shall study the approximation
problems, and in the more general framework of locally convex spaces in which we
shall study the optimization problems.

A (real) linear space is a set $X$ in which there are defined two “vector oper-
ations,” namely, an internal binary operation, called “addition,” which associates,
to each pair of elements $x, y \in X$ an element $x + y \in X$, and an external binary
operation, called “multiplication by a scalar,” which associates, to each pair $(a, x)$
consisting of a real number $a \in \mathbb{R}$ and an element $x \in X$, an element $ax \in X,$
with these operations satisfying the following conditions, for all $x, y, z \in X$ and
$a, b \in \mathbb{R}$:

1. $x + y = y + x,$
2. $x + (y + z) = (x + y) + z,$
3. $x + y = x + z \Rightarrow y = z,$
4. $a(x + y) = ax + ay,$
5. $(a + b)x = ax + bx,$
6. $a(bx) = (ab)x,$
7. $1x = x.$

From these “axioms” one deduces easily that there exists a unique element $0 \in
X$ such that $x + 0 = 0 + x = 0$ for all $x \in X.$ The “opposite element” of any
1. Preliminaries

$x \in X$ is defined by $-x := (-1)x$, and "subtraction" of elements is defined by $x - y := x + (-y)$ ($x, y \in X$).

For example, the set $\mathbb{R}^n$ of all ordered $n$-tuples of real numbers $x = (x_1, \ldots, x_n)$, where $1 \leq n < +\infty$, with componentwise vector operations

$$x + y = (x_1 + y_1, \ldots, x_n + y_n), \quad ax = (ax_1, \ldots, ax_n),$$

where $a \in \mathbb{R}$, is a linear space.

A normed linear space is a linear space $X$ in which to each element $x \in X$ there is associated a real number $\|x\|$, called the "norm" of $x$, satisfying the following conditions, for all $x, y \in X$ and $a \in \mathbb{R}$:

1. $\|0\| = 0$ and $\|x\| > 0$ for each $x \neq 0$,
2. $\|x + y\| \leq \|x\| + \|y\|$,  
3. $\|ax\| = |a| \|x\|$.

In a normed linear space $X$ a sequence $\{x_n\}$ is said to "converge" to an element $x \in X$, in symbols, $x_n \rightarrow x$, or $\lim_{n \rightarrow \infty} x_n = x$, if $\lim_{n \rightarrow \infty} \|x_n - x\| = 0$. A normed linear space is called complete, if for every sequence $\{x_n\} \subset X$ satisfying $\lim_{n, m \rightarrow \infty} \|x_n - x_m\| = 0$ there exists an element $x \in X$ such that $\lim_{n \rightarrow \infty} x_n = x$. A complete normed linear space is also called a Banach space.

Here are some important examples of Banach spaces, to which we shall refer later:

(i) The space $l^\infty$, i.e., the linear space $\mathbb{R}^n$ endowed with the norm

$$\|x\| = \max_{1 \leq j \leq n} |x_j|.$$  \hspace{1cm} (1.2)

(ii) The space $l^1$, i.e., the linear space $\mathbb{R}^n$ endowed with the norm

$$\|x\| = \sum_{i=1}^n |x_i|.$$  \hspace{1cm} (1.3)

(iii) The "Euclidean space" $l^2_2$, i.e., the linear space $\mathbb{R}^n$ endowed with the norm

$$\|x\| = \sqrt{\sum_{i=1}^n |x_i|^2}.$$  \hspace{1cm} (1.4)

When $n = 2$ or $n = 3$, it is easy to visualize the "unit ball"

$$B_X = \{x \in X | \|x\| \leq 1\}$$  \hspace{1cm} (1.5)

of these spaces (see Figure 1.1).

While the above examples are of finite dimension $n$, here are some infinite-dimensional versions:
(iv) The space $l^\infty$ of all bounded sequences of real numbers $x = (x_i)_{i=1}^\infty = (x_i)$, with the componentwise vector operations

$$x + y = (x_1 + y_1, \ldots, x_n + y_n, \ldots), \quad ax = (ax_1, \ldots, ax_n, \ldots),$$

where $a \in R$, and with the norm

$$\|x\| = \sup_{1 \leq i < +\infty} |x_i|. \quad (1.7)$$

(v) The space $l^1$ of all summable sequences of real numbers $x = (x_i)_{i=1}^\infty = (x_i)$ (i.e., such that $\sum_{i=1}^{\infty} |x_i| < +\infty$), with the componentwise vector operations (1.6), and with the norm

$$\|x\| = \sum_{i=1}^{\infty} |x_i|. \quad (1.8)$$

(vi) The space $l^2$ of all square summable sequences of real numbers $x = (x_i)_{i=1}^\infty = (x_i)$ (i.e., such that $\sum_{i=1}^{\infty} |x_i|^2 < +\infty$), with the componentwise vector operations (1.6), and with the norm

$$\|x\| = \sqrt{\sum_{i=1}^{\infty} |x_i|^2}. \quad (1.9)$$

Let us also mention the following examples of function spaces, corresponding to the above examples.

(vii) The space $C([a, b])$ of all continuous functions $x : [a, b] \to R$, on a closed interval $[a, b]$, with the pointwise vector operations

$$(x + y)(t) = x(t) + y(t), \quad (ax)(t) = ax(t) \quad (t \in [a, b]), \quad (1.10)$$
where \( a \in \mathbb{R} \), and with the norm
\[
\| x \| = \max_{t \in [a, b]} |x(t)|.
\] (1.11)

(viii) The space \( L^1([a, b]) \) of all measurable functions on \([a, b]\) such that \( \int_a^b |x(t)| \, dt < +\infty \) (the integral being taken in the sense of Lebesgue), where two functions that coincide almost everywhere are considered identical (i.e., represent the same element of \( X \)), endowed with the pointwise vector operations (1.10), and with the norm
\[
\| x \| = \int_a^b |x(t)| \, dt.
\] (1.12)

(ix) The space \( L^2([a, b]) \) defined similarly to example (viii), with the condition \( \int_a^b |x(t)| \, dt < +\infty \) and the norm (1.12) replaced, respectively, by \( \int_a^b |x(t)|^2 \, dt < +\infty \) and
\[
\| x \| = \left( \int_a^b |x(t)|^2 \, dt \right)^{1/2}.
\] (1.13)

Thus, the elements of \( L^1([a, b]) \) and \( L^2([a, b]) \) are equivalence classes of functions.

We recall that a **Hilbert space** is a Banach space \( X \) in which for any \( x, y \in X \) there is defined a number \( (x, y) \in \mathbb{R} \), called the “inner product” of \( x \) and \( y \), such that for all \( x, y, z \in X \) and \( a \in \mathbb{R} \) we have \( (x, x) \geq 0 \), \( (x, x) = 0 \) if and only if \( x = 0 \), \( (x, y) = (y, x) \), \( (x + y, z) = (x, z) + (y, z) \), \( (ax, y) = a(x, y) \), and in which the norm of \( x \in X \) is defined by \( \| x \| := \sqrt{(x, x)} \). The spaces \( l^2_\mathbb{R} \), \( l^2 \), and \( L^2([a, b]) \) of examples (iii), (vi), and (ix) are Hilbert spaces, with the usual inner products
\[
(x, y) = \sum_{i=1}^{n} x_i y_i, \quad (x, y) = \sum_{i=1}^{\infty} x_i y_i, \quad (x, y) = \int_a^b x(t)y(t) \, dt,
\] (1.14)
respectively.

For any two elements \( x, y \) in a normed linear space \( X \) one defines the **distance** between \( x \) and \( y \) by
\[
\text{dist}(x, y) := \| x - y \|.
\] (1.15)

In the framework of normed linear spaces, the “best approximation” amounts to the minimization of a distance, and the “worst approximation” amounts to the maximization of a distance. This fact permits us to use geometric intuition (but rigorous analytic proofs), and the connections of the phenomena become clearer and the arguments simpler than those occurring in the various concrete spaces. Thus, this general framework constitutes both a unified foundation for the classical approximation theory in various concrete normed linear spaces (which treats the problems in each space with ad hoc methods), and a powerful tool for obtaining new results (see [210], [211]).
1.1 Some preliminaries from convex analysis

We recall that a subset $C$ of a linear space $X$ is said to be convex if the relations $c_1, c_2 \in C$ imply that $\theta c_1 + (1 - \theta) c_2 \in C$ for all scalars $\theta$ satisfying $0 < \theta < 1$. Geometrically, this means that along with any two points $c_1, c_2 \in C$, the set $C$ contains the whole “segment” $[c_1, c_2] := \{ \theta c_1 + (1 - \theta) c_2 \mid 0 \leq \theta \leq 1 \}$ joining them (or with “endpoints” $c_1, c_2$). For any set $C$ in a linear space $X$, we shall denote by $\text{co } C$ the convex hull of $C$, i.e., the intersection of all convex sets containing $C$.

Assuming that the reader knows some elements of general topology, we recall that a linear space $X$ endowed with a Hausdorff topology $T$ on $X$ (i.e., such that for each pair of distinct elements $x, y$ there are neighborhoods $U_x$ of $x$ and $U_y$ of $y$ such that $U_x \cap U_y = \emptyset$) is called a topological linear space if the vector operations are continuous, that is, if the mappings $(x, y) \mapsto x + y$ from $X \times X$ into $X$, and $(a, x) \mapsto ax$ from $R \times X$ into $X$, are continuous. A topological linear space $X$ is called locally convex if every neighborhood of any element $x \in X$ contains a convex neighborhood of $x$.

Remark 1.1. (a) We shall not recall here the theories of normed linear spaces and locally convex spaces. These can be found, e.g., in the standard books of Dunford and Schwartz [49], Bourbaki [24], Day [40], Schaefer [204], Kelley and Namioka [117], Holmes [107], and Köthe [121], but the elements are also recalled briefly in some books on convex analysis (e.g., in Ioffe and Tikhomirov [111], Barbu and Precupanu [14]).

(b) We shall not consider here best approximation in “seminorms”, i.e., real numbers satisfying (9) and (10) above (see, e.g., Bourbaki [24]) or in “asymmetric norms,” i.e., such that (10) is replaced by $\|ax\| = a\|x\|$ for all $x \in X$ and all $a \geq 0$ (see, e.g., Krein [123]), or in the more general “norms” introduced in [223].

(c) Unless specified otherwise, we shall consider every normed linear space $X$ as being endowed with the “norm topology” (i.e., in which for any element $x \in X$ a fundamental system of neighborhoods of $x$ is the family of all open balls $\{ z \in X \mid \| z - x \| < r \}$ ($r > 0$) with center $x$); this space is locally convex. Therefore, the results given later for locally convex spaces are valid, in particular, in normed linear spaces (endowed with the norm topology), and hence the reader who does not want to work in locally convex spaces, may consider those results only in normed linear spaces. Moreover, for simplicity, the reader may consider the subsequent results only in the usual Euclidean space $l_p^2$, i.e., in $\mathbb{R}^n$ endowed with the norm (1.4) (with the exception of a few specifically infinite-dimensional counterexamples, which the reader can omit).

If $X$ is a locally convex space, a function $f : X \to \overline{\mathbb{R}} := [-\infty, +\infty]$ is said to be

(a) lower semicontinuous at $x_0 \in X$ if for each $k \in R$ with $k < f(x_0)$ there exists a neighborhood $U$ of $x_0$ such that $k < f(x)$ for all $x \in U$, or, equivalently, if

$$f(x_0) \leq \liminf_{y \to x_0} f(y), \quad (1.16)$$

where

$$\liminf_{y \to x_0} f(y) := \sup_{U \in \mathcal{U}(x_0)} \inf_{y \in U} f(y), \quad (1.17)$$
with \( \mathcal{U}(x_0) \) denoting the set of all neighborhoods of \( x_0 \);

(b) \textit{lower semicontinuous}, if it is lower semicontinuous at each \( x_0 \in X \);

(c) \textit{upper semicontinuous} at \( x_0 \in X \) (respectively, \textit{upper semicontinuous}) if the function \( -f \) is lower semicontinuous at \( x_0 \) (respectively, lower semicontinuous);

(d) \textit{continuous} at \( x_0 \in X \) (respectively, \textit{continuous}), if it is both lower semicontinuous and upper semicontinuous at \( x_0 \) (respectively, both lower semicontinuous and upper semicontinuous).

For any subset \( C \) of a locally convex space \( X \), we shall denote by \( \overline{C}, \text{int} \, C, \text{bd} \, C, \text{and} \, \emptyset C \) the closure, the interior, the boundary, and the complement \( X \setminus C \), of \( C \), respectively. The complement \( \emptyset C \) of any bounded open convex set \( C \) in a normed linear space is called a \textit{cavern}.

**Lemma 1.1.** Let \( X \) be a topological space, \( C \subseteq X \) and \( f: X \to \overline{R} \) a function.

(a) If \( f \) is lower semicontinuous, then

\[
\sup f(\overline{C}) = \sup f(C).
\]  

(b) If \( f \) is upper semicontinuous, then

\[
\inf f(\overline{C}) = \inf f(C).
\]

**Proof.** (a) For (1.18) see, e.g., Bourbaki [23, Ch. 4, Section 6.2, Exercise 5].

(b) If \( f \) is upper semicontinuous, then \(-f\) is lower semicontinuous, whence by part (a) and the equality \( \inf f(C) = - \sup (-f)(C) \), we obtain (1.19). \( \square \)

**Lemma 1.2.** Let \( X \) be a topological space, and \( C \subseteq X \). Then

\[
\overline{\text{C}} = \text{C} \left( \text{int} \, C \right).
\]

**Proof.** See, e.g., Kuratowski [124, Ch. 2, Section 6, Part I]. \( \square \)

For the study of semicontinuity and continuity (and of many other properties) it is convenient to introduce the following sets associated with any function \( f: X \to \overline{R} \):

(i) the \textit{epigraph} of \( f \), i.e., the subset epi \( f \) of \( X \times R \) defined by

\[
epi f := \{(x, d) \in X \times R \mid f(x) \leq d\}.
\]

(ii) the (sub)level sets \( S_d(f) \) and \( A_d(f) \) of \( f \), defined by

\[
S_d(f) := \{y \in X \mid f(y) \leq d\} \quad (d \in \overline{R}),
\]

\[
A_d(f) := \{y \in X \mid f(y) < d\} \quad (d \in \overline{R}).
\]

Then, a function \( f: X \to \overline{R} \) is lower semicontinuous if and only if its epigraph \( \text{epi} \, f \) is closed in \( X \times R \), or, equivalently, all level sets \( S_d(f) \) \((d \in R)\) are closed subsets of \( X \); hence \( f: X \to \overline{R} \) is upper semicontinuous if and only if all sets \( A_d(f) \) are open (since \(-f\) is lower semicontinuous if and only if all complements \( \emptyset A_d(f) = \{x \in X \mid f(x) \geq d\} = S_{-d}(-f) \) are closed).
If $X$ is a linear space, a function $f : X \to \mathbb{R} := (-\infty, +\infty)$ is said to be
(a) linear if it is additive and homogeneous, i.e.,

$$f(x + y) = f(x) + f(y) \quad (x, y \in X), \quad (1.24)$$

$$f(ax) = af(x) \quad (x \in X, a \in \mathbb{R}); \quad (1.25)$$

(b) affine if $f = \Phi + c$, where $\Phi : X \to \mathbb{R}$ is a linear function and $c \in \mathbb{R}$.

We shall denote the zero function by $0$; that is, $0(x) = 0$ for all $x \in X$.

The conjugate space $X^*$ of a locally convex space is the set of all continuous linear functions $\Phi : X \to \mathbb{R}$ endowed with the pointwise vector operations

$$(\Phi + \Psi)(x) = \Phi(x) + \Psi(x) \quad (x \in X), \quad (1.26a)$$

$$(a\Phi)(x) = a\Phi(x) \quad (x \in X, a \in \mathbb{R}). \quad (1.26b)$$

The "weak topology" $\sigma(X, X^*)$ on a locally convex space $X$ (in which for any element $x \in X$ a fundamental system of neighborhoods of $x$ is the family of all sets of the form $\{y \in X| \max_{1 \leq i \leq m} |\Phi_i(x) - \Phi_i(y)| < \varepsilon\} (1 \leq m < +\infty, \Phi_1, \ldots, \Phi_m \in X^*, \varepsilon > 0))$ is also a locally convex topology. We recall the following fact (see, e.g., Holmes [107, p. 157]), which will often be used later:

**Lemma 1.3.** If $C$ is a weakly compact subset of a locally convex space $X$ (i.e., compact for the weak topology $\sigma(X, X^*)$), then every $\Phi \in X^*$ attains its supremum on $C$ (i.e., for every $\Phi \in X^*$ there exists an element $c_0 \in C$ such that $\Phi(c_0) = \sup \Phi(C)$).

The conjugate space of any locally convex space, endowed with the "weak* topology" $\sigma(X^*, X)$ (in which for any element $\Phi \in X^*$ a fundamental system of neighborhoods of $\Phi$ is the family of all sets of the form $\{\Psi \in X^*| \max_{1 \leq i \leq m} |\Phi(x_i) - \Psi(x_i)| < \varepsilon\} (1 \leq m < +\infty, x_1, \ldots, x_m \in X, \varepsilon > 0))$ is a locally convex space.

If $X$ is a normed linear space, then $X^*$ is endowed with the pointwise vector operations (1.26) and with the norm

$$\|\Phi\| = \sup_{\|x\| \leq 1} |\Phi(x)| \quad (\Phi \in X^*), \quad (1.27)$$

with which it becomes a Banach space.

For a Cartesian product $X \times Z$ of two locally convex spaces, we shall use the canonical identification of the conjugate space $(X \times Z)^*$ with the Cartesian product $X^* \times Z^*$; namely, for any $\Phi \in X^*$ and $\Psi \in Z^*$, one defines $(\Phi, \Psi) \in (X \times Z)^*$ by

$$(\Phi, \Psi)(x, z) := \Phi(x) + \Psi(z) \quad (x \in X, z \in Z); \quad (1.28)$$

we shall often apply this to the particular case $Z = \mathbb{R}$.

In a locally convex space $X$, any set $H$ of the form

$$H = H_{\Phi, d} := \{y \in X| \Phi(y) = d\}, \quad (1.29)$$
where $\Phi \in X^* \setminus \{0\}$ and $d \in R$, is called a (closed) hyperplane. The sets

$$V_{\Phi,d}^\ge := \{y \in X| \Phi(y) \ge d\}, \quad V_{\Phi,d}^\le := \{y \in X| \Phi(y) \le d\}, \quad (1.30)$$

where $\Phi \in X^* \setminus \{0\}$ and $d \in R$, are called closed half-spaces (the closed half-spaces determined by the hyperplane $H = H_{\Phi,d}$ of (1.29)), and the sets

$$U_{\Phi,d}^\ge := \{y \in X| \Phi(y) > d\}, \quad U_{\Phi,d}^\le := \{y \in X| \Phi(y) < d\}, \quad (1.31)$$

where $\Phi \in X^* \setminus \{0\}$ and $d \in R$, are called open half-spaces (the open half-spaces determined by the hyperplane $H = H_{\Phi,d}$ of (1.29)).

Given two sets $C_1$ and $C_2$ in a locally convex space $X$, the hyperplane $H = H_{\Phi,d}$ of (1.29) is said to separate the sets $C_1$ and $C_2$ if $C_1$ is contained in one of the closed half-spaces determined by $H$ and $C_2$ is contained in the other closed half-space determined by $H$, that is, if either $C_1 \subseteq V_{\Phi,d}^\le$ and $C_2 \subseteq V_{\Phi,d}^\ge$ (if such $d \in R$ exists, or, equivalently, if $\sup \Phi(C_1) \le \inf \Phi(C_2)$, we say that $H$ or the function $\Phi \in X^* \setminus \{0\}$ separates $C_1$ from $C_2$), or $C_2 \subseteq V_{\Phi,d}^\le$ and $C_1 \subseteq V_{\Phi,d}^\ge$ (if such $d \in R$ exists, or, equivalently, if $\sup \Phi(C_2) \le \inf \Phi(C_1)$, we say that $H$ or $\Phi \in X^* \setminus \{0\}$ separates $C_2$ from $C_1$). The hyperplane $H = H_{\Phi,d}$ of (1.29) is said to strictly separate the sets $C_1$ and $C_2$ if $C_1$ is contained in one of the open half-spaces determined by $H$ and $C_2$ is contained in the other open half-space determined by $H$, that is, if either $C_1 \subseteq U_{\Phi,d}^\le$ and $C_2 \subseteq U_{\Phi,d}^\ge$ (if such $d \in R$ exists, we say that $H$ or $\Phi \in X^* \setminus \{0\}$ strictly separates $C_1$ from $C_2$), or $C_2 \subseteq U_{\Phi,d}^\le$ and $C_1 \subseteq U_{\Phi,d}^\ge$ (if such $d \in R$ exists, we say that $H$ or $\Phi \in X^* \setminus \{0\}$ strictly separates $C_2$ from $C_1$).

Clearly, $\Phi \in X^* \setminus \{0\}$ separates (respectively, strictly separates) $C_1$ from $C_2$ if and only if $-\Phi$ separates (respectively, strictly separates) $C_2$ from $C_1$.

Let us recall now some well-known theorems on separation of convex sets, which will be important tools later. The following is simply called “the separation theorem” (or “the basic separation theorem”):

**Theorem 1.1.** Let $C_1$ and $C_2$ be nonempty convex subsets of a topological linear space $X$, with $\text{int} \ C_1 \neq \emptyset$. Then $C_1$ can be separated from $C_2$ by a hyperplane, or, equivalently, there exists $\Phi \in X^* \setminus \{0\}$ such that

$$\sup \Phi(C_1) \le \inf \Phi(C_2). \quad (1.32)$$

if and only if

$$C_2 \cap \text{int} \ C_1 = \emptyset. \quad (1.33)$$

For the proof see, e.g., Ioffe and Tikhomirov [111, Ch. 3, Section 3.1]. The main part is, of course, the “if” part, but the “only if” part will also be used later.

**Remark 1.2.** In the particular case when $C_2$ is a singleton, from Theorem 1.1 we obtain the following: Let $C$ be a nonempty convex subset of a topological linear space $X$, with $\text{int} \ C \neq \emptyset$. There exists a hyperplane $H$ passing through $x$, such that $C$ lies in one of the closed half-spaces determined by $H$, if and only if $x \notin \text{int} \ C$. 
In locally convex spaces we have the following "strict separation theorem":

**Theorem 1.2.** If $C$ is a nonempty closed convex subset of a locally convex space $X$, and $x \notin C$, then there exists $\Phi \in X^* \setminus \{0\}$ that strictly separates $C$ from $x$, or, equivalently, such that

$$\sup \Phi(C) < \Phi(x). \quad (1.34)$$

We shall also use the following less standard result, which complements the separation theorems 1.1 and 1.2.

**Theorem 1.3.** Let $X$ be a normed linear space, $C$ a bounded convex subset of $X$, and $z_0 \in \overline{C}$. If either $\text{int} \ C \neq \emptyset$ or $C$ is closed, then there exists $\Psi \in X^* \setminus \{0\}$ such that

$$\Psi(z_0) = \sup \Psi(C). \quad (1.35)$$

**Proof.** Let us first observe that it is sufficient to consider the second case, i.e., that $C$ is closed. Indeed, if we are in the first case, i.e., $\text{int} \ C \neq \emptyset$, and if $z_0 \in \overline{C} \cap C^c = \text{bd} \ C$, then, by the separation theorem, there exists $\Psi \in X^* \setminus \{0\}$ satisfying (1.35); hence, we may assume that $z_0 \in \overline{C}^c$, and then, by the second case applied to the bounded closed convex set $C$, there exists $\Psi \in X^* \setminus \{0\}$ satisfying $\Psi(z_0) = \sup \Psi(C) = \sup \Psi(C)$.

Thus, assume that $C$ is a bounded closed convex set. Then, by $z_0 \in \overline{C}$ and the strict separation theorem,

$$\{ \Phi \in X^* | \sup \Phi(C) < \Phi(z_0) \} \neq \emptyset. \quad (1.36)$$

We shall show that every boundary point $\Psi$ (in the norm-topology of $X^*$) of the set $A := \{ \Phi \in X^* \setminus \{0\} | \sup \Phi(C) \leq \Phi(z_0) \}$ satisfies (1.35). To this end, it will be sufficient to prove that every point $\Phi_0$ of the set (1.36) is an interior point (in the norm-topology of $X^*$) of $A$. Let $\sup \Phi_0(C) < \Phi_0(z_0)$ and let $\varepsilon > 0$. Then, for $\| \Phi - \Phi_0 \|$ sufficiently small, we have $|\Phi_0(z_0) - \Phi_0(z_0)| < \varepsilon/2$. Furthermore, since the mapping $\Phi \mapsto \sup \Phi(C)$ is sublinear on $X^*$ (see (1.93) and (1.94) below) and $C$ is bounded, we have

$$|\sup \Phi_0(C) - \sup \Phi(C)| \leq \| \sup (\Phi_0 - \Phi) \sup C \| < +\infty,$$

and hence, for $\| \Phi - \Phi_0 \|$ sufficiently small, we have $|\sup \Phi_0(C) - \sup \Phi(C)| < \varepsilon/2$. Consequently, taking $\varepsilon > 0$ sufficiently small, and using that $\sup \Phi_0(C) < \Phi_0(z_0)$ and

$$\Phi(z_0) - \sup \Phi(C)$$

$$= [\Phi(z_0) - \Phi_0(z_0)] + [\Phi_0(z_0) - \sup \Phi_0(C)] + [\sup \Phi_0(C) - \sup \Phi(C)],$$

we obtain that for $\| \Phi - \Phi_0 \|$ sufficiently small, $\Phi(z_0) - \sup \Phi(C) > 0$. □
It is a well-known consequence of the strict separation theorem that a set $C$ in a locally convex space $X$ is closed and convex if and only if for every $x \notin C$ there exists $\Phi \in X^*$ satisfying (1.34), or, equivalently, a closed half-space $V$ such that $C \subseteq V$, $x \notin V$. In other words, $C$ is closed and convex if and only if it is an intersection of a family of closed half-spaces. For any set $C$ in a locally convex space $X$, we shall denote by $\overline{\text{co}} C$ the \textit{closed convex hull} of $C$, i.e., the intersection of all closed convex sets containing $C$.

A subset $C$ of a locally convex space $X$ is said to be \textit{evenly convex} if for each $x \in \overline{\text{co}} C$ there exists $\Phi \in X^*$ such that
\[
\Phi(c) < \Phi(x) \quad (c \in C),
\]
(1.37)
or, equivalently, an open half-space $U$ such that $C \subseteq U$, $x \notin U$. In other words, $C$ is evenly convex if and only if it is an intersection of a family of open half-spaces. It is also known that a connected set $C$ is evenly convex if and only if for every $x \notin C$ there exists a (closed) hyperplane $H$ with $x \in H$, $C \cap H = \emptyset$. For example, any open convex set and any closed convex set are evenly convex. For any set $C$ in a locally convex space $X$, we shall denote by $\text{eco} C$ the \textit{evenly convex hull} of $C$, i.e., the intersection of all evenly convex sets containing $C$.

A subset $C$ of a locally convex space $X$ is said to be \textit{evenly coaffine}, if for each $x \in \overline{\text{co}} C$ there exists $\Phi \in X^*$ such that
\[
\Phi(x) \notin \Phi(C)
\]
(1.38)
(i.e., such that $\Phi(x) \neq \Phi(c)$ ($c \in C$)); or, equivalently, $C$ is an intersection of a family of complements of hyperplanes; that is, $C$ is evenly coaffine if and only if it is the complement of the union of a family of hyperplanes. It is also known (see, e.g., [254, Ch. 2, Proposition 2.2]) that every evenly convex set is evenly coaffine, and every connected evenly coaffine set is evenly convex; hence ([254, Ch. 2, Corollary 2.2]) a set $C$ is evenly convex if and only if it is convex and evenly coaffine. For any set $C$ in a locally convex space $X$, we shall denote by $\text{eca} C$ the \textit{evenly coaffine hull} of $C$, i.e., the intersection of all evenly coaffine sets containing $C$.

For a set $C \subseteq X$ and for its closed convex hull $\overline{\text{co}} C$, we have
\[
\overline{\text{co}} C = \cap_{(\Phi, d) \in (X^* \setminus \{0\}) \times R} S_d(\Phi)
\]
\[
\sup \Phi(C) \leq d
\]
\[
= \{ x \in X \mid \exists (\Phi, d) \in (X^* \setminus \{0\}) \times R, \sup \Phi(C) \leq d < \Phi(x) \}. \quad (1.39)
\]
Furthermore, for a set $C \subseteq X$ and for its evenly convex hull $\text{eco} C$, we have
\[
\text{eco} C = \cap_{(\Phi, d) \in (X^* \setminus \{0\}) \times R} A_d(\Phi)
\]
\[
\Phi(c) < d \quad (c \in C)
\]
\[
= \{ x \in X \mid \exists (\Phi, d) \in (X^* \setminus \{0\}) \times R, \Phi(c) < d \leq \Phi(x) \quad (c \in C) \}. \quad (1.40)
\]
We recall that a hyperplane
1.1 Some preliminaries from convex analysis

\[ H = \{ y \in X | \Phi(y) = d \}, \quad (1.41) \]

where \( \Phi \in X^* \setminus \{0\} \), \( d \in R \), is a **support hyperplane** of a set \( C \subseteq X \), or, briefly, \( H \) **supports** \( C \) if \( C \cap H \neq \emptyset \) and \( C \) is contained in one of the two closed half-spaces determined by \( H \), that is, either

\[ C \subseteq \{ y \in X | \Phi(y) \leq d \} \quad (1.42) \]

or

\[ C \subseteq \{ y \in X | \Phi(y) \geq d \}. \quad (1.43) \]

Let us also introduce the following more general concept, in which we do not require \( C \cap H \neq \emptyset \):

**Definition 1.1.** Let \( X \) be a locally convex space. We shall say that a (closed) hyperplane \( H \) is a **quasi-support hyperplane** of a set \( C \subseteq X \), or that \( H \) quasi-supports \( C \), if we have either (1.42) and

\[ C \nsubseteq \{ y \in X | \Phi(y) < d \} - \varepsilon \quad (1.44) \]

or (1.43) and

\[ C \nsubseteq \{ y \in X | \Phi(y) > d + \varepsilon \} \quad (1.45) \]

Clearly, if also \( C \cap H \neq \emptyset \), then (1.42) and (1.43) imply, respectively, (1.44) and (1.45), and \( H \) is a support hyperplane of \( C \). Thus, every support hyperplane is a quasi-support hyperplane. However, the converse is not true, as shown by the following example.

**Example 1.1.** Let \( X = c_0 \), the Banach space of all sequences \( x = (x_n) \) converging to 0, with the norm (1.7), \( C = B_X := \{ y \in X | \|y\| \leq 1 \} \) (the unit ball of \( X \)), and let us consider the function \( \Phi_0 \in X^* \) defined by

\[ \Phi_0(x) = \sum_{i=1}^{\infty} \frac{1}{2^i} x_i \quad (x = (x_n) \in c_0). \quad (1.46) \]

Then \( C \subseteq X \) is a bounded closed convex set, \( \Phi_0 \in X^* \), \( \|\Phi_0\| = \sup \Phi_0(B_X) = 1 \), and there exists no \( x \in B_X \) such that \( \Phi_0(x) = 1 \). Hence, the hyperplane

\[ H = \{ y \in X | \Phi_0(y) = 1 \} \quad (1.47) \]

quasi-supports the unit ball \( B_X \) (by Lemma 1.4 below), but it does not support \( B_X \).

One can also give such examples in finite-dimensional spaces, but only with unbounded closed convex sets \( C \), as shown by Lemma 1.4 below (since a bounded closed set \( C \) in a finite-dimensional space \( X \) is compact, and hence every \( \Phi \in X^* \) attains its supremum and its infimum on \( C \)). Indeed:
Example 1.2. Let $X = \mathbb{R}^2$, with any one of the norms \((1.2)-(1.4)\), $C := \{x = (x_1, x_2) \in \mathbb{R}^2 | x_1 > 0, x_1 x_2 < -1\}$, and let us consider the function
\[
\Phi_0(x) = x_2 \quad (x = (x_1, x_2) \in \mathbb{R}^2). \tag{1.48}
\]
Then $C$ is a closed convex set, $\Phi_0 \in X^*$, $\|\Phi_0\| = 1$, and $\Phi_0(c) < \sup \Phi_0(C) = 0$ ($c \in C$). Hence, the hyperplane $H = \{y \in \mathbb{R}^2 | \Phi_0(y) = 0\}$ quasi-supports the set $C$, but it does not support $C$.

Remark 1.3. (a) From \((1.42)\) or \((1.43)\) it follows that if a hyperplane $H$ quasi-supports a set $C$, then $H$ quasi-supports also $\overline{c0} C$.

(b) Theorem 1.3 admits the following geometric interpretation: under the assumptions of Theorem 1.3, for each $z_0 \in \overline{C} C$ there exists a quasi-support hyperplane of $C$ passing through $z_0$ (see Figure 1.2).

![Figure 1.2.](image)

Now we shall present some results on quasi-support hyperplanes and support hyperplanes, since they will play an important role in the geometric interpretations of the duality results in approximation and optimization.

Lemma 1.4. A hyperplane \((1.41)\) in a locally convex space $X$ quasi-supports a set $C \subseteq X$ if and only if either $d = \sup C$ or $d = \inf C$.

Proof. Assume that \((1.41)\) quasi-supports $C$, say, \((1.42)\) and \((1.44)\) hold. Then, by \((1.42)\), we have $\sup \Phi(C) \leq d$. On the other hand, by \((1.44)\), for each $\varepsilon > 0$ there exists $c_\varepsilon \in C$ such that $\Phi(c_\varepsilon) > d - \varepsilon$, whence $\sup \Phi(C) \geq d - \varepsilon$ for all $\varepsilon > 0$, so $\sup \Phi(C) \geq d$. Hence, finally, $d = \sup \Phi(C)$. The case of \((1.43)\), \((1.45)\) is similar.

Conversely, assume now that $d = \sup \Phi(C)$, that is, $H = \{y \in X | \Phi(y) = \sup \Phi(C)\}$. Then clearly we have \((1.42)\). Furthermore, for each $\varepsilon > 0$ there exists $c_\varepsilon \in C$ such that $\Phi(c_\varepsilon) > \sup \Phi(C) - \varepsilon = d - \varepsilon$, which shows that we also have \((1.44)\). The case $d = \inf \Phi(C)$ is similar, mutatis mutandis. \(\square\)

Corollary 1.1. Every hyperplane $H$ quasi-supporting $C$ can be written in the form
\[
H = \{y \in X | \Phi(y) = \sup \Phi(C)\}; \tag{1.49}
\]
where $\Phi \in X^* \setminus \{0\}$, $\sup \Phi(C) \in \mathbb{R}$, and conversely, every hyperplane $H$ of the form \((1.49)\), where $\Phi \in X^* \setminus \{0\}$, $\sup \Phi(C) \in \mathbb{R}$, quasi-supports the set $C$. 


Proof. By Lemma 1.4, it is enough to consider the case when \( H = \{ y \in X \mid \Phi_0(y) = \inf \Phi_0(C) \} \). In this case, by \( \inf \Phi_0(C) = - \sup( - \Phi_0)(C) \), we have \( H = \{ y \in X \mid - \Phi_0(y) = \sup( - \Phi_0)(C) \} \), i.e., (1.49) for \( \Phi = - \Phi_0 \).

Let us consider now the particular case that \( X \) is a normed linear space. We recall that for any two elements \( x, y \in X \) one defines the distance between \( x \) and \( y \) by (1.15). For any nonempty set \( C \subseteq X \) and any \( x_0 \in X \), one defines the distance from \( x_0 \) to \( C \) by

\[
\text{dist}(x_0, C) := \inf_{c \in C} \| x_0 - c \|. \tag{1.50}
\]

The distance from \( x_0 \) to the empty set is defined by the following convention:

\[
\text{dist}(x_0, \emptyset) = +\infty. \tag{1.51}
\]

By (1.50), for any nonempty set \( C \subseteq X \) we have

\[
\text{dist}(x_0, C) = \text{dist}(x_0, \overline{C}); \tag{1.52}
\]

by (1.51) and \( \emptyset = \emptyset \), the equality (1.52) remains valid also for \( C = \emptyset \).

The distance between two subsets \( C_1, C_2 \) of a normed linear space \( X \) is defined by

\[
\text{dist}(C_1, C_2) := \inf_{c_1 \in C_1, c_2 \in C_2} \| c_1 - c_2 \|. \tag{1.53}
\]

We shall need the following lemma, which gives a useful formula for the distance to a hyperplane.

**Lemma 1.5.** Let \( X \) be a normed linear space, and let \( H = \{ x \mid \Phi_0(x) = \Phi_0(C) \} \) be a hyperplane (1.29), where \( \Phi \in X^* \setminus \{0\}, d \in \mathbb{R} \). Then, for any \( x_0 \notin H \), we have

\[
\text{dist}(x_0, H) = \frac{|\Phi(x_0) - d|}{\| \Phi \|}. \tag{1.54}
\]

**Proof.** For any \( y \in H \),

\[
\| x_0 - y \| \geq \frac{|\Phi(x_0 - y)|}{\| \Phi \|} = \frac{|\Phi(x_0) - d|}{\| \Phi \|},
\]

whence

\[
\text{dist}(x_0, H) \geq \frac{|\Phi(x_0) - d|}{\| \Phi \|}. \tag{1.55}
\]

On the other hand, if \( 0 < \varepsilon < \| \Phi \| \), there exists \( z \in X \) such that \( |\Phi(z)| > (\| \Phi \| - \varepsilon) \| z \| \). Multiplying this inequality by \( |\Phi(x_0) - d| / \Phi(z) \) and putting

\[
y := x_0 - \frac{\Phi(x_0) - d}{\Phi(z)} z, \tag{1.56}
\]
we obtain
\[
|\Phi(x_0) - d| > (\|\Phi\| - \varepsilon) \left| \frac{\Phi(x_0) - d}{\Phi(z)} \right| ||z|| = (\|\Phi\| - \varepsilon) \|x_0 - y\| ,
\]
whence, since \( y \in H_{\Phi,d} \),
\[
\text{dist}(x_0, H_{\Phi,d}) \leq \|x_0 - y\| < \frac{|\Phi(x_0) - d|}{\|\Phi\| - \varepsilon} . \tag{1.57}
\]
Since \( \varepsilon > 0 \) was arbitrary, we obtain the opposite inequality to (1.55), and hence the equality (1.54). \( \square \)

Using the distance (1.53) and applying Lemma 1.5, one can give the following characterization of quasi-support hyperplanes in normed linear spaces:

**Proposition 1.1.** Let \( X \) be a normed linear space. For a set \( C \subseteq X \) and a hyperplane (1.41), the following statements are equivalent:

1°. \( H \) quasi-supports the set \( C \).

2°. We have (1.42) or (1.43), and
\[
\text{dist}(C, H) = \inf_{c \in C, h \in H} \|c - h\| = 0 . \tag{1.58}
\]

**Proof.** 1° \( \Rightarrow \) 2°. Assume 1°. Then we have either (1.42) or (1.43), so we may assume (1.42) (in the case of (1.43), replacing \( \Phi \) by \( -\Phi \), we arrive at the case (1.42)). Then, condition (1.44) means that for each \( \varepsilon > 0 \) there exists \( c_\varepsilon \in C \) such that \( \Phi(c_\varepsilon) > d - \varepsilon \). Using (1.42) as well, we obtain
\[
d \geq \sup \Phi(C) \geq \Phi(c_\varepsilon) > d - \varepsilon \quad (\varepsilon > 0),
\]
whence \( \sup \Phi(C) = d \), so \( H = \{ y \in X | \Phi(y) = \sup \Phi(C) \} \). Then, by Lemma 1.5,
\[
\text{dist}(c, H) = \frac{1}{\|\Phi\|} |\Phi(c) - \sup \Phi(C)| = \frac{1}{\|\Phi\|} \{ \sup \Phi(C) - \Phi(c) \} \quad (c \in C),
\]
whence
\[
\text{dist}(C, H) = \frac{1}{\|\Phi\|} \inf_{c \in C} \text{dist}(c, H) = \frac{1}{\|\Phi\|} \inf_{c \in C} \{ \sup \Phi(C) - \Phi(c) \} = 0 .
\]

2° \( \Rightarrow \) 1°. Assume 2° and (1.42) (in the case of (1.43), replacing \( \Phi \) by \( -\Phi \), we arrive at the case (1.42)), and let \( \varepsilon > 0 \). Then by (1.58), there exist \( c_\varepsilon \in C \) and \( h_\varepsilon \in H \) such that \( \|c_\varepsilon - h_\varepsilon\| < \frac{1}{\|\Phi\|} \varepsilon \). Hence,
\[
\Phi(h_\varepsilon) - \Phi(c_\varepsilon) = \Phi(h_\varepsilon - c_\varepsilon) \leq \|\Phi\| \|h_\varepsilon - c_\varepsilon\| < \varepsilon ,
\]
and thus \( \Phi(c_\varepsilon) > \Phi(h_\varepsilon) - \varepsilon = d - \varepsilon \), which proves (1.44). \( \square \)