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 "$\widetilde{\mathrm{BC}}$–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

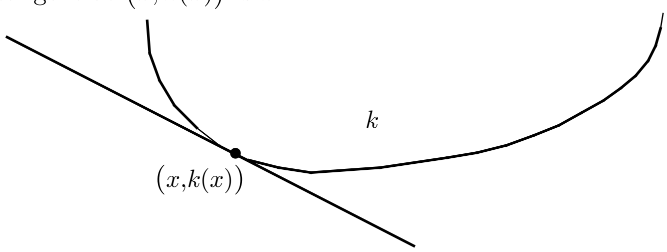
II Fenchel duality

7 A sharp version of the Fenchel Duality theorem

This section is similar in spirit to Section 6, but this time we apply the Hahn–Banach–Lagrange theorem to problems in convex analysis. So suppose that E is a nonzero normed space with dual E^*, $k \in \mathcal{PC}(E)$ and $x \in E$. Then the *subdifferential* of k at x is defined by

$$\partial k(x) := \{z^* \in E^*\colon \quad y \in E \Longrightarrow k(x) + \langle y - x, z^* \rangle \le k(y)\}.$$

We first consider the question of when $\partial k(x) \neq \emptyset$. This question is answered in Example 7.1 below. The justification, using Theorem 1.11, follows similar lines to those used in Section 6(exercise!). In the sketch below, the "slope" of the subtangent at $\big(x, k(x)\big)$ is z^*.

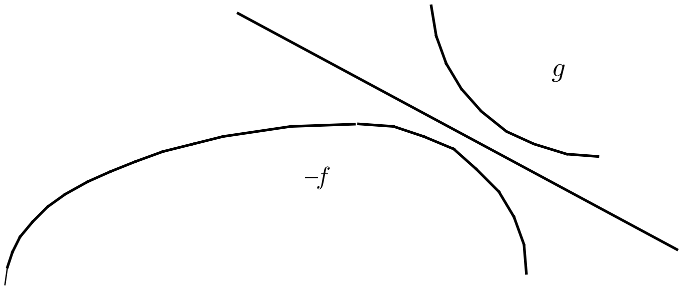


Example 7.1. *$\partial k(x) \neq \emptyset$ if, and only if, $x \in \operatorname{dom} k$ and there exists $M \ge 0$ such that*

$$y \in E \quad \Longrightarrow \quad k(x) - M\|y - x\| \le k(y).$$

We now come to a more interesting example. Again, suppose that E is a nonzero normed space with dual E^*. When can a concave function and a convex function be separated by a continuous affine function? More precisely, given $f, g \in \mathcal{PC}(E)$, when do there exist $z^* \in E^*$ and $\beta \in \mathbb{R}$ such that

$$-f \le z^* + \beta \le g \quad \text{on} \quad E? \tag{7.1}$$



Now there exists $\beta \in \mathbb{R}$ such that $-f \le z^* + \beta \le g$ on E if, and only if, for all $x, y \in E$, $f(x) + g(y) + \langle x - y, z^* \rangle \ge 0$. The same technique as above (using Theorem 1.11, with $C = E \times E$, $k(x,y) = f(x) + g(y)$ and $j(x,y) = x - y$) then leads rapidly to the result below (exercise!):

Example 7.2. *Let E be a nonzero normed space with dual E^*, and $f, g \in \mathcal{PC}(E)$. Then there exist $z^* \in E^*$ and $\beta \in \mathbb{R}$ such that* (7.1) *is satisfied if, and only if,*

$$\left.\begin{array}{l} \text{there exists } M \ge 0 \text{ such that,} \\ \qquad x,\ y \in E \quad \Longrightarrow \quad f(x) + g(y) + M\|x - y\| \ge 0. \end{array}\right\} \tag{7.2}$$

In this case, there exist z^* and β satisfying (7.1) *such that $\|z^*\| \le M$.*

Remark 7.3. We note that (7.1) can also be split up into the two statements "$f^*(-z^*) \le \beta$" and "$g^*(z^*) \le -\beta$", where the *Fenchel conjugate* f^* is defined by

$$f^*(x^*) := \sup\nolimits_E (x^* - f).$$

It follows that (7.2) is equivalent to:

$$\text{there exists } z^* \in E^* \text{ such that} \quad f^*(-z^*) + g^*(z^*) \le 0. \tag{7.3}$$

This is an old condition in convex analysis, due to Fenchel in the finite dimensional case. For this reason, we will say that $z^* \in E^*$ is a *Fenchel functional for f and g* if $f^*(-z^*) + g^*(z^*) \le 0$. Theorem 7.4 below, the *sharp Fenchel duality theorem*, contains a result for Fenchel functionals analogous to the result proved for Lagrange multipliers in Theorem 6.4.

The reader will note that we have defined Fenchel conjugate and Fenchel functional for proper convex functions on a normed space. This presents an impediment for some situations that will arise in our later discussions on multifunctions. Accordingly, we will redefine Fenchel conjugate and Fenchel functional *with respect to a bilinear form* in Section 8, where these issues will be discussed in greater detail. These later definitions will be entirely consistent with what we have introduced above.

Theorem 7.4. *Let E be a nonzero normed space with dual E^*, and f, $g \in \mathcal{PC}(E)$. Then:*
(a) *f and g have a Fenchel functional if, and only if,* (7.2) *is satisfied.*
(b) *If $z^* \in E^*$ is a Fenchel functional for f and g then*

$$\sup_{x,\, y\in E,\ x\neq y} \frac{-f(x)-g(y)}{\|x-y\|} \le \|z^*\| < \infty.$$

(c) *If $f+g \ge 0$ on E and* $\sup_{x,\, y\in E,\ x\neq y} \dfrac{-f(x)-g(y)}{\|x-y\|} < \infty$ *then*

$$\left.\begin{aligned} &\min\left\{\|z^*\|\colon\ z^* \text{ is a Fenchel functional for } f \text{ and } g\right\}\\ &\qquad\qquad = \sup_{x,\, y\in E,\ x\neq y} \frac{-f(x)-g(y)}{\|x-y\|} \vee 0. \end{aligned}\right\} \tag{7.4}$$

Proof. (a) is a restatement of Example 7.2.

(b) The analysis preceding Example 7.2 implies that

$$x, y \in E \quad\Longrightarrow\quad -f(x)-g(y) \le \langle x-y, z^*\rangle \le \|x-y\|\|z^*\|,$$

which gives the required result.

(c) The inequality "$\ge$" in (7.4) follows from (b) and the fact that $\|z^*\| \ge 0$. Now write M for the right hand side of (7.4). Then $M \ge 0$. Let $x, y \in E$. We have

$$x \neq y \quad\Longrightarrow\quad \frac{-f(x)-g(y)}{\|x-y\|} \le M \quad\Longrightarrow\quad f(x)+g(y)+M\|x-y\| \ge 0,$$

and

$$x = y \quad\Longrightarrow\quad f(x)+g(y)+M\|x-y\| = (f+g)(x) \ge 0.$$

The result now follows from Example 7.2. □

Example 7.5. The purpose of this example is to show that the "$\vee$" in (7.4) is necessary. Let $E = \mathbb{R}$ and $f = g = |\cdot|$. Then (exercise!) $f+g \ge 0$ on $\mathbb{R}$ and $\sup_{x,\, y\in\mathbb{R},\ x\neq y} \dfrac{-|x|-|y|}{|x-y|} = -1.$ However, there cannot exist $z^* \in E^*$ such that $\|z^*\| = -1$.

Remark 7.6. In this remark, we discuss a geometric interpretation of Theorem 7.4(c). We write $\mathcal{CA}(E)$ for the set of all continuous affine real functions on E. If $a \in \mathcal{CA}(E)$ then a can be written uniquely in the form $a = z^* + \beta$, where $z^* \in E^*$ and $\beta \in \mathbb{R}$. Since z^* is the derivative of a in any reasonable sense, we shall write $z^* = a'$. Turning now to Theorem 7.4(c), write $h = -f$, so that h is proper and concave and $h \le g$ on E. Let $a \in \mathcal{CA}(E)$. Remark 7.3 tells us that if $h \le a \le g$ on E then $f^*(-a') + g^*(a') \le 0$ and, conversely, if $f^*(-z^*) + g^*(z^*) \le 0$ then there exists $a \in \mathcal{CA}(E)$ such that $a' = z^*$ and $h \le a \le g$ on E. Furthermore,

$$\sup_{x,\, y \in E,\ x \neq y} \frac{-f(x) - g(y)}{\|x - y\|} = \sup_{x,\, y \in E,\ x \neq y} \frac{h(x) - g(y)}{\|x - y\|}$$

Now suppose, in addition, that $\sup_E h > \inf_E g$, to avoid the "$\vee$" in (7.4). Then the conclusion of Theorem 7.4(c) is that

$$\min \big\{ \|a'\| \colon\ a \in \mathcal{CA}(E),\ h \le a \le g \text{ on } E \big\} = \sup_{x,\, y \in E,\ x \neq y} \frac{h(x) - g(y)}{\|x - y\|}.$$

The quotient on the right hand side of the equality above is, of course, the slope of the line–segment going from the point $\big(y, g(y)\big)$ on the graph of g to the point $\big(x, h(x)\big)$ on the graph of h.

8 Fenchel duality with respect to a bilinear form — locally convex spaces

Let E and E^* be nonzero real vector spaces, and $\langle \cdot, \cdot \rangle \colon E \times E^* \to \mathbb{R}$ be a bilinear form that separates the points of E and also separates the points of E^*. $\big($This means that if $x \in E \setminus \{0\}$ then there exists $x^* \in E^*$ such that $\langle x, x^* \rangle \neq 0$ and that if $x^* \in E^* \setminus \{0\}$ then there exists $x \in E$ such that $\langle x, x^* \rangle \neq 0.\big)$ If $f \in \mathcal{PC}(E)$, the *Fenchel conjugate* f^* *with respect to* $\langle \cdot, \cdot \rangle$ is defined by

$$f^*(x^*) := \sup\nolimits_{x \in E} \big[\langle x, x^* \rangle - f(x) \big]. \tag{8.1}$$

We note that this definition implies the *Fenchel–Young inequality*

$$(x, x^*) \in E \times E^* \quad \Longrightarrow \quad \langle x, x^* \rangle \le f(x) + f^*(x^*). \tag{8.2}$$

If k: $E^* \to]-\infty, \infty]$ is convex in the sense of Definition 1.8, the function ${}^*\!k$: $E \to [-\infty, \infty]$ is defined by

$${}^*\!k(x) := \sup\nolimits_{x^* \in E^*} \big[\langle x, x^* \rangle - k(x^*) \big]. \tag{8.3}$$

If $f, g \in \mathcal{PC}(E)$, a *Fenchel functional for* f *and* g is an element z^* of E^* such that $f^*(-z^*) + g^*(z^*) \le 0$. The definitions of f^* and "Fenchel functional" are compatible with those introduced in Remark 7.3 for normed spaces, if we take $\langle \cdot, \cdot \rangle$ to be the canonical bilinear form on $E \times E^*$. If $E^* = E$, we will write $f^{@}$ instead of f^*. (See Definition 19.1.)

This is an appropriate point to make some comments about our formulation of the definition of Fenchel conjugate and Fenchel functional. It will become clear in Lemma 22.1 that when we consider the theory of monotone multifunctions on nonreflexive Banach spaces, we need a version of the Fenchel duality theorem that falls outside the scope of Theorem 7.4(a). This need can be met by proving such a version for locally convex spaces. However, an inspection of (8.1) shows that it is the bilinear form $\langle \cdot, \cdot \rangle$ that is important for the definition of f^*, and not the topology on E that gives E^* as dual. This is why we have opted to make the definition as above. In some cases, the bilinear form is determined by a given topology on E and, in other cases, the topology is determined by a given bilinear form.

We say that a locally convex topology $\mathcal{T}$ on E is *E^*–compatible* if the $\mathcal{T}$–dual of E is exactly $\big\{\langle\cdot, x^*\rangle\colon\ x^* \in E^*\big\}$. In this case, we write $\mathcal{S}(E,\mathcal{T})$ for the family of all $\mathcal{T}$–continuous seminorms on E. The facts that we shall need about locally convex spaces are that (i) *if $z^* \in E^*$ then $|z^*| \in \mathcal{S}(E,\mathcal{T})$*, (ii) *if $P \in \mathcal{S}(E,\mathcal{T})$ and L is a linear functional on E such that $L \le P$ on E then $L \in \big\{\langle\cdot, x^*\rangle\colon\ x^* \in E^*\big\}$* and (iii) *the sets $\{x \in E\colon\ Q(x) < 1\}$ $\big(Q \in \mathcal{S}(E,\mathcal{T})\big)$ form a $\mathcal{T}$–base for the neighborhoods of* 0.

The main results of this section are Theorem 8.1 and Theorem 8.4. In Theorem 8.1, we show how Theorem 1.11 leads to a necessary and sufficient condition for there to exist a Fenchel functional for f and g in this context, while in Theorem 8.4 we give a sufficient condition (in which neither function satisfies a semicontinuity condition) implying the results that are used in practice. Corollary 8.5 is a special case (which will be bootstrapped in Theorem 10.1) that leads to Corollary 8.6, a classical result due to Rockafellar. We will also show in Theorem 15.1 how to deduce the Attouch-Brezis version of the Fenchel duality theorem from Theorem 8.4.

We emphasize that Theorem 8.1 and Theorem 7.4 give *necessary and sufficient* conditions for the existence of Fenchel functionals, and not merely *sufficient* conditions.

Theorem 8.1. *Let f, $g \in \mathcal{PC}(E)$ and $\mathcal{T}$ be an E^*–compatible topology on E. Then there exists a Fenchel functional for f and g if, and only if,*

$$\left.\begin{array}{l}\text{there exists } P \in \mathcal{S}(E,\mathcal{T}) \text{ such that}\\ \qquad x, y \in E \quad\Longrightarrow\quad f(x) + g(y) + P(x-y) \ge 0.\end{array}\right\} \tag{8.4}$$

Proof. Suppose first that z^* is a Fenchel functional for f and g. Then, for all x, $y \in E$, $\langle x, -z^*\rangle - f(x) + \langle y, z^*\rangle - g(y) \le f^*(-z^*) + g^*(z^*) \le 0$. Consequently, $f(x) + g(y) + \langle x - y, z^*\rangle \ge 0$, and (8.4) follows with $P := |z^*|$. See the remarks preceding Example 7.2 for an indication of how to prove the converse (using the Hahn–Banach–Lagrange theorem, Theorem 1.11). □

Remark 8.2. In this remark we sketch an "intrinsic" version of Theorem 8.1, i.e., a version which depends only on the duality and does not involve an additional topology, $\mathcal{T}$. We have the following result: *Let f, $g \in \mathcal{PC}(E)$. Then there exists a Fenchel functional z^* for f and g if, and only if, there exists a nonempty convex $w(E^*,E)$–compact subset K of E^* such that*

$$x, y \in E \quad\Longrightarrow\quad f(x) + g(y) + \sup\langle x - y, K\rangle \ge 0.$$

"Only if" is obvious by taking $K := \{z^*\}$, and the converse follows by using the minimax theorem, Theorem 3.2, on the function defined on the set

$\big(\operatorname{dom} f \times \operatorname{dom} g\big) \times K$ by $\big((x, y), z^*\big) \mapsto f(x) + g(y) + \langle x - y, z^*\rangle$. Readers familiar with the theory of the Mackey topology will recognize the connection between this result and Theorem 8.1.

Notation 8.3. Let E be a nonzero vector space and $f, g \in \mathcal{PC}(E)$. If $w \in E$, we write $(f \ominus g)(w) := \inf_{z\in E} \big[f(z) + g(z - w)\big]$. Readers familiar with the definition of episum (= inf–convolution) will recognize that the definition of $f \ominus g$ is the same as the definition of the episum of f and g except for a change of sign in the argument of g. We note that

$$\big\{x \in E\colon\ (f \ominus g)(x) < \infty\big\} = \operatorname{dom} f - \operatorname{dom} g. \tag{8.5}$$

Before embarking on Theorem 8.4, we should explain in broad terms what it achieves. Theorem 8.1 gives a *necessary and sufficient* condition for there to exist a Fenchel functional in terms of certain expressions in f and g being bounded *below*. Theorem 8.4 transforms this into a (more useful) *sufficient* condition for there to exist a Fenchel functional in terms of certain expressions in f and g being bounded *above*. Theorem 8.4 is sometimes known as a *decoupling* result. We will use Theorem 8.4 explicitly in Corollary 8.5, Theorem 10.1 and Theorem 15.1.

Theorem 8.4. *Let* $f,\ g \in \mathcal{PC}(E)$, $f + g \ge 0$ *on* E *and*

$$F := \bigcup\nolimits_{\lambda>0} \lambda(\operatorname{dom} g - \operatorname{dom} f) \ni 0.$$

Suppose that $\mathcal{T}$ *is an* E^**–compatible topology on* E *and* $f \ominus g$ *is (finitely) bounded above in some* $\mathcal{T}$*–neighborhood of* 0 *relative to* F*. Then:*
(a) (8.4) *is satisfied.*
(b) *There exists a Fenchel functional for* f *and* g.

Proof. Choose $Q \in \mathcal{S}(E, \mathcal{T})$ and $M \ge 0$ such that

$$w \in F \text{ and } Q(w) < 1 \quad\Longrightarrow\quad (f \ominus g)(w) < M. \tag{8.6}$$

We shall prove that (8.4) is satisfied with $P := MQ \in \mathcal{S}(E, \mathcal{T})$. Now the inequality in (8.4) is immediate if $x \notin \operatorname{dom} f$ or $y \notin \operatorname{dom} g$, so we can and will assume that $x \in \operatorname{dom} f$ and $y \in \operatorname{dom} g$. Let $\lambda > Q(x - y) \ge 0$ and $w := (y - x)/\lambda \in F$. Since $Q(w) < 1$, (8.6) gives $z \in E$ such that

$$f(z) + g(z - w) < M. \tag{8.7}$$

Now $x + \lambda z = y + \lambda(z - w)$, hence, since $f + g \ge 0$ on E,

$$f\left(\frac{x + \lambda z}{1 + \lambda}\right) + g\left(\frac{y + \lambda(z - w)}{1 + \lambda}\right) \ge 0.$$

Thus, using the convexity of f and g,

$$f(x) + \lambda f(z) + g(y) + \lambda g(z - w) \ge 0.$$

Combining this with (8.7), we derive that

$$f(x) + g(y) + M\lambda \geq 0.$$

Letting $\lambda \to Q(x-y)$ gives $\quad f(x) + g(y) + MQ(x-y) \geq 0, \quad$ that is to say

$$f(x) + g(y) + P(x-y) \geq 0.$$

This completes the proof of (a), and (b) then follows from Theorem 8.1. □

Corollary 8.5. *Let $f,\, g \in \mathcal{PC}(E)$, $f + g \geq 0$ on E, $\mathcal{T}$ be an E^*–compatible topology on E, and g be (finitely) bounded above in some $\mathcal{T}$–neighborhood of a point of* $\operatorname{dom} f$. *Then there exists a Fenchel functional for f and g.*

Proof. In this case, $\bigcup_{\lambda>0} \lambda(\operatorname{dom} g - \operatorname{dom} f) = E \ni 0$. Choose $z \in \operatorname{dom} f$, $N \in \mathbb{R}$ and $Q \in \mathcal{S}(E, \mathcal{T})$ such that $\quad w \in E$ and $Q(w) < 1 \Longrightarrow\ g(z-w) < N$, and define

$$M := f(z) + N > f(z) + g(z) = (f+g)(z) \geq 0.$$

If $w \in E$ and $Q(w) < 1$ then $(f \ominus g)(w) \leq f(z) + g(z-w) < M$, and so (8.6) is satisfied. The result now follows from Theorem 8.4(b). □

Corollary 8.6 is an immediate consequence of Corollary 8.5 (see Rockafellar, [77, Theorem 1, pp. 82–83], Zălinescu, [119, Theorem 2.8.3(iii), p. 123] or Borwein–Zhu, [24, Sections 4.3.1–2, pp. 127–129]). We will use this result explicitly in Theorem 9.3, the transversality theorem, Theorem 19.16, Lemma 22.1 and Lemma 35.5. We refer the reader to Remark 15.3 for a comparison of Corollary 8.6 with the Attouch–Brezis version of the Fenchel duality theorem, Theorem 15.1.

Corollary 8.6. *Let $f,\, g \in \mathcal{PC}(E)$, $f + g \geq 0$ on E, $\mathcal{T}$ be an E^*–compatible topology on E, and g be finite and $\mathcal{T}$–continuous at a point of* $\operatorname{dom} f$. *Then there exists a Fenchel functional for f and g.*

We have presented Corollary 8.6 as a consequence of Corollary 8.5. However, they are in fact equivalent, as is evident from the following result, which will be used in Lemma 13.3 and Definition 38.1. The argument is an adaptation to the seminorm case of that of Phelps, [68, Proposition 1.6, p. 4] or Borwein–Zhu, [24, Section 4.1.2, pp. 112–113].

Theorem 8.7. *Let E be a nonzero vector space, $f \in \mathcal{PC}(E)$, $z_0 \in E$, $K \in \mathbb{R}$, and $P\colon E \to \mathbb{R}$ be a seminorm such that*

$$z \in E \text{ and } P(z - z_0) \leq 1 \quad \Longrightarrow \quad f(z) \leq K. \tag{8.8}$$

Then

$$\left.\begin{aligned} x, y \in E, P(x - z_0) \leq \tfrac{1}{2} \text{ and } P(y - z_0) \leq \tfrac{1}{2} \quad \Longrightarrow \qquad\qquad& \\ \big|f(x) - f(y)\big| \leq 4\big(K - f(z_0)\big)P(x-y).& \end{aligned}\right\} \tag{8.9}$$

Proof. Let $x, y \in E$, $P(x - z_0) \le \frac{1}{2}$ and $P(y - z_0) \le \frac{1}{2}$. (8.8) implies that $f(x), f(y) \in \mathbb{R}$, and we can and will suppose that $f(x) - f(y) \ge 0$. Let $\lambda > 2P(x - y) \ge 0$, and $z := x + (x - y)/\lambda$. Then

$$P(z - z_0) \le P(x - z_0) + P(x - y)/\lambda \le \tfrac{1}{2} + \tfrac{1}{2} = 1,$$

and thus (8.8) gives

$$f(z) \le K. \tag{8.10}$$

On the other hand, since $P\big((2z_0 - y) - z_0\big) = P(y - z_0) \le \frac{1}{2}$, (8.8) also gives $f(2z_0 - y) \le K$, from which

$$f(z_0) \le \tfrac{1}{2} f(y) + \tfrac{1}{2} f(2z_0 - y) \le \tfrac{1}{2} f(y) + \tfrac{1}{2} K.$$

Thus $f(y) \ge 2f(z_0) - K$. If we combine this with (8.10), we obtain

$$f(z) - f(y) \le K - \big(2f(z_0) - K\big) = 2\big(K - f(z_0)\big). \tag{8.11}$$

Now $x = (y + \lambda z)/(1 + \lambda)$, from which $f(x) \le \big(f(y) + \lambda f(z)\big)/(1 + \lambda)$ and so, combining with (8.11),

$$(1 + \lambda)\big(f(x) - f(y)\big) \le \lambda\big(f(z) - f(y)\big) \le 2\lambda\big(K - f(z_0)\big).$$

Thus, since $1 \le 1 + \lambda$ and $f(x) - f(y) \ge 0$,

$$f(x) - f(y) \le 2\lambda\big(K - f(z_0)\big).$$

(8.9) now follows by letting $\lambda \to 2P(x - y)$. □

We conclude this section with a result (the "bipolar theorem for locally convex spaces") which we will use in Theorem 45.12. The proof is an obvious adaptation of that of Theorem 4.4.

Theorem 8.8. *Let C be a nonempty convex subset of E, $\mathcal{T}$ be an E^*–compatible topology on E, $C^{\mathcal{T}}$ be the closure of C with respect to $\mathcal{T}$ and $x \in E$. Then $x \in C^{\mathcal{T}}$ if, and only if,*

$$x^* \in E^* \quad \Longrightarrow \quad \langle x, x^* \rangle \le \sup\langle C, x^* \rangle. \tag{8.12}$$

Proof. "Only if" is immediate since all the functions $\big\{\langle \cdot, x^* \rangle\colon\ x^* \in E^*\big\}$ are $\mathcal{T}$–continuous on E. Suppose, conversely, that $x \in E \setminus C^{\mathcal{T}}$. Then there exists $P \in \mathcal{S}(E, \mathcal{T})$ such that $\inf_{c \in C} P(x - c) > 0$. From the Mazur–Orlicz theorem, Lemma 1.6, with $D := x - C$, there exists a linear function L on E such that $L \le P$ on E and $\inf_D L = \inf_D P > 0$. Thus there exists $x^* \in E^*$ such that $\inf_{c \in C} \langle x - c, x^* \rangle > 0$, from which $\langle x, x^* \rangle > \sup\langle C, x^* \rangle$. So (8.12) fails, completing the proof of the theorem. □

9 Some properties of $\frac{1}{2}\|\cdot\|^2$

The main result of this section is Theorem 9.3, a sharp version of the Fenchel duality theorem for normed spaces that has proved to be very useful in the investigation of monotone multifunctions. This result was first established in Simons–Zălinescu, [109, Theorem 2.1, pp. 5–6] with a proof that is more direct but also much more computational. We start off with two preliminary lemmas.

Lemma 9.1. *Let E be a nonzero normed space with dual E^*, $x, y \in E$ and $-\infty < c \le \frac{1}{2}\|x\|^2$. Then*

$$c - \tfrac{1}{2}\|y\|^2 \le \|x-y\|\left[\|x\| - \sqrt{\|x\|^2 - 2c}\right].$$

Proof. The triangle inequality gives $\big|\|x\| - \|x-y\|\big| \le \|y\|$. Squaring and dividing by 2, $\frac{1}{2}\|x\|^2 + \frac{1}{2}\|x-y\|^2 - \|x-y\|\|x\| \le \frac{1}{2}\|y\|^2$, from which

$$\tfrac{1}{2}\|x\|^2 + \tfrac{1}{2}\|x-y\|^2 - \tfrac{1}{2}\|y\|^2 \le \|x-y\|\|x\|.$$

Thus

$$\begin{aligned} c - \tfrac{1}{2}\|y\|^2 &\le c - \tfrac{1}{2}\|y\|^2 + \tfrac{1}{2}\left[\sqrt{\|x\|^2 - 2c} - \|x-y\|\right]^2 \\ &= \tfrac{1}{2}\|x\|^2 + \tfrac{1}{2}\|x-y\|^2 - \tfrac{1}{2}\|y\|^2 - \|x-y\|\sqrt{\|x\|^2 - 2c} \\ &\le \|x-y\|\|x\| - \|x-y\|\sqrt{\|x\|^2 - 2c}. \end{aligned}$$

□

Lemma 9.2. *Let E be a nonzero normed space with dual E^*, $x \in E$ and $-\infty \le c \le \frac{1}{2}\|x\|^2$. Then*

$$\sup_{y \in E\setminus\{x\}} \frac{c - \frac{1}{2}\|y\|^2}{\|x-y\|} \vee 0 = \left[\|x\| - \sqrt{\|x\|^2 - 2c}\right] \vee 0. \tag{9.1}$$

Proof. Since both sides of the above equation are zero when $c = -\infty$, we can and will suppose that $c \in \mathbb{R}$. Furthermore, the inequality "$\le$" in (9.1) follows easily from Lemma 9.1. Define $g\colon E \to \mathbb{R}$ by $g(z) := \frac{1}{2}\|z\|^2$ and let

$$M := \sup_{y \in E\setminus\{x\}} \frac{c - \frac{1}{2}\|y\|^2}{\|x-y\|} \vee 0.$$

Then $0 \le M < \infty$ and $y \ne x \implies M\|x-y\| + g(y) \ge c$. Since this inequality holds trivially if $y = x$, the Hahn–Banach–Lagrange theorem, Theorem 1.11, gives $z^* \in E^*$ such that $\|z^*\| \le M$ and

$$y \in E \implies \langle x - y, z^*\rangle + g(y) \ge c \iff 2\langle x, z^*\rangle \ge 2\langle y, z^*\rangle - 2g(y) + 2c.$$

Taking the supremum of the latter inequality over $y \in E$ and using the (well known) fact that $g^*(z^*) = \frac{1}{2}\|z^*\|^2$ (exercise!), we have

$$2\langle x, z^*\rangle \geq 2g^*(z^*) + 2c = \|z^*\|^2 + 2c.$$

Thus

$$\|x\|^2 - 2\|x\|\|z^*\| + \|z^*\|^2 \leq \|x\|^2 - 2\langle x, z^*\rangle + \|z^*\|^2 \leq \|x\|^2 - 2c.$$

Taking the square root gives $\|x\| - \|z^*\| \leq \sqrt{\|x\|^2 - 2c}$, and so

$$M \geq \|z^*\| \geq \|x\| - \sqrt{\|x\|^2 - 2c}.$$

The inequality "$\geq$" in (9.1) now follows immediately. $\square$

Our next result will be used explicitly in Theorem 21.4.

Theorem 9.3. *Let E be a nonzero normed space with dual E^*, $f \in \mathcal{PC}(E)$ and*

$$x \in E \quad\Longrightarrow\quad f(x) + \tfrac{1}{2}\|x\|^2 \geq 0.$$

Then:
(a) *There exists $x^* \in E^*$ such that*

$$f^*(x^*) + \tfrac{1}{2}\|x^*\|^2 \leq 0.$$

(b)

$$\min\left\{\|x^*\|\colon\ x^* \in E^*,\ f^*(x^*) + \tfrac{1}{2}\|x^*\|^2 \leq 0\right\} = \sup_{x,\, y \in E,\ x \neq y} \frac{-f(x) - \frac{1}{2}\|y\|^2}{\|x - y\|} \vee 0.$$

(c) *There exists $x^* \in E^*$ such that* $\quad f^*(x^*) + \tfrac{1}{2}\|x^*\|^2 \leq 0 \quad$ *and*

$$\|x^*\| = \sup_{x \in E}\left[\|x\| - \sqrt{2f(x) + \|x\|^2}\right] \vee 0.$$

(d) *Let* $\quad x^* \in E^* \quad$ *and* $\quad f^*(x^*) + \tfrac{1}{2}\|x^*\|^2 \leq 0. \quad$ *Then*

$$\sup_{x \in E}\left[\|x\| - \sqrt{2f(x) + \|x\|^2}\right] \vee 0 \leq \|x^*\| \leq \inf_{x \in E}\left[\|x\| + \sqrt{2f(x) + \|x\|^2}\right].$$

Proof. (a) Define $g\colon E \to \mathbb{R}$ by $\quad g(z) := \tfrac{1}{2}\|z\|^2$. Since g is continuous and convex, Rockafellar's version of the Fenchel duality theorem, Corollary 8.6, implies that there exists a Fenchel functional, z^*, for f and g, and (a) follows with $x^* := -z^*$.

(b) This follows from Theorem 7.4(b,c).

(c) It is clear that

$$\sup_{x,\, y \in E,\ x \neq y} \frac{-f(x) - \frac{1}{2}\|y\|^2}{\|x - y\|} = \sup_{x \in E}\ \sup_{y \in E \setminus \{x\}} \frac{-f(x) - \frac{1}{2}\|y\|^2}{\|x - y\|}.$$

For a given $x \in E$, we use Lemma 9.2 with $c = -f(x)$, then take the supremum over $x \in E$ and appeal to (b).

(d) The Fenchel–Young inequality, (8.2), implies that, for all $x \in E$,

$$\big(\|x^*\| - \|x\|\big)^2 = \|x^*\|^2 - 2\|x\|\|x^*\| + \|x\|^2 \le \|x^*\|^2 + 2\langle x, x^*\rangle + \|x\|^2$$
$$\le 2f(x) + 2f^*(x^*) + \|x^*\|^2 + \|x\|^2 \le 2f(x) + \|x\|^2,$$

thus $\|x\| - \sqrt{2f(x) + \|x\|^2} \le \|x^*\| \le \|x\| + \sqrt{2f(x) + \|x\|^2}$, and (d) is now immediate from the fact that $\|x^*\| \ge 0$. □

10 The conjugate of a sum in the locally convex case

As in Section 8, E and E^* are nonzero real vector spaces, and $\langle\cdot,\cdot\rangle\colon E \times E^* \to \mathbb{R}$ is a bilinear form that separates the points of E and also separates the points of E^*. We now bootstrap Theorem 8.4(b) to obtain Theorem 10.1. The conclusion of Theorem 10.1 and its two corollaries is that $(f+g)^*$ is the *exact episum* or *exact inf–convolution* of f^* and g^*. Corollary 10.4 will be applied later to the existence of autoconjugates in SSDB spaces.

Theorem 10.1. *Let $f, g \in \mathcal{PC}(E)$ and $F := \bigcup_{\lambda>0} \lambda\big[\operatorname{dom} g - \operatorname{dom} f\big] \ni 0$. Let $\mathcal{T}$ be a E^*–compatible topology on E, $x^* \in E^*$ and $(f - x^*) \ominus g$ be (finitely) bounded above in some $\mathcal{T}$–neighborhood of* 0 *relative to F. Then*

$$(f+g)^*(x^*) = \min\nolimits_{z^* \in E^*} \big[f^*(x^* - z^*) + g^*(z^*)\big]. \tag{10.1}$$

Proof. If $z^* \in E^*$ and $x \in E$ then

$$\langle x, x^*\rangle - (f+g)(x) = \langle x, x^* - z^*\rangle - f(x) + \langle x, z^*\rangle - g(x) \le f^*(x^* - z^*) + g^*(z^*).$$

Taking the supremum over $x \in E$, we have $(f+g)^*(x^*) \le f^*(x^* - z^*) + g^*(z^*)$. Consequently,

$$(f+g)^*(x^*) \le \inf\nolimits_{z^* \in E^*} \big[f^*(x^* - z^*) + g^*(z^*)\big].$$

In order to prove the opposite inequality, we can and will suppose that $(f+g)^*(x^*) \in \mathbb{R}$. The Fenchel–Young inequality, (8.2), implies that

$$\big(f - x^* + (f+g)^*(x^*)\big) + g = (f+g) + (f+g)^*(x^*) - x^* \ge 0 \text{ on } E.$$

Since $\operatorname{dom}\big(f - x^* + (f+g)^*(x^*)\big) = \operatorname{dom} f$, Theorem 8.4(b) implies that there exists a Fenchel functional, z^*, for $f - x^* + (f+g)^*(x^*)$ and g. However,

$$\big(f - x^* + (f+g)^*(x^*)\big)^*(-z^*) = f^*(x^* - z^*) - (f+g)^*(x^*),$$

and so $f^*(x^* - z^*) - (f+g)^*(x^*) + g^*(z^*) \le 0$, that is to say

$$f^*(x^* - z^*) + g^*(z^*) \le (f+g)^*(x^*).$$

This completes the proof of (10.1). □

Corollary 10.2. *Let $f, g \in \mathcal{PC}(E)$, $\mathcal{T}$ be a E^*–compatible topology on E, and g be (finitely) bounded above in some $\mathcal{T}$–neighborhood of a point of* $\operatorname{dom} f$ *and $x^* \in E^*$. Then*

$$(f+g)^*(x^*) = \min_{z^* \in E^*} \big[f^*(x^* - z^*) + g^*(z^*)\big]. \tag{10.1}$$

Proof. In this case, $F = E \ni 0$. Choose $u \in \operatorname{dom} f$, $N \in \mathbb{R}$ and $Q \in \mathcal{S}(E, \mathcal{T})$ such that $\quad w \in E$ and $Q(w) < 1 \Longrightarrow \ g(u - w) < N, \quad$ and define

$$M := f(u) - \langle u, x^* \rangle + N.$$

If $w \in E$ and $Q(w) < 1$ then $\big((f - x^*) \ominus g\big)(w) \le f(u) - \langle u, x^* \rangle + g(u - w) < M$, and so the result now follows from Theorem 10.1. □

Corollary 10.3 is an immediate consequence of Corollary 10.2 $\big($see Rockafellar, [77, Theorem 3(a), p. 85] or Zălinescu, [119, Theorem 2.8.7(iii), p. 127]$\big)$:

Corollary 10.3. *Let $f, g \in \mathcal{PC}(E)$, $\mathcal{T}$ be a E^*–compatible topology on E, and g be finite and $\mathcal{T}$–continuous at a point of* $\operatorname{dom} f$*. Then, for all $x^* \in E^*$,*

$$(f+g)^*(x^*) = \min_{z^* \in E^*} \big[f^*(x^* - z^*) + g^*(z^*)\big]. \tag{10.1}$$

The following "symmetric" result will be used in our discussion of the existence of autoconjugates in SSDB spaces in Theorem 21.10. It is based on some of the results established by Bauschke–Wang in [13] for "kernel averages" in spaces of the form $E \times E^*$(where E is a reflexive Banach space).

Corollary 10.4. *Let $f_1, f_2, g \in \mathcal{PC}(E)$, $\mathcal{T}$ be a E^*–compatible topology on E, and g be finite and $\mathcal{T}$–continuous at a point of* $\operatorname{dom} f_1 - \operatorname{dom} f_2$*. Suppose that, for all $x \in E$,*

$$h(x) := \inf_{z \in E} \big[\tfrac{1}{2} f_1(x+z) + \tfrac{1}{2} f_2(x-z) + \tfrac{1}{4} g(2z)\big] > -\infty.$$

It is easily seen that $h \in \mathcal{PC}(E)$. Then, for all $x^* \in E^*$,

$$h^*(x^*) = \min_{z^* \in E^*} \big[\tfrac{1}{2} {f_1}^*(x^* + z^*) + \tfrac{1}{2} {f_2}^*(x^* - z^*) + \tfrac{1}{4} g^*(-2z^*)\big].$$

Proof. Define $f_3, g_1 \colon E \times E \to]-\infty, \infty]$ by

$$f_3(x_1, x_2) := \tfrac{1}{2} f_1(x_1) + \tfrac{1}{2} f_2(x_2) \quad \text{and} \quad g_1(x_1, x_2) := \tfrac{1}{4} g(x_1 - x_2).$$

Then

$$h^*(x^*) = \sup_{x, z \in E} \big[\langle x, x^* \rangle - f_3(x+z, x-z) - \tfrac{1}{4} g(2z)\big].$$

Now make the substitution $(x_1, x_2) = (x+z, x-z)$, so that $x = \frac{1}{2}(x_1 + x_2)$ and $2z = x_1 - x_2$. Thus

$$\begin{aligned}
h^*(x^*) &= \sup_{(x_1,x_2)\in E\times E} \left[\langle x_1+x_2, \tfrac{1}{2}x^*\rangle - f_3(x_1,x_2) - \tfrac{1}{4}g(x_1-x_2)\right] \\
&= \sup_{(x_1,x_2)\in E\times E} \left[\big\langle (x_1,x_2), (\tfrac{1}{2}x^*, \tfrac{1}{2}x^*)\big\rangle - f_3(x_1,x_2) - g_1(x_1,x_2)\right] \\
&= (f_3+g_1)^*(\tfrac{1}{2}x^*, \tfrac{1}{2}x^*), \qquad (10.2)
\end{aligned}$$

where the conjugate is computed using the bilinear form defined on $(E\times E)\times(E^*\times E^*)$ by $\big((x_1,x_2),(x_1^*,x_2^*)\big) \mapsto \langle x_1,x_1^*\rangle + \langle x_2,x_2^*\rangle$. We note then that the topology $\mathcal{T}\times\mathcal{T}$ on $E\times E$ is $E^*\times E^*$–compatible. We now compute ${f_3}^*$ and ${g_1}^*$. For all $(y_1^*,y_2^*)\in E^*\times E^*$, we have

$$\begin{aligned}
{f_3}^*(y_1^*,y_2^*) &= \sup_{(x_1,x_2)\in E\times E} \left[\langle x_1,y_1^*\rangle + \langle x_2,y_2^*\rangle - \tfrac{1}{2}f_1(x_1) - \tfrac{1}{2}f_2(x_2)\right] \\
&= \tfrac{1}{2}\sup_{x_1,x_2\in E} \left[\langle x_1,2y_1^*\rangle + \langle x_2,2y_2^*\rangle - f_1(x_1) - f_2(x_2)\right] \\
&= \tfrac{1}{2}{f_1}^*(2y_1^*) + \tfrac{1}{2}{f_2}^*(2y_2^*)\big] \qquad (10.3)
\end{aligned}$$

and, for all $(z_1^*,z_2^*)\in E^*\times E^*$, we have

$$\begin{aligned}
{g_1}^*(z_1^*,z_2^*) &= \sup_{(x_1,x_2)\in E\times E} \left[\langle x_1,z_1^*\rangle + \langle x_2,z_2^*\rangle - \tfrac{1}{4}g(x_1-x_2)\right] \\
&= \sup_{x_1,x_3\in E} \left[\langle x_1,z_1^*\rangle + \langle x_1-x_3,z_2^*\rangle - \tfrac{1}{4}g(x_3)\right] \\
&= \sup_{x_1,x_3\in E} \left[\langle x_1,z_1^*+z_2^*\rangle + \langle -x_3,z_2^*\rangle - \tfrac{1}{4}g(x_3)\right] \\
&= \sup_{x_1\in E} \Big[\langle x_1,z_1^*+z_2^*\rangle + \sup\nolimits_{x_3\in E}\left[\langle x_3,-z_2^*\rangle - \tfrac{1}{4}g(x_3)\right]\Big] \\
&= \sup_{x_1\in E} \langle x_1,z_1^*+z_2^*\rangle + \tfrac{1}{4}g^*(-4z_2^*) \\
&= \begin{cases} \tfrac{1}{4}g^*(-4z_2^*), & \text{if } z_1^*+z_2^*=0; \\ \infty, & \text{otherwise.} \end{cases} \qquad (10.4)
\end{aligned}$$

Since $\operatorname{dom} f_3 = \operatorname{dom} f_1 \times \operatorname{dom} f_2$, g_1 is $\mathcal{T}\times\mathcal{T}$–continuous at a point of $\operatorname{dom} f_3$, and so we derive from Corollary 10.3 and (10.4) that

$$(f_3+g_1)^*(\tfrac{1}{2}x^*, \tfrac{1}{2}x^*) = \min_{z_1^*+z_2^*=0} \left[{f_3}^*\big(\tfrac{1}{2}x^*-z_1^*, \tfrac{1}{2}x^*-z_2^*\big) + \tfrac{1}{4}g^*(-4z_2^*)\right].$$

If we now put $z_1^* = -\tfrac{1}{2}z^*$ and $z_2^* = \tfrac{1}{2}z^*$, we obtain from (10.2) and (10.3) that

$$\begin{aligned}
h^*(x^*) &= \min_{z^*\in E^*} \left[{f_3}^*\big(\tfrac{1}{2}x^*+\tfrac{1}{2}z^*, \tfrac{1}{2}x^*-\tfrac{1}{2}z^*\big) + \tfrac{1}{4}g^*(-2z^*)\right] \\
&= \min_{z^*\in E^*} \left[\tfrac{1}{2}{f_1}^*(x^*+z^*) + \tfrac{1}{2}{f_2}^*(x^*-z^*) + \tfrac{1}{4}g^*(-2z^*)\right]. \qquad \square
\end{aligned}$$

11 Fenchel duality vs the conjugate of a sum

The results of Section 10 indicate the (well known) fact that results on the conjugate of a sum are very close to the Fenchel duality theorem. The purpose of this section is to draw a distinction between these two kinds of result. Examples 11.1 and 11.2 were worked out in collaboration with Regina Burachik, and Example 11.3 is based on a suggestion of Jonathan Borwein. There is an example similar to Example 11.1 in Boţ–Wanka, [27, pp. 2798–2799]. Let E be a nonzero Banach space and $f, g \in \mathcal{PC}(E)$. We say that f and g *satisfy Fenchel duality* if there exists $z^* \in E^*$ such that

$$f^*(-z^*) + g^*(z^*) = (f+g)^*(0).$$

Example 11.1. We give an example of proper, convex lower semicontinuous functions f and g on $\mathbb{R}^2$ that satisfy Fenchel duality but, for most $r \in \left(\mathbb{R}^2\right)^* = \mathbb{R}^2$, it is not true that there exist $p, q \in \mathbb{R}^2$ such that $p + q = r$ and $f^*(p) + g^*(q) = (f+g)^*(r)$.

Let $C = \{x \in \mathbb{R}^2 \colon \|x\| \le 1\}$ and $x_0 = (1,0) \in \mathbb{R}^2$. Write $A := x_0 - C$, $B := C - x_0$, $f := \mathbb{I}_A$ and $g := \mathbb{I}_B$, where $\mathbb{I}_X$ is the *indicator function of* X, that is to say

$$\mathbb{I}_X(x) := \begin{cases} 0, & \text{if } x \in X; \\ \infty, & \text{otherwise.} \end{cases}$$

We note then that $f + g = \mathbb{I}_{\{0\}}$. Since $f^*(0) = g^*(0) = (f+g)^*(0) = 0$, f and g satisfy Fenchel duality.

Now, for all $p, q \in \mathbb{R}^2$, $f^*(p) = \|p\| + p_1$ and $g^*(q) = \|q\| - q_1$. Consequently

$$f^*(p) \ge 0 \quad \text{and} \quad \big(f^*(p) = 0 \iff p_1 \le 0 \text{ and } p_2 = 0\big), \tag{11.1}$$

and

$$g^*(q) \ge 0 \quad \text{and} \quad \big(g^*(q) = 0 \iff q_1 \ge 0 \text{ and } q_2 = 0\big). \tag{11.2}$$

If $p, q \in \mathbb{R}^2$ are such that $p + q = r$ and $f^*(p) + g^*(q) = (f+g)^*(r)$ then, since $(f+g)^*(r) = 0$, (11.1) and (11.2) imply that $f^*(p) = 0$ and $g^*(q) = 0$, consequently $p_2 = 0$ and $q_2 = 0$, from which $r_2 = 0$. Thus if $r_2 \ne 0$ then there do not exist $p, q \in \mathbb{R}^2$ such that $p + q = r$ and $f^*(p) + g^*(q) = (f+g)^*(r)$.

We can look at this another way: if $r \in \mathbb{R}^2$, and f and $g - r$ satisfy Fenchel duality then there exists $p \in \mathbb{R}^2$ such that

$$(f + g - r)^*(0) = f^*(p) + (g - r)^*(-p),$$

that is to say $f^*(p) + g^*(r - p) = 0$, and the analysis above shows that $r_2 = 0$. This argument can easily be reversed: if $r \in \mathbb{R}^2$ and $r_2 = 0$ then there exist $p, q \in \mathbb{R}^2$ such that $p + q = r$ and $f^*(p) + g^*(q) = (f+g)^*(r)$, and f and $g - r$ satisfy Fenchel duality. At any rate, f and g fail "stable Fenchel–Rockafellar duality" in the sense of [34, Theorem 3.2(i)].

Example 11.2. [34, Theorem 3.2(ii)] tells us that $\operatorname{epi} f^* + \operatorname{epi} g^*$ is not closed in $\mathbb{R}^2 \times \mathbb{R}$ in Example 11.1. We now confirm this by giving an explicit description of this set. If $p_1 < 0 < q_1$ then

$$f^*(p) = \frac{\|p\|^2 - {p_1}^2}{\|p\| - p_1} \le \frac{{p_2}^2}{2|p_1|} \quad \text{and} \quad g^*(q) = \frac{\|q\|^2 - {q_1}^2}{\|q\| + q_1} \le \frac{{q_2}^2}{2|q_1|}. \tag{11.3}$$

Let r be an arbitrary element of $\mathbb{R}^2$ and $n > |r_1|$. Then, from (11.3),

$$f^*\left(\frac{r}{2} - ne_1\right) \le \frac{{r_2}^2}{4(2n - r_1)} \quad \text{and} \quad g^*\left(\frac{r}{2} + ne_1\right) \le \frac{{r_2}^2}{4(2n + r_1)}.$$

Thus

$$\left(\frac{r}{2} - ne_1, \frac{{r_2}^2}{4(2n - r_1)}\right) \in \operatorname{epi} f^* \quad \text{and} \quad \left(\frac{r}{2} + ne_1, \frac{{r_2}^2}{4(2n + r_1)}\right) \in \operatorname{epi} g^*,$$

and so

$$\left(r, \frac{{r_2}^2}{4(2n - r_1)} + \frac{{r_2}^2}{4(2n + r_1)}\right) \in \operatorname{epi} f^* + \operatorname{epi} g^*.$$

Since $\operatorname{epi} f^* + \operatorname{epi} g^*$ recedes vertically, it follows by letting $n \to \infty$ that

$$\big\{(r_1, r_2, \lambda)\colon\ r_2 = 0, \lambda \ge 0\big\} \cup \big\{(r_1, r_2, \lambda)\colon\ r_2 \ne 0, \lambda > 0\big\} \subset \operatorname{epi} f^* + \operatorname{epi} g^*.$$

It is also clear from (11.1) and (11.2) that $\operatorname{epi} f^* + \operatorname{epi} g^* \subset \mathbb{R}^2 \times \mathbb{R}^+$. Suppose now that $(r, 0) \in \operatorname{epi} f^* + \operatorname{epi} g^*$. Then there exist $(p, \lambda) \in \operatorname{epi} f^*$ and $(q, \mu) \in \operatorname{epi} g^*$ such that $(p+q, \lambda+\mu) = (r, 0)$. Then $0 = \lambda+\mu \ge f^*(p)+g^*(q)$ so, from (11.1) and (11.2), $f^*(p) = 0$ and $g^*(q) = 0$. Arguing as in Example 11.1, $r_2 = 0$.

Combining all this together, we have

$$\operatorname{epi} f^* + \operatorname{epi} g^* = \big\{(r_1, r_2, \lambda)\colon\ r_2 = 0, \lambda \ge 0\big\} \cup \big\{(r_1, r_2, \lambda)\colon\ r_2 \ne 0, \lambda > 0\big\}$$

(which is obviously not closed).

We now investigate an even more unstable case of Fenchel duality. However, the analysis is a little more technical. Let E be a nonzero Banach space and $f, g \in \mathcal{PC}(E)$. We shall say that the pair f, g is *totally Fenchel unstable* if f and g satisfy Fenchel duality but

$$y^*, z^* \in E^* \text{ and } f^*(y^*) + g^*(z^*) = (f+g)^*(y^* + z^*) \quad \Longrightarrow \quad y^* + z^* = 0.$$

Example 11.3. We recall that if C is a convex subset of a Banach space E and $x \in C$ then x is a *support point* of C if there exists $x^* \in E^* \setminus \{0\}$ such that $\langle x, x^* \rangle = \sup\langle C, x^* \rangle$. We will give an example below of a nonempty $w(E, E^*)$–compact convex subset C of a Banach space E (actually ℓ^2) such that there exists an extreme point x_0 of C which is not a support point of C. Again, write $A := x_0 - C$, $B := C - x_0$, $f := \mathbb{I}_A$ and $g := \mathbb{I}_B$. The fact that x_0 is an extreme point of C implies that $f + g = \mathbb{I}_{\{0\}}$. As in Example 11.1, f and g satisfy Fenchel duality.

Now, for all $y^*, z^* \in E^*$,

$$f^*(y^*) = \langle x_0, y^* \rangle - \inf\langle C, y^* \rangle \ge 0 \quad \text{and} \quad g^*(z^*) = \sup\langle C, z^* \rangle - \langle x_0, z^* \rangle \ge 0.$$

Let $y^*, z^* \in E^*$ be such that

$$y^* + z^* = x^* \text{ and } f^*(y^*) + g^*(z^*) = (f + g)^*(x^*).$$

Thus $f^*(y^*) + g^*(z^*) = 0$, from which $f^*(y^*) = 0$ and $g^*(z^*) = 0$. Consequently, $\langle x_0, y^* \rangle = \inf\langle C, y^* \rangle$ and $\langle x_0, z^* \rangle = \sup\langle C, z^* \rangle$. Since x_0 is not a support point of C, $y^* = 0$ and $z^* = 0$, thus $x^* = y^* + z^* = 0$. So we have established that f and g are totally Fenchel unstable.

By analogy with the result established in Example 11.2, one is tempted to ask whether

$$\operatorname{epi} f^* + \operatorname{epi} g^* = \big\{(0,0)\big\} \cup \big(E^* \times]0, \infty[\big). \tag{11.4}$$

The inclusion "$\subset$" is clear from the discussion above, and it is also clear that $(0,0) = (0,0) + (0,0) \in \operatorname{epi} f^* + \operatorname{epi} g^*$. Thus (11.4) is equivalent to:

$$\operatorname{epi} f^* + \operatorname{epi} g^* \supset E^* \times]0, \infty[\,. \tag{11.5}$$

We now prove that this is the case, using an adaptation of a very nice argument provided by Radu Ioan Boţ (personal communication). Let $y^* \in E^*$. Let $h\colon E^* \to \mathbb{R}$ and $k\colon E^* \to \mathbb{R}$ be defined by $h := f^*$ and $k(z^*) := g^*(y^* - z^*)$. Since h and k are continuous and convex on E^*, it follows from Rockafellar's formula for the conjugate of a sum, Corollary 10.3, that

$$-\inf\nolimits_{E^*}\big[h + k\big] = (h + k)^*(0) = \min\nolimits_{z^{**} \in E^{**}}\big[h^*(z^{**}) + k^*(-z^{**})\big].$$

Since $\widehat{A}$ and $\widehat{B}$ are $w(E^{**}, E^*)$–compact and $w(E^{**}, E^*)$ is an E^*–compatible topology on E^{**}, it follows from Theorem 8.8 that, for all $z^{**} \in E^{**}$, $h^*(z^{**}) = \mathbb{I}_{\widehat{A}}(z^{**})$ and $k^*(-z^{**}) = \mathbb{I}_{\widehat{B}}(z^{**}) - \langle y^*, z^{**} \rangle$. Consequently, if $h^*(z^{**}) + k^*(-z^{**}) < \infty$ then $z^{**} \in \widehat{A} \cap \widehat{B}$, from which $z^{**} = 0$. Thus

$$-\inf\nolimits_{E^*}\big[h + k\big] = h^*(0) + k^*(-0) = 0,$$

and so, for all $\varepsilon > 0$, there exists $z^* \in E^*$ such that $h(z^*) + k(z^*) \le \varepsilon$, that is to say $f^*(z^*) + g^*(y^* - z^*) \le \varepsilon$. It is clear from this that $(y^*, \varepsilon) \in \operatorname{epi} f^* + \operatorname{epi} g^*$, which gives (11.5), as required.

Here is the promised example, which was suggested by Jonathan Borwein. Let $E = \ell^2$, $1 < p < 2$, and $C := \{x \in \ell^2\colon \|x\|_p \le 1\}$. Since the function $\|\cdot\|_p$ is lower semicontinuous on ℓ^2, C is closed, and obviously C is convex (and $C = -C$). Then x is an extreme point of C if, and only if, $\|x\|_p = 1$.

Let $x \in C$ and $\|x\|_p = 1$. We shall prove that *x is a support point of C if, and only if, $x \in \ell^{2(p-1)}$*. Suppose first that x is a support point of C. Then there exists $y \in \ell^2 = (\ell^2)^*$ such that $y \ne 0$ and (assuming that $\frac{1}{p} + \frac{1}{q} = 1$)

$$\langle x, y\rangle = \sup\langle C, y\rangle = \|y\|_q = \|x\|_p\|y\|_q.$$

Thus we have equality in Hölder's inequality, and so there exists $\lambda > 0$ such that, for all $n \ge 1$, $|y_n|^q = \big(\lambda|x_n|\big)^p$. Since $y \in \ell^2$, $\sum_{n\ge1}\big(\lambda|x_n|\big)^{2p/q} < \infty$, that is to say, $x \in \ell^{2(p-1)}$, as required. Suppose, conversely, that $x \in \ell^{2(p-1)}$. For all $n \ge 1$, let $y_n = \operatorname{sgn} x_n|x_n|^{p-1}$. Then $y \in \ell^2 = (\ell^2)^*$. Further,

$$\langle x, y\rangle = \sum_{n\ge1} x_n y_n = \sum_{n\ge1} x_n \operatorname{sgn} x_n|x_n|^{p-1} = \sum_{n\ge1}|x_n|^p = 1$$

and

$$\sup\langle C, y\rangle = \|y\|_q = \Big(\sum_{n\ge1}|x_n|^{q(p-1)}\Big)^{1/q} = \Big(\sum_{n\ge1}|x_n|^p\Big)^{1/q} = 1^{1/q} = 1,$$

so x is a support point of C. Since $2(p-1) < p$, there are plenty of extreme points of C that are not support points.

Remark 11.4. What we have actually shown above is that if C is a $w(E,E^*)$–compact convex subset of a Banach space E, x_0 is an extreme point of C, $f := \mathbb{I}_{x_0-C}$, $g := \mathbb{I}_{C-x_0}$, $y^* \in E^*$ and $\varepsilon > 0$ then there exists $z^* \in E^*$ such that $f^*(z^*) + g^*(y^* - z^*) \le \varepsilon$. This last inequality is equivalent to the statement that there exists $z^* \in E^*$ such that, for all $x, y \in E$, $f(x) + g(y) + \langle y - x, z^*\rangle \ge \langle y, y^*\rangle - \varepsilon$. From the Hahn–Banach–Lagrange theorem, Theorem 1.11, this is in turn equivalent to the statement that there exists $M \ge 0$ such that, for all $x, y \in E$, $f(x) + g(y) + M\|y - x\| \ge \langle y, y^*\rangle - \varepsilon$, that is to say there exists $M \ge 0$ such that, for all $u, v \in C$, $M\|u + v - 2x_0\| \ge \langle v - x_0, y^*\rangle - \varepsilon$. This observation leads to the following problem (which only makes sense if E is not reflexive):

Problem 11.5. Let C be a bounded closed convex subset of a Banach space E, x_0 be an extreme point of C, $y^* \in E^*$ and $\varepsilon > 0$. Then does there always exist $M \ge 0$ such that, for all $u, v \in C$, $M\|u + v - 2x_0\| \ge \langle v - x_0, y^*\rangle - \varepsilon$? If the answer to this question is in the affirmative then

$$\operatorname{epi}\big(\mathbb{I}_{x_0-C}\big)^* + \operatorname{epi}\big(\mathbb{I}_{C-x_0}\big)^* \supset E^* \times\,]0,\infty[\,.$$

Problem 11.6. Do there exist a nonzero *finite dimensional* Banach space E and $f, g \in \mathcal{PC}(E)$ such that the pair f, g is totally Fenchel unstable?

12 The restricted biconjugate and Fenchel–Moreau points

We now return to the more general considerations of Section 8. Let E and E^* be nonzero real vector spaces, and $\langle\cdot,\cdot\rangle\colon E\times E^*\to\mathbb{R}$ be a bilinear form that separates the point of E and also separates the points of E^*. We define *the restricted biconjugate* of f to be ${}^*(f^*)$: $E\to[-\infty,\infty]$ $\big($see (8.3)$\big)$. To simplify notation, we shall abbreviate this to ${}^*f^*$. It follows easily from the definition of f^* in (8.1) that, for all $x\in E$,

$$f(x)\geq {}^*f^*(x). \tag{12.1}$$

One of the fundamental results in convex analysis is the *Fenchel–Moreau theorem* that if $f\in\mathcal{PC}(E)$ is lower semicontinuous with respect to a E^*–compatible topology on E then $f={}^*f^*$ on E. We will revisit this result in Corollary 12.4.

Now suppose that $\mathcal{T}$ is a E^*–compatible topology on E and f is not necessarily $\mathcal{T}$–lower semicontinuous. Let us say that $x\in E$ is a *Fenchel–Moreau point of* f if equality holds in (12.1). It is very tempting to speculate that every point of $\mathcal{T}$–lower semicontinuity of f is a Fenchel–Moreau point of f. Example 12.1 below shows that this is false. However, we establish in Theorem 12.2 that every point of $\mathcal{T}$–lower semicontinuity of f is a Fenchel–Moreau point provided that f is bounded below in a $\mathcal{T}$–neighborhood of at least one point in its effective domain. Putting this another way, if there is a point of $\mathcal{T}$–lower semicontinuity of f that is not a Fenchel–Moreau point then f is unbounded below in every $\mathcal{T}$–neighborhood of every point of $\operatorname{dom} f$.

Example 12.1. Let E be an infinite–dimensional normed space. Fix $x^*\in E^*\setminus\{0\}$ and a discontinuous linear functional L on E. Define

$$f(x):=\begin{cases}\infty, & \text{if } \langle x,x^*\rangle<1;\\ L(x), & \text{if } \langle x,x^*\rangle\geq 1.\end{cases}$$

Clearly, $f\in\mathcal{PC}(E)$ and f is lower semicontinuous at 0. Let y^* be an arbitrary element of E^*. Since x^* and y^*-L are linearly independent, there exist $y,z\in E$ such that

$$\langle y,x^*\rangle=1,\ \langle z,x^*\rangle=0,\ (y^*-L)(y)=0,\ \text{and } (y^*-L)(z)=1.$$

Let $\lambda\in\mathbb{R}$, and set $x:=y+\lambda z$. Then $\langle x,x^*\rangle=\langle y,x^*\rangle=1$, and so $f(x)=L(x)$. Thus

$$f^*(y^*)\geq\langle x,y^*\rangle-f(x)=(y^*-L)(x)=\lambda(y^*-L)(z)=\lambda.$$

Since this holds for all $\lambda\in\mathbb{R}$, $f^*(y^*)=\infty$. Thus we have

$$f(0)=\infty>-\infty=\sup\nolimits_{y^*\in E^*}\big[\langle 0,y^*\rangle-f^*(y^*)\big],$$

and so 0 is not a Fenchel–Moreau point of f. (This example can also be justified using Corollary 10.3.)

Theorem 12.2 contains a positive result on Fenchel–Moreau points. The subtlety in its proof is that we can do arithmetic with the expression $f(x) - f(z)$, but we cannot do arithmetic with the expression $f(x) - f(y)$, which may well have the value $-\infty$.

Theorem 12.2. *Let $f \in \mathcal{PC}(E)$ be (finitely) bounded below in a $\mathcal{T}$–neighborhood of an element z of* $\operatorname{dom} f$*, and f be $\mathcal{T}$–lower semicontinuous at an element y of E. Then y is a Fenchel–Moreau point of f, and $f^* \in \mathcal{PC}(E^*)$.*

Proof. Let $\lambda \in \mathbb{R}$ and $\lambda < f(y)$. Choose $\nu \in \mathbb{R}$ and $Q \in \mathcal{S}(E, \mathcal{T})$ such that

$$Q(x - z) \le 1 \quad \Longrightarrow \quad f(x) > \nu \tag{12.2}$$

and

$$Q(x - y) \le 1 \quad \Longrightarrow \quad f(x) > \lambda. \tag{12.3}$$

Write $\rho := f(z) - \nu > 0$. We first prove that

$$x \in E \quad \Longrightarrow \quad f(x) + \rho Q(x - y) \ge \nu - \rho Q(y - z). \tag{12.4}$$

To this end, let x be an arbitrary element of E. If $Q(x - z) \le 1$ then (12.2) implies that

$$f(x) + \rho Q(x - z) \ge f(x) > \nu \ge \nu - \rho Q(y - z).$$

If, on the other hand, $Q(x - z) > 1$, let $\gamma := 1/Q(x - z) \in]0, 1[$ and put $u := \gamma x + (1 - \gamma)z$. Then $Q(u - z) = \gamma Q(x - z) = 1$ and so, from the convexity of f, and (12.2) with x replaced by u,

$$\gamma f(x) + (1 - \gamma)f(z) \ge f\big(\gamma x + (1 - \gamma)z\big) = f(u) > \nu,$$

thus the definition of ρ implies that $\gamma\big(f(x) - f(z)\big) + \rho \ge 0$. Substituting in the formula for γ and clearing of fractions yields $f(x) + \rho Q(x - z) \ge f(z)$. Consequently, using (12.2) with $x = z$ and the fact that $Q(z - z) \le 1$,

$$\begin{aligned} f(x) + \rho Q(x - y) &\ge f(x) + \rho Q(x - z) - \rho Q(y - z) \\ &\ge f(z) - \rho Q(y - z) > \nu - \rho Q(y - z). \end{aligned}$$

This completes the proof of (12.4). Now let $\sigma := \big[\lambda - \nu + \rho Q(y - z)\big] \vee 0 \ge 0$. We will prove that

$$x \in E \quad \Longrightarrow \quad f(x) + (\rho + \sigma)Q(x - y) \ge \lambda. \tag{12.5}$$

To this end, let x be an arbitrary element of E. If $Q(x - y) \le 1$ then (12.3) implies that $f(x) + (\rho + \sigma)Q(x - y) \ge f(x) > \lambda$. If, on the other hand, $Q(x - y) > 1$ then, from (12.4),

$$\begin{aligned} f(x) + (\rho + \sigma)Q(x - y) &= f(x) + \rho Q(x - y) + \sigma Q(x - y) \\ &\ge \nu - \rho Q(y - z) + \sigma \ge \lambda. \end{aligned}$$

This completes the proof of (12.5). It now follows from the Hahn–Banach–Lagrange theorem, Theorem 1.11 that there exists a linear functional L on E such that $L \le (\rho + \sigma)Q$ on E and

$$x \in E \quad \Longrightarrow \quad f(x) + L(x - y) \ge \lambda.$$

Let $z^* = -L \in E^*$. Then, for all $x \in E$, $\langle y, z^*\rangle - \big[\langle x, z^*\rangle - f(x)\big] \ge \lambda$. Taking the infimum over $x \in E$, $\langle y, z^*\rangle - f^*(z^*) \ge \lambda$. It follows by letting $\lambda \to f(y)$ that y is a Fenchel–Moreau point of f. Now $f^*\colon\ E^* \to]-\infty, \infty]$ is obviously convex. If z^* is a functional constructed as above for some $\lambda < f(y)$ then the inequality $\langle y, z^*\rangle - f^*(z^*) \ge \lambda$ implies that $f^*(z^*) \in \mathbb{R}$, and so $f^* \in \mathcal{PC}(E^*)$. □

Definition 12.3. If E is a nonzero Hausdorff locally convex space, we write $\mathcal{PCLSC}(E)$ for the set

$$\{f \in \mathcal{PC}(E)\colon\ f \text{ is lower semicontinuous on } E\}.$$

Corollary 12.4 is the original Fenchel–Moreau result, which follows immediately from Theorem 12.2. See Moreau, [64, Section 5–6, pp. 26–39] or Zălinescu, [119, Theorem 2.3.3, pp. 77–78]. Corollary 12.4 will be used explicitly in Theorem 18.7, (19.9), Lemma 35.1, Lemma 45.9, Theorem 48.4 and Lemma 48.9.

Corollary 12.4. *Let $f \in \mathcal{PCLSC}(E, \mathcal{T})$. Then $f^* \in \mathcal{PC}(E^*)$ and ${}^*f^* = f$ on E.*

13 Surrounding sets and the dom lemma

In this and the next section, we collect together some results on convex lower semicontinuous functions that we shall need for our later work. In this section, we give the "dom lemma", Lemma 13.3, which is a "quantitative" result, and the "dom corollary", Corollary 13.5, which is a "qualitative" result. The dom lemma will be of use in Lemma 22.7. Both the dom lemma and the dom corollary are subsumed by the results of the next section — we have treated them independently for essentially pedagogical reasons.

Let E be a Banach space, $x \in E$ and $A \subset E$. A is said to be *absorbing* if $\bigcup_{\lambda>0} \lambda A = E$. Any neighborhood of 0 is absorbing (exercise!). We write "$x \in \operatorname{sur} A$" and say that "$A$ surrounds x" if, for each $w \in E \setminus \{0\}$, there exists $\delta > 0$ such that $x + \delta w \in A$. The statement "$x \in \operatorname{sur} A$" is related to x being an "absorbing point" of A (see Phelps, [68, Definition 2.27(b), p. 28]), but differs in that we do not require that $x \in A$. We also note that, if A is convex then $\operatorname{sur} A \subset A$, and so $\operatorname{sur} A$ is identical with the "core" or algebraic interior of A. In particular:

$$\text{if } A \text{ is convex then} \qquad (0 \in \operatorname{sur} A \iff A \text{ is absorbing}). \tag{13.1}$$

In terms of these concepts, we have the following useful algebraic result about convex functions:

Lemma 13.1. *Let E be a nonzero vector space, $f \in \mathcal{PC}(E)$ and $\operatorname{dom} f$ surround 0. Then there exists $n \geq 1$ such that $\{z \in E\colon\ f(z) \leq n\}$ is absorbing.*

Proof. From (13.1),

$$\operatorname{dom} f \quad \text{is absorbing.} \tag{13.2}$$

In particular, $0 \in \operatorname{dom} f$. Let $n \geq f(0) \vee 0 + 1$. We will show that n has the required property. To this end, let y be an arbitrary element of E. (13.2) now provides $\lambda > 0$ and $x \in \operatorname{dom} f$ such that $\lambda y = x$. Choose $\mu \in]0, 1]$ so that $\mu\big(f(x) - n + 1\big) \leq 1$. Then

$$\begin{aligned} f(\mu\lambda y) = f(\mu x) &\leq \mu f(x) + (1-\mu) f(0) \\ &\leq \mu f(x) + (1-\mu)(n-1) \\ &= \mu\big(f(x) - n + 1\big) + n - 1 \leq n. \end{aligned}$$

Consequently, $\{z \in E\colon\ f(z) \leq n\}$ is absorbing, as required. □

Our next result depends ultimately on Baire's theorem:

Lemma 13.2. *Let E be a nonzero Banach space and C be a closed convex absorbing set in E. Then C is a neighborhood of 0.*

Proof. Let $D := C \cap -C$. Then D is closed, convex and absorbing (exercise!) and $D = -D$, i.e., D is a "barrel". The result follows by applying Kelley–Namioka, [52, p. 104] to D. □

Lemma 13.3. *Let E be a nonzero Banach space, $f \in \mathcal{PCLSC}(E)$ and $\operatorname{dom} f$ surround 0. Then there exist $\eta > 0$ and $n \geq 1$ such that*

$$z \in E \text{ and } \|z\| \leq \eta \quad \Longrightarrow \quad f(z) \leq n. \tag{13.3}$$

Furthermore, f is continuous at 0.

Proof. Choose $n \geq 1$ as in Lemma 13.1. Lemma 13.2 now implies that $\{z \in E\colon\ f(z) \leq n\}$ is a neighborhood of 0, and it follows from Theorem 8.7 that f is continuous at 0. □

Remark 13.4. The dom lemma, Lemma 13.3, can also be deduced from Rockafellar, [76, Corollary 7C, p. 61] (see also Moreau, [64, Proposition 5.f, p. 30] for a simpler proof of Rockafellar's result).

Corollary 13.5. *Let E be a nonzero Banach space and $f \in \mathcal{PCLSC}(E)$. Then*

$$\operatorname{sur}(\operatorname{dom} f) = \operatorname{int}(\operatorname{dom} f).$$

Proof. Exercise!

Remark 13.6. The classical "uniform boundedness theorem" can easily be deduced from the dom lemma. Here are the details: *Let E be a nonzero Banach space, F be a normed space and $\mathcal{B}$ be a nonempty pointwise bounded set of continuous linear operators from E into F. Then $\mathcal{B}$ is bounded in norm.*

Proof. Define $f\colon\ E \to \mathbb{R}$ by

$$f(x) := \sup_{T \in \mathcal{B}} \|Tx\|.$$

Since $\operatorname{dom} f = E$, we can apply the dom lemma. It then follows from (13.3) that

$$T \in \mathcal{B} \quad \Longrightarrow \quad \|T\| \le \frac{n}{\eta}. \qquad \square$$

The proof of the uniform boundedness theorem given above can be found in Holmes, [51, §17, p. 134]. Lemma 13.3 also implies the result that a convex lower semicontinuous function is locally bounded on the interior of its domain. (See, for instance, Phelps, [68, Proposition 3.3, p. 39.])

14 The ⊖–theorem

We now come to the "⊖–theorem", Theorem 14.2, which will be crucial for our analysis of the sums of maximally monotone operators in reflexive spaces. The ⊖–theorem is a "quantitative" result that also has a "qualitative" version, the "⊖–corollary", Corollary 14.3. Both of these results will have their uses, the ⊖–theorem in our proof of the Attouch–Brezis theorem, Theorem 15.1, and the ⊖–corollary in the local transversality theorem, Theorem 21.12, and also in Corollary 22.6. The ⊖–theorem, which generalizes the open mapping theorem (see Remark 14.4) can itself be generalized considerably. (In this connection, we refer the reader to Robinson, [75], Ursescu, [112], and Borwein, [16]). Here we confine our attention to what we will need in these notes. The idea for the proof of Lemma 14.1 is taken from Aubin–Ekeland, [3, Lemma 3.3.9, p. 136]. The dom lemma is an immediate consequence of the ⊖–theorem with $g := \mathbb{I}_{\{0\}}$.

Lemma 14.1 is the ⊖–theorem under more restrictive hypotheses. It will be bootstrapped in Theorem 14.2. We remind the reader that the function $f \ominus g\colon\ F \to [-\infty, \infty[$ was defined in Notation 8.3.

Lemma 14.1. *Let F be a nonzero Banach space, $f,\ g \in \mathcal{PCLSC}(F)$, $f \ge \|\cdot\|$ and $g \ge \|\cdot\|$ on F, and $\operatorname{dom} f - \operatorname{dom} g$ surround 0. Then*

$$f \ominus g \text{ is (finitely) bounded above in a neighborhood of } 0 \text{ in } F. \tag{14.1}$$

Proof. We first observe that, for all $w \in F$,

$(f \ominus g)(w) = \inf_{z\in E}\big[f(z)+g(z-w)\big] \ge \inf_{z\in E}\big[\|z\|+\|w-z\|\big] \ge \|w\| > -\infty,$

from which it follows easily that $f \ominus g \in \mathcal{PC}(F)$. (8.5) implies that $\operatorname{dom}(f\ominus g) = \operatorname{dom} f - \operatorname{dom} g$, and so $\operatorname{dom}(f\ominus g)$ surrounds 0. We now deduce from Lemma 13.1 that there exists $m > 1$ such that $\{w \in F\colon\ (f\ominus g)(w) < m\}$ is absorbing. Let $W := \{w \in F\colon\ (f \ominus g)(w) < m\}$. Since $\overline{W}$ is closed, convex and absorbing, Lemma 13.2 gives us that $\overline{W}$ is a neighborhood of 0 in F. Choose $\eta > 0$ so that

$$w \in F \text{ and } \|w\| \le 2\eta \quad\Longrightarrow\quad w \in \overline{W}. \tag{14.2}$$

We shall prove that

$$w \in F \text{ and } \|w\| \le \eta \quad\Longrightarrow\quad (f \ominus g)(w) \le m, \tag{14.3}$$

which will give (14.1). So let $w \in F$ and $\|w\| \le \eta$. Then, from (14.2), $2w \in \overline{W}$, consequently

$$\text{there exists } w_1 \in W \quad \text{such that} \quad \|2w - w_1\| \le \eta.$$

From (14.2) again, $4w - 2w_1 = 2(2w - w_1) \in \overline{W}$, thus

$$\text{there exists } w_2 \in W \quad \text{such that} \quad \|4w - 2w_1 - w_2\| \le \eta.$$

Continuing this argument, we find $w_1,\ w_2,\ w_3,\ \ldots \in W$ such that, for all $k \ge 1$,

$$\|2^k w - 2^{k-1}w_1 - \cdots - w_k\| \le \eta,$$

from which

$$\|w - 2^{-1}w_1 - \cdots - 2^{-k}w_k\| \le 2^{-k}\eta,$$

hence $\sum_{k=1}^{\infty} 2^{-k}w_k = w$. For all $n \ge 1$, since $w_n \in W$, we can choose $u_n \in F$ such that

$$f(u_n) + g(u_n - w_n) < m. \tag{14.4}$$

This implies that $\|u_n\| \le \|u_n\| + \|u_n - w_n\| < m$. Since F is complete, there exists $u \in F$ such that $\sum_{k=1}^{\infty} 2^{-k}u_k = u$, from which

$$\sum\nolimits_{k=1}^{\infty} 2^{-k}(u_k - w_k) = u - w.$$

(14.4) and the lower semicontinuity of f and g now imply that

$$f(u) + g(u - w) \le m,$$

from which $(f \ominus g)(w) \le m$. This completes the proof of (14.3), and hence also that of Lemma 14.1. □

Theorem 14.2. *Let F be a nonzero Banach space, $h, k \in \mathcal{PCLSC}(F)$, and $\operatorname{dom} h - \operatorname{dom} k$ surround 0. Then $h \ominus k$ is (finitely) bounded above in a neighborhood of 0 in F.*

Proof. This is immediate from Lemma 14.1 with $f = h \vee \|\cdot\|$ and $g = k \vee \|\cdot\|$, since $\operatorname{dom} f = \operatorname{dom} h$, $\operatorname{dom} g = \operatorname{dom} k$ and $h \ominus k \le f \ominus g$ on F. $\square$

Corollary 14.3. *Let F be a nonzero Banach space and $f,\ k \in \mathcal{PCLSC}(F)$. Then* $\operatorname{sur}(\operatorname{dom} f - \operatorname{dom} k) = \operatorname{int}(\operatorname{dom} f - \operatorname{dom} k)$, *and so* $\operatorname{sur}(\operatorname{dom} f - \operatorname{dom} k)$ *is open.*

Proof. We shall prove that

$$\operatorname{sur}(\operatorname{dom} f - \operatorname{dom} k) \subset \operatorname{int}(\operatorname{dom} f - \operatorname{dom} k). \tag{14.5}$$

This gives the desired result, since the reverse inclusion is trivial. So let x be an arbitrary element of $\operatorname{sur}(\operatorname{dom} f - \operatorname{dom} k)$. Define $h \in \mathcal{PCLSC}(F)$ by $h(y) := f(y+x) \quad (y \in F)$. Then $\operatorname{dom} h = \operatorname{dom} f - x$, which implies that $0 \in \operatorname{sur}(\operatorname{dom} h - \operatorname{dom} k)$. Theorem 14.2 now gives $\eta > 0$ and $m > 1$ such that if $w \in F$ and $\|w\| \le \eta$ then $(h \ominus k)(w) \le m$, from which $w \in \operatorname{dom} h - \operatorname{dom} k$. Thus we have proved that $0 \in \operatorname{int}(\operatorname{dom} h - \operatorname{dom} k)$. Since $\operatorname{dom} h - \operatorname{dom} k = \operatorname{dom} f - x - \operatorname{dom} k$, we have $x \in \operatorname{int}(\operatorname{dom} f - \operatorname{dom} k)$, which completes the proof of (14.5). $\square$

Remark 14.4. The classical "open mapping theorem" can easily be deduced from the $\ominus$–theorem. Here are the details. We first observe that *if C and D are closed convex subsets of a Banach space F and $C - D$ surrounds 0 then there exist $\eta > 0$ and $m > 1$ such that if $w \in F$ and $\|w\| \le \eta$ then*

there exist $c \in C$ and $d \in D$ such that $w = c - d$ and $\|c\| \le m$.

We obtain this by applying Theorem 14.2 with $h := \mathbb{I}_C \vee \|\cdot\|$ and $k := \mathbb{I}_D$. If now E and H are Banach spaces and $T \in B(E,H)$ is surjective then, for all $(x,y) \in E \times H$, there exists $z \in E$ such that that $y = Tz$, and consequently

$$(x,y) = (x,Tz) = (z,Tz) - (z-x,0) \in G(T) - \big(E \times \{0\}\big).$$

We now define $F := E \times H$ with norm $\|(x,y)\| := \sqrt{\|x\|^2 + \|y\|^2}$, $C := G(T)$ and $D := E \times \{0\}$. From the result above, there exist $\eta > 0$ and $m > 1$ such that if $y \in H$ and $\|y\| \le \eta$ then there exist $x, z \in E$ such that $(0,y) = (x,Tx) - (z,0)$ and $\|(x,Tx)\| \le m$. This implies that $Tx = y$ and $\|x\| \le m$, and it follows that T is an open mapping. $\square$

Thus the $\ominus$–theorem is both a generalization of the open mapping theorem and, in some sense, a "second order" generalization of the uniform boundedness theorem.

Remark 14.5. As we have observed, Lemma 14.1 is a generalization of Lemma 13.3. In this remark, we shall sketch a generalization of Lemma 13.3 in a totally different direction. Let E be a nonzero Banach space.
(a) *Let B be a nonmeager Borel set in E (that is, a Borel set of the second category). Then $B - B$ is a neighborhood of 0.*

(b) *Let D be a convex absorbing Borel set in E and D be* symmetric, *i.e.,* $D = -D$. *Then D is a neighborhood of* 0.
(c) *Let C be a convex absorbing Borel set in E. Then C is a neighborhood of* 0.
(d) *Let C be a convex Borel set in E. Then* $\operatorname{sur} C = \operatorname{int} C$.
(e) *Let $f \in \mathcal{PC}(E)$ be a Borel function and* $\operatorname{dom} f$ *surround* 0. *Then there exist $\eta > 0$ and $n \geq 1$ such that*

$$w \in E \text{ and } \|w\| \leq \eta \quad \Longrightarrow \quad f(w) \leq n.$$

Proof. (a) Any Borel set satisfies the "condition of Baire", that is to say, there exists an open set U such that $U \setminus B$ and $B \setminus U$ are meager, and so (a) follows from the "difference theorem". See Kelley–Namioka, [52, 10.4, p. 92] and the discussion preceding.

(b) It follows from Baire's theorem that E, being a complete metric space, is nonmeager. Since $\bigcup_{n \geq 1} nD = E$ there exists $n \geq 1$ such that nD is nonmeager, from which $\frac{1}{2}D$ is nonmeager. Since D is convex and symmetric,

$$D = \tfrac{1}{2}D + \tfrac{1}{2}D = \tfrac{1}{2}D - \tfrac{1}{2}D,$$

thus it follows from (a) that D is a neighborhood of 0.

(c) Let $D := C \cap -C$. Then D is a convex absorbing Borel set and $D = -D$. From (b), D is a neighborhood of 0, from which C is a neighborhood of 0 also.

(d) is immediate from (c), a translation argument and (13.1).

(e) From Lemma 13.1, there exists $n \geq 1$ such that $\{x \in E\colon\ f(x) \leq n\}$ is absorbing. The result now follows from (c).

Remark 14.6. Theorem 14.2 and Remark 14.5 suggest the following question:

Problem 14.7. Let F be a Banach space, $h,\ k \in \mathcal{PC}(F)$ be Borel functions and $\operatorname{dom} h - \operatorname{dom} k$ surround 0. Is $h \ominus k$ necessarily (finitely) bounded above in some neighborhood of 0 in F? In particular: Let C and D be convex Borel sets in F and $C - D$ be absorbing. Is $C - D$ necessarily a neighborhood of 0 in F?

15 The Attouch–Brezis theorem

This section is devoted to a single result, the Attouch–Brezis version of the Fenchel duality theorem, which we will use explicitly in Lemma 16.2 and the local transversality theorem, Theorem 21.12. As stated below, this result also follows from [1, Corollary 2.3, pp. 131–132] (a much more general result was established in [119, Theorem 2.8.6, pp. 125–126]):

Theorem 15.1. *Let E be a nonzero Banach space, $f,\ g \in \mathcal{PCLSC}(E)$,*

$$F := \bigcup\nolimits_{\lambda>0} \lambda\big[\operatorname{dom} f - \operatorname{dom} g\big] \quad \textit{be a closed subspace of } E$$

and

$$f + g \geq 0 \textit{ on } E.$$

Then there exists a Fenchel functional for f and g.

Proof. Since $0 \in F$, there exists $z \in \operatorname{dom} f \cap \operatorname{dom} g$. Define $h,\ k\colon E \to]{-\infty}, \infty]$ by $h(x) := f(x+z)$ and $k(x) := g(x+z) \quad (x \in E)$. Then $\operatorname{dom} h \subset F$, $\operatorname{dom} k \subset F$ and $\operatorname{dom} h - \operatorname{dom} k$ surrounds 0 in F. From the $\ominus$–theorem, Theorem 14.2, there exist $\eta > 0$ and $m > 1$ such that if $w \in F$ and $\|w\| \leq \eta$ then

$$\text{there exist } u, v \in F \text{ such that } w = u - v \text{ and } h(u) + k(v) \leq m.$$

But then $w = (u+z) - (v+z)$ and $f(u+z) + g(v+z) \leq m$, and so $(f \ominus g)(w) \leq m$. The result now follows from Theorem 8.4(b). □

Remark 15.2. Theorem 15.1 can easily be bootstrapped into the following result $\big($which is [1, Theorem 1.1, pp. 126–130]$\big)$: *Let E be a nonzero Banach space, $f,\ g \in \mathcal{PCLSC}(E)$ and $\bigcup_{\lambda>0} \lambda(\operatorname{dom} f - \operatorname{dom} g)$ be a closed subspace of E. Then, for all $x^* \in E^*$,*

$$(f+g)^*(x^*) = \min\nolimits_{z^* \in E^*} \big[f^*(x^* - z^*) + g^*(z^*)\big]. \tag{10.1}$$

Remark 15.3. It is often said that, in the normed case, Theorem 15.1 is a "generalization" of Rockafellar's version of the Fenchel duality theorem, Corollary 8.6. This is inaccurate, since Theorem 15.1 requires both f and g to be lower semicontinuous.

In the two cases in these notes in which Corollary 8.6 is used explicitly in a normed space (Theorem 9.3 and Lemma 35.5), we cannot substitute Theorem 15.1 because of the lack of this semicontinuity.

Corollary 8.6 is also used explicitly in a non–normed situation in the transversality theorem, Theorem 19.16, and also in Lemma 22.1.

The Attouch–Brezis theorem is, however, a very powerful result, which enables us to consider Fenchel duality in which $\operatorname{int}\operatorname{dom} f = \operatorname{int}\operatorname{dom} g = \emptyset$. We will investigate a bivariate version of the Attouch–Brezis theorem in the next section.

16 A bivariate Attouch–Brezis theorem

The main result of this section is the bivariate version of the Attouch–Brezis theorem that will appear in Theorem 16.4. Apart from some minor changes of notation, this result was first proved in Simons–Zălinescu [109, Theorem 4.2, pp. 9–10]. The proof given here using Lemma 16.2 is somewhat simpler, and first appeared in [106].

Notation 16.1. If E and F are nonzero Banach spaces, we norm $E \times F$ by

$$\|b\| := \sqrt{\|b_1\|^2 + \|b_2\|^2} \quad \big(b = (b_1, b_2) \in E \times F\big).$$

The dual of $E \times F$ is $F^* \times E^*$ under the pairing

$$\langle b, v\rangle := \langle b_1, v_2\rangle + \langle b_2, v_1\rangle \quad \big(b = (b_1, b_2) \in E \times F,\ v = (v_1, v_2) \in F^* \times E^*\big),$$

and the dual norm of $F^* \times E^*$ is given by $\big\|(v_1, v_2)\big\| = \sqrt{\|v_1\|^2 + \|v_2\|^2}$. We define the *projection maps* π_1, π_2 by $\pi_1(x, y) := x$ and $\pi_2(x, y) := y$.

Lemma 16.2 is a stepping–stone to Theorem 16.4. It will also be used explicitly in Theorem 46.3, in our proof of the maximal monotonicity of the sum of maximally monotone multifunctions with convex graph.

Lemma 16.2. *Let E and F be nonzero Banach spaces, $p, q \in \mathcal{PCLSC}(E \times F)$,*

$$L := \bigcup\nolimits_{\lambda>0} \lambda\big[\pi_1 \mathrm{dom}\, p - \pi_1 \mathrm{dom}\, q\big] \text{ be a closed subspace of } E$$

and

$$(x, y, z) \in E \times F \times F \quad \Longrightarrow \quad p(x, y) + q(x, z) \ge 0.$$

Then

$$\text{there exists } x^* \in E^* \text{ such that } p^*(0, -x^*) + q^*(0, x^*) \le 0.$$

Proof. For all $(x, y, z) \in E \times F \times F$, let $f(x, y, z) := p(x, y)$ and $g(x, y, z) := q(x, z)$. We first prove that

$$\bigcup\nolimits_{\lambda>0} \lambda\big[\mathrm{dom}\, f - \mathrm{dom}\, g\big] = L \times F \times F. \tag{16.1}$$

To this end, let $(x, y, z) \in L \times F \times F$. Then there exist $\lambda > 0$, $(a_1, a_2) \in \mathrm{dom}\, p$ and $(b_1, b_2) \in \mathrm{dom}\, q$ such that $x = \lambda(a_1 - b_1)$. Thus

$$(x, y, z) = \lambda\big[(a_1, a_2, b_2 + z/\lambda) - (b_1, a_2 - y/\lambda, b_2)\big] \in \lambda\big[\mathrm{dom}\, f - \mathrm{dom}\, g\big].$$

This establishes "$\supset$" in (16.1), and (16.1) now follows since the inclusion "$\subset$" is obvious. Also,

$$(x, y, z) \in E \times F \times F \quad \Longrightarrow \quad (f + g)(x, y, z) = p(x, y) + q(x, z) \ge 0.$$

Now represent the dual of $E \times F \times F$ by $E^* \times F^* \times F^*$ under the pairing $\big\langle (x, y, z), (x^*, y^*, z^*)\big\rangle := \langle x, x^*\rangle + \langle y, y^*\rangle + \langle z, z^*\rangle$. Since $L \times F \times F$ is a closed subspace of $E \times F \times F$, Theorem 15.1 gives $(x^*, y^*, z^*) \in E^* \times F^* \times F^*$ such that

$$f^*(-x^*, -y^*, -z^*) + g^*(x^*, y^*, z^*) \le 0. \tag{16.2}$$

So $f^*(-x^*, -y^*, -z^*) < \infty$, from which $f^*(-x^*, -y^*, -z^*) = p^*(-y^*, -x^*)$ and $z^* = 0$. Similarly, $g^*(x^*, y^*, z^*) = q^*(z^*, x^*)$ and $y^* = 0$. Thus (16.2) reduces to

$$p^*(0, -x^*) + q^*(0, x^*) \le 0. \qquad \square$$

Before discussing the promised bivariate version of the Attouch–Brezis theorem, we make some preliminary definitions:

Definition 16.3. Let E and F be nonzero Banach spaces, $B := E \times F$ and $f, g \in \mathcal{PC}(B)$. For all $b \in B$, let

$$(f \oplus_2 g)(b) := \inf\big\{f(a) + g(c)\colon\ a, c \in B,\ a_1 = c_1 = b_1,\ a_2 + c_2 = b_2\big\}.$$

So $(f \oplus_2 g)(x, \cdot)$ is the *inf–convolution* of $f(x, \cdot)$ and $g(x, \cdot)$. Similarly, for all $b \in B$, let

$$(f \oplus_1 g)(b) := \inf\big\{f(a) + g(c)\colon\ a, c \in B,\ a_1 + c_1 = b_1,\ a_2 = c_2 = b_2\big\}.$$

The bivariate version of the Attouch–Brezis theorem that appears in Theorem 16.4 below will be used explicitly in Lemma 22.9 and Theorem 35.8. This latter result on $\widetilde{\text{BC}}$–functions will be pivotal for our investigation of the different classes of maximally monotone multifunctions on a nonreflexive Banach space. The conclusion of Theorem 16.4(a) is that $(f \oplus_2 g)^*(y^*, \cdot)$ is the *exact inf–convolution* of $f^*(y^*, \cdot)$ and $g^*(y^*, \cdot)$. A similar comment can be made about Theorem 16.4(b).

Theorem 16.4. *Let E and F be nonzero Banach spaces, $B := E \times F$ and $f, g \in \mathcal{PCLSC}(B)$. Write $B^* = F^* \times E^*$*
(a) *Let*

$$\textstyle\bigcup_{\lambda>0} \lambda\big[\pi_1 \mathrm{dom}\, f - \pi_1 \mathrm{dom}\, g\big] \text{ be a closed subspace of } E$$

and, for all $b \in B$, $(f \oplus_2 g)(b) > -\infty$. Then, for all $v \in B^* = (E \times F)^*$,

$$(f \oplus_2 g)^*(v) = \min\big\{f^*(u) + g^*(w)\colon\ u, w \in B^*,\ u_1 = w_1 = v_1,\ u_2 + w_2 = v_2\big\}.$$

In particular, $(f \oplus_2 g)^* = f^* \oplus_2 g^*$ on B^*.
(b) *Let*

$$\textstyle\bigcup_{\lambda>0} \lambda\big[\pi_2 \mathrm{dom}\, f - \pi_2 \mathrm{dom}\, g\big] \text{ be a closed subspace of } F$$

and, for all $b \in B$, $(f \oplus_1 g)(b) > -\infty$. Then, for all $v \in B^* = (E \times F)^*$,

$$(f \oplus_1 g)^*(v) = \min\big\{f^*(u) + g^*(w)\colon\ u, w \in B^*,\ u_1 + w_1 = v_1,\ u_2 = w_2 = v_2\big\}.$$

In particular, $(f \oplus_1 g)^* = f^* \oplus_1 g^*$ on B^*.

Proof. Let $h := f \oplus_2 g$, Then h is convex and, since $\pi_1 \operatorname{dom} f \cap \pi_1 \operatorname{dom} g \neq \emptyset$, h is proper. Let $v \in B^*$. It is easy to see that

$$h^*(v) \le \inf\big\{f^*(u) + g^*(w)\colon\ u, w \in B^*,\ u_1 = w_1 = v_1,\ u_2 + w_2 = v_2\big\}.$$

So what we have to prove for (a) is that there exists $x^* \in E^*$ such that

$$f^*(v_1, v_2 - x^*) + g^*(v_1, x^*) \le h^*(v). \tag{16.3}$$

Since h is proper, $h^*(v) > -\infty$, so we can and will suppose that $h^*(v) \in \mathbb{R}$. Define $p, q \in \mathcal{PCLSC}(B)$ by $p(x, y) := h^*(v) + f(x, y) - \langle x, v_2\rangle - \langle y, v_1\rangle$ and $q(x, z) := g(x, z) - \langle z, v_1\rangle$. Then, for all $(x, y, z) \in E \times F \times F$, the Fenchel–Young inequality, (8.2), implies that

$$\begin{aligned} p(x, y) + q(x, z) &= h^*(v) + f(x, y) - \langle x, v_2\rangle - \langle y, v_1\rangle + g(x, z) - \langle z, v_1\rangle \\ &\ge h^*(v) + h(x, y + z) - \langle x, v_2\rangle - \langle y + z, v_1\rangle \ge 0. \end{aligned}$$

Lemma 16.2 now gives $x^* \in E^*$ such that $p^*(0, -x^*) + q^*(0, x^*) \le 0$. By direct computation,

$$p^*(0, -x^*) = f^*(v_1, v_2 - x^*) - h^*(v) \quad \text{and} \quad q^*(0, x^*) = g^*(v_1, x^*),$$

which implies (16.3), and completes the proof of (a). The proof of (b) is similar. $\square$