TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 245, November 1978

SYMMETRIC DUALITY FOR STRUCTURED CONVEX PROGRAMS[1,2]

BY

L. MCLINDEN

ABSTRACT. A fully symmetric duality model is presented which subsumes the classical treatments given by Duffin (1956), Eisenberg (1961) and Cottle (1963) for linear, homogeneous and quadratic convex programming. Moreover, a wide variety of other special objective functional structures, including homogeneity of any nonzero degree, is handled with equal ease. The model is valid in spaces of arbitrary dimension and treats explicitly systems of both nonnegativity and linear inequality constraints, where the partial orderings may correspond to nonpolyhedral convex cones. The approach is based on augmenting the Fenchel-Rockafellar duality model (1951, 1967) with cone structure to handle constraint systems of the type mentioned. The many results and insights from Rockafellar's general perturbational duality theory can thus be brought to bear, particularly on sensitivity analysis and the interpretation of dual variables. Considerable attention is devoted to analysis of suboptimizations occurring in the model, and the model is shown to be the projection of another model.

1. Introduction. Consider the problem of minimizing a function $f(x)$ subject to constraints of the form $x \geqslant 0$ and $Ax \geqslant b$, where the function f is convex, the transformation A is linear, and the partial orderings are determined by convex cones. Often it is important to take into account the sensitivity of this problem with respect to small changes in the vector b. In addition to the special form of the constraints, the function f may also have special structure, such as linearity, quadraticity, positive homogeneity, etc. This extra structure ought usually to be reflected rather explicitly in dual approaches to the problem. Further, it may be essential to consider such an optimization problem in some infinite-dimensional real vector space rather than in R^n, and to regard the inequalities as determined by order cones not necessarily finitely generated.

Received by the editors May 9, 1977.

AMS (MOS) subject classifications (1970). Primary 90C25.

Key words and phrases. Convex programming, suboptimizations, special structure, symmetric duality.

[1]Research sponsored in part by the National Science Foundation under grant number MPS75-08025 at the University of Illinois at Urbana-Champaign, and in part by the United States Army under grant number DAAG29-75-C-0024 at the Mathematics Research Center, University of Wisconsin-Madison.

[2]Presented to the IX International Symposium on Mathematical Programming (Hungarian Academy of Sciences, Budapest, Hungary, August 23–27, 1976) on August 23, 1976.

The aim of this paper is to provide a duality model which deals fully with each of the above considerations, and which is also symmetric, in the sense that the dual problem generated will enjoy structural characteristics of the same qualitative type as the original problem. The goal of such symmetry is motivated not simply by aesthetics or by the proof-theoretic power which it provides, but mainly by its algorithmic implications. By symmetry one is guaranteed in advance that the character of the dual problem given by the model will be qualitatively no worse than the original problem.

In 1956, R. J. Duffin [3] gave such a model to handle the fundamental case of f linear, though it was not apparent at that time just what the connection with sensitivity was. Duffin's model is the natural, infinite-dimensional extension of the celebrated Gale-Kuhn-Tucker [6] symmetric treatment of linear programming duality. Working in finite dimensions and using polyhedral order cones, E. Eisenberg [4] in 1961, and R. W. Cottle [2] in 1963, gave symmetric duality models for the basic nonlinear cases of f positively homogeneous and quadratic, respectively, and in 1964 R. T. Rockafellar [21] dealt similarly with the case of a general convex f. The case of a positive definite quadratic functional in Hilbert space occurs very implicitly in a 1965 paper by J.-J. Moreau [12], although there the concern is not with constraints per se. In 1967, Rockafellar [17] provided a symmetric duality treatment for completely general convex optimization problems, by broadening and extending the finite-dimensional model presented by W. Fenchel [5] in 1951. It was here that the issue of sensitivity under perturbations was first dealt with and its intimate connection with the dual problem explored. Constraints were covered, but only implicitly, by the presence of extended-real-valued functions and a highly useful linear transformation. In 1967 also, Rockafellar presented the outlines of his subsequent extremely broad, and symmetric, perturbational duality theory for convex optimization problems. Details of this, a wide variety of applications, and further references may be found in [20].

In the present paper, the basic Fenchel-Rockafellar model [5], [17] is augmented with explicit cone structure, so as to handle directly general constraint systems of the form $x \geqslant 0$ and $Ax \geqslant b$. This is done in a manner which maintains the complete symmetry of the original model. The resulting framework forms a direct bridge, as it were, between the Fenchel-Rockafellar model for general convex problems, on the one hand, and the treatments given by Duffin, Eisenberg, and Cottle for the particular classes of linear, homogeneous, and quadratic problems, on the other hand. We indicate how these cases, as well as other classes of problems having special structure, can be handled in the present framework. This requires providing conjugacy and subdifferential formulas for various special functional structures of interest.

This we do for two broad classes of functions. The first consists of generalized convex "distance" functions: Minkowski gauge functionals composed with Young's functions on the half-line. The second class is the natural, concave analogue of the first, consisting of what might be viewed as generalized concave "utility" functions. Together the two classes include, in particular, positive homogeneity structure of *any* nonzero degree. The Lagrangian minimax problem associated with our primal and dual problems has explicit "nonnegativity" constraints in each of the two arguments. (Additional, implicit constraints might also be built in via the extended-real-valued saddle function.) The present results can thus also be interpreted as bearing on such constrained two-person zero-sum games.

A topical outline follows. In §2 notation is established and the Fenchel-Rockafellar model recalled briefly. In §3 a simple principle is observed for introducing further structure into the problem in a symmetric fashion. Based on this, the three problems forming our cone-augmented model are presented. It is indicated here how the various classical models can be recaptured from the present one. In §4 the main results relating the primal and dual problems are developed, and the issue of suboptimization partially addressed. In §5 the extremality conditions and the associated minimax problem are treated, and the issue of suboptimization analyzed further. It is shown in particular that the trio of problems being treated, regarded both collectively and individually, can rightly be viewed as the projection of another trio having no suboptimizations but "twice as many" variables. This other trio corresponds, it turns out, to the symmetric primal-dual pair outlined by Rockafellar in a more restrictive setting **[21]**. In §6 the projection phenomenon is analyzed further and seen to be actually quite a general construction, suggesting further issues for investigation. In §7 conjugacy and subdifferential formulas are indicated for the generalized convex distance and concave utility functions mentioned above, enabling such functions to be employed freely in the model. In §8 are some concluding remarks about possible variations and refinements of the present framework.

Although the thrust of the present paper is towards (possibly infinite) linear inequality constraint systems over cones, we should mention that a symmetrized duality model is also available for constraint systems consisting of finitely many convex inequalities. This was first presented by E. L. Peterson in 1972 for closed problems in the finite-dimensional setting **[14]**, **[15]**, and has recently been broadened by McLinden **[11]** to cover nonclosed problems in general spaces, as well as polyhedral refinements of the model in the finite-dimensional case. In that model, too, symmetrization is seen to be intimately tied to suboptimizations. On this point those papers serve as a useful complement to the present one.

To highlight the many symmetries appearing throughout, we work in the setting of locally convex real topological vector spaces paired in duality (see, e.g., [1] or [7]). Readers unfamiliar with this may, for convenience, interpret all spaces as being Euclidean with the usual topology, or reflexive Banach under the norm topologies. We use freely standard facts from the theory of conjugate convex functions (see, e.g., Moreau [13] or Rockafellar [18], [20]), and also general facts about perturbational duality theory contained in Rockafellar [20].

2. The basic Fenchel-Rockafellar model. This involves the following ingredients:

$$\begin{tikzcd}
& X \arrow[r, "A"] \arrow[dl, "h"'] \arrow[dd, "\partial h"] & U \arrow[dr, "k"] & \\
{[-\infty, +\infty]} & & & {[-\infty, +\infty]} \\
& V \arrow[ul, "h^*"] & Y \arrow[l, "A^*"] \arrow[uu, "\partial k^*"'] \arrow[ur, "k^*"'] &
\end{tikzcd}$$

Here X and V are locally convex Hausdorff topological vector spaces over the real numbers R and are paired in duality by a real bilinear form $(x, v) \to \langle x, v \rangle$. The situation is analogous for the spaces U and Y. The functions h and k are extended-real-valued convex and concave, respectively, and h^* and k^* are the functions conjugate to them under the Fenchel transform (in the convex and concave sense, respectively). The transformations ∂h and ∂k^* are the subdifferentials of h and k^*, respectively. They send points into closed convex sets and are generalizations of gradient mappings, to which they reduce in the presence of differentiability. The transformation A is linear and continuous with adjoint A^*.

The model is designed to treat the initial problem of interest, the *primal problem*

$$\min_x \{h(x) - k(Ax)\}, \qquad (\mathcal{P}_0)$$

by means of its close interrelationships with two other optimization problems: the *dual problem*

$$\max_y \{k^*(y) - h^*(A^*y)\}, \qquad (\mathcal{D}_0)$$

and the *Lagrangian problem*

$$\underset{x \quad y}{\text{minimax}} \{h(x) + k^*(y) - \langle Ax, y \rangle\}. \qquad (\mathcal{L}_0)$$

The variables in $(\mathcal{L}_0)$ are restricted to lie in the product set dom $h \times$ dom k^*. It is helpful to think of $(\mathcal{L}_0)$ as a sort of bridge linking $(\mathcal{P}_0)$ with $(\mathcal{D}_0)$.

Throughout the paper we make the *nondegeneracy assumption* that the lower semicontinuous hull of h is finite somewhere and the upper semicontinuous hull of k is finite somewhere. (In the Euclidean case this is satisfied if h and k are merely proper convex and proper concave, respectively.) This implies that h, k, h^*, and k^* are proper and, moreover, satisfy

$$\operatorname{lsc} h = \operatorname{cl} h = h^{**}, \quad \operatorname{usc} k = \operatorname{cl} k = k^{**}.$$

From the properness, note that the extended arithmetic in $(\mathcal{P}_0)$, $(\mathcal{D}_0)$, $(\mathcal{L}_0)$ never involves adding $-\infty$ to $+\infty$.

The subdifferentials ∂h and ∂k^* serve as vehicles for expressing the extremality conditions, or abstract *Kuhn-Tucker conditions*, associated with this trio of problems:

$$Ax \in \partial k^*(y) \quad \text{and} \quad A^*y \in \partial h(x).$$

A pair (x, y) satisfies these conditions if and only if it solves $(\mathcal{L}_0)$, in which case x solves $(\mathcal{P}_0)$, y solves $(\mathcal{D}_0)$, and all three optimal values coincide. Conversely, under any of a variety of hypotheses ("constraint qualifications"), in order that a vector x solve $(\mathcal{P}_0)$, it is necessary that there exists a vector y such that (x, y) satisfies the Kuhn-Tucker conditions.

An important feature of the model is its relevance to sensitivity analysis. Briefly, the function being optimized in $(\mathcal{D}_0)$ precisely mirrors the sensitivity of $(\mathcal{P}_0)$ with respect to a certain class of perturbations, which here correspond to horizontal translations of the graph of k. This relationship between primal sensitivity and the dual problem was first observed in **[17]** and is explicated thoroughly in **[18]**, **[20]**. Details of this, as well as other general relationships concerning the three problems just introduced, may also be obtained by appropriate specialization of the cone-augmented version of the model which follows.

3. The underlying idea and the cone-augmented model. Our method of obtaining symmetry is based on a very simple idea. The idea is also quite natural, in view of the basic properties of the Fenchel transform.

We proceed in a quasi-formal manner, introducing two nonempty families, $\mathcal{A}$ and $\mathcal{B}$, of extended-real-valued functions. The members of $\mathcal{A}$ are required to be convex with lower semicontinuous hull somewhere finite, and those of $\mathcal{B}$ are required to be concave with upper semicontinuous hull somewhere finite. These families will serve to single out special types of problem structure such as linearity, quadraticity, and even cone constraint structure, as we shall see shortly. When the functions h and k in $(\mathcal{P}_0)$ satisfy $h \in \mathcal{A}$ and $k \in \mathcal{B}$, we say $(\mathcal{P}_0)$ *is of type* $(\mathcal{A}, \mathcal{B})$. For symmetry, conditions are needed which will imply that $(\mathcal{D}_0)$ is of the same type as $(\mathcal{P}_0)$, up to closures and minus signs. Clearly, the above restrictions on h and k imply that $h^* \in \mathcal{A}^*$ and $k^* \in \mathcal{B}^*$, where the notation $\mathcal{C}^*$ denotes the family obtained from $\mathcal{C}$ by forming the conjugates (in the appropriate sense) of the members of $\mathcal{C}$. (Similarly for the notation cl $\mathcal{C}$ and $-\mathcal{C}$ to follow.) Since this means that $(\mathcal{D}_0)$ is of type $(\mathcal{B}^*, \mathcal{Q}^*)$, the conditions needed to ensure that $(\mathcal{D}_0)$ is of the same type as $(\mathcal{P}_0)$ are simply that $\mathcal{Q}$ and $\mathcal{B}$ satisfy

$$\mathcal{Q}^* = -\mathrm{cl}\,\mathcal{B} \quad \text{and} \quad \mathcal{B}^* = -\mathrm{cl}\,\mathcal{Q}. \tag{3.1}$$

In this terminology, the general Fenchel-Rockafellar problems $(\mathcal{P}_0)$ and $(\mathcal{D}_0)$ are each of type $(\mathcal{Q}, \mathcal{B})$ for the largest possible choices of $\mathcal{Q}$ and $\mathcal{B}$, namely $\mathcal{B} = -\mathcal{Q}$, where $\mathcal{Q}$ is the family of *all* convex functions having lower semicontinuous hull somewhere finite. With this choice, conditions (3.1) follow immediately from the basic properties of the Fenchel transform. We shall see below, in terms of the cone-augmented model, how other, more special choices of the families $\mathcal{Q}$ and $\mathcal{B}$ yield symmetric duals having specific structure.

We now apply the above idea to see how to handle in a symmetric fashion systems of both nonnegativity and linear inequality constraints over cones. Consider, for example, the problem posed at the beginning of the Introduction, namely, to minimize $f(x)$ subject to $x \in P$ and $Ax - b \in Q$, where P and Q are nonempty convex cones determining the orderings. This can be cast as $(\mathcal{P}_0)$ by letting $h = f + \psi_P$ and $k = -\psi_{Q+b}$. (We write ψ_C to denote the convex indicator of a convex set C.) Notice that incorporating the constraint $Ax \geqslant b$ into the $k(Ax)$ term of $(\mathcal{P}_0)$ ensures that the dual problem will yield information on the particular sensitivity in $(\mathcal{P}_0)$ of interest to us, namely the effects of perturbing the right-hand side vector b by small amounts (cf. remark concerning sensitivity at the end of §2). Now how can the preceding $(\mathcal{Q}, \mathcal{B})$-development be brought to bear, at least heuristically? Since h is of the form $h = h_1 + h_2$, let us take for $\mathcal{Q}$ all the convex functions "having this form". Then the elements of $\mathcal{Q}^*$ have the form of $h^* = \mathrm{cl}(h_1^* \,\square\, h_2^*)$, under suitable circumstances, where the symbol $\square$ denotes the operation of infimal convolution on convex functions. In view of the first requirement in (3.1), this suggests choosing for $\mathcal{B}$ all concave functions k of the form $k = k_1 \,\square\, k_2$, where here $\square$ denotes the operation of supremal convolution on concave functions. The members of $\mathcal{B}^*$ then look like $k^* = k_1^* + k_2^*$, so that, again under suitable conditions, the other requirement in (3.1) is met. Now consider what happens when h_2 is of the special form $h_2 = \psi_P$ for a convex cone P. The fact that $h_2^* = \psi_{P^0}$, where P^0 is the polar cone, suggests taking k_2 (in elements of $\mathcal{B}$) to be of the special form $k_2 = -\psi_Q$ for some convex cone Q. This, of course, yields $k_2^* = -\psi_{Q^0}(-\cdot) = -\psi_{Q^*}$, where Q^* is the dual cone (i.e. negative of the polar).

This outlines a rather general scheme, in which the three optimization problems assume the form

$$\min_x \{(h_1 + h_2)(x) - (k_1 \,\square\, k_2)(Ax)\}, \tag{3.2}$$

$$\max_y \left\{(k_1^* + k_2^*)(y) - (h_1^* \,\square\, h_2^*)(A^*y)\right\}, \tag{3.3}$$

and

$$\operatorname*{minimax}_{x \quad y} \left\{(h_1 + h_2)(x) + (k_1^* + k_2^*)(y) - \langle Ax, y\rangle\right\}, \tag{3.4}$$

and the functions h_2, k_2, h_2^*, k_2^* can be restricted to be the indicators of certain cones. Of course, all this is only heuristic, since the above "derivation" glossed over key technical issues in several places. It does, however, serve to illuminate the origin of the cone-augmented model below.

To clarify further, consider once more the problem from the Introduction. We can apply the above scheme to it by choosing $h_1 = f$, $h_2 = \psi_P$, $k_1 = -\psi_{\{b\}}$, $k_2 = -\psi_Q$. In view of $(-\psi_b) \,\square\, (-\psi_Q) = -\psi_{Q+b}$, problem (3.2) is then just

$$\min_{x \geq 0, Ax \geq b} \{f(x)\},$$

and its "dual" problem should be something like that given by (3.3), i.e.

$$\max_y \left\{(\langle b, y\rangle - \psi_{Q^*}(y)) - (f^* \,\square\, \psi_{P^0})(A^*y)\right\}.$$

By drawing the inf-convolution to the outside, this can be rewritten as

$$\max_{y \geq 0} \sup_{v \geq A^*y} \{\langle b, y\rangle - f^*(v)\} \quad \text{(to find } y \text{ only)},$$

where the partial orders are the natural ones induced on V and Y by P^* and Q^*. The associated minimax problem ought to be (3.4), i.e.

$$\operatorname*{minimax}_{x \geq 0\, y \geq 0} \left\{f(x) + \langle b, y\rangle - \langle Ax, y\rangle\right\}.$$

Notice in the "dual" the appearance of a suboptimization over an auxiliary variable v. This reflects the fact that the original, primal problem we started with is actually asymmetric in a certain sense having to do with the particular class of perturbations involved. The statements of all the results below can easily be specialized to such asymmetric cases.

Now let nonempty convex cones $P \subset X$ and $Q \subset U$ be given and fixed, once and for all, and let h, k, and A be as in the beginning of §2. The cone-augmented extension of the Fenchel-Rockafellar model which we shall study consists of problems (3.2)–(3.4), where $h_1 = h$, $h_2 = \psi_P$, $k_1 = k$, $k_2 = -\psi_Q$. Thus, our *primal problem* is

$$\min_x \left\{(h + \psi_P)(x) - (k \,\square\, -\psi_Q)(Ax)\right\},$$

which, by drawing the sup-convolution to the outside, can be rewritten as

$$\min_{x \geq 0} \inf_{Ax \geq z} \left\{h(x) - k(z)\right\} \quad \text{(to find } x \text{ only)}. \tag{$\mathcal{P}$}$$

Our *dual problem* is

$$\max_y \left\{(k^* - \psi_{Q^*})(y) - (h^* \,\square\, \psi_{P^0})(A^*y)\right\},$$

which, by drawing the inf-convolution to the outside, can be rewritten as

$$\max_{y \geq 0} \sup_{w \geq A^*y} \{k^*(y) - h^*(w)\} \quad \text{(to find } y \text{ only).} \tag{$\mathfrak{D}$}$$

The associated *Lagrangian saddlepoint problem* is

$$\operatorname*{minimax}_{x \in C\, y \in D} \{h(x) + k^*(y) - \langle Ax, y\rangle\}, \tag{$\mathfrak{L}$}$$

where

$$C = P \cap \operatorname{dom} h, \quad D = Q^* \cap \operatorname{dom} k^*. \tag{3.5}$$

Notice first of all that with the choices $P = X$ and $Q = \{0\}$ these three problems coincide exactly with the problems ($\mathfrak{P}_0$), ($\mathfrak{D}_0$), and ($\mathfrak{L}_0$) treated by the Fenchel-Rockafellar model. The various hypotheses we shall invoke in §§4, 5 in proving results for the cone-augmented model all reduce, for this particular choice of cones, to the "standard" conditions required in the original Fenchel-Rockafellar model.

Next, notice that Duffin's model results when the cone-augmented model is restricted to problems ($\mathfrak{P}$) of type ($\mathcal{Q}$, $\mathfrak{B}$), where $\mathfrak{B} = -\mathcal{Q}^*$ and $\mathcal{Q}$ consists of all continuous linear functions. This follows from the fact that the (convex) conjugate of a function of the form $x \to \langle x, a\rangle$ has the form $v \to \psi_a(v)$. Thus, if we take h and k to be of the form $h(x) = \langle x, c\rangle$ and $k(u) = -\psi_b(u)$, then the three problems assume the special form

$$\min_{x \geq 0, Ax \geq b} \{\langle x, c\rangle\}, \qquad \max_{y \geq 0, c \geq A^*y} \{\langle b, y\rangle\},$$

and

$$\operatorname*{minimax}_{x \geq 0\, y \geq 0} \{\langle x, c\rangle + \langle b, y\rangle - \langle Ax, y\rangle\}.$$

The above two cases can be regarded as the extreme cases of the cone-augmented model. They are the *only* cases in which the suboptimizations "disappear" from *both* ($\mathfrak{P}$) and ($\mathfrak{D}$).

Symmetric duality for various types of nonlinear objective functional structure can be obtained by using the conjugacy (and subdifferential) formulas presented in §7, together with the results in §§4, 5 for ($\mathfrak{P}$), ($\mathfrak{D}$) and ($\mathfrak{L}$). For example, Eisenberg's treatment of the homogeneous case is extended by considering those problems ($\mathfrak{P}$) of type ($\mathcal{Q}$, $\mathfrak{B}$), where $\mathfrak{B} = -\mathcal{Q}^*$ and $\mathcal{Q}$ is chosen to be the family of (closed) gauge functionals. Corollary 13B in §7 provides the essential facts for this. As another example, Cottle's treatment of the quadratic case is extended by considering those problems ($\mathfrak{P}$) of type ($\mathcal{Q}$, $\mathfrak{B}$), where again $\mathfrak{B} = -\mathcal{Q}^*$ but this time $\mathcal{Q}$ is chosen to be the family of all functionals of the form $x \to \frac{1}{2}\gamma_C(x)^2$, where γ_C is the closed gauge associated with a "polar" set C. Convex quadratic forms or, more generally, pth powers of norms, can be represented by the appropriate specification of C. The essential conjugacy (and subdifferential) formulas are provided by

Corollary 13A of §7. A number of other interesting objective functional structures likewise admit symmetric duality treatment with the aid of Propositions 13 and/or 14 of §7, combined with the results which follow in §§4, 5.

4. Relationships between ($\mathcal{P}$) and ($\mathcal{D}$). In this section we establish mild conditions under which the duality between ($\mathcal{P}$) and ($\mathcal{D}$) implied in §3 is in fact the case. We shall see that in the absence of such conditions, the two problems bear only a weaker, subduality relationship to one another. The difficulty is that certain closure or semicontinuity properties may in general be lacking.

Our objective is to work towards placing our trio ($\mathcal{P}$), ($\mathcal{D}$), ($\mathcal{L}$) in the general perturbational duality framework developed by Rockafellar [18], [20]. Towards this end, we introduce functions F, G and K defined by

$$F(x, u) = \begin{cases} (h + \psi_P)(x) - (k \,\square - \psi_Q)(Ax + u) & \text{if } x \in C, \\ +\infty \quad \text{if } x \notin C, \end{cases}$$

$$G(y, v) = \begin{cases} (k^* - \psi_{Q^*})(y) - (h^* \,\square\, \psi_{P^0})(A^*y + v) & \text{if } y \in D, \\ -\infty \quad \text{if } y \notin D, \end{cases}$$

and

$$K(x, y) = \begin{cases} (h + \psi_P)(x) + (k^* - \psi_{Q^*})(y) - \langle Ax, y \rangle & \text{if } x \in C, \\ +\infty \quad \text{if } x \notin C, \end{cases}$$

where C and D are the sets in (3.5). (Any product spaces occurring will be assumed to be paired in the obvious manner. Thus, for example, $X \times U$ and $V \times Y$ are regarded as paired under the bilinear form $\langle (x, u), (v, y) \rangle = \langle x, v \rangle + \langle u, y \rangle$.) These functions can be reformulated as

$$F(x, u) = \begin{cases} \inf_{Ax+u \geq z} \{h(x) - k(z)\} & \text{if } x \geq 0, \\ +\infty \quad \text{otherwise}, \end{cases}$$

$$G(y, v) = \begin{cases} \sup_{w \geq A^*y+v} \{k^*(y) - h^*(w)\} & \text{if } y \geq 0, \\ -\infty \quad \text{otherwise}, \end{cases}$$

and

$$K(x, y) = \begin{cases} h(x) + k^*(y) - \langle Ax, y \rangle & \text{if } x \in C \text{ and } y \in D, \\ -\infty \quad \text{if } x \in C \text{ and } y \notin D, \\ +\infty \quad \text{if } x \notin C. \end{cases}$$

In terms of these functions, the problems ($\mathcal{P}$), ($\mathcal{D}$) and ($\mathcal{L}$) can be expressed as

$$\min_x F(x, 0), \quad \max_y G(y, 0), \quad \text{and} \quad \operatorname*{minimax}_{x \;\; y} K(x, y).$$

In view of the similarity in notation between the above and the development in [20], it would appear that we can apply the results from there immediately. That would be incorrect, however, as it has not been established that the present functions F, G and K bear the same relationships to each other as do Rockafellar's corresponding three functions. The relationships among F, G and K presumed by the development in [20] all stem from two key identities:

$$G(y, v) = -F^*(v, -y) \tag{4.1}$$

and

$$K(x, y) = \inf_u \{F(x, u) + \langle u, y\rangle\} \tag{4.2}$$

(cf. [20, (4.17) and (4.2)]).

It is easy to see that one of these, (4.2), holds here without any additional conditions. Indeed, from our definitions of F and K it is satisfied trivially when $x \notin C$, while for $x \in C$ we can compute that

$$\begin{aligned}
&\inf_u \{F(x, u) + \langle u, y\rangle\} \\
&\quad = \inf_u \{(h + \psi_P)(x) - (k \,\square\, -\psi_Q)(Ax + u) + \langle u, y\rangle\} \\
&\quad = (h + \psi_P)(x) + \inf_u \{\langle u, y\rangle - (k \,\square\, -\psi_Q)(Ax + u)\} \\
&\quad = (h + \psi_P)(x) + \inf_{u'} \{\langle u', y\rangle - (k \,\square\, -\psi_Q)(u')\} - \langle Ax, y\rangle \\
&\quad = (h + \psi_P)(x) + (k^* - \psi_{Q^*})(y) - \langle Ax, y\rangle.
\end{aligned}$$

The identity (4.1) is a more delicate matter. In general, the best one can obtain is the inequality in Proposition 1 below. For it, and for the remainder of the paper, it is convenient to make the following *nondegeneracy assumption*:

$$C \neq \varnothing \quad \text{and} \quad D \neq \varnothing. \tag{4.3}$$

This simply has the effect of eliminating from the discussion certain trivial situations which would be awkward always to carry along. It can be shown that $D = \varnothing$ if and only if the function $\operatorname{usc}(k \,\square\, -\psi_Q)$ is $+\infty$ throughout the set $\operatorname{cl}(\operatorname{dom} k + Q)$.

PROPOSITION 1. *The functions F and G satisfy*

$$G(y, v) \leqslant -F^*(v, -y) \quad \textit{and} \quad F(x, u) \geqslant -G^*(u, -x).$$

Furthermore, F is proper convex with lsc F *never* $-\infty$, *and G is proper concave with* usc G *never* $+\infty$.

PROOF. We give the proof for the second inequality only, as the proof of the first one is quite similar. Computation yields that

$$\begin{aligned}
-G^*(u,-x) &= -\inf_{y,v}\ \{\langle (u,-x),(y,v)\rangle - G(y,v)\} \\
&= \sup_{y,v}\ \{-\langle u,y\rangle + \langle x,v\rangle + G(y,v)\} \\
&= \sup_{y\in D}\ \sup_{v}\ \{-\langle u,y\rangle + \langle x,v\rangle + (k^* - \psi_{Q^*})(y) \\
&\qquad\qquad - (h^* \,\square\, \psi_{P^0})(A^*y+v)\} \\
&= \sup_{y\in D}\ \{(k^*-\psi_{Q^*})(y) - \langle u,y\rangle + \alpha(x,y)\},
\end{aligned}$$

where

$$\begin{aligned}
\alpha(x,y) &= \sup_{v}\ \{\langle x,v\rangle - (h^* \,\square\, \psi_{P^0})(A^*y+v)\} \\
&= \sup_{v'}\ \{\langle x,v'\rangle - (h^* \,\square\, \psi_{P^0})(v')\} - \langle x, A^*y\rangle \\
&= (h^* \,\square\, \psi_{P^0})^*(x) - \langle x, A^*y\rangle \\
&= (h^{**} + \psi_{P^\infty})(x) - \langle x, A^*y\rangle.
\end{aligned}$$

If $x \notin C$, then $F(x,u) = +\infty$ and the desired inequality is satisfied trivially. If $x \in C$, then $\alpha(x,y)$ is finite and we have

$$\begin{aligned}
-G^*(u,-x) &= (h^{**} + \psi_{P^\infty})(x) \\
&\qquad + \sup_{y\in D}\ \{(k^*-\psi_{Q^*})(y) - \langle u,y\rangle - \langle x, A^*y\rangle\} \\
&= (h^{**} + \psi_{P^\infty})(x) - \inf_{y\in D}\ \{\langle Ax+u, y\rangle - (k^*-\psi_{Q^*})(y)\} \\
&= (h^{**} + \psi_{P^\infty})(x) - (k^*-\psi_{Q^*})^*(Ax+u).
\end{aligned}$$

Since $h^{**} + \psi_{P^\infty} \leqslant h + \psi_P$ and $(k^* - \psi_{Q^*})^* \geqslant k \,\square\, -\psi_Q$, the desired inequality follows. Next, we claim that G cannot be identically $-\infty$. Indeed, $D \neq \varnothing$ implies that $G \equiv -\infty$ if and only if $h^* \,\square\, \psi_{P^0} \equiv +\infty$, which happens if and only if $h^* \equiv +\infty$. But the latter cannot occur, because lsc h is assumed finite somewhere. Now since $G \not\equiv -\infty$, the first inequality implies that $F^* \not\equiv +\infty$, from which it follows that lsc F is never $-\infty$. In a similar way, using $C \neq \varnothing$, one shows that $F \not\equiv +\infty$. From this and the second inequality it follows that $G^* \not\equiv -\infty$, so that usc G is never $+\infty$. Finally, F is proper because of $-\infty <$ lsc $F \leqslant F \not\equiv +\infty$, and G is proper because of $-\infty \not\equiv G \leqslant$ usc $G < +\infty$.

COROLLARY 1A. *One always has the estimate* $F(x,0) \geqslant G(y,0)$. *Moreover, equality is attained by a pair of vectors x and y if and only if x solves* $(\mathcal{P})$, *y solves* $(\mathcal{D})$, *and* $\min(\mathcal{P}) = \max(\mathcal{D})$.

PROOF. From either of the inequalities established in the proposition, one can obtain the general inequality

$$F(x, u) - \langle x, v\rangle \geqslant G(y, v) - \langle u, y\rangle.$$

The corollary follows by specializing this to the case $u = 0$, $v = 0$.

The inequalities in Proposition 1 say that F and G are always *subconjugates* of one another (up to minus signs), and thus one can always view ($\mathcal{P}$) and ($\mathcal{D}$) as *subduals* of each other. But in general ($\mathcal{D}$) will not furnish as tight a lower bound on ($\mathcal{P}$) as one would hope for, unless (4.1) is satisfied. In addition, most of the nicest duality relationships between ($\mathcal{P}$) and ($\mathcal{D}$) require that the function $F(x, u)$ be closed in the u argument at least, and some even require F to be closed in (x, u) jointly (cf. [**20**]). As the next proposition shows, these types of regularity depend on the following condition's being met:

$$(h + \psi_P)^* = h^* \,\square\, \psi_{P^0} \qquad (\mathcal{K}_1)$$

or

$$(k^* - \psi_{Q^*})^* = k \,\square - \psi_Q. \qquad (\mathcal{K}_2)$$

In the case of the basic Fenchel-Rockafellar model (i.e. $P = X$ and $Q = \{0\}$), the first of these is met trivially, while the second amounts simply to having k closed. A similar remark applies to conditions ($\mathcal{K}_i'$) and ($\mathcal{K}_i''$) used in §5. Proposition 8 at the end of this section furnishes a number of conditions implying ($\mathcal{K}_i$), as well as the ($\mathcal{K}_i'$) and ($\mathcal{K}_i''$) occurring later on, in the case where the order cone structure is nontrivial.

PROPOSITION 2. *If* ($\mathcal{K}_1$) *holds, then* G *is closed* (*jointly*) *and satisfies the identity*

$$G(y, v) = -F^*(v, -y). \tag{4.1}$$

If ($\mathcal{K}_2$) *holds, then* $F(x, u)$ *is closed in* u *for each* x, *and if, in addition,* h *and* P *are closed, then* F *is closed* (*jointly*) *and satisfies*

$$F(x, u) = -G^*(u, -x).$$

PROOF. We prove only the latter two assertions, as the proof of the first assertion is similar. If $x \notin C$, then $F(x,\cdot) \equiv +\infty$, which is trivially closed. If $x \in C$, then $(h + \psi_P)(x)$ is finite, and so the closedness of $F(x,\cdot)$ is equivalent to that of $(k \,\square - \psi_Q)$. But this function is closed, since by ($\mathcal{K}_2$) it is a conjugate function. To prove F is closed jointly, it suffices to establish the identity, since that exhibits F as a conjugate function. The identity itself follows from an examination of the proof of the corresponding inequality in Proposition 1. Indeed, if h and ψ_P are closed, then $h^{**} + \psi_{P^{00}} = h + \psi_P$, and from this it follows that

$$-G^*(u, -x) = +\infty = F(x, u), \qquad \forall x \notin C,$$

and

$$- G^*(u, -x) = (h + \psi_P)(x) - (k^* - \psi_{Q^*})(Ax + u), \qquad \forall x \in C.$$

Under the hypothesis $(\mathcal{H}_2)$, the desired identity then follows.

COROLLARY 2A. *If* $(\mathcal{H}_1)$ *holds, then*

$$G(y, v) = \inf_x \{K(x,y) - \langle x, v\rangle\},$$

and so, in particular,

$$\sup(\mathcal{D}) = \sup \inf(\mathcal{L}).$$

If $(\mathcal{H}_2)$ *holds, then*

$$F(x, u) = \sup_y \{K(x,y) - \langle u, y\rangle\},$$

and so, in particular,

$$\inf(\mathcal{P}) = \inf \sup(\mathcal{L}).$$

PROOF. If $(\mathcal{H}_1)$ holds, then the proposition yields (4.1). But also

$$\begin{aligned} - F^*(v, -y) &= \inf_{x,u} \{F(x, u) - \langle x, v\rangle + \langle u, y\rangle\} \\ &= \inf_x \left\{ \inf_u \{F(x, u) + \langle u, y\rangle\} - \langle x, v\rangle \right\} \\ &= \inf_x \{K(x,y) - \langle x, v\rangle\} \end{aligned}$$

by (4.2). This establishes the first identity. Next, observe (4.2) can be rewritten as

$$K(x,y) = -F(x, \cdot\,)^*(-y),$$

and so

$$\begin{aligned} \sup_y \{K(x,y) - \langle u,y\rangle\} &= \sup_y \{-F(x, \cdot\,)^*(-y) - \langle u, y\rangle\} \\ &= \sup_y \{\langle u, -y\rangle - F(x, \cdot\,)^*(-y)\} = F(x, \cdot\,)^{**}(u). \end{aligned}$$

Since the last expression equals $F(x, u)$ when $F(x, \cdot\,)$ is closed, we are done by the proposition when $(\mathcal{H}_2)$ holds.

In view of Proposition 2, we are in a position to harvest immediately from **[20]** five additional propositions containing a great deal of information concerning our trio of problems. These could, of course, be established "from scratch," working directly with the model's ingredients h, k, P, Q and A, but that would serve little purpose besides lengthening the paper.

The first of these results provides a great deal of information concerning the precise relationship between $(\mathcal{P})$ and $(\mathcal{D})$. It relates the optimal value function in one problem to the objective function in the other problem. The *primal objective function* is

$$f(x) = F(x,0) = \begin{cases} \inf\{h(x) - k(z) | Ax \geqslant z\} & \text{if } x \in C, \\ +\infty & \text{if } x \notin C, \end{cases}$$

the *dual objective function* is

$$g(y) = G(y,0) = \begin{cases} \sup\{k^*(y) - h^*(w) | w \geqslant A^*y\} & \text{if } y \in D, \\ -\infty & \text{if } y \notin D, \end{cases}$$

the *primal optimal value function* is

$$\phi(u) = \inf_x F(x,u) = \inf\{h(x) - k(z) | x \geqslant 0, Ax + u \geqslant z\}, \tag{4.4}$$

and the *dual optimal value function* is

$$\gamma(v) = \sup_y G(y,v) = \sup\{k^*(y) - h^*(w) | y \geqslant 0, w \geqslant A^*y + v\}. \tag{4.5}$$

Notice that

$$\operatorname{dom} \phi = \operatorname{dom} k + Q - AC \neq \varnothing \quad \text{and}$$
$$\operatorname{dom} \gamma = \operatorname{dom} h^* - P^* - A^*D \neq \varnothing,$$

where the nonemptiness follows from the nondegeneracy assumption (4.3).

PROPOSITION 3. (a) *Assume* $(\mathcal{H}_1)$ *holds. Then* $g = (-\phi)^*$ *and* $-g^* = \operatorname{cl} \phi$. *In particular,*

$$\sup(\mathcal{D}) = \liminf_{u \to 0} \phi(u), \tag{4.6}$$

except in the case where $0 \notin \operatorname{cl} \operatorname{dom} \phi$ *and the function* $\operatorname{lsc} \phi$ *is nowhere finite. (In the exceptional case,* $\sup(\mathcal{D})$ *is* $-\infty$, *while the limit is* $+\infty$ *and* $\operatorname{lsc} \phi \equiv -\infty$ *on* $\operatorname{dom} \operatorname{lsc} \phi = \operatorname{cl} \operatorname{dom} \phi$.)

(b) *Assume* $(\mathcal{H}_2)$ *holds and that* h *and* P *are closed. Then* $f = (-\gamma)^*$ *and* $-f^* = \operatorname{cl} \gamma$. *In particular,*

$$\inf(\mathcal{P}) = \limsup_{v \to 0} \gamma(v), \tag{4.7}$$

except in the case where $0 \notin \operatorname{cl} \operatorname{dom} \gamma$ *and the function* $\operatorname{usc} \gamma$ *is nowhere finite. (In the exceptional case,* $\inf(\mathcal{P})$ *is* $+\infty$, *while the limit is* $-\infty$ *and* $\operatorname{usc} \gamma \equiv +\infty$ *on* $\operatorname{dom} \operatorname{usc} \gamma = \operatorname{cl} \operatorname{dom} \gamma$.)

(c) *Assume* $(\mathcal{H}_1)$ *and* $(\mathcal{H}_2)$ *both hold and that* h *and* P *are closed. Then both* (4.6) *and* (4.7) *hold, except in the degenerate case in which all of the following properties are present*:

$$\phi(0) = +\infty, \quad \textit{and} \quad \operatorname{lsc} \phi \equiv -\infty \qquad \textit{on } \operatorname{dom} \operatorname{lsc} \phi = \operatorname{cl} \operatorname{dom} \phi,$$
$$\gamma(0) = -\infty, \quad \textit{and} \quad \operatorname{usc} \gamma \equiv +\infty \qquad \textit{on } \operatorname{dom} \operatorname{usc} \gamma = \operatorname{cl} \operatorname{dom} \gamma.$$

PROOF. Parts (a) and (b) follow from [**20**, Theorems 7 and 7′], together with Proposition 2. Part (c) then follows from (a) and (b), with the aid of [**20**, Theorem 4].

The next result involves in part the extremality conditions associated with the trio $(\mathcal{P})$, $(\mathcal{D})$, $(\mathcal{L})$. By virtue of the saddlepoint characterization of them which obtains usually (i.e. under a suitable constraint qualification), they can be introduced as the *abstract Kuhn-Tucker conditions*:

$$(0, 0) \in \partial K(x, y) \tag{4.8}$$

(cf. [**20**, p. 39]). The pairs (x, y) satisfying (4.8) are precisely the saddlepoints of K, i.e. the solutions of $(\mathcal{L})$.

In the case of the basic Fenchel-Rockafellar model, the Lagrangian saddle function is

$$(x, y) \to \begin{cases} h(x) + k^*(y) - \langle Ax, y\rangle & \text{if } x \in \operatorname{dom} h, \\ +\infty \quad \text{if } x \notin \operatorname{dom} h, \end{cases} \tag{4.9}$$

and from this it is easy to see, using the definitions, that conditions (4.8) are equivalent to the conditions

$$A^*y \in \partial h(x) \quad \text{and} \quad Ax \in \partial k^*(y). \tag{4.10}$$

Now in the cone-augmented model, observe that the Lagrangian saddle function K coincides with what one obtains by substituting $h + \psi_P$ and $k \,\square\, - \psi_Q$ in place of h and k, respectively, in (4.9). It follows by simple substitution in (4.10), then, that the Kuhn-Tucker conditions for the cone-augmented model are equivalent to the conditions

$$A^*y \in \partial (h + \psi_P)(x) \quad \text{and} \quad Ax \in \partial (k^* - \psi_{Q^*})(y).$$

Corollary 12A in §5 will provide a further breakdown of these conditions.

PROPOSITION 4. *If* $(\mathcal{H}_1)$ *holds, then the implications* (a) $\Leftrightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d) *hold among the conditions*:

(a) $\inf(\mathcal{P}) = \sup(\mathcal{D})$;

(b) $\phi(0) = \operatorname{cl} \phi(0)$;

(c) *the saddle value of the Lagrangian K exists*;

(d) $\gamma(0) = \operatorname{cl} \gamma(0)$.

If $(\mathcal{H}_2)$ *holds, then* (b) *and* (c) *are equivalent. If both* $(\mathcal{H}_1)$ *and* $(\mathcal{H}_2)$ *hold and* h, k, P, Q *are all closed, then* (a)–(d) *are all equivalent. Furthermore, if* $(\mathcal{H}_1)$ *holds, then the implication* (e) $\Rightarrow$ (f) *holds between the following conditions, with actual equivalence when* $(\mathcal{H}_2)$ *is also satisfied*:

(e) *x solves $(\mathcal{P})$, y solves $(\mathcal{D})$, and* $\inf(\mathcal{P}) = \sup(\mathcal{D})$;

(f) *the pair (x, y) satisfies the Kuhn-Tucker conditions.*

PROOF. By [**20**, Theorem 15], together with Proposition 2.

COROLLARY 4A. *Assume* $(\mathcal{H}_1)$ *holds and that* $\inf(\mathcal{P}) = \max(\mathcal{D})$ *(i.e.,* $\inf(\mathcal{P}) = \sup(\mathcal{D})$ *and there actually exists a solution to* $(\mathcal{D})$*). Then a necessary condition in order that x solve $(\mathcal{P})$ is that there exists a y such that (x, y) satisfies the Kuhn-Tucker conditions. This condition is also sufficient when* $(\mathcal{H}_2)$

holds or, more generally, when $F(x, 0) = \mathrm{cl}_u F(x, 0)$.

PROOF. All but the very last remark follows directly from the proposition. The sufficiency under the assumption $F(x, 0) = \mathrm{cl}_u F(x, 0)$ follows upon a closer examination of the proof of [**20**, Theorem 15].

PROPOSITION 5. *Assume* $(\mathcal{H}_1)$ *holds. Then the following conditions on a vector* y *are equivalent*:

(a) y *solves* $(\mathcal{D})$ *and* $\inf(\mathcal{P}) = \sup(\mathcal{D})$;

(b) $-y \in \partial\phi(0)$;

(c) $\inf(\mathcal{P}) = \inf_x K(x, y)$.

PROOF. By [**20**, Theorem 16], together with Proposition 2.

COROLLARY 5A. *Assume* $(\mathcal{H}_1)$ *holds and that* $\phi(0) = \inf(\mathcal{P})$ *is finite. Then the following are equivalent*:

(a) $\inf(\mathcal{P}) = \max(\mathcal{D})$;

(b) $\liminf_{u' \to u} \phi'(0; u')$ *is finite for some* u.

PROOF. By the equivalence between (a) and (b) of the proposition, together with [**20**, Theorems 11(b) and 4] applied to the function $\theta(u) = \phi'(0; u)$.

PROPOSITION 6. *Assume* $(\mathcal{H}_1)$ *holds and that* ϕ *is bounded above on a neighborhood of* 0. *Then each of the following holds*:

(a) $\inf(\mathcal{P}) = \max(\mathcal{D})$;

(b) *In fact, for every real* β *the set* $\{y \mid g(y) \geqslant \beta\}$ *is closed, bounded and convex, actually equicontinuous and hence weakly compact* (*in the weak topology induced on* Y *by* U). *Thus, every maximizing sequence for* $(\mathcal{D})$ *has weak cluster points, and every such cluster point solves* $(\mathcal{D})$.

(c) *If the common optimal value in* (a) *is not* $-\infty$, *then*

$$\phi'(0; u) = \max\{\langle u, -y\rangle \mid y \text{ solves } (\mathcal{D})\}.$$

(d) *A vector* y *solves* $(\mathcal{D})$ *uniquely if and only if* $y = -\nabla\phi(0)$, *that is,* $\phi'(0; u) = -\langle u, y\rangle$, *and in this event every maximizing sequence for* $(\mathcal{D})$ *converges weakly to* y.

(e) *The maximizing sequences for* $(\mathcal{D})$ *all actually converge in the designated topology on* Y *if and only if* ϕ *is differentiable at* 0 *in relation to that topology.*

PROOF. By [**20**, Theorem 17] and Proposition 2.

PROPOSITION 7. *Each of the following conditions is sufficient for* ϕ *to be bounded above on a neighborhood of* 0 (*and hence continuous at* 0):

(a) *There exists an* x *such that the function* $u \to F(x, u)$ *is bounded above on a neighborhood of* 0. (*Or more generally, for some continuous mapping* θ: $U \to X$ *the function* $u \to F(\theta(u), u)$ *is bounded above on a neighborhood of* 0.)

(b) $U = R^n = Y$ *and* $0 \in$ core dom ϕ (*where* "core" *denotes the algebraic interior*).

(c) X *and* U *are each Banach spaces* (*in the designated topologies*), $0 \in$ core dom ϕ, h *and* P *are closed, and* $(\mathcal{H}_2)$ *holds*.

(d) $U = R^n = Y$, *at least one of the level sets* $\{y \mid g(y) \geq \beta\}$ *is nonempty and bounded, and* $(\mathcal{H}_1)$ *holds*.

(e) *Both* $(\mathcal{H}_1)$ *and* $(\mathcal{H}_2)$ *hold,* h *and* P *are closed, and there exist a neighborhood* N *of* 0 *in* V *and a real number* β *such that the set* $\{y \mid \exists v \in N, G(y, v) > \beta\}$ *is nonempty and equicontinuous*.

PROOF. By [**20**, Theorem 18] and Proposition 2. Concerning (c), the fact that it suffices to assume just that X is Banach, rather than V, is shown in [**16**, Corollary 1].

COROLLARY 7A. *Suppose that condition* (e) *of Proposition* 7 *holds, but with equicontinuity replaced by the assumption that the closure of* $\{y \mid \exists v \in N, G(y, v) > \beta\}$ *is weakly compact. Then* ϕ *is bounded above in a neighborhood of* 0 *relative to the Mackey topology on* U, *and hence all the conclusions of Proposition* 6 *are valid if interpreted in that topology*.

PROOF. Analogous to that of [**20**, Corollary 18′A].

Dual versions of all these results from Proposition 5 onwards could also be stated. One would simply impose the blanket hypothesis that both $(\mathcal{H}_1)$ and $(\mathcal{H}_2)$ hold and that h and P are closed. These assumptions guarantee, by Proposition 2, that F and G bear the needed relationship to each other and that F is actually closed.

In Propositions 2 through 7, heavy use has been made of the conditions $(\mathcal{H}_1)$ and $(\mathcal{H}_2)$. We now give a number of sufficient conditions for these to hold. It so happens that the sufficient conditions to be given actually ensure considerably more. In particular, they ensure attainment in the convolutions appearing in $(\mathcal{H}_1)$ and $(\mathcal{H}_2)$. They also ensure closely related subdifferential formulas which, as we will see from Corollary 12A in §5, come into play in simplifying the extremality conditions for our trio of problems.

In order not to take up excessive space, we leave to the reader the easy task of adapting to the present context those conditions in the next proposition which are of the most relevance to him. To make the proposition as easily applicable to $(\mathcal{H}_2)$ as to $(\mathcal{H}_1)$ notationally, we formulate it in terms of "neutral" paired spaces Z and W. Notice that in order to guarantee $(\mathcal{H}_2)$ via these conditions, one needs k and Q to be closed. In specific instances, though, it just might be possible to argue that $(\mathcal{H}_2)$ holds by ad hoc methods not requiring this.

Conditions (f) through (i) below are stated using the recession functions $f_i^* 0^+$; these are given by

$$(f_i^*0^+)(w) = \sup\{f_i^*(w' + w) - f_i^*(w') \mid f_i^*(w') < +\infty\},$$

and serve to describe the growth behavior, or asymptotic nature, of f_i^* (cf. [**18**, §8]).

PROPOSITION 8. *Let f_1 and f_2 be extended-real-valued proper convex functions on Z. Then*

$$(f_1 + f_2)^*(w) = \min\{f_1^*(w_1) + f_2^*(w_2) \mid w = w_1 + w_2\}$$

and

$$\partial(f_1 + f_2)(z) = \partial f_1(z) + \partial f_2(z)$$

hold whenever any one of the following conditions is fulfilled:

(a) *there exists a $z \in \operatorname{dom} f_1$ in a neighborhood of which f_2 is bounded above*;

(b) $Z = R^n = W$ *and* $0 \in \operatorname{core} S$, *where*

$$S = \left\{(z_1, z_2) \mid z_i \in Z \text{ and } \varnothing \neq \bigcap_{i=1}^{2} (z_i + \operatorname{dom} f_i)\right\};$$

(c) *Z is a Banach space (in the designated topology compatible with the pairing), f_1 and f_2 are closed, and* $0 \in \operatorname{core} S$ *for S as in* (b);

(d) $Z = R^n = W$, *and for some w and some real α the set*

$$\{(w_1, w_2) \mid w_i \in \operatorname{dom} f_i^*, w = w_1 + w_2, f_1^*(w_1) + f_2^*(w_2) \leq \alpha\}$$

is nonempty and bounded;

(e) *f_1 and f_2 are closed, and for some open set M in W the set*

$$\{(w_1, w_2) \mid w_i \in \operatorname{dom} f_i^*, w_1 + w_2 \in M, f_1^*(w_1) + f_2^*(w_2) < \alpha\}$$

is nonempty and equicontinuous;

(f) $Z = R^n = W$ *and* $\varnothing \neq \bigcap_{i=1}^2 \operatorname{ri} \operatorname{dom} f_i$, *where the relative interior* "ri" *may be deleted for either index i for which f_i may happen to be polyhedral*;

(g) $Z = R^n = W$, *and for all w the condition*

$$(f_1^*0^+)(w) + (f_2^*0^+)(-w) \leq 0$$

implies

$$(f_1^*0^+)(-w) + (f_2^*0^+)(w) \leq 0;$$

(h) $Z = R^n = W$, *f_1 is polyhedral, and whenever w satisfies* $(f_1^*0^+)(w) + (f_2^*0^+)(-w) \leq 0$ *and* $(f_1^*0^+)(-w) + (f_2^*0^+)(w) > 0$ *it follows that*

$$(f_1^*0^+)(w) = (f_2^*0^+)(w);$$

(i) $Z = R^n = W$, *f_1 and f_2 are polyhedral, and*

$$(f_1^*0^+)(w) + (f_2^*0^+)(-w) \geq 0, \quad \forall w \in W.$$

PROOF. The sufficiency of conditions (a) through (e) is proved in [**20**, Theorem 20]. The refinement that reflexivity is unnecessary in (c) follows from [**16**, Corollary 1]. The proof under condition (f) follows by combining

Theorems 16.4, 20.1 and 23.8 of **[18]**. Finally, conditions (g), (h) and (i) can be seen to be the dualized versions of the three conditions contained in (f). For (g), one uses Corollary 16.2.2 of **[18]** directly. For (h), one uses Theorem 13.3 and Corollary 20.2.1 of **[18]**, reformulating in slightly weaker form the condition given in Corollary 20.2.1. (The condition as stated there is only sufficient, not necessary and sufficient.) For (i), we shall sketch the proof that, when both f_1 and f_2 are polyhedral, the condition $\varnothing \neq \operatorname{dom} f_1 \cap \operatorname{dom} f_2$ is equivalent to $(f_1^* 0^+)(w) + (f_2^* 0^+)(-w) \geqslant 0$. By using the same technique employed in the proof of Corollary 16.2.2 of **[18]**, we see that it suffices to establish a polyhedral version of Lemma 16.2 of **[18]**. That is, it suffices to show that if L is a subspace of R^n and f is a proper polyhedral convex function on R^n, then

$$\varnothing \neq L \cap \operatorname{dom} f \quad \Leftrightarrow \quad (f^* 0^+)(w) \geqslant 0, \quad \forall w \in L^{\perp}.$$

But this can be established by modifying the proof of Lemma 16.2 of **[18]**, appealing to the polyhedral separation theorem [**18**, Theorem 20.2] in place of the general finite-dimensional one [**18**, Theorem 11.3]. This concludes the proof of the proposition.

5. The meaning of the suboptimizations over auxiliary variables. In this section we treat in some detail the issue of the suboptimizations occurring in $(\mathcal{P})$ and $(\mathcal{D})$. It will be shown that $(\mathcal{P})$, $(\mathcal{D})$ and even $(\mathcal{L})$ actually arise as the projections, both individually and as an optimization trio collectively, of another optimization trio. The precise connection between the optimal values and solutions for the three pairs of corresponding problems is given, as well as a comparison of the respective extremality conditions. The primal and dual problems of the new trio involve no suboptimizations. This feature comes at the expense, however, of having essentially twice as many problem variables and perturbation variables. Still, since the new, expanded problems involve a fuller class of perturbations, and hence entail additional sensitivity information, there may be situations in which one might prefer to work with the new problem trio instead.

The new primal problem is

$$\min_{x \geqslant 0, Ax \geqslant z} \{h(x) - k(z)\} \quad \text{(to find } both\ x \text{ and } z), \qquad (\mathcal{P}_1)$$

and the new dual problem is

$$\max_{y \geqslant 0, w \geqslant A^* y} \{k^*(y) - h^*(w)\} \quad \text{(to find } both\ y \text{ and } w). \qquad (\mathcal{D}_1)$$

These are the symmetric dual problems introduced in 1964 by Rockafellar **[21]** in the more restrictive setting of finite-dimensional spaces and polyhedral order cones.

First, we establish the connections between these and our earlier problems

$(\mathcal{P})$ and $(\mathcal{D})$, in which the variables z and w are merely auxiliary. Half of this correspondence requires the following strengthened forms of $(\mathcal{H}_1)$ and $(\mathcal{H}_2)$:

$$(h + \psi_P)^*(v) = \min\{h^*(v_1) + \psi_{P^0}(v_2) | v_1 + v_2 = v\}, \qquad (\mathcal{H}_1')$$

$$(k^* - \psi_{Q^*})(u) = \max\{k(u_1) - \psi_Q(u_2) | u_1 + u_2 = u\}. \qquad (\mathcal{H}_2')$$

Notice that a variety of conditions sufficient for each of these to hold is furnished by Proposition 8.

PROPOSITION 9. *The optimal values satisfy* $\inf(\mathcal{P}) = \inf(\mathcal{P}_1)$ *and* $\sup(\mathcal{D}) = \sup(\mathcal{D}_1)$. *If* (x, z) *solves* $(\mathcal{P}_1)$, *then* x *solves* $(\mathcal{P})$. *Conversely, if* x *solves* $(\mathcal{P})$ *and if* $(\mathcal{H}_2')$ *holds, then there exists a* z *such that* (x, z) *solves* $(\mathcal{P}_1)$. *If* (y, w) *solves* $(\mathcal{D}_1)$, *then* y *solves* $(\mathcal{D})$. *Conversely, if* y *solves* $(\mathcal{D})$ *and if* $(\mathcal{H}_1')$ *holds, then there exists a* w *such that* (y, w) *solves* $(\mathcal{D}_1)$.

PROOF. It is straightforward to check that

$$\inf_{x \geqslant 0} \left\{ \inf_{Ax \geqslant z} \{h(x) - k(z)\} \right\} = \inf_{x \geqslant 0, Ax \geqslant z} \{h(x) - k(z)\},$$

so the optimal values in $(\mathcal{P})$ and $(\mathcal{P}_1)$ agree. It follows that if (x, z) yields attainment of the infimum on the right, then x yields attainment of the outer infimum on the left. Now suppose, conversely, that $\bar{x}$ yields attainment of the outer infimum on the left and that $(\mathcal{H}_2')$ holds. Then

$$\begin{aligned} \inf(\mathcal{P}_1) &= \inf_{A\bar{x} \geqslant z} \{h(\bar{x}) - k(z)\} = h(\bar{x}) - \sup_{A\bar{x} \geqslant z} \{k(z)\} \\ &= h(\bar{x}) - \sup_z \{k(A\bar{x} - z) - \psi_Q(z)\}. \end{aligned}$$

By $(\mathcal{H}_2')$, the latter supremum is actually attained at some $\tilde{z}$, where without loss of generality (considering the possibility that the supremum is $-\infty$ trivially) we can assume that $\tilde{z} \in Q$. Hence,

$$\inf(\mathcal{P}_1) = h(\bar{x}) - k(A\bar{x} - \tilde{z})$$

for some $\tilde{z} \in Q$. This shows that the pair $(\bar{x}, \bar{z})$, where $\bar{z} = A\bar{x} - \tilde{z}$, solves $(\mathcal{P}_1)$. The assertions concerning $(\mathcal{D})$ and $(\mathcal{D}_1)$ are established similarly.

COROLLARY 9A. *If* $\min(\mathcal{P}_1) = \max(\mathcal{D}_1)$, *then* $\min(\mathcal{P}) = \max(\mathcal{D})$. *The converse is valid when both* $(\mathcal{H}_1')$ *and* $(\mathcal{H}_2')$ *hold.*

Our next aim is to show that $(\mathcal{D}_1)$ is indeed a dual of $(\mathcal{P}_1)$ in the sense of Rockafellar's perturbational duality theory **[18]**, **[20]**, and to identify, moreover, the saddlepoint problem $(\mathcal{L}_1)$ corresponding to $(\mathcal{P}_1)$ and $(\mathcal{D}_1)$. For this, it is necessary to parametrize $(\mathcal{P}_1)$ "convexly" in such a way that the general theory in **[20]** yields $(\mathcal{D}_1)$ as the dual problem. To do this, we shall exhibit functions F_1, G_1 and K_1 satisfying identities analogous to the key identities (4.1) and (4.2).

Define F_1 on $X \times U \times X \times U$, G_1 on $V \times Y \times V \times Y$, and K_1 on $X \times U \times V \times Y$ by means of

$$F_1(x, z, s, u) = h(x) + \psi_Q(z) + \psi_P(x + s) - k(Ax - z + u),$$
$$G_1(w, y, v, t) = k^*(y) - \psi_{P^*}(w) - \psi_{Q^*}(y - t) - h^*(A^*y + w + v),$$

and

$$K_1(x, z, w, y) = [h(x) + \psi_Q(z)] + [k^*(y) - \psi_{P^*}(w)] - [\langle x, w\rangle + \langle Ax, y\rangle - \langle z, y\rangle]$$

if $(x, z) \in \operatorname{dom} h \times Q$ and $K_1(x, z, w, y) = +\infty$ if $(x, z) \notin \operatorname{dom} h \times Q$. It is easy to see that

$$(\mathcal{P}_1) \quad \text{is} \quad \min_{x,z} F_1(x, z, 0, 0), \quad (\mathcal{D}_1) \quad \text{is} \quad \max_{y,w} G_1(w, y, 0, 0),$$

and that K_1 determines a saddlepoint problem over the product set $(\operatorname{dom} h \times Q) \times (P^* \times \operatorname{dom} k^*)$. We define this saddlepoint problem to be $(\mathcal{L}_1)$.

The fact that $(\mathcal{D}_1)$ is actually the dual of $(\mathcal{P}_1)$ and that $(\mathcal{L}_1)$ is the associated Lagrangian minimax problem follows from **[20]** and the two identities to be established now.

PROPOSITION 10. *The functions F_1, G_1 and K_1 satisfy the identities*

$$G_1(w, y, v, t) = -F_1^*(v, t, -w, -y)$$

and

$$K_1(x, z, w, y) = \inf_{s,u} \{F_1(x, z, s, u) + \langle (s, u), (w, y)\rangle\}.$$

PROOF. These can be derived most easily in stages. First, suppose q is some function convex on $X \times U$ and that $B: X \to U$ is a continuous linear operator. If H is defined as

$$H(x, u) = q(x, u + Bx),$$

it is routine to check that

$$-H^*(v, -y) = -q^*(v + B^*y, -y)$$

and

$$\inf_u \{H(x, u) + \langle u, y\rangle\} = \inf_u \{q(x, u) + \langle u, y\rangle\} - \langle Bx, y\rangle.$$

Now apply these identities with the choices

$$q(x, z; s, u) = h(x) + \psi_Q(z) + \psi_P(s) - k(u)$$

and

$$B(x, z) = (x, Ax - z).$$

It is routine also to check that q^* and B^* are given by

$$q^*(v, t; w, y) = h^*(v) + \psi_{Q^\circ}(t) + \psi_{P^\circ}(w) - k^*(-y)$$

and

$$B^*(w,y) = (A^*y + w, -y).$$

Upon substitution we see that H becomes F_1, etc., and the desired identities follow from those already noted for H.

COROLLARY 10A. *One always has the estimate* $F_1(x, z, 0, 0) \geq G_1(w, y, 0, 0)$. *Moreover, equality is attained by pairs* (x, z) *and* (w, y) *if and only if* (x, z) *solves* $(\mathcal{P}_1)$, (w, y) *solves* $(\mathcal{D}_1)$, *and* $\inf(\mathcal{P}_1) = \sup(\mathcal{D}_1)$.

PROOF. Similar to that of Corollary 1A.

We turn now to the connection between the two Lagrangian problems $(\mathcal{L})$ and $(\mathcal{L}_1)$.

PROPOSITION 11. *Assume* P *is closed. Then*

$$\sup\inf K_1 \leq \sup\inf K \leq \inf\sup K \leq \inf\sup K_1.$$

If, in addition, $(\mathcal{H}_1)$ *and* $(\mathcal{H}_2)$ *hold, then the two lower saddle values agree and the two upper saddle values agree.*

PROOF. The middle inequality is always true, as can be easily verified. The assertions about the "sup inf" expressions follow from the computations

$$\begin{aligned}
\sup_{w,y}\inf_{x,z} K_1 &= \sup_{\substack{w\in P^*\\ y\in\operatorname{dom}k^*}}\ \inf_{\substack{z\in Q\\ x\in\operatorname{dom}h}} \{h(x) + k^*(y) - x\cdot w - Ax\cdot y + z\cdot y\}\\
&= \sup_{\substack{w\in P^*\\ y\in\operatorname{dom}k^*}} \left\{k^*(y) - \sup_{x\in\operatorname{dom}h}\left\{x\cdot(A^*y + w) - h(x) + \sup_{z\in Q} z\cdot(-y)\right\}\right\}\\
&= \sup_{y\in Q^*\cap\operatorname{dom}k^*}\ \sup_{w\in P^*} \{k^*(y) - h^*(A^*y + w)\}\\
&= \sup_{y\in D} \{(k^* - \psi_{Q^*})(y) - (h^* \,\square\, \psi_{P^0})(A^*y)\}
\end{aligned}$$

and

$$\begin{aligned}
\sup_y\inf_x K &= \sup_{y\in D}\ \inf_{x\in C} \{h(x) + k^*(y) - \langle Ax, y\rangle\}\\
&= \sup_{y\in D}\left\{k^*(y) - \sup_{x\in C}\{\langle x, A^*y\rangle - h(x)\}\right\}\\
&= \sup_{y\in D}\{(k^* - \psi_{Q^*})(y) - (h + \psi_P)^*(A^*y)\}
\end{aligned}$$

(where we have abbreviated the bilinear pairing function by the "dot" notation to shorten the formulas). Since in general $(h + \psi_P)^* \leq (h^* \,\square\, \psi_{P^0})$, it follows that $\sup\inf K_1 \leq \sup\inf K$ always holds, while if $(\mathcal{H}_1)$ holds we actually have equality. For the "inf sup" assertions we proceed similarly, calculating that

$$\begin{aligned}
\inf_{x,z}\sup_{w,y} K_1 &= \inf_{\substack{z\in Q\\ x\in\operatorname{dom} h}} \sup_{\substack{w\in P^*\\ y\in\operatorname{dom} k^*}} \{h(x)+k^*(y)-x\cdot w - Ax\cdot y + z\cdot y\}\\
&= \inf_{\substack{z\in Q\\ x\in\operatorname{dom} h}} \left\{h(x) - \inf_{y\in\operatorname{dom} k^*}\left\{(Ax-z)\cdot y - k^*(y) + \inf_{w\in P^*} x\cdot w\right\}\right\}\\
&= \inf_{x\in(\operatorname{cl} P)\cap\operatorname{dom} h}\ \inf_{z\in Q}\{h(x)-(\operatorname{cl} k)(Ax-z)\}\\
&= \inf_{x\in(\operatorname{cl} P)\cap\operatorname{dom} h}\{(h+\psi_{\operatorname{cl} P})(x) - (\operatorname{cl} k\,\square - \psi_Q)(Ax)\}
\end{aligned}$$

and

$$\begin{aligned}
\inf_x\sup_y K &= \inf_{x\in C}\sup_{y\in D}\{h(x)+k^*(y)-\langle Ax,y\rangle\}\\
&= \inf_{x\in C}\left\{h(x) - \inf_{y\in D}\{\langle Ax,y\rangle - k^*(y)\}\right\}\\
&= \inf_{x\in C}\{(h+\psi_P)(x) - (k^*-\psi_{Q^*})^*(Ax)\}.
\end{aligned}$$

Now in general one has

$$(k^*-\psi_{Q^*})^* = \operatorname{cl}(k\,\square - \psi_Q) \geqslant (\operatorname{cl} k\,\square - \psi_Q) \geqslant (k\,\square - \psi_Q).$$

Thus, if P is closed it follows that

$$\inf\sup K_1 \geqslant \inf\sup K.$$

If, in addition, $(\mathcal{H}_2)$ holds, then $(k^*-\psi_{Q^*})^* = (\operatorname{cl} k\,\square - \psi_Q)$, and so we actually have equality.

COROLLARY 11A. *Assume P is closed. If the saddle value in $(\mathcal{L}_1)$ exists, then so does the saddle value in $(\mathcal{L})$ and the two values coincide. Conversely, if the saddle value of $(\mathcal{L})$ exists and $(\mathcal{H}_1)$ and $(\mathcal{H}_2)$ hold, then the saddle value of $(\mathcal{L}_1)$ exists and the two values coincide.*

The connection between $\inf(\mathcal{P})$, $\sup(\mathcal{D})$ and the lower and upper saddle values of $(\mathcal{L})$ was noted in Corollary 2A. The parallel assertions concerning $(\mathcal{P}_1)$, $(\mathcal{D}_1)$ and $(\mathcal{L}_1)$ are that

$$\inf(\mathcal{P}_1) = \inf\sup(\mathcal{L}_1) \quad \text{if } P \text{ and } k \text{ are closed,}$$

while

$$\sup(\mathcal{D}_1) = \sup\inf(\mathcal{L}_1)$$

always holds. These facts follow from the identities of Proposition 10 and [**20**, pp. 18–19].

Now consider the abstract Kuhn-Tucker conditions corresponding to $(\mathcal{P}_1)$, $(\mathcal{D}_1)$ and $(\mathcal{L}_1)$: $(0,0,0,0)\in\partial K_1(x,z,y,w)$. These are characterized and contrasted with those corresponding to $(\mathcal{P})$, $(\mathcal{D})$ and $(\mathcal{L})$ in the next result and its corollary.

Proposition 12. *Vectors x and y satisfy* $(0, 0) \in \partial K(x, y)$ *if and only if*

$$A^*y \in \partial(h + \psi_P)(x) \quad \text{and} \quad Ax \in \partial(k^* - \psi_{Q^*})(y). \tag{5.1}$$

Vectors x, z, w, y satisfy $(0, 0, 0, 0) \in \partial K_1(x, z, w, y)$ *if and only if*

$$x \in \operatorname{cl} P, \quad w \in P^*, \quad \langle x, w\rangle = 0, \tag{5.2}$$

$$z \in Q, \quad y \in Q^*, \quad \langle z, y\rangle = 0, \tag{5.3}$$

$$A^*y + w \in \partial h(x) \quad \text{and} \quad Ax - z \in \partial k^*(y). \tag{5.4}$$

Proof. The first assertion was established following (4.10). The same approach used there can be used to obtain the second assertion. Namely, in (4.10) replace h by the function $(x, z) \to h(x) + \psi_Q(z)$, replace k by the function $(s, u) \to -\psi_P(s) + k(u)$, and replace A by the linear transformation $(x, z) \to (x, Ax - z)$. Then A^* becomes the transformation $(w, y) \to (A^*y + w, -y)$ and k^* becomes the function $(w, y) \to (-\psi_P)^*(w) + k^*(y)$. Substitution of these ingredients into (4.10) yields that $(0, 0, 0, 0) \in \partial K_1(x, z, w, y)$ if and only if

$$A^*y + w \in \partial h(x), \quad Ax - z \in \partial k^*(y),$$
$$-y \in \partial\psi_Q(z), \quad x \in \partial(-\psi_P)^*(w).$$

Finally, it is routine to show that $x \in \partial(-\psi_P)^*(w)$ is the same as (5.2) and that $-y \in \partial\psi_Q(z)$ is the same as (5.3).

The converse part of the following corollary establishes hypotheses under which the extremality conditions (5.1) can be broken down into (5.2)–(5.4). Its statement utilizes the conditions

$$\partial(h + \psi_P)(x) = \partial h(x) + \partial\psi_P(x), \tag{$\mathcal{H}_1''$}$$

$$\partial(k^* - \psi_{Q^*})(y) = \partial k^*(y) - \partial\psi_{Q^*}(y). \tag{$\mathcal{H}_2''$}$$

Notice that Proposition 8 provides a variety of sufficient conditions ensuring that these hold. (In fact, it can be shown that, for $i = 1$ and 2, condition $(\mathcal{H}_i')$ implies condition $(\mathcal{H}_i'')$.)

Corollary 12A. *If (x, z, w, y) solves $(\mathcal{L}_1)$ and P is closed, then (x, y) solves $(\mathcal{L})$. Conversely, if (x, y) solves $(\mathcal{L})$, Q is closed, and both $(\mathcal{H}_1'')$ and $(\mathcal{H}_2'')$ hold, then there exist vectors z and w such that (x, z, w, y) solves $(\mathcal{L}_1)$.*

Proof. Suppose first that (x, z, w, y) solves $(\mathcal{L}_1)$. Then the proposition yields (5.2)–(5.4). Now (5.2) is the same as $-w \in \partial\psi_P(x)$ when P is closed, and (5.3) implies $-z \in \partial\psi_{Q^*}(y)$. In view of the general inclusions

$$\partial h(x) + \partial\psi_P(x) \subset \partial(h + \psi_P)(x), \quad \partial k^*(y) - \partial\psi_{Q^*}(y) \subset \partial(k^* - \psi_{Q^*})(y),$$

(5.4) then yields (5.1). Now suppose conversely that (x, y) solves $(\mathcal{L})$. If $(\mathcal{H}_1'')$ and $(\mathcal{H}_2'')$ both hold, then by (5.1) we know there exist vectors z and w such that $-w \in \partial\psi_P(x)$, $-z \in \partial\psi_{Q^*}(y)$, and (5.4) hold. Now $-w \in \partial\psi_P(x)$ implies

(5.2), and $-z \in \partial\psi_{Q^*}(y)$ is the same as (5.3) when Q is closed. Hence the proposition implies, in the presence of the assumptions mentioned, that (x, z, w, y) solves $(\mathcal{L}_1)$.

It is interesting to examine the primal and dual optimal value functions associated with the duality between $(\mathcal{P}_1)$ and $(\mathcal{D}_1)$:

$$\begin{aligned}\phi_1(s, u) &= \inf_{x,z} F_1(x, z, s, u)\\ &= \inf\{h(x) - k(z) | x + s \geqslant 0, Ax + u \geqslant z\}\end{aligned}$$

and

$$\begin{aligned}\gamma_1(v, t) &= \sup_{w,y} G_1(w, y, v, t)\\ &= \sup\{k^*(y) - h^*(w) | y \geqslant t, w \geqslant A^*y + v\}.\end{aligned}$$

Comparison of these with ϕ and γ, given in (4.4) and (4.5), shows that the new trio of problems, involving more explicit variables, is based on perturbing the "nonnegativity" constraints in addition to the "linear inequality" constraints. Thus, the sensitivity information given by the model with more (primal perturbation) variables is more comprehensive, though at the price of having to deal with additional (dual problem) variables.

One might reasonably inquire, therefore, which model problem, $(\mathcal{P})$ or $(\mathcal{P}_1)$, is the more appropriate on which to focus the main attention. The answer would seem to depend mainly on the role being played by the nonnegativity constraints in the actual problem being studied. If, as is quite often the case, the nonnegativity constraints are deemed relatively inviolable, then the original cone-augmented problem $(\mathcal{P})$ seems indicated. On the other hand, if one has some need for sensitivity information concerning discrepancies in satisfying the nonnegativity constraints, and is at the same time willing to admit into the calculations additional explicit dual problem variables, then the model involving $(\mathcal{P}_1)$ seems indicated. Either $(\mathcal{P})$ or $(\mathcal{P}_1)$ will, of course, yield sensitivity information concerning the general linear inequality constraints.

The reader can formulate with little difficulty the parallel versions of Propositions 3 through 7 for the trio $(\mathcal{P}_1)$, $(\mathcal{D}_1)$, $(\mathcal{L}_1)$.

6. One optimization trio as the image of another. Results in the last section show that $(\mathcal{P})$, $(\mathcal{D})$ and $(\mathcal{L})$, taken individually, are the "projections" of $(\mathcal{P}_1)$, $(\mathcal{D}_1)$ and $(\mathcal{L}_1)$. Now we show that, in a precise sense, the entire optimization *trio* $(\mathcal{P})$, $(\mathcal{D})$, $(\mathcal{L})$ is the image of the trio $(\mathcal{P}_1)$, $(\mathcal{D}_1)$, $(\mathcal{L}_1)$ under a certain type of projection transformation. This will yield an alternate formulation of the question whether identity (4.1) holds, one which reveals much more clearly the connection with the suboptimizations occurring in F and G. The transformations involved are more general than the ordinary projection linear transformations; namely, they are particularly simple types of convex processes (see [18, §39] for definitions). The "self-dual" form of these transformations will help to explain why the extreme symmetry possessed by the trio $(\mathcal{P}_1)$, $(\mathcal{D}_1)$, $(\mathcal{L}_1)$ is passed on intact to the image trio $(\mathcal{P})$, $(\mathcal{D})$, $(\mathcal{L})$.

Consider first the convex process $M: X \times U \times X \times U \to X \times U$ defined by

$$M(x, z, s, u) = \begin{cases} \{(x, u)\} & \text{if } s = 0, \\ \varnothing & \text{if } s \neq 0, \end{cases}$$

where we assign M the supremum orientation. It is not hard to see that the sup-oriented convex process $M^{*-1}: V \times Y \times V \times Y \to V \times Y$ is given by

$$M^{*-1}(v, t, w, y) = \begin{cases} \{(v, y)\} & \text{if } t = 0, \\ \varnothing & \text{if } t \neq 0. \end{cases}$$

Now let N be the convex process having the same graph as M^{*-1} but which is assigned instead the infimum orientation. Then the image of F_1 under M is F, while the image of G_1 under N is G. Indeed,

$$(MF_1)(x, u) = \inf_{M^{-1}(x,u)} F_1 = \inf_z \{F_1(x, z, 0, u)\} = F(x, u),$$

and

$$(NG_1)(v, y) = \sup_{N^{-1}(v,y)} G_1 = \sup_w \{G_1(w, y, v, 0)\} = G(y, v),$$

where the last equalities in each case are straightforward calculations. From these identities, we see that it is the self-dual structure of M that permits the symmetry between F_1 and G_1 to be passed on to F and G.

One also has that

$$\begin{aligned} -G(-y, v) &= -\sup_w \{G_1(w, -y, v, 0)\} = \inf_w \{-G_1(-w, -y, v, 0)\} \\ &= \inf_w \{F_1^*(v, 0, w, y)\} = \inf_{M^*(v,y)} F_1^* = (M^{*-1}F_1^*)(v, y). \end{aligned}$$

Together with the earlier identity $MF_1 = F$, this shows that the underlying, central issue of

$$G(y, v) \overset{?}{=} -F^*(v, -y)$$

is actually equivalent to the question

$$M^{*-1}F_1^* \overset{?}{=} (MF_1)^*.$$

The problem of providing general conditions ensuring that such a duality identity holds between a convex function and a convex process was first dealt with by Rockafellar [18, Theorem 39.7]. The infinite-dimensional case is treated in McLinden [10], where the associated subdifferential formula is also derived.

Finally, let us see how the passage from $(\mathfrak{L}_1)$ to $(\mathfrak{L})$ can be viewed in similar terms. For this, regard the ordinary projection $L_1(x, z) = x$ as a sup-oriented convex process and likewise regard the projection $L_2(w, y) = y$ as an inf-oriented convex process. It can be checked that each of the two saddle functions

$$(x,y) \to \inf_{L_1^{-1}x} \sup_{L_2^{-1}y} K_1 \quad \text{and} \quad (x,y) \to \sup_{L_2^{-1}y} \inf_{L_1^{-1}x} K_1$$

has effective domain $C \times D$ and agrees with K there. It follows that each of these two possible ways of forming the image of K_1 under the "product convex process" $L = L_1 \times L_2$ gives rise to $(\mathfrak{L})$. Notice also that M and N can be expressed in terms of L_1 and L_2 as

$$M = L_1 \times L_2^{*-1} \quad \text{and} \quad N = L_1^{*-1} \times L_2,$$

provided we interpret L_2^{*-1} as sup-oriented and L_1^{*-1} as inf-oriented.

The above furnishes an outline which could clearly be extended to product convex processes more general than M, resulting in a general treatment of the effects of "projecting" one trio of optimization problems onto another. Of particular interest would be the associated suboptimizations (the counterpart of the z's and w's here) and also the ways in which various classes of perturbations would be transformed. For another instance of one entire optimization trio projecting onto another, see [11, especially §6].

7. Special functional structure. In this section we give conjugacy and subdifferential formulas for certain useful objective function structure. This information, when combined with the symmetric cone constrained duality model developed above, provides the basis for treating symmetrically a variety of specific classes of model problems. In §3 it was indicated how one would obtain quadratic and homogeneous programming. The same approach applies, using the tools which follow, to provide duality treatment for functions homogeneous of *any* nonzero degree. (Of course in general, one need not necessarily pick the family $\mathfrak{B}$ to satisfy $\mathfrak{B} = -\mathfrak{A}^*$. See §3.)

The class of functions to be treated first (in Proposition 13 below) may be heuristically viewed as a rather broad generalization of convex "distance" functions, such as the function $x \to \frac{1}{2}\|x\|^2$ in a normed space. Here, the role of the norm will be played by the gauge of a convex set containing the origin, and the role of the quadratic "scaling" function will be taken over by a generalized Young's function on the halfline. Before giving the result, we recall from Rockafellar [18] certain relevant material.

The *polar* of a nonempty convex subset C of X is

$$C^0 = \{v | \langle x, v\rangle \leqslant 1, \forall x \in C\}.$$

It is a basic fact that C^{00} is the closure of the convex hull of C and the origin. Consequently, if C is itself closed and contains the origin, then $C^{00} = C$. We shall call such a set a *polar set*. A *gauge* on X is a positively homogeneous proper convex function on X which is nonnegative and vanishes at the origin. The *polar* of a gauge γ is the function

$$\gamma^0(v) = \inf\{0 \leqslant \mu < +\infty \mid \langle x, v\rangle \leqslant \mu \cdot \gamma(x), \forall x\}.$$

This is a closed gauge on V, and one has $\gamma^{00} = \operatorname{cl} \gamma$. The closed gauges on X are in one-to-one correspondence with the polar sets in X via

$$\gamma(x) = \inf\{0 \leqslant \mu < +\infty \mid x \in \mu C\} \quad \text{and} \quad C = \{x \mid \gamma(x) \leqslant 1\}.$$

It is convenient to write γ_C for the gauge corresponding in this way to a polar set C, and one has $(\gamma_C)^0 = \gamma_{C^0}$. A function f: $X \to [-\infty, +\infty]$ is called *gaugelike* provided $f(0) = \inf_X f < +\infty$ and the various level sets $\{x \mid f(x) \leqslant \alpha\}$, $f(0) < \alpha < +\infty$, are all proportional (i.e. can all be expressed as positive scalar multiples of a single set). If ϕ: $[0, +\infty) \to (-\infty, +\infty]$ is a nondecreasing convex function which is finite at zero, then its *monotone conjugate* is the function ϕ^+: $[0, +\infty) \to (-\infty, +\infty]$ given by

$$\phi^+(\tau) = \sup_{0 \leqslant \sigma < +\infty} \{\sigma\tau - \phi(\sigma)\}.$$

This is another such function, and moreover it is lower semicontinuous. One has $\phi^{++} = \operatorname{lsc} \phi$, the lower semicontinuous hull of ϕ. If ϕ is nonconstant and finite somewhere on $(0, +\infty)$, then ϕ^+ has these same properties. From the definition of ϕ^+ it follows that one always has

$$\phi(\sigma) + \phi^+(\tau) \geqslant \sigma\tau, \qquad \forall \sigma, \tau \in [0, +\infty).$$

The *subdifferential* of ϕ is the multivalued mapping $\partial\phi$: $[0, +\infty) \to [0, +\infty)$ defined by

$$\tau \in \partial\phi(\sigma) \quad \Leftrightarrow \quad \phi(\sigma) + \phi^+(\tau) = \sigma\tau.$$

The set $\partial\phi(\sigma)$ amounts to simply the derivative at points σ where ϕ is differentiable. Monotone conjugacy for functions on $[0, +\infty)$ may be viewed as the natural generalization of the classical facts concerning Young's functions. We are now in a position to present the first result.

PROPOSITION 13. *A function f is a gaugelike closed proper convex function if and only if it can be expressed in the form $f(x) = \phi(\gamma(x))$, where γ is a closed gauge and ϕ is a nondecreasing, lower semicontinuous convex function on* $[0, +\infty)$ *which is nonconstant and finite somewhere on* $(0, +\infty)$. (*Here $\phi(+\infty)$ is to be interpreted as $+\infty$ in the formula for f.*) *In this case f^* is gaugelike, too, and in fact*

$$f^*(v) = \phi^+(\gamma^0(v)),$$

where ϕ^+ satisfies the same conditions as ϕ. Moreover, one then has $v \in \partial f(x)$ if and only if

$$\langle x, v\rangle = \gamma(x)\cdot\gamma^0(v) \quad \textit{and} \quad \gamma^0(v) \in \partial\phi(\gamma(x)).$$

Also, in this event the sets

$$\{x|\gamma(x) \leqslant 1\}, \quad \{v|\gamma^0(v) \leqslant 1\}$$

are polar to each other.

PROOF. See Rockafellar [**18**, Theorem 15.3] for the original, finite-dimensional presentation of this result. The proof given there extends, in broad outline form, to the general case; see McLinden [**9**] for details in the infinite-dimensional case.

"Quadratic" convex functions are included in the case $p = 2$ of the following corollary. The case in which γ is a norm is also quite important.

COROLLARY 13A. *A function f is closed proper convex and positively homogeneous of degree p, where $1 < p < +\infty$, if and only if it is of the form $f(x) = (1/p)\gamma(x)^p$ for some closed gauge γ. In this case f^* is positively homogeneous of degree q, where $1/p + 1/q = 1$; in fact,*

$$f^*(v) = (1/q)\gamma^0(v)^q.$$

One then has

$$\langle x, v\rangle \leqslant \left[pf(x)\right]^{1/p}\left[qf^*(v)\right]^{1/q}, \qquad \forall x \in \operatorname{dom} f, \forall v \in \operatorname{dom} f^*,$$

and the sets $\{x|f(x) \leqslant 1/p\}$, $\{v|f^*(v) \leqslant 1/q\}$ are polar to each other. Moreover, $v \in \partial f(x)$ if and only if

$$\langle x, v\rangle = \gamma(x)\cdot\gamma^0(v) \quad \textit{and} \quad \gamma^0(v) = \gamma(x)^{p-1}.$$

PROOF. If $f(x) = (1/p)\gamma(x)^p$ for some $1 < p < +\infty$ and closed gauge γ, then $f = \phi \circ \gamma$, where $\phi(\sigma) = (1/p)\sigma^p$ has the properties described in the proposition, and hence f is closed proper convex. On the other hand, if f is closed proper convex and positively homogeneous of degree p for $1 < p < +\infty$, one can deduce that $0 = f(0) = \inf_X f$, the set $C = \{x|f(x) \leqslant 1/p\}$ is closed convex containing the origin, and that

$$\{x|f(x) \leqslant \alpha\} = (\alpha p)^{1/p}C = \{x|\phi \circ \gamma_C(x) \leqslant \alpha\}$$

for every $0 < \alpha < +\infty$, where $\phi(\sigma) = (1/p)\sigma^p$. It follows that $f = \phi \circ \gamma_C$. The remainder of the corollary now follows from the proposition, upon noticing that for this ϕ one has $\phi^+(\tau) = (1/q)\tau^q$ for $1/p + 1/q = 1$.

Ordinary positive homogeneity (i.e. of degree $p = 1$) for convex functions is handled by the basic correspondence relating indicator functions to support functions, which is readily accessible from the general conjugacy correspondence. (See also [**18**, Theorem 13.2].) The next corollary deals with nonnegativity *combined* with positive homogeneity. The case in which C and C^0 are the unit balls of polar norms is of particular importance.

COROLLARY 13B. *Let C be a polar set. Then $\gamma_C = \psi^*_{C^0}$ and $\gamma^*_C = \psi_{C^0}$. Moreover, $v \in \partial\gamma_C(x)$ if and only if $\langle x, v\rangle = \gamma_C(x) \cdot \gamma_{C^0}(v)$ and either $\gamma_C(x) = 0$ and $\gamma_{C^0}(v) \leqslant 1$ or else $\gamma_C(x) > 0$ and $\gamma_{C^0}(v) = 1$.*

PROOF. The proposition applied to $f = \phi \circ \gamma_C$, where $\phi(\sigma) = \sigma$, yields $f^* = \phi^+ \circ \gamma_{C^0}$. Since $\phi^+(\tau) = \psi_{[0,1]}(\tau)$, this means $\gamma^*_C = \psi_{C^0}$. Taking conjugates then yields the first identity, since γ_C is closed. The subdifferential characterization follows by the proposition and the fact that $\tau \in \partial\phi(\sigma)$ if and only if either $\sigma = 0$ and $0 \leqslant \tau \leqslant 1$ or else $\sigma > 0$ and $\tau = 1$.

We have promised coverage of positive homogeneity of any nonzero degree p. For $p < 1$, this clearly puts us most naturally into the realm of concave functions. What we shall do is present a natural, concave analogue of Proposition 13. It turns out that the class of functions it treats can be heuristically viewed as concave "utility" functions, as contrasted with the convex "distance" functions of the previous discussion. To formulate this concave analogue (Proposition 14), we require appropriate analogues of the concepts and facts used earlier. For more on the following material, see McLinden [**9**].

Let C be a nonempty convex subset of X such that $0 \notin \text{cl}\, C$. The *antipolar* of C is

$$C^0 = \{v | \langle x, v\rangle \geqslant 1, \forall x \in C\}.$$

(Here, and below, we rely on the context to indicate whether the symbols ^{0}, $^+$, *, and cl are to be interpreted in the earlier, convex sense or in the present, concave sense.) It is a basic fact that C^{00} is the set $\text{cl}\{\lambda x | \lambda \geqslant 1, x \in C\}$. Hence, $C^{00} = C$ if and only if C is a nonempty closed convex set which excludes the origin and satisfies the condition

$$\lambda x \in C \quad \text{for all } \lambda \geqslant 1 \text{ and } x \in C.$$

We call such a set an *antipolar set*. (A particularly nice example of an antipolar set in $X = R^n$ is the "hyperbolic" set $C = \{(\xi_1, \ldots, \xi_n) | \xi_i > 0 \text{ each } i, \text{ and } \xi_1\xi_2 \cdots \xi_n \geqslant n^{-n/2}\}$. This has the self-dual property $C^0 = C$.) An *antigauge* on X is a positively homogeneous proper concave function on X which vanishes at the origin and is nonnegative but not identically zero on its effective domain. The *antipolar* of an antigauge γ is the function

$$\gamma^0(v) = \sup\{0 \leqslant \mu < +\infty | \langle x, v\rangle \geqslant \mu \cdot \gamma(x), \forall x \in \text{dom}\, \gamma\},$$

where we use the convention $\sup \varnothing = -\infty$. This function is a closed antigauge on V, and one has $\gamma^{00} = \text{cl}\, \gamma$. The closed antigauges on X are in one-to-one correspondence with the antipolar sets in X via

$$\gamma(x) = \begin{cases} \sup\{0 \leqslant \mu < +\infty \mid x \in \mu C\} & \text{if } x \in \operatorname{cone} C, \\ 0 & \text{if } x \in \operatorname{asym} C, \\ -\infty & \text{if } x \notin \operatorname{rec} C \end{cases}$$

and

$$C = \{x \mid \gamma(x) \geqslant 1\}.$$

Here *rec* C denotes the recession cone of C (i.e. rec $C = \{z \mid x + \lambda z \in C, \forall x \in C, \forall \lambda \geqslant 0\}$), *cone* C denotes the projecting cone of C (i.e. cone $C = \{\lambda x \mid x \in C, \lambda > 0\}$), and *asym* C consists of what might be thought of as the asymptotes of C (with respect to the origin), namely asym C = rec $C \setminus$ cone C. (Some feel for asym C may be gained by considering a few simple examples in the plane. If $C_1 = \{(\xi_1, \xi_2) \mid \xi_2 \geqslant 1\}$, then asym $C_1 = \{(\xi_1, 0) \mid \xi_1 \in R\}$. If $C_2 = \{(\xi_1, \xi_2) \mid \xi_2 \geqslant 1, \xi_1 \geqslant 0\}$, then asym $C_2 = \{(\xi_1, 0) \mid \xi_1 \geqslant 0\}$. If $C_3 = \{(\xi_1, \xi_2) \mid 0 < \xi_1 \leqslant 1 \leqslant \xi_1\xi_2$, or else $1 \leqslant \xi_1, \xi_2\}$, then asym $C_3 = \{(\xi_1, 0) \mid \xi_1 \geqslant 0\} \cup \{(0, \xi_2) \mid \xi_2 \geqslant 0\}$. The set $D_1 = \{(\xi_1, \xi_2) \mid \xi_2 \geqslant e^{\xi_1}\}$ is not antipolar, but $C_4 = D_1^{00}$ is. One can check that $C_4 = \{(\xi_1, \xi_2) \mid \xi_2 \geqslant e^{\xi_1}$ and $\xi_1 \leqslant 1$, or else $\xi_2 \geqslant e\xi_1$ and $1 \leqslant \xi_1\}$, so that asym $C_4 = \{(\xi_1, 0) \mid \xi_1 \leqslant 0\}$. The set $D_2 = \{(\xi_1, \xi_2) \mid \xi_2 \geqslant 1 + \xi_1^2\}$ is not antipolar, but $C_5 = D_2^{00}$ is. One has $C_5 = \{(\xi_1, \xi_2) \mid \xi_2 \geqslant 1 + \xi_1^2$ and $|\xi_1| \leqslant 1$, or else $\xi_2 \geqslant 2|\xi_1|$ and $|\xi_1| \geqslant 1\}$ and asym $C_5 = \{(0, 0)\}$.) It is convenient to write γ_C for the antigauge corresponding in the above manner to an antipolar set C, and one has $(\gamma_C)^0 = \gamma_{C^0}$. We shall call a function f: $X \to [-\infty, +\infty]$ *antigaugelike* (admittedly, a verbal monstrosity!) provided that, on the set $\{x \mid f(x) > -\infty\}$, f is nonconstant and bounded below by $f(0)$, and that the level sets

$$\{x \mid f(x) \geqslant \alpha\}, \quad f(0) < \alpha < \sup_X f$$

are all proportional (i.e. can all be expressed as positive scalar multiples of a single set). If ϕ: $[0, +\infty) \to [-\infty, +\infty)$ is a nondecreasing concave function not identically $-\infty$, then its *monotone conjugate* is the function ϕ^+: $[0, +\infty) \to [-\infty, +\infty)$ given by

$$\phi^+(\tau) = \inf_{0 \leqslant \sigma < +\infty} \{\sigma\tau - \phi(\sigma)\}.$$

This is another such function, and moreover it is upper semicontinuous. One has ϕ^{++} = usc ϕ, the upper semicontinuous hull of ϕ. If ϕ is nonconstant, then ϕ is also nonconstant. From the definition of ϕ^+ it follows that one always has

$$\phi(\sigma) + \phi^+(\tau) \leqslant \sigma\tau, \qquad \forall \sigma, \tau \in [0, +\infty).$$

The *subdifferential* of ϕ is the multivalued mapping $\partial\phi$: $[0, +\infty) \to [0, +\infty)$ defined by

$$\tau \in \partial\phi(\sigma) \quad \Leftrightarrow \quad \phi(\sigma) + \phi^+(\tau) = \sigma\tau.$$

As before, $\partial\phi(\sigma)$ amounts to simply the ordinary derivative at points σ where ϕ is differentiable. Equipped with these notions and facts, we can proceed to the concave analogue of Proposition 13.

PROPOSITION 14. *A function f is an antigaugelike closed proper concave function if and only if it can be expressed in the form $f(x) = \phi(\gamma(x))$, where γ is a closed antigauge and ϕ is a nondecreasing, nonconstant, upper semicontinuous concave function on $[0, +\infty)$. (Here $\phi(-\infty)$ is to be interpreted as $-\infty$ in the formula for f.) In this case f^* is antigaugelike, too, and in fact*

$$f^*(v) = \phi^+(\gamma^0(v)),$$

where ϕ^+ satisfies the same conditions as ϕ. Moreover, one then has $v \in \partial f(x)$ if and only if

$$\langle x, v\rangle = \gamma(x)\cdot\gamma^0(v) \quad \textit{and} \quad \gamma^0(v) \in \partial\phi(\gamma(x)).$$

Also, in this event the sets

$$\{x | \gamma(x) \geqslant 1\}, \qquad \{v | \gamma^0(v) \geqslant 1\}$$

are antipolar to each other.

PROOF. See McLinden [9].

The following corollary completes the treatment of positive homogeneity of nonzero degree p. (The case $-\infty < p < 0$ corresponds to the dual version of the corollary.)

COROLLARY 14A. *A function f is closed proper concave, nonconstant on its effective domain, and positively homogeneous of degree p, where $0 < p < 1$, if and only if it is of the form $f(x) = (1/p)\gamma(x)^p$ for some closed antigauge γ. In this case f^* is positively homogeneous of degree q, where $1/p + 1/q = 1$; in fact,*

$$f^*(v) = \begin{cases} (1/q)\gamma^0(v)^q & \textit{if } \gamma^0(v) > 0, \\ -\infty & \textit{otherwise}. \end{cases}$$

One then has

$$\langle x, v\rangle \geqslant [pf(x)]^{1/p}[qf^*(v)]^{1/q} \geqslant 0, \qquad \forall x \in \operatorname{dom} f, \forall v \in \operatorname{dom} f^*,$$

and the sets $\{x | f(x) \geqslant 1/p\}$, $\{v | f^*(v) \geqslant 1/q\}$ are antipolar to each other. Moreover, $v \in \partial f(x)$ if and only if $\gamma(x) > 0$, $\gamma^0(v) > 0$,

$$\langle x, v\rangle = \gamma(x)\cdot\gamma^0(v) \quad \textit{and} \quad \gamma^0(v) = \gamma(x)^{p-1}.$$

PROOF. We proceed in a manner quite similar to that for Corollary 13A. First, suppose f is of the form $f = (1/p)\gamma^p$ for some $0 < p < 1$ and closed antigauge γ. Then it is clearly nonconstant and positively homogeneous of degree p, and furthermore, for $\phi(\sigma) = (1/p)\sigma^p$ it is of the form treated by the proposition, and hence is closed proper concave. On the other hand, suppose f is closed proper concave, nonconstant on its effective domain, and positively homogeneous of degree p for some $0 < p < 1$. One can deduce routinely that $f(0) = 0$, that $f(x) > -\infty$ implies $f(x) \geqslant 0$, and that $C = \{x | f(x) \geqslant 1/p\}$ is an antipolar set. (The nonconstancy assumption is needed to show C is nonempty.) Then for each $0 < \alpha < +\infty$ one can show that

$$\{x | f(x) \geqslant \alpha\} = (\alpha p)^{1/p} C = \{x | \phi \circ \gamma_C(x) \geqslant \alpha\},$$

where $\phi(\sigma) = (1/p)\sigma^p$. From this it follows that $f = \phi \circ \gamma_C$. The remainder of the corollary now follows from the proposition, since $\phi^+(\tau) = (1/q)\tau^q$ for $0 < \tau < +\infty$ and $\phi^+(0) = -\infty$.

COROLLARY 14B. *Let C be an antipolar set. Then $\gamma_C = (-\psi_{C^0})^*$ and $\gamma_C^* = -\psi_{C^0}$. Moreover, $v \in \partial\gamma_C(x)$ if and only if $\langle x, v\rangle = \gamma_C(x) \cdot \gamma_{C^0}(v)$ and either $\gamma_C(x) = 0$ and $\gamma_{C^0}(v) \geqslant 1$ or else $\gamma_C(x) > 0$ and $\gamma_{C^0}(v) = 1$.*

PROOF. The proposition applied to $f = \phi \circ \gamma_C$, where $\phi(\sigma) = \sigma$, yields $f^* = \phi^+ \circ \gamma_{C^0}$. Since $\phi^+(\tau) = -\psi_{[1,+\infty)}(\tau)$, this means $\gamma_C^* = -\psi_{C^0}$. Taking conjugates then yields the first identity, since γ_C is closed. The subdifferential characterization follows by the proposition and the fact that $\tau \in \partial\phi(\sigma)$ if and only if either $\sigma = 0$ and $\tau \geqslant 1$ or else $\sigma > 0$ and $\tau = 1$.

The situation treated in Corollary 14B amounts to the concave version of positive homogeneity of degree $p = 1$.

In terms of viewing the functions in Proposition 14 as concave "utility" functions, still other choices of the "scaling" function ϕ besides pth powers are of interest. We shall cite just three, leaving it to the reader to formulate the corollaries corresponding to these choices as well as others. These three each have the nice feature that $\phi^+ = \phi + \alpha$, where α is some real constant. Upon replacing ϕ by $\phi - \alpha/2$, one could thus obtain complete self-duality, i.e. $(\phi + \alpha/2)^+ = \phi + \alpha/2$. Of course, $\partial(\phi + \alpha/2) = \partial\phi$. In general, the ϕ's which are self-dual (up to an additive constant as above) are precisely those whose subdifferentials are symmetric upon reflection through the axis $\sigma = \tau$ in R^2_+.

EXAMPLE 1. Take $\phi(\sigma) = \ln\sigma$ for $0 < \sigma < +\infty$ and $\phi(0) = -\infty$. Then $\phi^+ = \phi + 1$. The graph of the subdifferential $\partial\phi$ here is $\{(\sigma, \tau) \in R^2_+ | \sigma\tau = 1\}$. By comparing this graph with the graphs of the subdifferentials of the pth powers appearing above, one sees this as the natural limiting case corresponding to $p = 0$. See McLinden [9] for more on this point. The subgradient formula one obtains from Proposition 14 for this choice is $v \in \partial(\phi \circ \gamma_C)(x)$ if and only if $\langle x, v\rangle = \gamma_C(x) \cdot \gamma_{C^0}(v) = 1$.

EXAMPLE 2. Take $\phi(\sigma) = \sigma$ for $0 \leqslant \sigma \leqslant 1$ and $\phi(\sigma) = 1$ for $1 \leqslant \sigma < +\infty$.

Then $\phi^+ = \phi - 1$. Here the "utility" function $\phi \circ \gamma_C$ increases *linearly* as x moves from the origin toward C and becomes constant once x gets inside C.

EXAMPLE 3. Take $\phi(\sigma) = \sigma - \sigma^2/2$ for $0 \leqslant \sigma \leqslant 1$ and $\phi(\sigma) = \frac{1}{2}$ for $1 \leqslant \sigma < +\infty$. Then $\phi^+ = \phi - \frac{1}{2}$. Here the "utility" function $\phi \circ \gamma_C$ increases *quadratically* as x moves from the origin towards C and becomes constant once x is inside C.

8. Concluding comments. We wish to emphasize the open-ended nature of the present framework. For each of our model problem's elements–the cones P and Q, the functions h and k, and the transformation A–there are, of course, many specific, basic structural forms of potential interest and importance. When one contemplates further to consider intermingling these various structural forms, the possibilities become virtually endless. In view of this, we have chosen in this paper to concentrate on developing a general framework (§§3–6) broad and versatile enough to encompass a variety of specific problem types of known value, and secondarily to indicate (§7) the essential facts necessary for applying the framework in the broad cases of convex "distance" and/or concave "utility" structure in h and k.

Note, too, that our entire approach has in effect been based on a simple but very general "symmetry principle" (§3). This principle might be kept in mind as a potentially powerful conceptual aid in deciding how to formulate a class of problems initially so as to achieve various desirable duality effects.

We close by mentioning a few of the other possibilities which might be combined fruitfully with the present framework.

First, altering h or k by affine or "conjugate affine" terms does not materially alter the framework. It simply has the effect of removing the normalizations (with respect to the origins) which we imposed for purely notational convenience. See [**18**, Theorem 12.3].

Second, additive separability in h or k and A is a powerful feature, on which decomposition into smaller subproblems might be based. See [**8**] for an illustration of how the decomposition principle can be formulated in the context of Fenchel-Rockafellar duality. Quasiseparability could also be useful, particularly in conjunction with faithful convexity. See [**19**], where both these notions are defined and studied in the context of finitely many convex inequality constraints.

Third, structure present in the cones P and Q might be exploited. Finitely generated or polyhedral cones would admit sharpenings of some results (by Proposition 8(f), for example). Intersection or finite sum structure could also be useful, at least insofar as computing P^* and Q^*.

Finally, the type of transformation A treated in the model can be broadened considerably with the aid of slightly stronger assumptions. In [**10**] it is shown how the present model extends quite naturally to handle A's which are closed convex processes (see [**18**, §39] for the definition). In particular, densely defined single-valued linear A's having closed graph are covered there.

References

1. N. Bourbaki, *Espaces vectoriels topologiques*, Chapters I, II, Hermann, Paris, 1966.

2. R. W. Cottle, *Symmetric dual quadratic programs*, Quart. Appl. Math. **21** (1963), 237–243.

3. R. J. Duffin, *Infinite programs*, Ann. of Math. Studies, no. 38, Princeton Univ. Press, Princeton, N.J., 1956, pp. 157–170.

4. E. Eisenberg, *Duality in homogeneous programming*, Proc. Amer. Math. Soc. **12** (1961), 783–787.

5. W. Fenchel, *Convex cones, sets and functions*, Mimeographed lecture notes, Princeton Univ., Princeton, N.J., 1951.

6. D. Gale, H. W. Kuhn and A. W. Tucker, *Linear programming and the theory of games*, Activity Analysis of Production and Allocation (T. C. Koopmans, Editor), Wiley, New York, 1951, pp. 317–329.

7. J. L. Kelley, I. Namioka et al., *Linear topological spaces*, Van Nostrand, Princeton, N.J., 1963.

8. L. McLinden, *A decomposition principle for minimax problems*, Decomposition of Large-Scale Problems (D. M. Himmelblau, Editor), North-Holland, Amsterdam, 1973, pp. 427–435.

9. ______, *Operators obeying Hölder's inequality* (in preparation).

10. ______, *Extremum problems involving a convex process* (in preparation).

11. ______, *Symmetrized separable convex programming*, Trans. Amer. Math. Soc. (to appear).

12. J.-J. Moreau, *Proximité et dualité dans un espace Hilbertien*, Bull. Soc. Math. France **93** (1965), 273–299.

13. ______, *Fonctionnelles convexes*, Mimeographed lecture notes, Collège de France, Paris, 1967.

14. E. L. Peterson, *Generalization and symmetrization of duality in geometric programming*, Discussion paper, Northwestern Univ., Evanston, Ill., 1972.

15. ______, *Geometric programming*, SIAM Rev. **18** (1976), 1–51.

16. S. M. Robinson, *Regularity and stability for convex multivalued functions*, Math. Operations Res. **1** (1976), 130–143.

17. R. T. Rockafellar, *Duality and stability in extremum problems involving convex functions*, Pacific J. Math. **21** (1967), 167–187.

18. ______, *Convex analysis*, Princeton Univ. Press, Princeton, N.J., 1970.

19. ______, *Some convex programs whose duals are linearly constrained*, Nonlinear Programming (J. B. Rosen, O. L. Mangasarian, and K. Ritter, Editors), Academic Press, New York, 1970, pp. 293–322.

20. ______, *Conjugate duality and optimization*, CBMS Regional Conf. Ser. Appl. Math., no. 16, SIAM Publ., Philadelphia, Pa., 1974.

21. ______, *Duality theorems for convex functions*, Bull. Amer. Math. Soc. **70** (1964), 189–192.

Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801