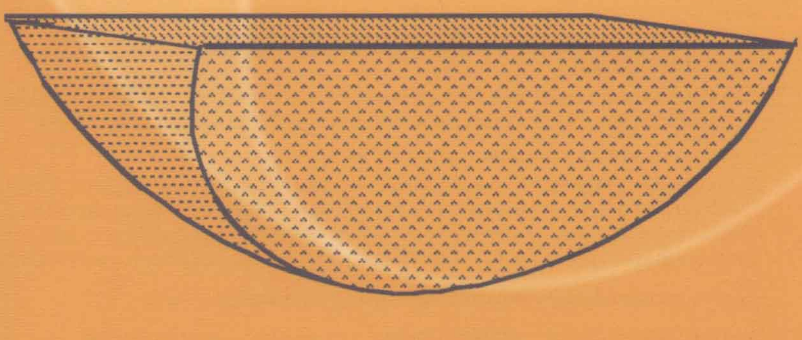
Lecture Notes in Mathematics

1693

Stephen Simons

From Hahn-Banach to Monotonicity

Second Edition



Springer

Stephen Simons

From Hahn-Banach to Monotonicity

2nd, expanded edition

Springer

Stephen Simons
Department of Mathematics
University of California
Santa Barbara, CA 93105-3080
USA
simons@math.ucsb.edu
http://www.math.ucsb.edu/~simons

1st edition 1998 LNM 1693: Minimax and Monotonicity

ISBN 978-1-4020-6918-5 e-ISBN 978-1-4020-6919-2

DOI 10.1007/978-1-4020-6919-2

Lecture Notes in Mathematics ISSN print edition: 0075-8434
ISSN electronic edition: 1617-9692

Library of Congress Control Number: 2007942159

Mathematics Subject Classification (2000): 46A22, 49J35, 47N10, 49J52, 47H05

Cover design: *WMXDesign GmbH*

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

For Jacqueline,
whose support and patience
are unbounded.

Preface

A more accurate title for these notes would be: "The Hahn–Banach–Lagrange theorem, Convex analysis, Symmetrically self–dual spaces, Fitzpatrick functions and monotone multifunctions".

The Hahn–Banach–Lagrange theorem is a version of the Hahn–Banach theorem that is admirably suited to applications to the theory of monotone multifunctions, but it turns out that it also leads to extremely short proofs of the standard existence theorems of functional analysis, a minimax theorem, a Lagrange multiplier theorem for constrained convex optimization problems, and the Fenchel duality theorem of convex analysis.

Another feature of the Hahn–Banach–Lagrange theorem is that it can be used to transform problems on the existence of continuous linear functionals into problems on the existence of a single real constant, and then obtain a sharp lower bound on the norm of the linear functional satisfying the required condition. This is the case with both the Lagrange multiplier theorem and the Fenchel duality theorem applications mentioned above.

A multifunction from a Banach space into the subsets of its dual can, of course, be identified with a subset of the product of the space with its dual. Simon Fitzpatrick defined a convex function on this product corresponding with any such multifunction. So part of these notes is devoted to the rather special convex analysis for the product of a Banach space with its dual.

The product of a Banach space with its dual is a special case of a "symmetrically self–dual space". The advantage of going to this slightly higher level of abstraction is not only that it leads to more general results but, more to the point, it cuts the length of each proof approximately in half which, in turn, gives a much greater insight into the nature of the processes involved. Monotone multifunctions then correspond to subsets of the symmetrically self–dual space that are "positive" with respect to a certain quadratic form.

We investigate a particular kind of convex function on a symmetrically self–dual space, which we call a "BC–function". Since the Fitzpatrick function of a maximally monotone multifunction is always a BC–function, these BC–functions turn out to be very successful for obtaining results on maximally monotone multifunctions on reflexive spaces.

The situation for nonreflexive spaces is more challenging. Here, it turns out that we must consider two symmetrically self–dual spaces, and we call the corresponding convex functions "$\widetilde{BC}$–functions". In this case, a number of different subclasses of the maximally monotone multifunctions have been introduced over the years — we give particular attention to those that are "of type (ED)". These have the great virtue that all the common maximally monotone multifunctions are of type (ED), and maximally monotone multifunctions of type (ED) have nearly all the properties that one could desire. In order to study the maximally monotone multifunctions of type (ED), we have to introduce a weird topology on the bidual which has a number of very nice properties, despite that fact that it is not normally compatible with its vector space structure.

These notes are somewhere between a sequel to and a new edition of [99]. As in [99], the essential idea is to reduce questions on monotone multifunctions to questions on convex functions. In [99], this was achieved using a "big convexification" of the graph of the multifunction and the "minimax technique" for proving the existence of linear functionals satisfying certain conditions. The "big convexification" is a very abstract concept, and the analysis is quite heavy in computation. The Fitzpatrick function gives another, more concrete, way of associating a convex functions with a monotone multifunction. The problem is that many of the questions on convex functions that one obtains require an analysis of the special properties of convex functions on the product of a Banach space with its dual, which is exactly what we do in these notes. It is also worth noting that the minimax theorem is hardly used here.

We envision that these notes could be used for four different possible courses/seminars:

- An introductory course in functional analysis which would, at the same time, touch on minimax theorems and give a grounding in convex Lagrange multiplier theory and the main theorems in convex analysis.
- A course in which results on monotonicity on general Banach spaces are established using symmetrically self–dual spaces and Fitzpatrick functions.
- A course in which results on monotonicity on reflexive Banach spaces are established using symmetrically self–dual spaces and Fitzpatrick functions.
- A seminar in which the the more technical properties of maximal monotonicity on general Banach spaces that have been established since 1997 are discussed.

We give more details of these four possible uses at the end of the introduction.

I would like to express my sincerest thanks to Heinz Bausckhe, Patrick Combettes, Michael Crandall, Carl de Boor, Radu Ioan Boţ, Juan Enrique Martínez-Legaz, Xianfu Wang and Constantin Zălinescu for reading preliminary versions of parts of these notes, making a number of excellent suggestions and, of course, finding a number of errors.

Of course, despite all the excellent efforts of the people mentioned above, these notes doubtless still contain errors and ambiguities, and also doubtless have other stylistic shortcomings. At any rate, I hope that there are not too many of these. Those that do exist are entirely my fault.

Stephen Simons
September 23, 2007
Santa Barbara
California

Table of Contents

III Multifunctions, SSD spaces, monotonicity and Fitzpatrick functions

IV Monotone multifunctions on general Banach spaces

V Monotone multifunctions on reflexive Banach spaces

VI Special maximally monotone multifunctions

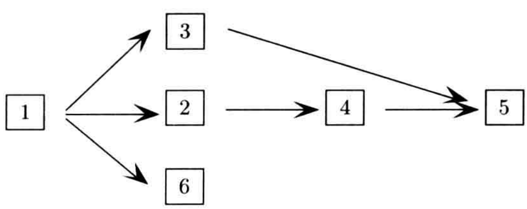
VII The sum problem for general Banach spaces

Introduction

These notes fall into three distinct parts. In Chapter I, we discuss the "Hahn–Banach–Lagrange theorem", a new version of the Hahn–Banach theorem, which gives very efficient proofs of the main existence theorems in functional analysis, optimization theory, minimax theory and convex analysis. In Chapter II, we zero in on the applications to convex analysis. In the remaining five chapters, we show how the results of the first two chapters can be used to obtain a large number of results on monotone multifunctions, many of which have not yet appeared in print.

Chapter I: The main result in Chapter I is the "Hahn-Banach-Lagrange" theorem, which first appeared in [103]. We prove this result in Theorem 1.11, discuss the classical functional analytic applications in Section 2 (namely the "Sandwich theorem" in Corollary 2.1, the "extension form of the Hahn–Banach theorem" in Corollary 2.2, and the "one dimensional form of the Hahn–Banach theorem" in Corollary 2.4) and give an application to a classical minimax theorem in Section 3. In Section 4, we introduce the results from classical Banach space theory that we shall need. In Section 5, we prove, among other things, a minimax criterion for a subset of a Banach space to be weakly compact using the concepts of "excess" and "duality gap". The contents of this section first appeared in [102].

In Section 6, we give a necessary and sufficient condition for the existence of Lagrange multipliers for constrained convex optimization problems (generalizing the classical sufficient "Slater condition"), with a sharp lower bound on the norm of the multiplier. We also prove a similar result for Karush–Kuhn–Tucker problems for functions with convex Gâteaux derivatives. Some of the results on Lagrange multipliers first appeared in [104]. In the flowchart below, we show the dependencies of the sections in Chapter I. We note, in particular, that Section 6 does not depend on Sections 2–5.



Chapter II: As explained above, Chapter II is about convex analysis. We start our discussion in Section 7 by using the Hahn-Banach-Lagrange theorem to obtain a necessary *and sufficient* condition for the Fenchel duality theorem to hold for two convex functions on a normed space, with a sharp lower bound on the norm of the functional obtained. (Incidentally, this approach avoids the aggravating problem of the "vertical hyperplane" that so destroys the elegance of the usual approach through the Eidelheit separation theorem.) This sharp version of the Fenchel duality theorem is in Theorem 7.4, and it is explained in Remark 7.6 how the lower bound obtained is of a very geometric character.

While the concept of Fenchel conjugate is introduced in Section 7 with reference to a convex function on a normed space, in fact this causes no end of confusion when dealing with monotone multifunctions on a nonreflexive Banach space. The way out of this problem (as has been observed by many authors) is to define Fenchel conjugates with respect to a dual pair of spaces. This is what we do in Section 8, and it enables a painless transition to the locally convex case. As we will see in Section 22, this is exactly what we need for our discussion of monotone multifunctions on a nonreflexive Banach space. We present a necessary and sufficient condition for the Fenchel duality theorem to be true in this sense in Theorem 8.1, and in Theorem 8.4 we present a unifying sufficient condition that implies the results that are used in practice, the versions due to Rockafellar and Attouch–Brezis. Theorem 8.4 uses the binary operation $\ominus$ defined in Notation 8.3.

In Section 9, we return to the normed case and give some results of a more numerical character, in which we explore the properties of the function $\frac{1}{2}\|\cdot\|^2$. These results will enable us to give a precise expression for the minimum norm of the resolvent of a maximally monotone multifunction on a reflexive Banach space in Theorem 29.5.

We bootstrap Theorem 8.4 in Section 10, and obtain sufficient conditions for the "inf–convolution" formula for the conjugate of a sum to hold, and give as application in Corollary 10.4 a consequence that will be applied in Theorem 21.10 to the existence of autoconjugates in SSDB spaces. This bootstrapping operation exhibits the well known fact that results on the conjugate of a sum are very close to the Fenchel duality theorem. However, these concepts are not interchangeable, and in Section 11 we give examples which should serve to distinguish them (giving examples of the failure of "stability" in duality).

In Section 12, we introduce the concepts of the *biconjugate* of a convex function, and the *Fenchel–Moreau points* of a convex function on a locally convex space. We deduce the Fenchel–Moreau formula in Corollary 12.4 in the case where the function is lower semicontinuous. Some of these results first appeared in [103].

We collect together in Sections 13 and 14 various results on convex functions that depend ultimately on Baire's theorem. The "dom lemma", Lemma 13.3, is a generalization to convex functions of the classical uniform boundedness (Banach Steinhaus) theorem (see Remark 13.6) and the "$\ominus$–theorem", Theorem 14.2 (which uses the operation $\ominus$ already mentioned) is a generalization to convex functions of the classical open mapping theorem (see Remark 14.4). Both of these results will be applied later on to obtain results on monotonicity. We can think of the dom lemma and the $\ominus$–theorem as "quantitative" results, since their main purpose is to provide numerical bounds. Associated with them are two "qualitative" results, the "dom corollary", Corollary 13.5, and the "$\ominus$–corollary", Corollary 14.3, from which the numerics have been removed. The $\ominus$–corollary will also be of use to us later on. In Remark 14.5, we give a brief discussion of convex Borel sets and functions.

In Theorem 15.1, we show how the $\ominus$–theorem leads to the Attouch–Brezis version of the Fenchel duality theorem, which we will use (via the local transversality theorem, Theorem 21.12, and Theorem 30.1) to prove various surjectivity results, including an abstract Hammerstein theorem; and in Theorem 16.4 we obtain a bivariate version of the Attouch-Brezis theorem, which we will use in Theorem 24.1 (via Lemma 22.9) and in Theorem 35.8 (a result that is fundamental for the understanding of maximally monotone multifunctions on nonreflexive Banach spaces). This bivariate version of the Attouch–Brezis theorem first appeared in [109].

Chapter III: In Chapter III, we will discuss the basic result on monotonicity. Section 17 starts off with a conventional discussion of multifunctions, monotonicity and maximal monotonicity. Remark 17.1 is a bridge in which we show that if E is a Banach space then there is a vector space B and a quadratic form q on B such that if $S\colon\ E \rightrightarrows E^*$ is a multifunction then there is a subset A of B such that S is monotone if and only if $b, c \in A \Longrightarrow q(b-c) \geq 0$. Actually $B = E \times E^*$ and $A = G(S)$, but this paradigm leads to a strict generalization of monotonicity, in which the proofs are much more concise.

In Section 18, we digress a little from the general theory in order to give a short proof of Rockafellar's fundamental result that the subdifferential of a proper convex lower semicontinuous function on a Banach space is maximally monotone. In Theorem 18.1 and Theorem 18.2, we give the formula for the subdifferential of the sum of convex functions under two different hypotheses, in Corollary 18.5 and Theorem 18.6, we show how to deduce the Brøndsted–Rockafellar theorem from Ekeland's variational principle and the Hahn–Banach–Lagrange theorem, and then we come finally to our proof of the maximal monotonicity of subdifferentials in Theorem 18.7, which is based on the very elegant one found recently by M. Marques Alves and B. F. Svaiter in [60]. We also give in Corollary 18.3 and Theorem 18.10 two results about normal cones that will be useful later on. Readers who are familiar with the formula for the subdifferential of the sum of convex functions and the Brøndsted–Rockafellar theorem should be able to understand this section without having to read any of the previous sections.

We return to our development of the general theory in Sections 19–21. In Section 19, we introduce the concept of a SSD (symmetrically self–dual) space, a nonzero real vector space with a symmetric bilinear form which separates points. This bilinear form defines a quadratic form, q, in the obvious way. In general, this quadratic form is not positive, but we isolate certain subsets of a SSD space that we will call "q–positive". Appropriate convex functions on the SSD space define q–positive sets. We zero in on a subclass of the convex functions on a SSD space which we call "BC–functions". Critical to this enterprise is the self–dual property, because the conjugate of a convex function has the same domain of definition as the original convex function. Lemma 19.12 contains an unexpected result on BC–functions, but the most important result on BC–functions is undoubtedly the transversality theorem, Theorem 19.16, which leads (via Theorem 21.4) to generalizations of Rockafellar's classical surjectivity theorem for maximally monotone multifunctions on a reflexive Banach space (see Theorem 29.5) together with a sharp lower bound on the norm of solutions in terms of the Fitzpatrick function (see Theorem 29.6), and to sufficient conditions for the sum of maximally monotone multifunctions on a reflexive Banach space to be maximally monotone (see Theorem 24.1). Section 19 concludes with a discussion of how every q–positive set, A, gives rise to a convex function, Φ_A (this construction is an abstraction of the construction of the "Fitzpatrick function" that we will consider in Section 23). In Section 20, we introduce maximally q–positive sets, and show that the convex function determined by a maximally q–positive set is a BC–function.

In Section 21, we introduce the SSDB spaces, which are SSD spaces with an appropriate Banach norm. Roughly speaking, the additional structure that SSDB spaces possess over SSD spaces is ultimately what accounts for the fact that maximally monotone multifunctions on reflexive Banach spaces are much more tractable than maximally monotone multifunctions on general Banach spaces. That is not to say that the SSD space determined by a nonreflexive Banach space does not have a norm structure, the problem is that this norm structure is not "appropriate". Apart from Theorem 21.4, which we have already mentioned, the other important results in this section are Theorem 21.10 on the existence of autoconjugates, Theorem 21.11, which gives a formula for a maximally q–positive superset of a given nonempty q–positive set, and the local transversality theorem, Theorem 21.12, which leads ultimately to a number of surjectivity results, including an abstract Hammerstein theorem in Section 30.

We start considering in earnest the special SSD space $E \times E^*$ (where E is a nonzero Banach space) in Section 22. We first prove some preliminary results which depend ultimately on Rockafellar's version of the Fenchel duality theorem introduced in Corollary 8.6. It is important to realize that, despite the fact that E is a Banach space, we need Corollary 8.6 for *locally convex spaces* since the topology we are using for this result is the topology

$\mathcal{T}_{\| \ \|}(E) \times w(E^*, E)$ on $E \times E^*$. Theorem 22.5 has a precise description of the projection on E of the domain of the conjugate of a proper convex function on $E \times E^*$ in terms of a related convex function on E. In Theorem 22.8, we establish the equality of six sets determined by certain proper convex functions on $E \times E^*$, and in Lemma 22.9 we prove a result which will be critical for our treatment of sum theorems for maximally monotone multifunctions in Theorem 24.1.

In Section 23, we show how the concepts introduced in Sections 19–21 specialize to the situation considered in Section 22. The q–positive sets introduced in Section 19 then become the graphs of monotone multifunctions, the maximally q–positive sets introduced in Section 20 then become the graphs of maximally monotone multifunctions, and the function Φ_A determined by a q–positive set A introduced in Section 19 becomes the Fitzpatrick function, φ_S, determined by a monotone multifunction S. The Fitzpatrick function was originally introduced in [42] in 1988, but lay dormant until it was rediscovered by Martínez-Legaz and Théra in [63] in 2001.

This is an appropriate place for us to make a comparison between the analysis presented in these notes with the analysis presented in [99]. In both cases, the essential idea is to reduce questions on monotone multifunctions to questions on convex functions. In [99], this was achieved using a "big convexification" of the graph of the multifunction and the "minimax technique" for proving the existence of linear functionals satisfying certain conditions. This technique is very successful for working back from conjectures, and finding conditions under which they hold. On the other hand, the "big convexification" is a very abstract concept, and the analysis is quite heavy in computation. Now the Fitzpatrick function gives another way of associating a convex functions with a monotone multifunction, and this can also be used to reduce questions on monotone multifunctions to questions on convex functions. The problem is that many of the questions on convex functions that one obtains require an analysis of the special properties of convex functions on $E \times E^*$. This is exactly the analysis that we perform in Section 22, and later on in Section 35. As already explained, the SSD spaces introduced in Sections 19–21 give us a strict generalization of monotonicity. More to the point, the fact that the notation is more concise enables us to get a much better grasp of the underlying structures. A good example of this is Theorem 35.8, a relatively simple result with far–reaching applications to the classification of maximally monotone multifunctions on nonreflexive spaces. Another example is provided by Section 46 on maximally monotone multifunctions with convex graph.

We now return to our discussion of Section 23. We also introduce the "fitzpatrification", S_φ, of a monotone multifunction, S. This is a multifunction with convex graph which is normally much larger than the graph of S. S_φ is, in general, not monotone, but it is very useful since its use shortens the statements of many results considerably. The final result of this section, Lemma 23.9, will be used in our discussion of the sum problem in Theorem 24.1 and the Brezis–Haraux condition in Theorem 31.4.

In Section 24, we give sufficient conditions for the sum of maximally monotone multifunctions to be maximally monotone. These results will be extended in the reflexive case in Section 32, and we will discuss the nonreflexive case in Chapter VII.

Chapter IV: In Chapter IV, we use results from Sections 4, 12, 18 and 23 to establish a number of results on monotone multifunctions on general Banach spaces. Section 25 is devoted to the single result that a maximally monotone multifunction with bounded range has full domain, and in Section 26, we prove a local boundedness theorem for any (not necessarily maximally) monotone multifunction on a Banach space. Specifically, we prove that *a monotone multifunction, S, is locally bounded at any point surrounded by $D(S_\varphi)$.*

In Section 27, we prove the "six set theorem", Theorem 27.1, that if S is maximally monotone then the six sets $\operatorname{int} D(S)$, $\operatorname{int}(\operatorname{co} D(S))$, $\operatorname{int} D(S_\varphi)$, $\operatorname{sur} D(S)$, $\operatorname{sur}(\operatorname{co} D(S))$ and $\operatorname{sur} D(S_\varphi)$ coincide, and its consequence, the "nine set theorem", Theorem 27.3, that, if $\operatorname{sur} D(S_\varphi) \neq \emptyset$, then the nine sets $\overline{D(S)}$, $\overline{\operatorname{co} D(S)}$, $\overline{D(S_\varphi)}$, $\overline{\operatorname{int} D(S)}$, $\overline{\operatorname{int}(\operatorname{co} D(S))}$, $\overline{\operatorname{int} D(S_\varphi)}$, $\overline{\operatorname{sur} D(S)}$, $\overline{\operatorname{sur}(\operatorname{co} D(S))}$ and $\overline{\operatorname{sur} D(S_\varphi)}$ coincide. ("Sur" is defined in Section 13.) The six set theorem and the nine set theorem not only extend the results of Rockafellar that $\operatorname{int} D(S)$ is convex and that, if $\operatorname{int}(\operatorname{co} D(S)) \neq \emptyset$ then $\overline{D(S)}$ is convex, but also answer in the affirmative a question raised by Phelps, namely whether an absorbing point of $D(S)$ is necessarily an interior point. In Theorem 27.5 and Theorem 27.6, we give sufficient conditions that $\overline{D(S)} = \overline{D(S_\varphi)}$ and $\overline{R(S)} = \overline{R(S_\varphi)}$ — these conditions do not have any interiority hypotheses.

Section 28 contains the results that if S is maximally monotone then the closed linear hull of $D(S_\varphi)$ is identical with the closed linear hull of $D(S)$, and the closed affine hull of $D(S_\varphi)$ is identical with the closed affine hull of $D(S)$. The arguments here are quite simple, which is in stark contrast with the similar question for closed convex hulls. This section also contains some results for pairs of multifunctions, which will be used in our analysis of bootstrapped sum theorems for reflexive spaces in Section 32. We also give some results on the "restriction" of a monotone multifunction to a closed subspace. The results in this section depend ultimately on the result of Lemma 20.4 on q–positive sets that are "flattened" by certain elements of a SSD space.

Chapter V: Chapter V is concerned with maximally monotone multifunctions on reflexive Banach spaces. In Section 29, we use the theory of SSDB spaces developed in Section 21 to obtain various criteria for a monotone multifunctions on a reflexive Banach space to be maximally monotone. We deduce in Theorem 29.5 and Theorem 29.6 Rockafellar's surjectivity theorem, together with a sharp lower bound on the norm of solutions in terms of the Fitzpatrick function. Theorem 29.8 contains an expression for a max-