Mathematical Programming 57 (1992) 15-48
North-Holland

Partially finite convex programming, Part I: Quasi relative interiors and duality theory

J.M. Borwein*

Department of Mathematics, Statistics and Computing Science, Dalhousie University, Halifax, N.S., Canada B3H 3J5

A.S. Lewis

Department of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, Waterloo, Ont., Canada N2L 3G1

Received 17 February 1988
Revised manuscript received 31 June 1989; 25 January 1990

We study convex programs that involve the minimization of a convex function over a convex subset of a topological vector space, subject to a finite number of linear inequalities. We develop the notion of the quasi relative interior of a convex set, an extension of the relative interior in finite dimensions. We use this idea in a constraint qualification for a fundamental Fenchel duality result, and then deduce duality results for these problems despite the almost invariable failure of the standard Slater condition. Part II of this work studies applications to more concrete models, whose dual problems are often finite-dimensional and computationally tractable.

AMS 1985 Subject Classifications: Primary 90C25, 49B27; Secondary 90C48, 52A07, 65K05.

Key words: Convex programming, duality, constraint qualification, Fenchel duality, semi-infinite programming.

1. Introduction

Duality theorems are central to the study of constrained optimization problems. From the point of view of problems arising in practice, their usefulness is twofold: under appropriate conditions (constraint qualifications), first they enable us to check whether a given feasible solution is optimal, and secondly we can often find the optimal solution by first solving the corresponding dual problem. When the primal problem is infinite-dimensional two difficulties become apparent. The most straightforward constraint qualification (the "Slater" condition) is frequently not met in practice, since it requires the existence of an interior point of a convex set which often has empty interior. Furthermore, since the dual problem will generally also be infinite-dimensional, it may be very hard to solve.

In this paper we shall primarily be concerned with problems of the form

$$\text{(P)}\qquad \begin{array}{ll} \inf & f(x) \\ \text{subject to} & Ax \leq b, \\ & x \in C, \end{array}$$

* Present address: Department of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, Waterloo, Ont., Canada N2L 3G1.

where C is some convex subset of a vector space X, $f: X \to]-\infty, \infty]$ is convex, $A: X \to \mathbb{R}^n$ is linear and $b \in \mathbb{R}^n$. This convex model covers a wide variety of interesting problems which arise in practice. Examples we shall consider include constrained approximation problems, variants of semi-infinite linear programming, and semi-infinite transportation problems.

We shall construct a powerful duality theory for this class of problems which has two main advantages over standard duality theorems in infinite-dimensional optimization. First, the new constraint qualification required to apply the duality result is much weaker than the standard Slater condition, and equally easy to check: rather than requiring an interior point of C to be feasible, we only need a feasible point in the 'quasi relative interior' of C, an extension of the idea of relative interior in $\mathbb{R}^n$. The quasi relative interior will frequently be nonempty even when the interior of C is empty. Secondly, under suitable conditions on the space X and when the function f and the set C are sufficiently simple, the corresponding dual problem is finite-dimensional and of very simple form: it is frequently computationally tractable, and by first solving it we can often calculate a primal optimal solution.

An outline of the paper is as follows. We begin in Sections 2 and 3 by studying the notion of the quasi relative interior of a convex set. We develop a number of its elementary properties, largely analogous to results for relative interior in $\mathbb{R}^n$. An important property of the relative interior of a convex set in $\mathbb{R}^n$ is that it is nonempty; we give a parallel result for quasi relative interiors. We also compute the quasi relative interior of some important sets, including the positive cones of some standard Banach spaces.

The next section (4) presents the fundamental result, a Fenchel duality theorem. It is from this result that we derive the duality theory described above. We also obtain as corollaries a subgradient formula and a minimax theorem.

The calculation and uniqueness of primal solutions, and the numerical treatment of the dual problem depend on the differentiability of the dual objective function. Many of the problems we consider are posed in vector lattices, so Section 5 is devoted to studying the differentiability of certain commonly arising real-valued convex functions in vector lattices. We also derive some properties of the *absolute kernel* of a positive linear functional on a vector lattice, which we require in what follows.

In Part II of this work we study the application of these ideas to more concrete models. We concentrate on two special cases: the case when C is a cone (Section 6, 7) and the case when C is of the form

$$(*) \qquad \left\{ (x_1, \ldots, x_m) \in E^m \;\middle|\; \sum_{i=1}^{m} x_i = e, x_i \geq 0 \text{ for each } i \right\},$$

where $e \geq 0$ is some fixed element of the partially ordered vector space E (Section 8, 9).

With C a cone we obtain duality results for semi-infinite linear programming and variants, and for certain quadratic programs in the Hilbert space of square integrable functions, $L^2(T, \mu)$. When the function f is a norm we obtain the constrained interpolation problems considered in, for example, Irvine, Marin and Smith (1986), and spectral estimation problems (see Ben-Tal, Borwein and Teboulle, 1988 and 1989).

In the second case, where C is of the form (*) (and $X = E^m$), when the function f is linear we obtain an analogue of many of the features of classical linear programming theory. As examples we consider semi-infinite linear programming with an upper bound constraint, and L^1-approximation. In Section 9 we study two more examples: the semi-infinite transportation problem considered in Kortanek and Yamasaki (1982), and the generalized market area problem (see Lowe and Hurter, 1976).

Throughout this work we shall address the standard questions of constrained optimization: the formulation of a dual problem, conditions ensuring the equality of primal and dual values, attainment and uniqueness in the primal and dual problem, optimality conditions and complementary slackness, and the derivation of a primal optimum from the solution of the dual problem. We shall for the most part defer consideration of computational aspects to a later paper.

The convex analytic terminology we use, when not explicit, is that of Rockafellar (1970). We also use ideas and terminology from the theory of vector lattices extensively in what follows. Vector lattices provide an extremely useful unifying framework for many of our problems. However, in order to maintain accessibility, where possible we have given elementary or convex analytic proofs of the results we use. The reader will find that the calculations we perform in lattice notation may be easily followed through in concrete spaces, with no knowledge of vector lattices.

2. Quasi relative interiors

We begin with a simple result concerning convex sets in $\mathbb{R}^n$, which will motivate our definition of the quasi relative interior of a set. Given a vector space X and $C \subset X$, we denote the cone generated by C by $\mathbb{P}C = \{\lambda x \mid x \in C, \lambda \in \mathbb{R}, \lambda \geqslant 0\}$.

Proposition 2.1. *Suppose $C \subset \mathbb{R}^n$ is convex. Then $\hat{x} \in \operatorname{ri} C$ if and only if $\mathbb{P}(C - \hat{x})$ is a subspace.*

Proof. Suppose $\hat{x} \in \operatorname{ri} C$, so for some neighbourhood N of $\hat{x}$, $N \cap \operatorname{aff} C \subset C$. It follows that $\mathbb{P}(C - \hat{x}) = (\operatorname{aff} C) - \hat{x}$, which is a subspace. On the other hand, suppose $\hat{x} \notin \operatorname{ri} C$, so we can properly separate $\hat{x}$ from C: for some $y \in \mathbb{R}^n$, $y^{\mathrm{T}}\hat{x} \leqslant y^{\mathrm{T}}x$ for all $x \in C$ with strict inequality for some $\bar{x} \in C$ (Rockafellar, 1970, 11.3). Thus $y^{\mathrm{T}}z \geqslant 0$ for all $z \in \mathbb{P}(C - \hat{x})$, and $\bar{x} - \hat{x} \in \mathbb{P}(C - \hat{x})$, but $\hat{x} - \bar{x} \notin \mathbb{P}(C - \hat{x})$, so $\mathbb{P}(C - \hat{x})$ is not a subspace. $\square$

Lemma 2.2. *Suppose $C \subset \mathbb{R}^n$ is convex. Then C is a subspace if and only if* cl C *is a subspace.*

Proof. If C is a subspace then $\mathrm{cl}\, C = C$. On the other hand, if $C \neq \mathrm{cl}\, C$ and $\mathrm{cl}\, C$ is a subspace then C lies in a closed halfspace of $\mathrm{cl}\, C$, by Rockafellar (1970, 11.5.2), which is clearly impossible. □

Thus in Proposition 2.1 we could replace $\mathbb{P}(C - \hat{x})$ with its closure. This motivates the following definition. Henceforth, X will be a Hausdorff topological vector space.

Definition 2.3. For convex $C \subset X$, the *quasi relative interior* of C (qri C) is the set of those $x \in C$ for which $\mathrm{cl}\, \mathbb{P}(C - x)$ is a subspace. (We also write $\tau -\mathrm{qri}\, C$, where τ denotes the relevant topology.)

The above observation shows that in $\mathbb{R}^n$ the notions of relative interior and quasi relative interior coincide:

Proposition 2.4. *Let X be a normed vector space, and $C \subset X$ convex and finite-dimensional. Then* $\mathrm{qri}\, C = \mathrm{ri}\, C$.

Proof. Restricting attention to aff C, we are reduced to the $\mathbb{R}^n$ case (since all norms on $\mathbb{R}^n$ are equivalent). The result now follows from Proposition 2.1 and Lemma 2.2. □

Notice that, unlike the finite-dimensional case, we may not have $\mathrm{qri}\, C = \mathrm{qri}(\mathrm{cl}\, C)$. For instance, if C is a dense subspace of X then $\mathrm{qri}\, C = C$ while $\mathrm{qri}(\mathrm{cl}\, C) = X$.

The following result is trivial.

Proposition 2.5. *Suppose $X = \prod_{i=1}^m X_i$, where the X_i's are topological vector spaces, and that $C_i \subset X_i$ is convex for each i. Then* $\mathrm{qri}(\prod_{i=1}^m C_i) = \prod_{i=1}^m (\mathrm{qri}\, C_i)$. □

Proposition 2.6. *In a normed vector space the quasi relative interiors of a convex set in the weak and norm topologies are identical.*

Proof. The result follows immediately from the fact that in a normed vector space, a convex set has identical weak and norm closure. □

The following result is extremely important in what follows. It will be strengthened in Proposition 2.10.

Proposition 2.7. *Suppose $C \subset X$ is convex, and $A: X \to \mathbb{R}^n$ is a continuous linear map. Then $A(\text{qri } C) \subset \text{ri}(AC)$.*

Proof. Let $x \in \text{qri } C$, so $\text{cl } \mathbb{P}(C-x)$ is a subspace. It follows that $A(\text{cl } \mathbb{P}(C-x)) \subset \mathbb{R}^n$ is a subspace, and hence also closed. But

$$\begin{aligned} A(\text{cl } \mathbb{P}(C-x)) &\subset \text{cl}(A\mathbb{P}(C-x)) \quad \text{(since } A \text{ continuous)} \\ &\subset \text{cl}(A[\text{cl } \mathbb{P}(C-x)]) \\ &= A(\text{cl } \mathbb{P}(C-x)). \end{aligned}$$

Hence $\text{cl}(A\mathbb{P}(C-x))$ is a subspace, and therefore, by Lemma 2.2, so is $A\mathbb{P}(C-x) = \mathbb{P}(AC - Ax)$. The result now follows by Proposition 2.1. □

The next result gives what is often the most useful tool in determining the quasi relative interior of sets. X^* denotes the topological dual of X, and $N_C(\hat{x}) \subset X^*$ denotes the normal cone to C at $\hat{x} \in C$:

$$N_C(\hat{x}) = \{\phi \in X^* \mid \phi(x - \hat{x}) \leq 0 \ \forall x \in C\}.$$

Proposition 2.8. *Let X be locally convex, $C \subset X$ be convex, and $\hat{x} \in C$. Then $\hat{x} \in \text{qri } C$ if and only if $N_C(\hat{x})$ is a subspace of X^*.*

Proof. For $K \subset X$, the polar of K is given by

$$K^\circ = \{\phi \in X^* \mid \phi(x) \leq 1 \ \forall x \in K\} = \{\phi \in X^* \mid \phi(x) \leq 0 \ \forall x \in K\},$$

if K is a cone. Similarly, for a cone $L \subset X^*$,

$${}^\circ L = \{x \in X \mid \phi(x) \leq 0 \ \forall \phi \in L\}.$$

It is immediate that if K is a subspace, so is K°, and if L is a subspace, so is ${}^\circ L$.

Now for $\phi \in X^*$, $\phi(x - \hat{x}) \leq 0$ for all $x \in C$ if and only if $\phi(u) \leq 0$ for all $u \in \text{cl } \mathbb{P}(C - \hat{x})$, by the continuity of ϕ. Thus $N_C(\hat{x}) = (\text{cl } \mathbb{P}(C - \hat{x}))^\circ$.

On the other hand,

$${}^\circ N_C(\hat{x}) = {}^\circ((\text{cl } \mathbb{P}(C - \hat{x}))^\circ) = \text{cl } \mathbb{P}(\{0\} \cup \text{cl } \mathbb{P}(C - \hat{x})) = \text{cl } \mathbb{P}(C - \hat{x}),$$

by the bipolar theorem (Holmes, 1975). The result now follows. □

Lemma 2.9. *Let X be locally convex, $C \subset X$ be convex, and suppose $x_1 \in \text{qri } C$ and $x_2 \in C$. Then $\lambda x_1 + (1-\lambda)x_2 \in \text{qri } C$, for all $0 < \lambda \leq 1$.*

Proof. By convexity, $\lambda x_1 + (1-\lambda)x_2 \in C$, and we can assume $\lambda < 1$. Suppose $\phi \in N_C(\lambda x_1 + (1-\lambda)x_2)$, so $\phi(x - \lambda x_1 - (1-\lambda)x_2) \leq 0 \ \forall x \in C$. Putting $x = x_1$ gives $\phi(x_1 - x_2) \leq 0$, while $x = x_2$ gives $\phi(x_2 - x_1) \leq 0$, so $\phi(x_1) = \phi(x_2)$. Thus $\phi(x - x_1) \leq 0$ for all $x \in C$, or $\phi \in N_C(x_1)$.

But $x_1 \in \text{qri } C$, so by Proposition 2.8, $-\phi \in N_C(x_1)$ also. It follows that

$$-\phi(x - \lambda x_1 - (1-\lambda)x_2) \leq 0 \quad \text{for all } x \in C,$$

or $-\phi \in N_C(\lambda x_1 + (1-\lambda)x_2)$. Thus $N_C(\lambda x_1 + (1-\lambda)x_2)$ is a subspace, and the result follows by Proposition 2.8. □

Notice that, unlike in the finite-dimensional case (Rockafellar, 1970, 6.1), this result can fail if we allow $x_2 \in \text{cl}\, C$. For example, let $X = C[0, 1]$, with C being the dense subspace of all polynomials. Then qri $C = C$, cl $C = X$, and the result clearly fails.

We are now ready to strengthen Proposition 2.7.

Proposition 2.10. *Let X be locally convex, $C \subset X$ be convex, and $A: X \to \mathbb{R}^n$ a continuous linear map. If* qri $C \neq \emptyset$ *then* $A(\text{qri}\, C) = \text{ri}(AC)$.

Proof. By Proposition 2.7, $A(\text{qri}\, C) \subset \text{ri}(AC)$. Suppose $y \in \text{ri}(AC)$. By assumption, there exists $x_1 \in \text{qri}\, C$. For some $\varepsilon > 0$, $y - \varepsilon(Ax_1 - y) \in AC$, so for some $x_2 \in C$, $y = A((\varepsilon/1+\varepsilon)x_1 + (1/1+\varepsilon)x_2) \in A(\text{qri}\, C)$, by Lemma 2.9. □

Proposition 2.11. *Let X be locally convex, and $C \subset X$ be convex. Then* qri C *is convex.*

Proof. Immediate, by Lemma 2.9. □

We shall see in the next section that in the case where C is the positive cone in a topological lattice, the concept of quasi relative interior is very close to the concept of quasi-interior in Schaefer (1974). Schaefer shows in particlar in this case that if X is locally convex and qri $C \neq \emptyset$ then qri C is dense in C, that if int $C \neq \emptyset$ then qri $C = \text{int}\, C$, and that if X is separable and complete metrizable then qri $C \neq \emptyset$. We shall prove analogous results to these.

Proposition 2.12. *Let X be locally convex and $C \subset X$ be convex. If* qri $C \neq \emptyset$ *then* $\text{cl}(\text{qri}\, C) = \text{cl}\, C$.

Proof. Clearly $\text{cl}(\text{qri}\, C) \subset \text{cl}\, C$. Pick $\hat{x} \in \text{qri}\, C$, and suppose $x \in C$. By Lemma 2.9, $\lambda\hat{x} + (1-\lambda)x \in \text{qri}\, C$, for all $0 < \lambda \leq 1$. Letting $\lambda \to 0$ it follows that $x \in \text{cl}(\text{qri}\, C)$. Thus $C \subset \text{cl}(\text{qri}\, C)$, so $\text{cl}\, C \subset \text{cl}(\text{qri}\, C)$, which completes the proof. □

Theorem 2.13. *Let X be locally convex and $C, D \subset X$ be convex. If $C \cap \text{int}\, D \neq \emptyset$, then* $(\text{qri}\, C) \cap (\text{int}\, D) = \text{qri}(C \cap D)$.

Proof. Let $\hat{x} \in C \cap \text{int}\, D$. Note first that $\text{qri}(C \cap D) \subset \text{int}\, D$. To see this, suppose $0 \in \text{qri}(C \cap D)$, but $0 \notin \text{int}\, D$. Then we can separate (Holmes, 1975): for some nonzero $\phi \in X^*$, $\phi(x) \leq 0$, for all $x \in D$. In particular, $\phi \in N_{C \cap D}(0)$, but $\hat{x} \in C \cap D$ and $\phi(\hat{x}) < 0$, since $\hat{x} \in \text{int}\, D$. Thus $-\phi \notin N_{C \cap D}(0)$, so $N_{C \cap D}(0)$ is not a subspace, which is a contradiction by Proposition 2.8.

Now suppose $0 \in C \cap \text{int}\, D$. Then $N_{C \cap D}(0) = N_C(0)$, since if $\phi \in N_{C \cap D}(0)$ then for any $x \in C$ there exists $\varepsilon > 0$ with $\varepsilon x \in C \cap D$, so $\phi(\varepsilon x) \leq 0$, and thus $\phi(x) \leq 0$: therefore $\phi \in N_C(0)$. It follows that $(\text{int}\, D) \cap \text{qri}(C \cap D) = (\text{int}\, D) \cap (\text{qri}\, C)$, by Proposition 2.8. The result now follows. □

Corollary 2.14. *Let X be locally convex and $D \subset X$ be convex. If* $\operatorname{int} D \neq \emptyset$ *then* $\operatorname{qri} D = \operatorname{int} D$.

Proof. Put $C = X$ in Theorem 2.13. □

The next result gives a geometric interpretation of the idea of a quasi relative interior point. First, a definition from Peressini (1967).

Definition 2.15. Let $C \subset X$ be convex, and $x \in C$. Then x is a *nonsupport point* of C if every closed supporting hyperplane to C at x contains C.

Proposition 2.16. *Let X be locally convex and $C \subset X$ be convex. Then the quasi relative interior of C is exactly the set of nonsupport points of C.*

Proof. By Proposition 2.8, $\hat{x} \notin \operatorname{qri} C$ if and only if for some $\phi \in X^*$, $\phi(x - \hat{x}) \leq 0$ for all $x \in C$, with strict inequality for some $\bar{x} \in C$, and this ϕ then defines a closed supporting hyperplane to C at $\hat{x}$ which does not contain $\bar{x}$. □

The importance of the quasi relative interior of a set, as we have noted already, is that it is often nonempty even when C has empty interior. Our next main result gives conditions ensuring $\operatorname{qri} C \neq \emptyset$. We begin with two lemmas.

Lemma 2.17. *Let X be be a separable normed space. Then every subset of X^* is $\sigma(X^*, X)$-separable.*

Proof. Let $C \subset X^*$. The unit ball B_{X^*} is $\sigma(X^*, X)$-compact by the Banach–Alaoglu Theorem (Jameson, 1974) and $\sigma(X^*, X)$-metrizable (Jameson, 1974, 27.8). It follows that nB_{X^*} is $\sigma(X^*, X)$-separable, since a compact metric space is separable (Jameson, 1974, 10.3). Since every subset of a separable metric space is separable, $(nB_{X^*}) \cap C$ is $\sigma(X^*, X)$-separable (Jameson, 1974, 6.15), and therefore $C = \bigcup_{n=1}^{\infty} ((nB_{X^*}) \cap C)$ is $\sigma(X^*, X)$-separable. □

We recall the following definition from Jameson (1974). We say that a set $C \subset X$ is (τ-) CS-closed if for any $\lambda_n \geq 0$ with $\sum_{n=1}^{\infty} \lambda_n = 1$ and any $x_n \in C$, $n = 1, 2, \ldots$, for which $\sum_{n=1}^{N} \lambda_n x_n \xrightarrow{\tau}$ some $\hat{x}$, $\hat{x} \in C$. Clearly every CS-closed set is convex. In a Banach space, all convex sets which are closed, open, finite-dimensional or G_δ's are CS-closed (see Borwein, 1981).

Lemma 2.18. *Suppose (X, τ) is a topological vector space with either*

(a) *(X, τ) separable and complete metrizable, or*

(b) *$X = Y^*$ with Y a separable normed space, and $\tau = \sigma(Y^*, Y)$.*

Suppose that $C \subset X$ is τ-CS-closed. Then there exists a τ-dense sequence $(x_n)_1^{\infty} \subset C$ and $\alpha_n > 0$, $n = 1, 2, \ldots$, with $\sum_{n=1}^{\infty} \alpha_n = 1$ such that $\sum_{n=1}^{N} \alpha_n x_n \xrightarrow{\tau}$ some $x \in C$ as $N \to \infty$.

Proof. (a) C is separable, since a subset of a separable metric space is separable (Jameson, 1974, 6.15). Pick $(x_n)_1^\infty$ dense in C. Let ρ be a pseudo-norm on X such that the metric $(u, v) \mapsto \rho(u-v)$ generates the topology τ (Schaefer, 1971, I.6.1). Then ρ is clearly continuous, so for each n we can pick $0<\beta_n<2^{-n}$ such that $\rho(\beta_n x_n)<2^{-n}$. The sequence $(\sum_{m=1}^n \beta_m x_m)_1^\infty$ is Cauchy, so by completeness $\sum_{m=1}^n \beta_m x_m \xrightarrow{\tau}$ some $\bar{x}$ as $n\to\infty$. Now put $\alpha_m = (\sum_{n=1}^\infty \beta_n)^{-1}\beta_m$, so $\alpha_m>0$ for each m, $\sum_{m=1}^\infty \alpha_m = 1$ and $\sum_{m=1}^n \alpha_m x_m \xrightarrow{\tau}$ some $\hat{x} = (\sum_{n=1}^\infty \beta_n)^{-1}\bar{x}$. By CS-closure, $\hat{x}\in C$.

(b) By Lemma 2.17, C is $\sigma(Y^*, Y)$-separable, Pick $(x_n)_1^\infty$, $\sigma(Y^*, Y)$-dense in C. Choose $0<\beta_n<2^{-n}$ such that $\beta_n\|x_n\|_*<2^{-n}$ for each n.

Then $(\sum_{m=1}^n \beta_n x_n)_1^\infty$ is Cauchy and $(Y^*, \|\cdot\|_*)$ is complete, so $\sum_{m=1}^n \beta_m x_m \xrightarrow{\|\cdot\|_*}$ some $\bar{x}$ as $n\to\infty$. Now put $\alpha_m = (\sum_{n=1}^\infty \beta_n)^{-1}\beta_m$, so $\alpha_m>0$ for each m, $\sum_{m=1}^\infty \alpha_m = 1$, and $\sum_{m=1}^n \alpha_m x_m \xrightarrow{\|\cdot\|_*}$ some $\hat{x} = (\sum_{n=1}^\infty \beta_n)^{-1}\bar{x}$, as $n\to\infty$. It follows that $\sum_{m=1}^n \alpha_m x_m \xrightarrow{\sigma(Y^*,Y)} \hat{x}$, as $n\to\infty$, and by CS-closure $\hat{x}\in C$. □

Theorem 2.19. *Suppose (X, τ) is a topological vector space with either*

(a) *(X, τ) a separable Fréchet space, or*

(b) *$X = Y^*$ with Y a separable normed space, and $\tau = \sigma(Y^*, Y)$.*

Suppose that $C\subset X$ is CS-closed. Then τ-qri $C\neq\emptyset$.

Proof. In either case (X, τ) is locally convex, and by Lemma 2.18 there exists dense $(x_n)_1^\infty$ in C and $\alpha_n>0$, $n=1, 2, \ldots$, with $\sum_{n=1}^\infty \alpha_n = 1$ such that $\sum_{n=1}^N \alpha_n x_n \to \hat{x}\in C$ as $N\to\infty$. We claim $\hat{x}\in \operatorname{qri} C$.

Suppose $\phi\in N_C(\hat{x})$, so $\phi(x-\hat{x})\leq 0$, for all $x\in C$. In particular, $\phi(x_n)\leq\phi(\hat{x})$ for each n. Suppose that $\phi(x_m)<\phi(\hat{x})$, for some m. Then we would have

$$
\begin{aligned}
\phi(\hat{x}) &= \phi\left(\lim_{N\to\infty}\sum_{n=1}^N \alpha_n x_n\right) \\
&= \lim_{N\to\infty} \phi \sum_{n=1}^N \alpha_n x_n \quad \text{(since } \phi \text{ continuous)} \\
&= \lim_{N\to\infty}\left(\alpha_m\phi(x_m) + \sum_{\substack{n=1\\ n\neq m}}^N \alpha_n\phi(x_n)\right) \\
&< \lim_{N\to\infty}\left(\alpha_m\phi(\hat{x}) + \sum_{\substack{n=1\\ n\neq m}}^N \alpha_n\phi(\hat{x})\right) \quad \text{(since } \alpha_m>0) \\
&= \phi(\hat{x}) \lim_{N\to\infty}\sum_{n=1}^N \alpha_n \\
&= \phi(\hat{x}),
\end{aligned}
$$

which is a contradiction. It follows that $\phi(x_n) = (\hat{x})$ for each n, and since $(x_n)_1^\infty$ is dense in C and ϕ is continuous, $\phi(x)=\phi(\hat{x})$, for all $x\in C$. Therefore $-\phi\in N_C(\hat{x})$, so $N_C(\hat{x})$ is a subspace. Proposition 2.8 now gives the result. □

We conclude this section with an example which illustrates the computation of a quasi relative interior, and which is useful for constructing counterexamples.

Example 2.20. *Define a convex set* $C \subset \ell_2(\mathbb{N})$ *by*

$$C := \{x \in \ell_2 \mid \|x\|_1 \leq 1\}.$$

Then

$$\operatorname{qri} C = C \backslash \{x \in \ell_2 \mid \|x\|_1 = 1,\ x_n = 0\ \forall n > \textit{some } N\}.$$

Proof. Suppose $\hat{x} \in C$. For any $y \in \ell_2$, $y \in N_C(\hat{x})$ if and only if $\langle x - \hat{x}, y\rangle \leq 0$ for all $x \in C$, or

$$\sup\left\{ \sum_{n=1}^{\infty} x_n y_n \,\middle|\, \sum_n |x_n| \leq 1,\ x \in \ell_2 \right\} = \sum_n \hat{x}_n y_n.$$

This is equivalent to $\|y\|_\infty = \langle \hat{x}, y\rangle$, and in this case

$$\|y\|_\infty = \sum_n \hat{x}_n y_n \leq \sum_n |\hat{x}_n|\,|y_n| \leq \|y\|_\infty \sum_n |\hat{x}_n| = \|y\|_\infty \|\hat{x}\|_1 \leq \|y\|_\infty, \tag{2.1}$$

implying equality throughout.

Now suppose $\hat{x} \notin \operatorname{qri} C$, so certainly there exists $0 \neq y \in N_C(\hat{x})$. It follows from (2.1) that $\|\hat{x}\|_1 = 1$ and that $|y_n| = \|y\|_\infty > 0$ whenever $\hat{x}_n \neq 0$. Since $y \in \ell_2$ we must have $\hat{x}_n = 0$ for all n sufficiently large.

Conversely, suppose $\hat{x} \in \ell_2$ with $\|\hat{x}\|_1 = 1$ and $\hat{x}_n = 0$ for all $n > N$. Define $y \in \ell_2$ by $y_n = \operatorname{sign} \hat{x}_n$, where

$$\operatorname{sign} t = \begin{cases} 1, & t > 0, \\ 0, & t = 0, \\ -1, & t < 0. \end{cases}$$

Then $\|y\|_\infty = 1$ and $\langle \hat{x}, y\rangle = \sum_n \hat{x}_n y_n = \sum_n |x_n| = \|x\|_1 = 1$, so by (2.1), $y \in N_C(\hat{x})$.

However, $\langle 0 - \hat{x}, y\rangle = -1 < 0$, so $-y \notin N_C(\hat{x})$. Thus by Proposition 2.8, $\hat{x} \notin \operatorname{qri} C$. □

Now consider the following example. Let $A : \ell_1 \to \ell_2$ be the (continuous) injection (notice $\|x\|_2 \leq \|x\|_1$ always). Let B be the unit ball in ℓ_1. By Corollary 2.14, $\operatorname{qri} B = \operatorname{int} B$, so $A(\operatorname{qri} B) = \{x \in \ell_2 \mid \|x\|_1 < 1\}$. But as the above example shows,

$$\operatorname{qri}(AB) = \{x \in \ell_2 \mid \|x\|_1 < 1\} \cup \{x \in \ell_2 \mid \|x\|_1 = 1 \text{ and } x_n \neq 0 \text{ infinitely often}\}.$$

In particular $(1/2^n)_1^\infty \in \operatorname{qri}(AB) \backslash A(\operatorname{qri} B)$, so $A(\operatorname{qri} B) \neq \operatorname{qri}(AB)$. Thus Proposition 2.10 fails to extend to the case where A has infinite-dimensional range.

On the other hand, Proposition 2.7 will extend, providing the spaces involved are locally convex.

Proposition 2.21. *Let X and Y be locally convex with $C \subset X$ convex and $A: X \to Y$ continuous and linear. Then $A(\mathrm{qri}\ C) \subset \mathrm{qri}(AC)$.*

Proof. Suppose $\hat{x} \in \mathrm{qri}\ C$, so by Proposition 2.8, $N_C(\hat{x})$ is a subspace. Suppose $\phi \in N_{AC}(A\hat{x})$, so $\phi(Ax - A\hat{x}) \leq 0$, for all $x \in C$, or $A^{\mathrm{T}}\phi \in N_C(\hat{x})$. Then $-A^{\mathrm{T}}\phi \in N_C(\hat{x})$ also, or $-\phi \in N_{AC}(A\hat{x})$. Thus $N_{AC}(A\hat{x})$ is a subspace, so by Proposition 2.8, $A\hat{x} \in \mathrm{qri}(AC)$. □

Example 2.20 also generates the following counterexample to the conjecture that $\mathrm{qri}\ C_1 \cap \mathrm{qri}\ C_2 \subset \mathrm{qri}(C_1 \cap C_2)$ (cf. Theorem 2.13). Let $X = \ell_2(\mathbb{N})$, and define $\bar{x} := (1/2^n)_1^\infty \in \ell_2$. Set $C_1 := \{x \mid \|x\|_1 \leq 1\}$ and $C_2 := \{\alpha\bar{x} \mid \alpha \in \mathbb{R}\}$. Then from Example 2.20 it is easy to see that $\mathrm{qri}\ C_1 \cap \mathrm{qri}\ C_2 = \{\alpha\bar{x} \mid -1 \leq \alpha \leq 1\}$, whereas $\mathrm{qri}(C_1 \cap C_2) = \{\alpha\bar{x} \mid -1 < \alpha < 1\}$. Thus in particular, $\bar{x} \in (\mathrm{qri}\ C_1 \cap \mathrm{qri}\ C_2) \backslash \mathrm{qri}(C_1 \cap C_2)$.

3. Examples of quasi relative interiors in vector lattices

We now turn to the computation of the quasi relative interior of some specific sets which will play a large role in what follows. We begin with the case where C is a cone. We will be particularly interested in the case where C is the positive cone in a normed vector lattice. We begin therefore with some lattice-theoretic ideas (see Schaefer, 1974 and Peressini, 1967).

We shall use the following definitions from Peressini (1967). We shall write the lattice operations as $x \vee y$ and $x \wedge y$ for the supremum and infimum of x and y respectively. As usual, $x^+ = x \vee 0$, $x^- = (-x)^+$, and $|x| = x \vee (-x)$. Let X be a vector lattice. A subspace I of X is an *ideal* if $y \in I$ whenever $x \in I$ and $|y| \leq |x|$. If X is the direct sum of two subspaces, I and J ($X = I \oplus J$) then this sum is *order direct* if $x \geq 0$, $x = x_1 + x_2$ with $x_1 \in I$, $x_2 \in J$ implies x_1, $x_2 \geq 0$; this is always the case if I and J are ideals (Schaefer, 1974, II.2.7). X is *Archimedean* if $nx \leq y$ for all $n \in \mathbb{N}$ implies $x \leq 0$, and (stronger) X is (*countably*) *order complete* if the supremum of every (countable) majorized subset of X exists in X. A subset B of X is *order bounded* if B is contained in some *order interval* $[x_1, x_2] = \{x \in X \mid x_1 \leq x \leq x_2\}$, and a linear functional $\phi: X \to \mathbb{R}$ is *order bounded* if it maps order bounded sets in X to order bounded sets in $\mathbb{R}$. The space of all order bounded linear functionals on X, ordered by the cone of all positive linear functionals, is called the *order dual* of X (X^{b}).

A net $\{x_\alpha \mid \alpha \in K\}$ in X *order converges* to $x_0 \in X$ if it is order bounded and there exists a net $\{y_\alpha \mid \alpha \in K\}$ in X with $\inf y_\alpha = 0$ and $|x_\alpha - x_0| \leq y_\alpha \leq y_\beta$ for all $\alpha \leq \beta$. A linear functional $\phi: X \to \mathbb{R}$ is *order continuous* if the net $\{\phi(x_\alpha) \mid \alpha \in K\}$ order converges to 0 in $\mathbb{R}$ whenever $\{x_\alpha \mid \alpha \in K\}$ order converges to 0 in X. If in this definition we replace 'net' by 'sequence' then we call ϕ *sequentially order continuous*. The speces of all order continuous and sequentially order continuous linear functionals (ordered by the cone of positive linear functionals) we denote by X_{00}^{b} and X_0^{b} respectively. Clearly $X_{00}^{\mathrm{b}} \subset X_0^{\mathrm{b}}$. If X is Archimedean then $X_0^{\mathrm{b}} \subset X^{\mathrm{b}}$ (Peressini, 1967, I5.15), and in this case X_0^{b} and X_{00}^{b} are ideals in X^{b} (Wong and Ng, 1973, 10.17).

Similar definitions of order convergence and order continuity may be found in Schaefer (1974), in terms of filters rather than nets. It should however be noted that Schaefer's definitions may not be equivalent to ours if X is not order complete (see Schaefer, 1974, p. 141). The definition of order continuity in Wong and Ng (1973) again differs slightly from ours, but coincides at least when X is Archimedean.

Finally, a vector lattice X equipped with a norm $\|\cdot\|$ is called a *normed lattice* if $|x| \leqslant |y|$ implies $\|x\| \leqslant \|y\|$; if $(X, \|\cdot\|)$ is complete then X is a *Banach lattice*. Any normed lattice is Archimedean. We denote the topological dual of $(X, \|\cdot\|)$ by X^*. X_+ will always denote the positive cone in X.

Proposition 3.1. *Let X be a vector lattice. Then X^{b} is an order complete vector lattice, with lattice operations given by (for $x \geqslant 0$)*

$$(\phi_1 \vee \phi_2)(x) = \sup\{\phi_1(x_1) + \phi_2(x_2) \mid x_1 + x_2 = x,\ x_1, x_2 \geqslant 0\},$$

$$(\phi_1 \wedge \phi_2)(x) = \inf\{\phi_1(x_1) + \phi_2(x_2) \mid x_1 + x_2 = x,\ x_1, x_2 \geqslant 0\}.$$

Proof. Schaefer (1974, II.4.2). □

Proposition 3.2. *Let X be a normed lattice. Then X^* is an ideal in X^{b}, and an order complete Banach lattice under its dual norm and ordering. Furthermore, if X is actually a Banach lattice then $X^* = X^{\mathrm{b}}$.*

Proof. Schaefer (1974, II.5.5). □

If X is a vector lattice and M is a subspace of X, ordered by the cone $X_+ \cap M$, then let us denote the lattice operations in M by $x_1 \wedge_M x_2$ and $x_1 \vee_M x_2$ (if they exist), for $x_1, x_2 \in M$. Recall that M is a *sublattice* of X if $x_1 \wedge x_2, x_1 \vee x_2 \in M$ for all $x_1, x_2 \in M$.

Lemma 3.3. *If X is a vector lattice and $M \subset X$ a sublattice, then $x_1 \wedge x_2 = x_1 \wedge_M x_2$ and $x_1 \vee x_2 = x_1 \vee_M x_2$, for all $x_1, x_2 \in M$.*

Proof. By definition, $x_1 \wedge x_2 \in M$, and $x_1 \wedge x_2 \leqslant x_1, x_2$ so certainly $x_1 \wedge x_2 \leqslant_M x_1, x_2$. Suppose $x \in M$ and $x \leqslant_M x_1, x_2$. Then $x \leqslant x_1, x_2$, so $x \leqslant x_1 \wedge x_2$ and therefore $x \leqslant_M x_1 \wedge x_2$. It follows that $x_1 \wedge_M x_2$ exists and equals $x_1 \wedge x_2$. Similarly for $x_1 \vee x_2$. □

Proposition 3.4. *Let X be a vector lattice and M an ideal of X. Then M is a sublattice of X.*

Proof. Easy to verify (Schaefer, 1974, p. 56). □

Theorem 3.5. *Under evaluation, every normed lattice X is lattice isomorphic to a normed sublattice of its bidual X^{**}.*

Proof. Schaefer (1974, II.5.5, Corollary 2). □

Notice in particular that this implies $\widehat{x^+} = (\hat{x})^+$, where $\hat{\ }: X \to X^{**}$ denotes the evaluation map.

Recall that if (X, Y) is a dual pair of vector spaces, and $S \subset X$ is a convex cone, then the dual cone $S^+ \subset Y$ is given by

$$S^+ := \{y \in Y \mid \langle x, y\rangle \geqslant 0 \text{ for all } x \in S\}.$$

Definition 3.6. Suppose (X, Y) is a dual pair of vector spaces, with X and Y partially ordered by convex cones S_X and S_Y respectively. We call $((X, S_X), (Y, S_Y))$ a *dual lattice pair* if $S_Y = S_X^+$, (X, S_X) is a vector lattice, and (Y, S_Y) is a sublattice of the order dual of (X, S_X), X^{b}.

Examples 3.7.
(i) *X a normed lattice, $Y = X^*$, with the lattice cones.*
(ii) *Y a normed lattice, $X = Y^*$, with the lattice cones.*
In both cases (X, Y) is a dual lattice pair.

Proof. (i) Propositions 3.2 and 3.4.
(ii) Regarded as a subspace of $X^* = Y^{**}$, Y is a sublattice of X^* by Theorem 3.5, and X^* is a sublattice of X^{b} by Propositions 3.2 and 3.4. The result follows. □

If (X, Y) is a dual lattice pair then the lattice operations in Y are given by the formulae of Proposition 3.1, by Lemma 3.3.

We shall now show that, in a normed lattice, the quasi relative interior of the positive cone is identical with the quasi-interior defined in Schaefer (1974). In a partially ordered vector space (X, S) (where S is the positive cone) we denote the order interval $\{x \mid x_1 \leqslant x \leqslant x_2\}$ by $[x_1, x_2]$. The following definition may be found in Peressini (1967).

Definition 3.8. Let (X, S) be a partially ordered topological vector space. Then $e \in S$ is a *quasi-interior point* of S, $e \in \operatorname{qi} S$, if $\operatorname{cl} \mathbb{P}[-e, e] = X$.

Theorem 3.9. *Let (X, Y) be a dual pair. Consider X with the $\sigma(X, Y)$ topology, and suppose S is a closed convex cone partially ordering X, Consider the following properties for a point $e \in S$:*
(i) $\operatorname{cl} \mathbb{P}[0, e] = S$.

(ii) $\mathrm{cl}\,\mathbb{P}[-e, e] = X$ ($e \in \mathrm{qi}\, S$).

(iii) $\mathrm{cl}\,\mathbb{P}(S-e) = X$ ($\Leftrightarrow e \in \mathrm{qri}\, S$, *when* $X = \mathrm{cl}(S-S)$).

Then we have the following:

(a) (i)$\Rightarrow$(ii) *if* $\mathrm{cl}(S-S) = X$.

(b) (ii)$\Rightarrow$(iii).

(c) (iii)$\Rightarrow$(i) *if* $((X, S), (Y, S^+))$ *is a dual lattice pair. In particular, if* X *is a normed lattice then* $\mathrm{qi}\, X_+ = \mathrm{qri}\, X_+$, *and if* $X = Y^*$ *is the dual of a normed lattice* Y, *then* $\sigma(Y^*, Y) - \mathrm{qi}\, X_+ = \sigma(Y^*, Y) - \mathrm{qri}\, X_+$.

Proof.

(a) $X = \mathrm{cl}(S-S) = \mathrm{cl}(\mathrm{cl}\,\mathbb{P}[0, e] - \mathrm{cl}\,\mathbb{P}[0, e]) = \mathrm{cl}(\mathbb{P}[0, e] - \mathbb{P}[0, e]) = \mathrm{cl}\,\mathbb{P}[-e, e]$.

(b) $[-e, e] \subset S - e$ gives the result.

(c) Suppose $e \in \mathrm{qri}\, S$, but that (i) fails: for some $\bar{x} \geq 0$, $\bar{x} \notin \mathrm{cl}\,\mathbb{P}[0, e]$. It follows by separation that for some $y \in Y$, $\langle \bar{x}, y\rangle > \langle x, y\rangle$, for all $x \in \mathrm{cl}\,\mathbb{P}[0, e]$. For this y we must have $\langle x, y\rangle \leq 0$, for all $x \in [0, e]$, so $\langle e, y^+\rangle = \sup\{\langle x, y\rangle \mid x \in [0, e]\} \leq 0$, by Proposition 3.1 and Lemma 3.3. Hence we have $\langle x - e, y^+\rangle \geq 0$, for all $x \geq 0$, whence $\langle x, y^+\rangle \geq 0$, for all $x \in \mathrm{cl}\,\mathbb{P}(S-e) = X$. Thus $y^+ = 0$, or $y \leq 0$. But we know $\bar{x} \geq 0$ and $\langle \bar{x}, y\rangle > 0$, which is a contradiction.

The final statement follows from Examples 3.7, the fact that weak and norm closures of convex sets in a normed space are identical and since in both cases $X = S - S$. □

Further, more technical conditions for the equivalence of these and various other notions may be found in Peressini (1967, p. 184).

The following result makes the quasi relative interior of the positive cone of various common vector lattices particularly easy to compute, cf. Schaefer (1974, II.6.3).

Theorem 3.10. *Let* X *be locally convex, partially ordered by a convex cone* S *with* $\mathrm{cl}(S-S) = X$, *and suppose* X^* *is partially ordered by* S^+. *Then* $e \in \mathrm{qri}\, S$ *if and only if* $\hat{e}$ ($\in X^{**}$) *is strictly positive on* X^*: $\phi(e) > 0$ *for all* $\phi \in S^+ \backslash \{0\}$.

Proof. Suppose first that $e \in \mathrm{qri}\, S$, but that for some nonzero $\phi \in S^+$, $\phi(e) = 0$. Then $-\phi(x-e) \leq 0$, for all $x \in S$, so $-\phi \in N_S(e)$. By Proposition 2.8, $\phi \in N_S(e)$, so $\phi(x - e) \leq 0$, for all $x \in S$, or $\phi(x) \leq 0$, for all $x \in S$. But $\phi \in S^+$, so $\phi(x) = 0$, for all $x \in S$, whence $\phi(x) = 0$, for all $x \in \mathrm{cl}(S-S) = X$. Thus $\phi = 0$, which is a contradiction.

Conversely, if $e \notin \mathrm{qri}\, S$ then by Proposition 2.8, for some $\phi \in X^*$ we have $-\phi(x - e) \leq 0$, for all $x \in S$, with strict inequality for some $\bar{x} \in S$. It follows that $\phi \neq 0$, and that $\phi \geq 0$. Also, if $\phi(e) > 0$ we would have $-\phi(\frac{1}{2}e - e) > 0$, which is a contradiction, so $\phi(e) = 0$. Thus $\hat{e}$ is not strictly positive on X^*. □

We now give as examples the quasi relative interiors of the positive cones in some of the standard normed lattices.

Examples 3.11. *All spaces are real, and unless specified otherwise, we work with respect to the norm topology.*

(i) $X = L^p(T, \mu)$ *with* (T, μ) *a σ-finite measure space and* $1 \leqslant p < \infty$,

$$\mathrm{qri}(X_+) = \{x \mid x(t) > 0 \text{ a.e.}\}.$$

(ii) $X = L^\infty(T, \mu)$ *with* (T, μ) *a σ-finite measure space,*

$$\|\cdot\| - \mathrm{qri}(X_+) = \{x \mid \operatorname{ess\,inf} x > 0\},$$

$$\sigma(L^\infty, L^1) - \mathrm{qri}(X_+) = \{x \mid x(t) > 0 \text{ a.e.}\}.$$

(iii) $X = \ell^2(\mathbb{R})$, *the square summable functions on* $\mathbb{R}$,

$$\mathrm{qri}(X_+) = \emptyset.$$

(iv) $X = C_0(T)$ *with* T *locally compact, the continuous functions vanishing at infinity,*

$$\mathrm{qri}(X_+) = \begin{cases} \{x \mid x(t) > 0, \forall t \in T\}, & \text{if } T \text{ countable at } \infty, \\ \emptyset, & \text{otherwise.} \end{cases}$$

(v) $X = C_b(T)$ *with* T *completely regular, the bounded continuous functions,*

$$\mathrm{qri}(X_+) = \{x \mid \inf x > 0\}.$$

(vi) $X = M(T)$ *with* T *compact Hausdorff, the regular Borel measures,*

$$\|\cdot\| - \mathrm{qri}(X_+) = \emptyset, \quad \text{if } T \text{ is uncountable},$$

$$\sigma(M(T), C(T)) - \mathrm{qri}(X_+) = \{\mu \geqslant 0 \mid \text{support } \mu = T\}.$$

(vii) $X = c_0(\mathbb{N})$, *the null sequences,*

$$\mathrm{qri}(X_+) = \{x \mid x_i > 0 \text{ for each } i\}.$$

(viii) $X = c(\mathbb{N})$, *the convergent sequences,*

$$\mathrm{qri}(X_+) = \left\{x \;\middle|\; \inf_i x_i > 0\right\}.$$

Proof. Many of these examples may be found in Schaefer (1974, p. 98) (applying Theorem 3.9).

(i) $(L^p(T, \mu))^* = L^q(T, \mu)$, where $(1/p) + (1/q) = 1$ (see for instance, Holmes, 1975).

Now apply Theorem 3.10.

(ii) The first result follows from Corollary 2.14. The second part follows from Theorem 3.10, since $(X, \sigma(X, Y))^* = Y$.

(iii) Exactly as (i): $\ell^2(\mathbb{R})$ is reflexive, and it is easy to see there can exist no strictly positive linear functionals.

(iv) See Schaefer (1974).

(v) See Schaefer (1974).

(vi) For the first part, see Schaefer (1974). For the second part, by Theorem 3.10, we need to show that a measure $\mu \geqslant 0$ is strictly positive as functional on $C(T)$ if and only if support $\mu = T$, in other words if and only if $\mu(T_1) > 0$ for all nonempty open $T_1 \subset T$.

Clearly, if support $\mu = T$ and $x \in C(T)_+$ with $x(t) > 0$ for some t, then by continuity $\int_T x \, d\mu > 0$, so μ is strictly positive. On the other hand suppose $\mu(T_1) = 0$ for some nonempty open $T_1 \subset T$. Then by Urysohn's Lemma (Jameson, 1974, 12.2) we can construct a nonnegative, nonzero $x \in C(T)$ which vanishes on T_1^c, and then $\int_T x \, d\mu = 0$, so μ is not strictly positive.

(vii) $(c_0(\mathbb{N}))^* = \ell^1(\mathbb{N})$ (Holmes, 1975), and the result follows by Theorem 3.10.

(viii) Follows from Corollary 2.14. $\square$

Notes. (a) Example (i) includes the case $X = \ell^p(\mathbb{N})$, $1 \leqslant p < \infty$:

$$\operatorname{qri}(X_+) = \{x \mid x_i > 0 \text{ for each } i\}.$$

(b) With regard to (iii), in fact if (T, μ) is not σ-finite then

$$\operatorname{qri}(L^p(T, \mu)_+) = \emptyset, \quad 1 \leqslant p < \infty \quad \text{(Schaefer, 1974).}$$

(c) (iv) and (v) both include the case $X = C(T)$ with T compact Hausdorff.

(d) With regard to (vii) and (viii), notice that $\operatorname{qri}((c_0)_+) \neq \operatorname{qri}(c_+) \cap c_0$.

We now turn to the second important special case which we wish to consider.

Theorem 3.12. *Let (X, Y) be a dual pair, with S a convex cone partially ordering X and $e \in S$. Define a set $F \subset X^m$ by*

$$F = \left\{ (x_1, \ldots, x_m) \,\middle|\, \sum_{i=1}^{m} x_i = e, x_i \geqslant 0, i = 1, \ldots, m \right\}. \tag{3.1}$$

In the $\sigma(X, Y)$ topology, if, for some $(\bar{x}_1, \ldots, \bar{x}_m) \in F$,

$$\operatorname{cl} \mathbb{P}[0, \bar{x}_i] = \operatorname{cl} \mathbb{P}[0, e] \quad \textit{for each } i = 1, \ldots, m, \tag{3.2}$$

then $(\bar{x}_1, \ldots, \bar{x}_m) \in \operatorname{qri} F$. If $((X, S), (Y, S^+))$ is a dual lattice pair then the converse is also true.

Proof. Suppose (3.2) holds and $(y_1, \ldots, y_m) \in N_F(\bar{x}_1, \ldots, \bar{x}_m)$, so $\sum_{i=1}^m \langle x_i - \bar{x}_i, y \rangle \leqslant 0$, for all $(x_1, \ldots, x_m) \in F$. For any $1 \leqslant j, k \leqslant m$ with $j \neq k$, setting $x_i = \bar{x}_i$ for $i \neq j, k$ gives that for all $x_j, x_k \geqslant 0$ with $x_j + x_k = \bar{x}_j + \bar{x}_k$, $\langle x_j - \bar{x}_j, y_j \rangle + \langle x_k - \bar{x}_k, y_k \rangle \leqslant 0$. We can rewrite this as $\langle x_j - \bar{x}_j, y_j - y_k \rangle \leqslant 0$, for all $x_j \in [0, \bar{x}_j + \bar{x}_k]$, or $\langle x, y_j - y_k \rangle \leqslant 0$, for all $x \in [-\bar{x}_j, \bar{x}_k]$. Thus $\langle x, y_j - y_k \rangle \leqslant 0$, for all $x \in [0, \bar{x}_k]$ and $\langle x, y_j - y_k \rangle \geqslant 0$, for all $x \in [0, \bar{x}_j]$, and by (3.2) both these inequalities must also hold for all $x \in [0, e]$. This gives $\langle x, y_j\rangle = \langle x, y_k\rangle$, for all $x \in [0, e]$. Since j and k were arbitrary it follows that $\sum_{i=1}^m \langle x_i - \bar{x}_i, y_i\rangle = \langle \sum_{i=1}^m (x_i - \bar{x}_i), y_1\rangle = 0$, for all $(x_1, \ldots, x_m) \in F$, so $-(y_1, \ldots, y_m) \in N_F(\bar{x}_1, \ldots, \bar{x}_m)$ also. Thus $(\bar{x}_1, \ldots, \bar{x}_m) \in \operatorname{qri} F$, by Proposition 2.8.

Conversely, suppose $((X, S), (Y, S^+))$ is a dual lattice pair and that (3.2) fails: for some i, $\operatorname{cl} \mathbb{P}[0, \bar{x}_i] \subset \operatorname{cl} \mathbb{P}[0, e]$ strictly. It follows from the bipolar theorem that $[0, e]^\circ \subset [0, \bar{x}_i]^\circ$ strictly, so for some $y \in Y$, $\langle x, y\rangle \leq 0$ for all $x \in [0, \bar{x}_i]$, but $\langle x_0, y\rangle > 0$, for some $x_0 \in [0, e]$. Thus $\langle \bar{x}_i, y^+\rangle \leq 0$ and $\langle e, y^+\rangle > 0$.

Now set $\bar{y}_j = 0$ for $j \neq i$, $\bar{y}_i = -y^+$. Then $\sum_{j=1}^m \langle x_j - \bar{x}_j, \bar{y}_j\rangle = \langle x_i - \bar{x}_i, -y^+\rangle \leq 0$, for all $x_i \geq 0$, so $(\bar{y}_1, \ldots, \bar{y}_m) \in N_F(\bar{x}_1, \ldots, \bar{x}_m)$. However, setting $x_j = 0$, for $j \neq i$ and $x_i = e$, $\sum_{j=1}^m \langle x_j - \bar{x}_j, -\bar{y}_j\rangle = \langle e - \bar{x}_i, y^+\rangle > 0$, so $-(\bar{y}_1, \ldots, \bar{y}_m) \notin N_F(\bar{x}_1, \ldots, \bar{x}_m)$. Thus by Proposition 2.8, $(\bar{x}_1, \ldots, \bar{x}_m) \notin \operatorname{qri} F$. □

The last two results identify examples where points in the quasi relative interior of the set F of (3.1) can be easy to recognize.

Corollary 3.13. *Let F be as in Theorem 3.12. Then for any $\lambda_1, \ldots, \lambda_m > 0$ with $\sum_{i=1}^m \lambda_i = 1$, $(\lambda_1 e, \ldots, \lambda_m e) \in \operatorname{qri} F$.*

Proof. Clearly $\operatorname{cl} \mathbb{P}[0, \lambda_i e] = \operatorname{cl} \mathbb{P}[0, e]$, for each i, so the result follows by Theorem 3.12. □

Corollary 3.14. *Let F be as in Theorem 3.12 and suppose $((X, S), (Y, S^+))$ is a dual lattice pair. Suppose further that $e \in \operatorname{qri} S$. Then $(\bar{x}_1, \ldots, \bar{x}_m) \in F$ lies in $\operatorname{qri} F$ if and only if $\bar{x}_i \in \operatorname{qri} S$, for each i.*

Proof. By Theorem 3.9, $\operatorname{cl} \mathbb{P}[0, x] = S$ if and only if $x \in \operatorname{qri} S$. The result now follows by Theorem 3.12. □

4. Fenchel duality

We now come to the principal result of this paper, a Fenchel duality theorem from which all our duality results will be derived. We begin with a simple weak duality result, followed by the main theorem. The significance of quasi relative interiors in the constraint qualification will then be made apparent. We shall also derive a subgradient formula and a minimax theorem from the main result.

Throughout this section we shall adopt the following notation. A convex function $f: X \to]-\infty, \infty]$ is proper if it is not identically $+\infty$.

$$\begin{aligned}
&X \text{ is a locally convex topological vector space,}\\
&g: X \to]-\infty, \infty], \text{ convex, proper,}\\
&h: \mathbb{R}^n \to]-\infty, \infty], \text{ convex, proper,}\\
&A: X \to \mathbb{R}^n \text{ continuous, linear.}
\end{aligned} \tag{4.1}$$

The following simple weak duality result is well-known (see for example Rockafellar, 1967). We include its proof for completeness.

Proposition 4.1. *With the notation of* (4.1),

$$\inf\{g(x)+h(Ax) \mid x \in X\} \geq \sup\{-g^*(A^{\mathrm{T}}\lambda)-h^*(-\lambda) \mid \lambda \in \mathbb{R}^n\}.$$

Proof. For all $x \in X$, $\lambda \in \mathbb{R}^n$ we have, from the definition of the Fenchel conjugate,

$$g(x)+g^*(A^{\mathrm{T}}\lambda)+h(Ax)+h^*(-\lambda) \geq (A^{\mathrm{T}}\lambda)(x)-\lambda^{\mathrm{T}}(Ax)=0,$$

so $g(x)+h(Ax) \geq -g^*(A^{\mathrm{T}}\lambda)-h^*(-\lambda)$. Taking the infimum over $x \in X$ and the supremum over $\lambda \in \mathbb{R}^n$ gives the result. $\square$

The next result is the fundamental Fenchel duality theorem from which all our duality results will be derived. The proof follows directly from the finite-dimensional Fenchel result in Rockafellar (1970), and produces a duality theorem requiring a constraint qualification which does not appear to be in a useful form. However, as will be made clear by the corollary following the next theorem, using our results on quasi relative interiors we can express the constraint qualification in a much more practical way.

Theorem 4.2. *With the notation of* (4.1), *suppose*
either (i) $\mathrm{ri}(A \operatorname{dom} g) \cap \mathrm{ri}(\operatorname{dom} h) \neq \emptyset$
or (ii) $\mathrm{ri}(A \operatorname{dom} g) \cap \operatorname{dom} h \neq \emptyset$, *and h polyhedral.*
Then

$$\inf\{g(x)+h(Ax) \mid x \in X\} = \max\{-g^*(A^{\mathrm{T}}\lambda)-h^*(-\lambda) \mid \lambda \in \mathbb{R}^n\},$$

appropriately read when $-\infty$.

Proof. Let

$$\begin{aligned}\mu &= \inf\{g(x)+h(Ax) \mid x \in X\}\\ &= \inf\{g(x)+h(y) \mid y=Ax,\ x \in X,\ y \in \mathbb{R}^n\}\\ &= \inf\{h(y)+p(y) \mid y \in \mathbb{R}^n\}\end{aligned}$$

where we define $p(y) := \inf\{g(x) \mid Ax=y,\ x \in X\}$.

Now $p : \mathbb{R}^n \to [-\infty, \infty]$ is convex. To see this, suppose $0<\gamma<1$, $p(y_1)<\alpha$ and $p(y_2)<\beta$. Then for some $x_1, x_2 \in X$ with $Ax_1=y_1$, $Ax_2=y_2$, we have $g(x_1)<\alpha$ and $g(x_2)<\beta$. But then

$$\begin{aligned}(1-\gamma)\alpha+\gamma\beta &> (1-\gamma)g(x_1)+\gamma g(x_2)\\ &\geq g((1-\gamma)x_1+\gamma x_2)\\ &\geq \inf\{g(x) \mid Ax=(1-\gamma)y_1+\gamma y_2\}\\ &= p((1-\gamma)y_1+\gamma y_2).\end{aligned}$$

It follows that p is convex (Rockafellar, 1970, 4.2).

We also have $\operatorname{dom} p = A \operatorname{dom} g$. Furthermore

$$\begin{aligned} p^*(\lambda) &= \sup\{\lambda^{\mathrm{T}} y - p(y) \mid y \in \mathbb{R}^n\} \\ &= \sup\{\lambda^{\mathrm{T}} y - g(x) \mid Ax = y,\, y \in \mathbb{R}^n,\, x \in X\} \\ &= \sup\{\lambda^{\mathrm{T}}(Ax) - g(x) \mid x \in X\} \\ &= g^*(A^{\mathrm{T}}\lambda), \end{aligned}$$

for all $\lambda \in \mathbb{R}^n$.

Now by assumption there is an $\hat{x} \in \operatorname{dom} g$ with $A\hat{x} \in \operatorname{ri}(A \operatorname{dom} g) \cap \operatorname{dom} h = \operatorname{ri}(\operatorname{dom} p) \cap \operatorname{dom} h$, so $\mu < \infty$. By Proposition 4.1 we can without loss of generality assume $\mu > -\infty$.

Suppose $p(A\hat{x}) = -\infty$. Then there would exist a sequence $(x_i)_1^\infty \subset X$ with $Ax_i = A\hat{x}$, for each i, and $g(x_i) \to -\infty$ as $i \to \infty$. But then we would have $g(x_i) + h(Ax_i) = g(x_i) + h(A\hat{x}) \to -\infty$, as $i \to \infty$ (since $A\hat{x} \in \operatorname{dom} h$), and this contradicts $\mu > -\infty$. Thus $p(A\hat{x})$ is finite with $A\hat{x} \in \operatorname{ri}(\operatorname{dom} p)$, and by Rockafellar (1970, 7.2), this is sufficient to ensure p is proper.

Now the assumptions say

either (a) $\operatorname{ri}(\operatorname{dom} p) \cap \operatorname{ri}(\operatorname{dom} h) \neq \emptyset$,

or (b) $\operatorname{ri}(\operatorname{dom} p) \cap \operatorname{dom} h \neq \emptyset$, and h polyhedral.

We can therefore apply the classical Fenchel duality theorem (Rockafellar, 1970, 31.1) to obtain

$$\begin{aligned} \mu &= \inf\{g(x) + h(Ax) \mid x \in X\} \\ &= \inf\{p(y) + h(y) \mid y \in \mathbb{R}^n\} \\ &= \max\{-p^*(\lambda) - h^*(-\lambda) \mid \lambda \in \mathbb{R}^n\} \\ &= \max\{-g^*(A^{\mathrm{T}}\lambda) - h^*(-\lambda) \mid \lambda \in \mathbb{R}^n\}. \qquad \square \end{aligned}$$

This result may essentially be found in Rockafellar (1967), using perturbational proof techniques (cf. Borwein and Lewis, 1991).

Comment. Suppose in fact that, in the notation of the above theorem, we have that there exists a $\hat{y} \in \operatorname{int}(A \operatorname{dom} g) \cap \operatorname{dom} h = \operatorname{int}(\operatorname{dom} p) \cap \operatorname{dom} h$. This holds in particular if A is onto and $A(\operatorname{int}(\operatorname{dom} g)) \cap \operatorname{dom} h \neq \emptyset$. Define $\hat{h}$ by $\hat{h}(y) = h(-y)$, so $h^*(-\lambda) = \hat{h}^*(\lambda)$. The dual objective function is given by $k = p^* + \hat{h}^* = (p \square \hat{h})^*$ (Rockafellar, 1970, 16.4) ('$\square$' denotes *infimal convolution*), so

$$\begin{aligned} 0 &\in \operatorname{int}((\operatorname{dom} p) - \hat{y}) \\ &\subset \operatorname{int}(\operatorname{dom} p - \operatorname{dom} h) = \operatorname{int}(\operatorname{dom} p + \operatorname{dom} \hat{h}) = \operatorname{int}(\operatorname{dom}(p \,\square\, \hat{h})) \\ &\subset \operatorname{int}(\operatorname{dom}(p \,\square\, \hat{h})^{**}) = \operatorname{int}(\operatorname{dom} k^*). \end{aligned}$$

It then follows by Rockafellar (1970, 14.2.2), that k has bounded level sets. Thus any descent method will stay bounded. Recession arguments such as this can be used to prove many of our duality results.

When $\operatorname{qri}(\operatorname{dom} g) \neq \emptyset$ we know from Proposition 2.10 that $\operatorname{ri}(A \operatorname{dom} g) = A(\operatorname{qri}(\operatorname{dom} g))$, so we obtain the following result.

Corollary 4.3. *With the notation of* (4.1), *suppose*

$$\text{either} \quad A(\operatorname{qri}(\operatorname{dom} g)) \cap \operatorname{ri}(\operatorname{dom} h) \neq \emptyset, \tag{4.2a}$$

$$\text{or} \quad A(\operatorname{qri}(\operatorname{dom} g)) \cap \operatorname{dom} h \neq \emptyset \text{ and } h \text{ polyhedral.} \tag{4.2b}$$

Then

$$\inf\{g(x) + h(Ax) \mid x \in X\} = \max\{-g^*(A^{\mathrm{T}}\lambda) - h^*(-\lambda) \mid \lambda \in \mathbb{R}^n\}.$$

Proof. Theorem 4.2 and Proposition 2.10. □

Expressing the constraint qualification in the form (4.2) makes the significance of the quasi relative interior clear. In order to check condition (i) in Theorem 4.2 we would have to compute the image of $(\operatorname{dom} g)$ under the map A. On the other hand, to check this condition rewritten in the form (4.2a) we simply need to find a single point x_0 (*exactly* as in the Slater condition) with $x_0 \in \operatorname{qri}(\operatorname{dom} g)$, and $Ax_0 \in \operatorname{ri}(\operatorname{dom} h)$. It is this simplification which will be exploited in Part II of this work, when we consider more concrete models.

As usual, we obtain a subgradient formula and a minimax theorem as direct consequences of the Fenchel result.

Corollary 4.4. *With the notation of* (4.1), *for any* $\bar{x} \in X$,

$$\partial(g + hA)(\bar{x}) \supset \partial g(\bar{x}) + A^{\mathrm{T}} \partial h(A\bar{x}),$$

with equality if (4.2a) *or* (4.2b) *holds.*

Proof. Take $\phi \in \partial g(\bar{x})$ and $\mu \in \partial h(A\bar{x})$, so

$$\phi(x - \bar{x}) \leq g(x) - g(\bar{x}), \quad \text{for all } x \in X,$$

and

$$\mu^{\mathrm{T}}(y - A\bar{x}) \leq h(y) - h(A\bar{x}), \quad \text{for all } y \in \mathbb{R}^n.$$

Therefore $(\phi + A^{\mathrm{T}}\mu)(x - \bar{x}) \leq (g + hA)(x) - (g + hA)(\bar{x})$, for all $x \in X$, so $\phi + A^{\mathrm{T}}\mu \in \partial(g + hA)(\bar{x})$.

On the other hand, suppose (4.2a) or (4.2b) holds and $\phi \in \partial(g + hA)(\bar{x})$, so $\phi(x - \bar{x}) \leq (g + hA)(x) - (g + hA)(\bar{x})$, for all $x \in X$. It follows that

$$\begin{aligned} (g - \phi)(\bar{x}) + h(A\bar{x}) &= \inf\{(g - \phi)(x) + h(Ax) \mid x \in X\} \\ &= -(g - \phi)^*(A^{\mathrm{T}}\lambda) - h^*(-\lambda) \end{aligned}$$

for some $\lambda \in \mathbb{R}^n$, by Corollary 4.3. We then have

$$\begin{aligned} 0 &= (g - \phi)(\bar{x}) + (g - \phi)^*(A^{\mathrm{T}}\lambda) + h(A\bar{x}) + h^*(-\lambda) \\ &\geq (A^{\mathrm{T}}\lambda)(\bar{x}) - \lambda^{\mathrm{T}}(A\bar{x}) = 0, \end{aligned}$$

so we have equality. Thus $A^{\mathrm{T}}\lambda \in \partial(g-\phi)(\bar{x}) = \partial g(\bar{x}) - \{\phi\}$, and $-\lambda \in \partial h(A\bar{x})$, so $\phi \in \partial g(\bar{x}) + A^{\mathrm{T}}\partial h(A\bar{x})$. $\square$

We denote the *indicator function* of a set C by $\delta(\cdot \mid C)$. Recall that dom $\delta^*(\cdot \mid C)$ is the *barrier cone* of C, which, for closed, convex C, is the polar of the recession cone of C (Rockafellar, 1970).

Corollary 4.5. *Let X be locally convex, $A : X \to \mathbb{R}^n$ continuous and linear, $C \subset X$ convex and $D \subset \mathbb{R}^n$ nonempty, closed and convex. Suppose*

$$A(\operatorname{qri} C) \cap \operatorname{ri}(\operatorname{dom} \delta^*(\cdot \mid D)) \neq \emptyset.$$

Then

$$\inf_{x \in C} \sup_{y \in D} y^{\mathrm{T}}(Ax) = \max_{y \in D} \inf_{x \in C} y^{\mathrm{T}}(Ax).$$

In particular, this holds if $0 \in \operatorname{qri} C$ and $\mathbb{P}AC = \mathbb{R}^n$.

Proof. Define $g := \delta(\cdot \mid C)$ and $h := \delta^*(\cdot \mid D)$. Then Corollary 4.3 applies and we obtain

$$\inf_x \{\delta(x \mid C) + \delta^*(Ax \mid D)\} = \max_y \{-\delta^*(A^{\mathrm{T}}y \mid C) - \delta^{**}(-y \mid D)\}.$$

Since D is closed, $\delta^{**}(\cdot|D) = \delta(\cdot \mid D)$ (Rockafellar, 1970, 12.2), and the desired result follows.

To see the last part, clearly the result follows if dom $\delta^*(\cdot \mid D) = \{0\}$. If this is not true then since dom $\delta^*(\cdot \mid D)$ is a convex cone in $\mathbb{R}^n$ we can find nonzero y such that $ky \in \operatorname{ri}(\operatorname{dom} \delta^*(\cdot \mid D))$, for all $k > 0$. Since $\mathbb{P}AC = \mathbb{R}^n$, for some $x \in C$, $Ax = \varepsilon y$ for some $\varepsilon > 0$. By Lemma 2.9, $\frac{1}{2}x \in \operatorname{qri} C$, but then $A(\frac{1}{2}x) = (\frac{1}{2}\varepsilon)y \in \operatorname{ri}(\operatorname{dom} \delta^*(\cdot \mid D))$, so the constraint qualification is satisfied and the result follows. $\square$

We will now specialize our Fenchel result to provide duality theorems for the convex models we shall use in the remainder of the paper.

Corollary 4.6. *With the notation of* (4.1), *suppose $C \subset \operatorname{dom} g$ and $D \subset \operatorname{dom} h$ are convex, and that the following constraint qualification holds*:

either (i) $(A \operatorname{qri} C) \cap \operatorname{ri} D \neq \emptyset$,

or (ii) $(A \operatorname{qri} C) \cap D \neq \emptyset$, *with h and D polyhedral.*

Consider the primal problem

$$\begin{array}{ll} \inf & g(x) + h(Ax) \\ \text{subject to} & Ax \in D, \\ & x \in C, \end{array}$$

and the dual problem

$$\begin{array}{ll} \max & -(g + \delta(\cdot \mid C))^*(A^{\mathrm{T}}\lambda) - (h + \delta(\cdot \mid D))^*(-\lambda) \\ \text{subject to} & \lambda \in \mathbb{R}^n. \end{array}$$

Then the primal and dual values are equal (*with dual attainment*).

Proof. In Corollary 4.3, replace g and h by $g+\delta(\cdot|C)$ and $h+\delta(\cdot|D)$ respectively. $\square$

The next result gives a duality theorem for the *convex model* (CM) we shall use most frequently in what follows.

We first need a lemma. Recall that, for a convex cone $Q\subset\mathbb{R}^n$, the dual cone $Q^+\subset\mathbb{R}^n$ is given by

$$Q^+:=\{y\in\mathbb{R}^n \mid y^T\lambda\geq 0 \text{ for all } \lambda\in Q\}.$$

Lemma 4.7. *Suppose $k:\mathbb{R}^n\to]-\infty,\infty]$ is convex, and $Q\subset\mathbb{R}^n$ is a convex cone. If k is differentiable at $\bar\lambda\in Q$ then $\bar\lambda$ is optimal for the problem* $\inf\{k(\lambda)\mid\lambda\in Q\}$ *if and only if*

(i) $\nabla k(\bar\lambda)\in Q^+$, *and*

(ii) $\bar\lambda^T\nabla k(\bar\lambda)=0$.

Proof. By Rockafellar (1970, 27.4), $\bar\lambda$ is optimal if and only if $-\nabla k(\bar\lambda)\in N_Q(\bar\lambda)$. Now $y\in N_Q(\bar\lambda)$ if and only if $y^T(\lambda-\bar\lambda)\leq 0$, for all $\lambda\in Q$. Putting $\lambda:=\frac{1}{2}\bar\lambda, \frac{3}{2}\bar\lambda$, this is equivalent to $y^T\bar\lambda=0$ and $y\in -Q^+$, and the result follows. $\square$

Corollary 4.8. *Let X be locally convex, $f:X\to]-\infty,\infty]$ convex, $C\subset\operatorname{dom} f$ convex, $A:X\to\mathbb{R}^n$ continuous and linear, $b\in\mathbb{R}^n$ and $P\subset\mathbb{R}^n$ a polyhedral cone. Consider the following dual pair of problems:*

$$\begin{array}{lll} \text{(CM)} & \inf & f(x) \\ & \text{subject to} & Ax\in b+P, \\ & & x\in C, \\ \text{(DCM)} & \max & -(f+\delta(\cdot|C))^*(A^T\lambda)+b^T\lambda \\ & \text{subject to} & \lambda\in P^+. \end{array}$$

If the following constraint qualification is satisfied,

(CQ) *there exists an $\hat x\in\operatorname{qri} C$ which is feasible for* (CM),

then the values of (CM) *and* (DCM) *are equal* (*with attainment in* (DCM)).

Suppose further that $f+\delta(\cdot|C)$ is closed. If $\bar\lambda$ is optimal for the dual, and $(f+\delta(\cdot|C))^*$ is differentiable at $A^T\bar\lambda$ with Gateaux derivative $\bar x\in X$, then $\bar x$ is optimal for (CM), *and is furthermore the unique optimal solution.*

Proof. In Corollary 4.6, put $h:=0$, and $D:=b+P$. Then we have,

$$\begin{aligned}(h+\delta(\cdot|D))^*(-\lambda)&=\sup\{-\lambda^Ty-h(y)\mid y\in D\}\\ &=\sup\{-\lambda^Ty\mid y-b\in P\}\\ &=\begin{cases}-\lambda^Tb, & \lambda\in P^+,\\ \infty, & \text{otherwise},\end{cases}\end{aligned}$$

which gives the duality result.

Suppose now that $f+\delta(\cdot|C)$ is closed, $\bar{\lambda}$ is dual optimal, and $\nabla(f+\delta(\cdot|C))^*(A^{\mathrm{T}}\bar{\lambda})=\bar{x}\in X$. By Ekeland and Temam (1976, I.5.3), $\partial(f+\delta(\cdot|C))^*(A^{\mathrm{T}}\bar{\lambda})=\{\bar{x}\}$, so $(f+\delta(\cdot|C))^*(A^{\mathrm{T}}\bar{\lambda})+(f+\delta(\cdot|C))^{**}(\bar{x})=(A^{\mathrm{T}}\bar{\lambda})(\bar{x})$. By Ekeland and Temam (1976, I.4.1), $(f+\delta(\cdot|C))^{**}=f+\delta(\cdot|C)$, and thus

$$(f+\delta(\cdot|C))(\bar{x})+(f+\delta(\cdot|C))^*(A^{\mathrm{T}}\bar{\lambda})=(A^{\mathrm{T}}\bar{\lambda})(\bar{x}),$$

so $\bar{x}\in C$ and $f(\bar{x})=\bar{\lambda}^{\mathrm{T}}(A\bar{x})-(f+\delta(\cdot|C))^*(A^{\mathrm{T}}\bar{\lambda})$. Also, by Lemma 4.7 and the chain rule, $A\bar{x}-b\in(P^+)^+=P$, and $\bar{\lambda}^{\mathrm{T}}(A\bar{x}-b)=0$. Thus $\bar{x}$ is feasible for (CM), with $f(\bar{x})=b^{\mathrm{T}}\bar{\lambda}-(f+\delta(\cdot|C))^*(A^{\mathrm{T}}\bar{\lambda})$, and therefore is optimal.

Any other optimal x must satisfy

$$(f+\delta(\cdot|C))(x)+(f+\delta(\cdot|C))^*(A^{\mathrm{T}}\bar{\lambda})=b^{\mathrm{T}}\bar{\lambda},$$

and $\bar{\lambda}^{\mathrm{T}}(Ax-b)\geq 0$. Thus

$$(f+\delta(\cdot|C))^{**}(x)+(f+\delta(\cdot|C))^*(A^{\mathrm{T}}\bar{\lambda})\leq(A^{\mathrm{T}}\bar{\lambda})(x),$$

so $x\in\partial(f+\delta(\cdot|C))^*(A^{\mathrm{T}}\bar{\lambda})=\{\bar{x}\}$. Therefore $\bar{x}$ is the unique primal optimal solution. $\square$

Notice that, in particular, if f and C are closed then $f+\delta(\cdot|C)$ is closed.

The above result clearly includes the equality constrained case. To conclude, we give a duality theorem for a convex program where we may relax the equality constraints.

Corollary 4.9. *With the notation of Corollary* 4.8, *suppose* $\|\cdot\|$ *is a norm on* $\mathbb{R}^n$ *with dual norm* $\|\cdot\|_*$, *and that* $\varepsilon>0$. *Consider the following dual pair of problems*:

$$\begin{array}{ll} \inf & f(x) \\ \text{subject to} & \|Ax-b\|\leq\varepsilon, \\ & x\in C, \\ \max & -(f+\delta(\cdot|C))^*(A^{\mathrm{T}}\lambda)+b^{\mathrm{T}}\lambda-\varepsilon\|\lambda\|_* \\ \text{subject to} & \lambda\in\mathbb{R}^n. \end{array}$$

If there exists an $\hat{x}\in\operatorname{qri} C$ *with* $\|A\hat{x}-b\|<\varepsilon$ *then the primal and dual values are equal* (*with dual attainment*).

Proof. Denote the unit ball $\{y\in\mathbb{R}^n\,|\,\|y\|\leq 1\}$ by B. Now in Corollary 4.6 put $h:=0$ and $D:=b+\varepsilon B$. Then

$$\begin{aligned}(h+\delta(\cdot|D))^*(-\lambda)&=\sup\{-\lambda^{\mathrm{T}}y\,|\,y-b\in\varepsilon B\}\\&=-\lambda^{\mathrm{T}}b+\sup\{-\lambda^{\mathrm{T}}u\,|\,u\in\varepsilon B\}\\&=-\lambda^{\mathrm{T}}b+\varepsilon\|\lambda\|_*,\end{aligned}$$

which gives the result, since $\operatorname{ri} D=b+\varepsilon(\operatorname{int} B)$. $\square$

If the norm $\|\cdot\|$ is polyhedral (e.g. ℓ^1 or ℓ^∞) then we can relax the constraint qualification to $\|A\hat{x}-b\|\leq\varepsilon$.

As we shall see, the convex model (CM) includes a wide variety of interesting problems. We have already seen how to identify the quasi relative interior of C in some common cases. A number of questions remain. How do we compute the function $(f+\delta(\cdot\,|\,C))^*$? When is the primal value attained, when is a primal optimal solution unique, and when can we compute it from a dual optimal solution? When is a dual optimum unique? How do we solve the dual problem? These questions will be addressed in the following sections.

One point is worth observing concerning separable problems. Suppose we are interested in a problem of the form

$$\begin{array}{ll} \inf & \sum_{i=1}^{m} f_i(x_i) \\ \text{subject to} & \sum_i A_i x_i \in b + P, \\ & x_i \in C_i, \quad i=1,\ldots,m, \end{array}$$

where $X_1,\ldots,X_m$ are each locally convex with convex $C_i \subset \operatorname{dom} f_i$, convex $f_i : X_i \to]-\infty,\infty]$, $A_i : X_i \to \mathbb{R}^n$ continuous and linear, each $i=1,\ldots,m$. Then it is easily checked that the corresponding dual problem is:

$$\begin{array}{ll} \max & -\sum_{i=1}^{m} (f_i+\delta(\cdot\,|\,C_i)^*(A_i^{\mathrm{T}}\lambda)+b^{\mathrm{T}}\lambda \\ \text{subject to} & \lambda \in P^+. \end{array}$$

By Proposition 2.5 the constraint qualification requires the existence of a primal feasible $(\hat{x}_1,\ldots,\hat{x}_n) \in \prod_{i=1}^m \operatorname{qri}(C_i)$.

It is natural to ask whether the duality theorem for the convex model (CM) can be extended to cases where the constraint map A has infinite-dimensional range. In fact the result can fail even when the objective function is linear, A is surjective and C is compact, as the following example shows.

Example 4.10. Let $d \in \ell_2(\mathbb{N})/\ell_1(\mathbb{N})$ (for example, $d_n=(1/n)$, $n=1,2,\ldots$). Consider the problem

$$\begin{array}{ll} \inf & t \\ \text{subject to} & u - td = 0, \\ & \|u\|_1 \leq 1,\ u \in \ell_2, \quad |t| \leq 1,\ t \in \mathbb{R}. \end{array}$$

We have $X=\ell_2\times\mathbb{R}$, $f: X\to\mathbb{R}$ defined by $f(u,t)=t$, $A: X\to\ell_2$ defined by $A(u,t)=u-td$, continuous, linear, and convex $C\subset X$, $C=\{(u,t)\in\ell_2\times\mathbb{R}\,|\,\|u\|_1\leq 1, |t|\leq 1\}$. A is clearly surjective and C is weakly compact (closed and bounded).

By Example 2.20, $(0,0)\in \operatorname{qri} C$ and is feasible, so the constraint qualification is satisfied. In fact it is easy to see that $(0,0)$ is the only feasible point, so the problem has value 0.

Now consider the dual problem. The adjoint map $A^{\mathrm{T}}: \ell_2 \to \ell_2 \times \mathbb{R}$ is given by $A^{\mathrm{T}}v = (v, -\langle d, v\rangle)$.

$$(f + \delta(\cdot \mid C))^*(w, s) = \sup\{\langle u, w\rangle + ts - t \mid \|u\|_1 \leq 1\ u \in \ell_2, |t| \leq 1, t \in \mathbb{R}\}$$
$$= \|w\|_\infty + |s - 1|, \quad \text{for } w \in \ell_2,\ s \in \mathbb{R}.$$

Thus the dual problem is

$$\begin{array}{ll} \max & -\|v\|_\infty - |-\langle d, v\rangle - 1| \\ \text{subject to} & v \in \ell_2. \end{array}$$

Clearly the dual objective function cannot attain the value 0.

5. Differentiation of convex functions in vector lattices

As we shall see in Part II, the dual problems for many of the examples in which we are interested involve certain simple, real-valued convex functions on vector lattices. The derivation of primal optimal solutions from a dual optimum, the existence and uniqueness of primal optima and computational approaches to solving the dual problem will depend on the differentiability of the dual objective function, as we saw in Corollary 4.8. In this section we shall therefore be interested in the differentiability of certain convex functions in a normed lattice.

We shall use the following notation. Suppose that X is a topological vector space and $f: X \to]-\infty, \infty]$ is convex.

$$\partial f(\bar{x}) = \{\phi \in X^* \mid \phi(x - \bar{x}) \leq f(x) - f(\bar{x})\ \forall x \in X\},$$

$$\partial^{\mathrm{a}} f(\bar{x}) = \{\phi \in X' \mid \phi(x - \bar{x}) \leq f(x) - f(\bar{x})\ \forall x \in X\},$$

where X^* and X' are respectively the topological and algebraic duals of X. Thus $\partial^{\mathrm{a}} f(\bar{x})$ is the set of algebraic subgradients while $\partial f(\bar{x})$ is the set of continuous subgradients. In particular, if (X, Y) is a dual pair then in the $\sigma(X, Y)$ topology, $\partial f(\bar{x}) \subset Y$.

If $\bar{x} \in \operatorname{core}(\operatorname{dom} f)$ then f has a (linear) Gateaux derivative $\nabla f(\bar{x})$ at $\bar{x}$ if and only if $\partial^{\mathrm{a}} f(\bar{x}) = \{\nabla f(\bar{x})\} \subset X'$ (see Borwein, 1984, Proposition 3.2). If f is actually continuous at $\bar{x}$ then $\partial^{\mathrm{a}} f(\bar{x}) = \partial f(\bar{x})$ (see Borwein, 1984, Corollary 2.5), so f has continuous Gateaux derivative $\nabla f(\bar{x})$ at $\bar{x}$ if and only if $\partial f(\bar{x}) = \{\nabla f(\bar{x})\}$. See also Holmes (1975, p. 29, 86).

The following result will be useful.

Lemma 5.1. *Let X be a topological vector space, and $f: X \to]-\infty, \infty]$ a positively homogeneous convex function. Then for $\phi \in X^*$, $\phi \in \partial f(\bar{x})$ if and only if $\phi \in \partial f(0)$ and $\phi(\bar{x}) = f(\bar{x})$.*

Proof. $\phi \in \partial f(\bar{x})$ if and only if $\phi(x-\bar{x}) \leqslant f(x)-f(\bar{x})$, for all $x \in X$. This clearly holds if $\phi \in \partial f(0)$ and $\phi(\bar{x})=f(\bar{x})$, and the converse follows by putting $x=0$ and $2\bar{x}$, and using $f(0)=0$ and $f(2\bar{x})=2f(\bar{x})$. □

The same result will hold for $\partial^{a} f$.

Proposition 5.2. *Let X be a normed lattice. Define $g: X \rightarrow \mathbb{R}$ by $g(x)=\|x\|$. For $\bar{x} \neq 0$,*

$$\partial g(\bar{x})=\{\phi \in X^* \mid \|\phi\|_*=1,\ \phi(\bar{x})=\|\bar{x}\|\},$$

$$\partial g(0)=\{\phi \in X^* \mid \|\phi\|_* \leqslant 1\}.$$

Proof. Immediate from the definition of the dual norm $\|\cdot\|_*$ and Lemma 5.1. □

Proposition 5.3. *Let X be a normed lattice. $x \mapsto \|x^+\|$ is a continuous, convex function.*

Proof. For $x_1, x_2 \in X$ and $0 \leqslant \lambda \leqslant 1$ we have $\lambda x_1^+ \geqslant \lambda x_1$, 0 and $(1-\lambda)x_2^+ \geqslant (1-\lambda)x_2$, 0, so adding we obtain $\lambda x_1^+ + (1-\lambda)x_2^+ \geqslant (\lambda x_1 + (1-\lambda)x_2)^+$. Since X is a normed lattice this gives

$$\begin{aligned}\|(\lambda x_1+(1-\lambda)x_2)^+\| &\leqslant \|\lambda x_1^+ + (1-\lambda)x_2^+\| \\ &\leqslant \lambda\|x_1^+\| + (1-\lambda)\|x_2^+\|,\end{aligned}$$

which shows the convexity.

The continuity follows from the continuity of the lattice operations, see Schaefer (1974, II.5.2). □

Corollary 5.4. *Let X be a normed lattice. Define $f: X \rightarrow \mathbb{R}$ by $f(x)=\|x^+\|$. Then $\partial f(x) \neq \emptyset$, for all $x \in X$.*

Proof. Proposition 5.3 and Holmes (1975, p. 84). □

Proposition 5.5. *Let X be a normed lattice. Define $f: X \rightarrow \mathbb{R}$ by $f(x)=\|x^+\|$. For $\bar{x} \notin -X_+$,*

$$\partial f(\bar{x})=\{\phi \in X^* \mid \phi \geqslant 0,\ \|\phi\|_*=1,\ \phi(\bar{x}^-)=0,\ \phi(\bar{x}^+)=\|\bar{x}^+\|\}.$$

For $\bar{x} \leqslant 0$,

$$\partial f(\bar{x})=\{\phi \in X^* \mid \phi \geqslant 0,\ \|\phi\|_* \leqslant 1,\ \phi(\bar{x})=0\}.$$

Proof. Since f is positively homogeneous, by Lemma 5.1 $\phi \in \partial f(\bar{x})$ if and only if $\phi \in \partial f(0)$ and $\phi(\bar{x})=f(\bar{x})$. Now $\phi \in \partial f(0)$ if and only if $\phi(x) \leqslant \|x^+\|$, for all $x \in X$, which is equivalent to $\phi \geqslant 0$ and $\|\phi\|_* \leqslant 1$. Thus $\phi \in \partial f(\bar{x})$ if and only if $\phi \geqslant 0$, $\|\phi\|_* \leqslant 1$ and $\phi(\bar{x})=\|\bar{x}^+\|$. This implies $\|\bar{x}^+\|=\phi(\bar{x}^+-\bar{x}^-) \leqslant \phi(\bar{x}^+) \leqslant \|\phi\|_*\|\bar{x}^+\| \leqslant \|\bar{x}^+\|$, so we have equality throughout, which gives the result. □

When x is not negative one can be more precise.

Proposition 5.6. *Let X be a normed lattice. Define $f, g : X \to \mathbb{R}$ by $f(x) = \|x^+\|$, $g(x) = \|x\|$ for all $x \in X$. Then for $\bar{x} \notin -X_+$, $\partial f(\bar{x}) = \partial g(\bar{x}^+) \cap \{\phi \in X^* \mid \phi \geqslant 0,\ \phi(\bar{x}^-) = 0\}$. In particular, for $\bar{x} \notin -X_+$, if the norm $\|\cdot\|$ is differentiable at $\bar{x}^+$ then f is differentiable at $\bar{x}$ with the same derivative.*

Proof. The formula for $\partial f(\bar{x})$ follows from Proposition 5.2 and Proposition 5.5. Also, if g is differentiable at $\bar{x}^+$ then $\partial g(\bar{x}^+)$ is a singleton. Therefore since $\partial f(\bar{x}) \neq \emptyset$ (Corollary 5.4) we must have $\partial f(\bar{x}) = \{\nabla g(\bar{x}^+)\}$. □

Examples 5.7. As before, $f(x) = \|x^+\|$, for $x \in X$.

(i) $X = L^p(T, \mu)$. By Holmes (1975, p. 170), for $x \neq 0$ we have

$$\nabla \|x\|_p = \|x\|_p^{1-p} |x|^{p-2} x.$$

Thus by Proposition 5.6, for $x \notin -X_+$, f is differentiable:

$$\nabla f(x) = \|x^+\|_p^{1-p} (x^+)^{p-1}.$$

(ii) $X = L^1(T, \mu)$. From Proposition 5.5 it follows that for any x, and $\phi \in L^\infty(T, \mu)$, $\phi \in \partial f(x)$ if and only if

$$\phi(t) \begin{cases} = 1 & \text{a.e. where } x(t) > 0, \\ = 0 & \text{a.e. where } x(t) < 0, \\ \in [0, 1] & \text{a.e. where } x(t) = 0. \end{cases}$$

Thus f is differentiable at x if and only if $|x(t)| > 0$ a.e., in which case $\nabla f(x) = \chi_{\{t \mid x(t) > 0\}}$.

(iii) $X = L^\infty(T, \mu)$, with the weak* topology. By Proposition 5.5, for $\phi \in L^1(T, \mu)$, and $x \notin -X_+$, $\phi = \partial f(x)$ if and only if $\phi \geqslant 0$, $\int_T \phi \, \mathrm{d}\mu = 1$, $\phi(t) = 0$ a.e. where $x(t) < 0$, and $\int_T \phi x^+ \, \mathrm{d}\mu = \operatorname{ess\,sup} x^+$, which is equivalent to $\phi \geqslant 0$, $\phi|_{T_x^c} = 0$ and $\int_{T_x} \phi \, \mathrm{d}\mu = 1$, where $T_x = \{t \in T \mid x(t) = \operatorname{ess\,sup} x\}$. It follows that if $x(t) < \operatorname{ess\,sup} x$ a.e. then $\partial f(x) = \emptyset$. Also, f is never differentiable at x if (T, μ) is nonatomic.

(iv) $X = C(T)$, with T a compact Hausdorff space. For $x \in X$, define $T_x = \{t \in T \mid x(t) = \max_\tau x(\tau)\}$. By Proposition 5.5, for $-x \notin X_+$, using the Riesz Representation Theorem (Rudin, 1966),

$$\partial f(x) = \{\phi \in M(T) \mid \phi \geqslant 0,\ \phi(T) = 1, \text{support } \phi \subset T_x\},$$

(in other words, probability measures supported on T_x). Thus for $x \notin -X_+$, f is differentiable at x if and only if T_x is a singleton, $\{t_0\}$ say, in which case $\nabla f(x) = \delta_{t_0}$, a unit mass concentrated at t_0. Such points are called *peak points*.

We now turn our attention to the convex function $x \mapsto \theta(x^+)$, where θ is some positive linear functional on X.

Proposition 5.8. *Let (X, Y) be a dual pair, with X a vector lattice and Y partially ordered by $(X_+)^+$. Suppose $y_0 \in Y$, $y_0 \geq 0$, and define $h: X \to \mathbb{R}$ by $h(x) = \langle x^+, y_0 \rangle$. Then h is a positively homogeneous convex function with, for any $x \in X$,*

$$\partial h(x) = \{y \in Y \mid 0 \leq y \leq y_0, \langle x^-, y \rangle = 0, \langle x^+, y_0 - y \rangle = 0\}.$$

If Y is an ideal in X^b then $\partial^a h(x) = \partial h(x)$.

Proof. h is clearly positively homogeneous. Take $x_1, x_2 \in X$ and $0 \leq \lambda \leq 1$. Then $\lambda x_1^+ \geq \lambda x_1$, 0 and $(1-\lambda)x_2^+ \geq (1-\lambda)x_2$, 0, so $(\lambda x_1 + (1-\lambda)x_2)^+ \leq \lambda x_1^+ + (1-\lambda)x_2^+$, and hence, since $y_0 \geq 0$, $\langle (\lambda x_1 + (1-\lambda)x_2)^+, y_0 \rangle \leq \lambda \langle x_1^+, y_0 \rangle + (1-\lambda)\langle x_2^+, y_0 \rangle$. Thus h is convex.

Now $y \in \partial h(0)$ if and only if $\langle x, y \rangle \leq \langle x^+, y_0 \rangle$, for all $x \in X$, which is equivalent to $y \in [0, y_0]$. Thus, by Lemma 5.1, $y \in \partial h(x)$ if and only if $y \in [0, y_0]$ and $\langle x, y \rangle = \langle x^+, y_0 \rangle$. But then

$$\langle x^+, y_0 \rangle \geq \langle x^+ y \rangle \geq \langle x, y \rangle = \langle x^+, y_0 \rangle,$$

so we have equality throughout, which gives the result.

Now for $\phi \in X'$, $\phi \in \partial^a h(0)$ if and only if $\phi(x) \leq \langle x^+, y_0 \rangle$, for all $x \in X$. This is equivalent to $\phi \geq 0$ and $y_0 - \phi \geq 0$. Thus $\phi \in X^b$ and $0 \leq \phi \leq y_0$, so $\phi \in Y$ as Y is an ideal. Therefore $\partial^a h(0) = \partial h(0)$. It follows that $\partial^a(x) = \partial h(x)$ for all $x \in X$. □

Proposition 5.9. *Let X be a normed lattice, $\theta \in X^*$, $\theta \geq 0$, and define $h: X \to \mathbb{R}$ by $h(x) = \theta(x^+)$. Then h is a continuous convex function and $\partial h(x) \neq \emptyset$ for all $x \in X$.*

Proof. We know h is convex by Proposition 5.8, and continuous by Schaefer (1974, II.5.2). Its subdifferentiability follows by Holmes (1975, p. 84). □

We now introduce some notation from Schaefer (1974). Let X be a vector lattice and suppose $\theta \in X^b$, $\theta \geq 0$. The *absolute kernel* of θ, $N(\theta) = \{x \in X \mid \theta(|x|) = 0\}$, is an ideal in X. If $\theta_1, \theta_2 \in X^b$, $\theta_1, \theta_2 \geq 0$, then

$$N(\theta_1 \vee \theta_2) = N(\theta_1) \cap N(\theta_2), \tag{5.1}$$

$$N(\theta_1 \wedge \theta_2) \supset N(\theta_1) + N(\theta_2). \tag{5.2}$$

In the following result, $\hat{}: X \to X^{**}$ denotes the evaluation map.

Theorem 5.10. *Let X be a normed lattice and suppose $0 \leq x_1, x_2 \in X$. Then in the lattice X^*, $N(\hat{x}_1 \wedge \hat{x}_2) = N(\hat{x}_1) + N(\hat{x}_2)$. Specifically, for any $0 \leq \theta \in X^*$ with $\theta(x_1 \wedge x_2) = 0$, there exist $0 \leq \theta_1, \theta_2 \in X^*$ with $\theta = \theta_1 + \theta_2$, $\theta_1(x_1) = \theta_2(x_2) = 0$.*

Proof. For any $x \in X$, $\hat{x} \in (X^*)^b$, by Proposition 3.2. Suppose first that $x_1 \wedge x_2 = 0$. Suppose $0 \leq \theta \in X^*$, and define $h: X \to \mathbb{R}$ by $h(x) = \theta(x^+)$. By Proposition 5.9, there exists $\theta_2 \in \partial h(x_1 - x_2)$, and by Proposition 5.8, $0 \leq \theta_2 \leq \theta$, $\theta_2(x_2) = 0$ and $(\theta - \theta_2)(x_1) = 0$. Setting $\theta_1 = \theta - \theta_2$ gives the result.

Now in the general case, we have $(x_1-x_2)^+ \wedge (x_1-x_2)^- = 0$, so by the above, for any $0 \leq \theta \in N(\hat{x}_1 \wedge \hat{x}_2)$, there exist θ_1, $\theta_2 \geq 0$ with $\theta = \theta_1 + \theta_2$ and $\theta_1((x_1-x_2)^+) = \theta_2((x_1-x_2)^-) = 0$. But $(\theta_1+\theta_2)(x_1 \wedge x_2) = 0$, so $\theta_1(x_1 \wedge x_2) = \theta_2(x_1 \wedge x_2) = 0$. But $(x_1-x_2)^+ + (x_1 \wedge x_2) = x_1$, so $\theta_1(x_1) = 0$. Similarly $\theta_2(x_2) = 0$, which gives the result. □

Theorem 5.10 is actually a special case of the following result of Schaefer (1974, II.4.11). Recall that we denote the space of all sequentially order continuous linear forms on a vector lattice Y by Y_0^{b}.

Theorem 5.11. *Let Y be a countably order complete vector lattice, and suppose $0 \leq \phi_1$, $\phi_2 \in Y_0^{\mathrm{b}}$. Then $N(\phi_1 \wedge \phi_2) = N(\phi_1) + N(\phi_2)$: for any $0 \leq y \in Y$, with $(\phi_1 \wedge \phi_2)(y) = 0$, there exist $0 \leq y_1, y_2 \in Y$ with $\phi_1(y_1) = \phi_2(y_2) = 0$, and $y_1 + y_2 = y$.*

Proof. The proof follows Schaefer (1974, II.4.11). □

For $0 \leq y \in Y$, we denote the *principal ideal* generated by y by $Y_y = \bigcup_{n=1}^{\infty} n[-y, y]$ (Schaefer, 1974, p. 57). We can now determine when the function $x \mapsto \theta(x^+)$ is differentiable.

Corollary 5.12. *Let (X, Y) be a dual pair, with Y a countably order complete vector lattice and X a sublattice of Y_0^{b}. For $0 \leq y_0 \in Y$, define convex $h: X \to \mathbb{R}$ by $h(x) = \langle x^+, y_0 \rangle$. Then $\partial h(x) \neq \emptyset$, for all $x \in X$.*

Suppose further that Y is an ideal in X^{b}. Then h is differentiable at x if and only if $\langle |x|, y \rangle > 0$ for all nonzero $y \in [0, y_0]$ in which case there is an order direct sum decomposition

$$Y_{y_0} = (Y_{y_0} \cap N(x^+)) \oplus (Y_{y_0} \cap N(x^-)),$$

and

$$\nabla h(x) = P_{Y_{y_0} \cap N(x^-)}(y_0),$$

where

$$P_{Y_{y_0} \cap N(x^-)}: Y_{y_0} \to Y_{y_0} \cap N(x^-)$$

is the natural projection.

Proof. By Proposition 5.8, h is convex with

$$\partial h(x) = \{y_1 \in Y \mid y_1 + y_2 = y_0, 0 \leq y_1 \in N(x^-), 0 \leq y_2 \in N(x^+)\}.$$

By Theorem 5.11, $N(x^+) + N(x^-) = N(x^+ \wedge x^-) = N(0) = Y$, so $\partial h(x) \neq \emptyset$, for all x.

If $\langle |x|, y \rangle > 0$ for all nonzero $y \in [0, y_0]$, then $\{0\} = N(|x|) \cap Y_{y_0} = N(x^+) \cap N(x^-) \cap Y_{y_0}$ by (5.1). Then

$$\partial h(x) = \{P_{Y_{y_0} \cap N(x^-)}(y_0)\}.$$

Since Y is an ideal in X^{b}, $\partial^{\mathrm{a}} h(x)=\partial h(x)$ by Proposition 5.8, so h is differentiable with the stated derivative.

Suppose on the other hand that for some nonzero $\bar{y}\in[0, y_0]$, $\langle |x|, \bar{y}\rangle=0$, so $\bar{y}\in N(x^-)\cap N(x^+)$. By Theorem 5.11, there exist $0\leqslant y_1\in N(x^-)$ and $0\leqslant y_2\in N(x^+)$ with $y_1+y_2=y_0-\bar{y}$. It then follows that both $y_1+\bar{y}$ and $y_1\in\partial h(x)$, so h is not differentiable at x. □

We shall discuss more concrete interpretations of the lattice theoretic ideas in this result at the end of this section.

Another similar convex function which arises in the dual problems in which we are interested is $k: X^m\to\mathbb{R}$, defined by $k(x_1,\ldots,x_m)=\theta\,(\bigvee_{i=1}^m x_i)$, for some positive linear functional θ.

Proposition 5.13. *Let (X, Y) be a dual pair, with X a vector lattice and Y partially ordered by $(X_+)^+$, Suppose $0\leqslant y_0\in Y$, and define $k: X^m\to\mathbb{R}$ by $k(x_1,\ldots,x_m)=\langle\bigvee_{i=1}^m x_i, y_0\rangle$. Then k is a positively homogeneous convex function with, for any $(x_1,\ldots,x_m)\in X^m$,*

$$\partial k(x_1,\ldots,x_m)$$
$$=\left\{(y_1,\ldots,y_m)\in Y^m \,\middle|\, \sum_{i=1}^m y_i=y_0,\ y_i\geqslant 0, \left\langle \bigvee_{k=1}^m x_k-x_i, y_i\right\rangle=0, \forall i\right\}.$$

If Y is an ideal in X^{b} then $\partial^{\mathrm{a}} k(x_1,\ldots,x_m)=\partial k(x_1,\ldots,x_m)$.

Proof. For $(u_1,\ldots,u_m)$, $(v_1,\ldots,v_m)\in X^m$ and $0\leqslant\lambda\leqslant 1$, $\lambda u_1\leqslant\lambda\bigvee_k u_k$ and $(1-\lambda)v_i\leqslant(1-\lambda)\bigvee_k v_k$, for each i, so $\bigvee_i(\lambda u_i+(1-\lambda)v_i)\leqslant\lambda\bigvee_i u_i+(1-\lambda)\bigvee_i v_i$. Thus since $y_0\geqslant 0$,

$$k(\lambda(u_1,\ldots,u_m)+(1-\lambda)(v_1,\ldots,v_m))\leqslant\lambda k(u_1,\ldots,u_m)+(1-\lambda)k(v_1,\ldots,v_m),$$

so k is convex. Also k is clearly positively homogeneous.

For $m=1$ the result is trivial, so suppose $m>1$. Now for $(y_1,\ldots,y_m)\in Y^m$, $(y_1,\ldots,y_m)\in\partial k(0)$ if and only if $\sum_{i=1}^m\langle x_i, y_i\rangle\leqslant\langle\bigvee_k x_k, y_0\rangle$, for all $(x_1,\ldots,x_m)\in X^m$. Putting $x_i:=x_0$, each i implies $\sum_i y_i=y_0$, and setting $x_i=x_0$ if $i=j$ and 0 otherwise gives $\langle x_0, y_j\rangle\leqslant\langle x_0^+, y_0\rangle$, for all $x_0\in X$, which implies $y_j\geqslant 0$, for each j. Thus

$$\partial k(0)=\left\{(y_1,\ldots,y_m)\in Y^m \,\middle|\, \sum_i y_i=y_0,\ y_i\geqslant 0\ \forall i\right\}.$$

The same proof shows that

$$\partial^{\mathrm{a}} k(0)=\left\{(\phi_1,\ldots,\phi_m)\in(X')^m \,\middle|\, \sum_i\phi_i=y_0,\ \phi_i\geqslant 0\ \forall i\right\}.$$

Thus when Y is an ideal in X^{b}, $\partial^{\mathrm{a}} k(0)\subset Y^m$, so $\partial^{\mathrm{a}} k(0)=\partial k(0)$.

Now by Lemma 5.1, for $(y_1,\ldots,y_m)\in Y^m$, $(y_1,\ldots,y_m)\in\partial k(x_1,\ldots,x_m)$ if and only if $(y_1,\ldots,y_m)\in\partial k(0)$ and $\sum_i\langle x_i, y_i\rangle=\langle\bigvee_k x_k, y_0\rangle$. But then $\sum_i\langle x_i, y_i\rangle\leqslant\sum_i\langle\bigvee_k x_k, y_i\rangle=\langle\bigvee_k x_k, y_0\rangle=\sum_i\langle x_i, y_i\rangle$, so we have equality, which gives $\langle\bigvee_k x_k-x_i, y_i\rangle=0$, for each i. The result follows.

Finally, when Y is an ideal in X^{b}, $\partial^{\mathrm{a}}k(0)=\partial k(0)$, so $\partial^{\mathrm{a}}k(x_1,\ldots,x_m)=\partial k(x_1,\ldots,x_m)$. $\square$

In an analogous fashion to (5.2) and Theorem 5.10, one has:

Proposition 5.14. *Let Y be a vector lattice, with $\phi_1,\ldots,\phi_m\in Y^{\mathrm{b}}$ $(m>1)$. Then*

$$N\Big(\bigvee_k \phi_k-\phi_i\Big)\cap\sum_{j\neq i}N\Big(\bigvee_k \phi_k-\phi_j\Big)\subset N\Big(\Big|\bigvee_{k\neq i}\phi_k-\phi_i\Big|\Big), \tag{5.3}$$

for each $i=1,\ldots,m$.

If Y is countably order complete and $\phi_1,\ldots,\phi_m\in Y_0^{\mathrm{b}}$ then we have equality in (5.3) and $Y=\sum_{i=1}^m N(\bigvee_k \phi_k-\phi_i)$, with positive elements having positive sums. Furthermore, if $I\subset Y$ is an ideal, then

$$\Big|\bigvee_{k\neq i}\phi_k-\phi_i\Big|>0\quad \text{on } I,\ \text{each } i=1,\ldots,m, \tag{5.4}$$

if and only if $I=\bigoplus_{i=1}^m(I\cap N(\bigvee_k \phi_k-\phi_i))$ is a direct sum of ideals.

Proof.

$$\begin{aligned}
&N\Big(\bigvee_k \phi_k-\phi_i\Big)\cap\sum_{j\neq i}N\Big(\bigvee_k \phi_k-\phi_j\Big)\\
&\quad\subset N\Big(\bigvee_k \phi_k-\phi_i\Big)\cap N\Big(\bigwedge_{j\neq i}\Big(\bigvee_k \phi_k-\phi_j\Big)\Big)\quad \text{(by (5.2))}\\
&\quad= N\Big(\bigvee_k \phi_k-\phi_i\Big)\cap N\Big(\bigvee_k \phi_k-\bigvee_{j\neq i}\phi_j\Big)\\
&\quad= N\Big(\Big(\bigvee_k \phi_k-\phi_i\Big)\vee\Big(\bigvee_k \phi_k-\bigvee_{j\neq i}\phi_j\Big)\Big)\quad \text{(by (5.1))}\\
&\quad= N\Big(\bigvee_k \phi_k-\Big(\phi_i\wedge\bigvee_{j\neq i}\phi_j\Big)\Big)\\
&\quad= N\Big(\Big|\bigvee_{k\neq i}\phi_k-\phi_i\Big|\Big),
\end{aligned}$$

so (5.3) follows.

If Y is countably order complete and $\phi_1,\ldots,\phi_m\in Y_0^{\mathrm{b}}$, then by Theorem 5.11,

$$\begin{aligned}
\sum_{i=1}^m N\Big(\bigvee_k \phi_k-\phi_i\Big)&=N\Big(\bigwedge_{i=1}^m\Big(\bigvee_k \phi_k-\phi_i\Big)\Big)\\
&=N\Big(\bigvee_k \phi_k-\bigvee_i \phi_i\Big)\\
&=N(0)\\
&=Y.
\end{aligned}$$

Finally, $\sum_{i=1}^{m}(I\cap N(\bigvee_k \phi_k-\phi_i))$ is a direct sum if and only if $I\cap N(|\bigvee_{k\neq i}\phi_k-\phi_i|)=I\cap N(\bigvee_k\phi_k-\phi_i)\cap\sum_{j\neq i}N(\bigvee_k\phi_k-\phi_j)=\{0\}$, for each i, which is easily seen to be equivalent to $|\bigvee_{k\neq i}\phi_k-\phi_i|>0$ on I. □

Notice that if $I=\bigoplus_{i=1}^{m} I_i$ is a direct sum of ideals then, by Proposition II.2.7 in Schaefer (1974), for any $0\leq y\in I$ there exist $0\leq y_i\in I_i$ with $\sum_{i=1}^{m} y_i=y$ (in other words, the sum is order direct).

We can now determine when the function $(x_1,\ldots,x_m)\mapsto\theta(\bigvee_k x_k)$ is differentiable. For $0\leq y\in Y$, we denote the principal ideal generated by y by $Y_y=\bigcup_{n=1}^{\infty} n[-y,y]$ (Schaefer, 1974, p. 57).

Corollary 5.15. *Let (X, Y) be a dual pair, with Y a countably order complete vector lattice and X a sublattice of Y_0^{b}. For $0\leq y_0\in Y$, define convex $k:X^m\to\mathbb{R}$ $(m>1)$ by $k(x_1,\ldots,x_m)=\langle\bigvee_k x_k, y_0\rangle$. Then $\partial k(x_1,\ldots,x_m)\neq\emptyset$, $(x_1,\ldots,x_m)\in X^m$. Moreover, if Y is an ideal in X^{b} then k is differentiable at $(x_1,\ldots,x_m)$ if and only if*

$$\left\langle\left|\bigvee_{k\neq i} x_k-x_i\right|, y\right\rangle>0, \tag{5.5}$$

for all nonzero $y\in[0,y_0]$, for each $i=1,\ldots,m$, in which case $Y_{y_0}=\bigoplus_{i=1}^{m}(Y_{y_0}\cap N(\bigvee_k x_k-x_i))$ is an order direct sum, and $(\nabla k(x_1,\ldots,x_m))_i=P_{Y_{y_0}\cap N(\bigvee_k x_k-x_i)}(y_0)$, each i where $P_{Y_{y_0}\cap N(\bigvee_k x_k-x_i)}: Y_{y_0}\to Y_{y_0}\cap N(\bigvee_k x_k-x_i)$ is the natural projection.

Proof. By Proposition 5.13, k is convex with

$$\partial k(x_1,\ldots,x_m)=\left\{(y_1,\ldots,y_m)\in Y^m\;\middle|\;\sum_i y_i=y_0,\,0\leq y_i\in N\left(\bigvee_k x_k-x_i\right),\forall i\right\}.$$

From Proposition 5.14 it follows that $\partial k(x_1,\ldots,x_m)\neq\emptyset$, for all $(x_1,\ldots,x_m)\in X^m$.

If (5.5) holds then $|\bigvee_{k\neq i}x_k-x_i|>0$ on Y_{y_0}, for each $i=1,\ldots,m$, so k has a unique continuous subgradient lying in Y^m, by Proposition 5.14. When Y is an ideal in X^{b} this is, by Proposition 5.13, the unique algebraic subgradient, and so k is differentiable at $(x_1,\ldots,x_m)$ with the stated derivative.

On the other hand, suppose (5.5) fails, so for some i and some nonzero $\bar{y}_i\in[0,y_0]$, $\bar{y}_i\in N(|\bigvee_{k\neq i}x_k-x_i|)$. Then by Proposition 5.14, $\bar{y}_i\in N(\bigvee_k x_k-x_i)$ and $\bar{y}_i\in\sum_{j\neq i}N(\bigvee_k x_k-x_j)$, so by Theorem 5.11 there exist $0\leq\bar{y}_j\in N(\bigvee_k x_k-x_j)$, $j\neq i$ with $\sum_{j\neq i}\bar{y}_j=\bar{y}_i$.

Now by Proposition 5.14 we can find $0\leq y_j\in N(\bigvee_k x_k-x_j)$ for each j, such that $\sum_{j=1}^{m}y_j=y_0-\bar{y}_i$, since this point is non-negative. It then follows that both $(u_1,\ldots,u_m)$, $(v_1,\ldots,v_m)\in\partial k(x_1,\ldots,x_m)$, where $u_i=y_i+\bar{y}_i$, $u_j=y_j$ for $j\neq i$, and $v_i=y_i$, $v_j=y_j+\bar{y}_j$, for $j\neq i$, so k is not differentiable at $(x_1,\ldots,x_m)$. □

The two differentiability results, Corollaries 5.12 and 5.15, are expressed in lattice theoretic notation. To conclude this section we shall interpret this notation for more concrete Banach spaces.

Theorem 5.16. *Let X be a normed lattice. Then for all $x \in X$, $\hat{x} \in X^{**}$ is order continuous on X^*.*

Proof. Wong and Ng (1973, 11.18). □

Theorem 5.17. *Under either of the following two conditions, Y is an order complete vector lattice, X is a sublattice of Y^{b}_{00} (and hence of Y^{b}_{0}), and Y is an ideal in X^{b}:*

(i) *X a normed lattice and $Y = X^*$.*

(ii) *Y a Banach lattice with $\sigma(Y, Y^*)$-compact order intervals, and $X = Y^*$.*

Proof. First note that in both cases Y is Archimedean, so $Y^{\mathrm{b}}_{00} \subset Y^{\mathrm{b}}$, by Peressini (1967, I.5.15), and is actually an ideal, by Wong and Ng (1973, 10.17).

(i) By Proposition 3.2, Y is order complete and $Y^{\mathrm{b}} = Y^* = X^{**}$. By Theorem 3.5, X is a sublattice of $X^{**} = Y^{\mathrm{b}}$, and by Theorem 5.16, $X \subset Y^{\mathrm{b}}_{00}$. The result now follows, as Y^{b}_{00} is an ideal in Y^{b}, and $Y = X^*$ is an ideal in X^{b} by Proposition 3.2.

(ii) By Proposition 3.2, $Y^* = Y^{\mathrm{b}}$, and we know Y^{b}_{00} is an ideal in Y^{b}. By Wong and Ng (1973, 13.5), Y is order complete and $Y^* \subset Y^{\mathrm{b}}_{00}$. Thus $X = Y^* = Y^{\mathrm{b}} = Y^{\mathrm{b}}_{00}$. Also, $X^{\mathrm{b}} = (Y^*)^{\mathrm{b}} = Y^{**}$, by Proposition 3.2, and Y is an ideal in Y^{**} (see Schaefer, 1974, II.5.10). □

Definition 5.18. A dual pair (X, Y) is called a *countably regular lattice pair* if X is a countably order complete vector lattice with Y a sublattice of X^{b}_{0}, and X an ideal in Y^{b}.

Any countably regular lattice pair is a dual lattice pair (Definition 3.6). By Theorem 5.17 if either Y is a normed lattice with $X = Y^*$ or X is a Banach lattice with $\sigma(X, X^*)$-compact order intervals and $Y = X^*$, then (X, Y) is a countably regular lattice pair.

Examples 5.19. Examples of Banach lattices Y with $\sigma(Y, Y^*)$-compact order intervals:

(i) $Y = c_0(T)$ for any index set T.

(ii) Y reflexive (for instance $Y = L^p(T, \mu)$ or $\ell^p(T)$ for $1 < p < \infty$ and (T, μ) a measure space).

(iii) Y an AL-space (Schaefer, 1974, II.8.1). For instance $L^1(T, \mu)$, $\ell^1(T)$ for a measure space (T, μ), and $M(T)$ for a compact Hausdorff space T.

(iv) More generally than (iii): Y weakly sequentially complete.

Proof. Schaefer (1974, p. 92). □

The Banach lattices $L^\infty(T, \mu)$, $\ell^\infty(T)$ and $C(T)$ (with T infinite) do not have weakly compact order intervals, since if they did, their unit balls would be weakly compact which would imply they were reflexive (Holmes, 1975, p. 126).

The only classical dual pair (X, Y) for which Theorem 5.17 (and Corollaries 5.12 and 5.15) fails to apply is $(X, Y)=(M(T), C(T))$, where T is a compact Hausdorff space (uncountable).

We conclude this section by illustrating Corollaries 5.12 and 5.15 for the cases $(X, Y)=(L^p(T, \mu), L^q(T, \mu))$, with $1 \leqslant p \leqslant \infty$, and $(C(T), M(T))$. By Theorem 5.17 both dual pairs satisfy the conditions of the Corollaries 5.12 and 5.15.

Let T be a compact Hausdorff space, and suppose $(X, Y)=(C(T), M(T))$. If $0 \leqslant y_0 \in M(T)$, then $\langle x, y\rangle=\int_T x(t)\, \mathrm{d}y(t)>0$, for all $0 \neq y \in[0, y_0]$ if and only if $\int_T x(t)u(t)\, \mathrm{d}y_0(t)>0$, for all $u \in L^1(T, \mu)$ with $0 \leqslant u(t) \leqslant 1$ y_0-a.e., by the Radon-Nikodym Theorem (Rudin, 1966), and this clearly holds if and only if $x(t)>0$, y_0-a.e.

For $x \in C(T)$, denote the set of zeros of x by $Z(x)=\{t \in T \mid x(t)=0\}$. Then for $0 \leqslant x \in C(T)$, $y \in N(x)$ if and only if $\int_T x(t)\, \mathrm{d}|y|(t)=0$, which is equivalent to $\operatorname{support}(y) \subset Z(x)$.

Since $Y_{y_0}=\bigcup_{n=1}^{\infty} n[-y_0, y_0]$, it follows again by the Radon-Nikodym Theorem that $Y_{y_0}=\{y \in M(T) \mid \mathrm{d}y/\mathrm{d}y_0 \in L^\infty(T, \mu)\}$. Thus

$$Y_{y_0} \cap N(x)=\left\{y \in M(T) \,\middle|\, \frac{\mathrm{d}y}{\mathrm{d}y_0} \in L^\infty(T, \mu), \frac{\mathrm{d}y}{\mathrm{d}y_0}=0 \text{ a.e. on } Z(x)^{\mathrm{c}}\right\}.$$

It follows that if $Y_{y_0}=\bigoplus_{i=1}^m (Y_{y_0} \cap N(x_i))$ is a direct sum of ideals then $\operatorname{support}(y_0)=\bigcup_{i=1}^m Z(x_i)$, and $Z(x_i) \cap Z(x_j)=\emptyset$ for $i \neq j$ (up to sets of y_0-measure 0), and so

$$P_{Y_{y_0} \cap N(x_i)}(y_0)=\begin{cases} y_0 & \text{on } Z(x_i), \\ 0 & \text{on } Z(x_i)^{\mathrm{c}}. \end{cases}$$

Finally, condition (5.5) may be interpreted as: the set $\{x_1(t), \ldots, x_m(t)\}$ has a unique largest element y_0-a.e.

Now let us consider the case $(X, Y)=(L^p(T, \mu), L^q(T, \mu))$ with (T, μ) a measure space and $1 \leqslant p \leqslant \infty$. For $x \in L^p(T, \mu)$, define $Z(x)=\{t \in T \mid x(t)=0\}$ (defined up to a set of μ-measure 0). If $0 \leqslant x \in L^p(T, \mu)$ and $0 \leqslant y_0 \in L^q(T, \mu)$, then $\langle x, y\rangle=\int_T x(t)y(t)\, \mathrm{d}\mu(t)>0$ for all $0 \neq y \in[0, y_0]$ if and only if $Z(x) \subset Z(y_0)$. To see this, suppose for some $0 \neq y \in[0, y_0]$, $\int_T x(t)y(t)\, \mathrm{d}\mu(t)=0$. Then for some $T_0 \subset T$ with $\mu(T_0)>0$, $y(t)>0$, μ-a.e. on T_0, so $T_0 \subset Z(x)$ but $t_0 \not\subset Z(y_0)$. On the other hand, if for some $T_0 \subset T$ with $\mu(T_0)>0$, $x(t)=0$, μ-a.e. on T_0 but $y_0(t)>0$, μ-a.e. on T_0, then $\int_T x(y_0\chi_{T_0})\, \mathrm{d}\mu(t)=0$.

Similarly, $y \in N(x)$ if and only if $\int_T x(t)|y(t)|\, \mathrm{d}\mu(t)=0$, which is equivalent to $(Z(y))^{\mathrm{c}} \subset Z(x)$.

Clearly we have $Y_{y_0}=\{uy_0 \mid u \in L^\infty(T, \mu)\}$. Thus $Y_{y_0} \cap N(x)=\{uy_0 \mid u \in L^\infty(T, \mu), Z(u)^{\mathrm{c}} \subset Z(x)\}$. It follows that if $Y_{y_0}=\bigoplus_{i=1}^m (Y_0 \cap N(x_i))$ is a direct sum of ideals then $Z(y_0)^{\mathrm{c}}=\bigcup_{i=1}^m Z(x_i)$, and $\mu(Z(x_i) \cap Z(x_j))=0$ for $i \neq j$, and so

$$P_{Y_{y_0} \cap N(x_i)}(y_0)=y_0\chi_{Z(x_i)}.$$

Finally, condition (5.5) may be interpreted as: the set $\{x_1(t), \ldots, x_m(t)\}$ has a unique largest element on $Z(y_0)^{\mathrm{c}}$, μ-a.e.

In Part II we will see these lattice notions in action.

References

A. Ben-Tal, J.M. Borwein and M. Teboulle, "A dual approach to multidimensional L_p spectral estimation problems," *SIAM Journal on Control and Optimization* 26 (1988) 985-996.

A. Ben-Tal, J.M. Borwein and M. Teboulle, "Spectral estimation via convex programming," to appear in: *Systems and Management Science by Extremal Methods: Research Honoring Abraham Charnes at Age 70* (Kluwer Academic Publishers, Dordrecht, 1992).

J.M. Borwein, "Convex relations in analysis and optimization," in: S. Schaible and W.T. Ziemba, eds., *Generalized Concavity in Optimization and Economics* (Academic Press, New York, 1981) pp. 335-377.

J.M. Borwein, "Continuity and differentiability properties of convex operators," *Proceedings of the London Mathematical Society* 44 (1982) 420-444.

J.M. Borwein, "Subgradients of convex operators," *Mathematische Operationsforschung und Statistik, Series Optimization* 15 (1984) 179-191.

J.M. Borwein and A.S. Lewis, "Practical conditions for Fenchel duality in infinite dimensions," in: J.-B. Baillon and M.A. Théra, eds., *Fixed Point Theory and Applications* (Longman, Harlow, UK, 1991) pp. 83-89.

I. Ekeland and R. Temam, *Convex Analysis and Variational Problems* (North-Holland, Amsterdam, 1976).

R.B. Holmes, *Geometric Functional Analysis and Applications* (Springer, New York, 1975).

L.D. Irvine, S.P. Marin and P.W. Smith, "Constrained interpolation and smoothing," *Constructive Approximation* 2 (1986) 129-151.

G.J.O. Jameson, *Topology and Normed Spaces* (Chapman and Hall, London, 1974).

K.O. Kortanek and M. Yamasaki, "Semi-infinite transportation problems," *Journal of Mathematical Analysis and Applications* 88 (1982) 555-565.

T.J. Lowe and A.P. Hurter Jr., "The generalized market area problem," *Management Science* 22 (1976) 1105-1115.

A.L. Peressini, *Ordered Topological Vector Spaces* (Harper and Row, New York, 1967).

R.T. Rockafellar, "Duality and stability in extremum problems involving convex functions," *Pacific Journal of Mathematics* 21 (1967) 167-187.

R.T. Rockafeller, *Convex Analysis* (Princeton University Press, Princeton, NJ, 1970).

H.H. Schaefer, *Topological Vector Spaces* (Springer, New York, 1971).

H.H. Schaefer, *Banach Lattices and Positive Operators* (Springer, Berlin, 1974).

Y.-C. Wong and K.-F. Ng, *Partially Ordered Topological Vector Spaces* (Clarendon Press, Oxford, 1973).