AN ANALOGUE OF MOREAU'S PROXIMATION THEOREM, WITH APPLICATION TO THE NONLINEAR COMPLEMENTARITY PROBLEM

L. McLinden

This paper concerns the problem of minimizing a convex function subject to nonnegativity constraints, an associated nonlinear complementarity problem, and a new approach to solving these problems. The approach involves solving a sequence of nicer problems which approximate the given ones better and better, and our focus is on certain natural "trajectories" of solutions of the approximating problems. Existence, characterization, and continuous dependence of the solutions is obtained by establishing a complete analogue of Moreau's Proximation Theorem. From this analogue also follow two new facts about the geometric nature of the graphs of subdifferentials in R^n, as well as new information about monotone conjugacy for coordinatewise nondecreasing convex functions on the nonnegative orthant. The largest part of the paper is then devoted to developing a number of rather strong properties of the solution trajectories, particularly as regards the nature of their convergence. Perhaps the most striking property is that these trajectories will locate a maximal strictly complementary solution which, furthermore, can be arranged to have a certain prescribed strong Pareto optimality property. The arithmetic-geometric mean inequality enters decisively at several key points, and the proofs generally rely strongly upon the techniques of finite-dimensional convex analysis.

1. Introduction. Consider the optimization problem

$$(P_0) \qquad \min_{x \in Q} \{f(x)\} ,$$

where f is a closed proper convex function on R^n and $Q = \{x \in R^n \,|\, x_k \geq 0 \; \forall k\}$. The simplest and most generally useful convex dual problem to (P_0) is

$$(P_0^*) \qquad \min_{y \in Q} \{f^*(y)\} ,$$

where f^* is the Fenchel conjugate of f, i.e.,

$$f^*(y) = \sup_x \{\langle x, y \rangle - f(x)\}$$

(here $\langle x, y \rangle$ denotes the usual dot product $\sum x_k y_k$). Associated with

(P_0) and (P_0^*) is the nonlinear complementarity problem

(C_0) find $(x, y) \in (Q \times Q) \cap \operatorname{graph} \partial f$ such that $\langle x, y \rangle = 0$,

which describes the extremality conditions that go with both (P_0) and (P_0^*). Here $\partial f: R^n \to R^n$ is the subdifferential of f, characterized by

$$y \in \partial f(x) \Longleftrightarrow f(x) + f^*(y) = \langle x, y \rangle .$$

It is known (see [6, Theorem 31.4]) that, under the very minimal assumption that

$$\varnothing \neq Q \cap \operatorname{ri}(\operatorname{dom} f) \quad \text{and} \quad \varnothing \neq Q \cap \operatorname{ri}(\operatorname{dom} f^*)$$

(here $\operatorname{dom} \varphi = \{z \mid \varphi(z) < +\infty\}$ and "ri" denotes relative interior), this trio of problems is well posed and statisfies several useful relationships, including: (i) all three problems have solutions; (ii) (x, y) solves (C_0) if and only if x solves (P_0) and y solves (P_0^*), and (iii) $\inf(P_0) + \inf(P_0^*) = 0$.

The present paper is concerned with developing relationships between the trio (P_0), (P_0^*), (C_0) and a certain closely related approximating trio of problems (P_z), (P_z^*), (C_z), where z is a parameter ranging over Q. These parametrized problems involve an auxiliary function $z \circ g: R^n \to [-\infty, +\infty)$ having a number of useful properties. While the analysis of the paper handles all z in Q, for notational simplicity in this introduction we define $z \circ g$ only for z in $P = \{x \in R^n \mid x_k > 0 \ \forall k\}$; namely, we set

$$(z \circ g)(x) = \begin{cases} \sum z_k \log x_k & \text{if } x \in P \\ -\infty & \text{if } x \notin P . \end{cases}$$

(The particular function $e \circ g$, where $e = (1, \cdots, 1)$, can be thought of as simply "g".) For each fixed $z \in P$, this function is (Proposition 1) closed proper concave in x with conjugate and subdifferential (both taken in the concave sense) given by

$$(z \circ g)^*(y) = (z \circ g)(y) + \sum (z_k - z_k \log z_k)$$

and

$$y \in \partial(z \circ g)(x) \Longleftrightarrow x \in Q, \ y \in Q \text{ and } x_k y_k = z_k \ \forall k .$$

The initial reason for interest in such a function $(z \circ g)(x)$ is that it acts as an interior penality term to approximate the constraint $x \in Q$ in (P_0). Indeed, as $z \to 0$ the function $-(z \circ g)(x)$ approaches the function $\psi_Q(x)$, where ψ_Q denotes the convex indicator function of Q (i.e., $\psi_Q(x)$ is 0 if $x \in Q$ and is $+\infty$ otherwise). Likewise, the formula for $(z \circ g)^*(y)$ shows that, as $z \to 0$, $-(z \circ g)^*(y)$ approaches

the function $\psi_Q(y)$ describing the constraint $y \in Q$ in (P_0^*). Also, since

$$y \in \partial(-\psi_Q)(x) \Longleftrightarrow x \in Q, \; y \in Q \text{ and } x_k y_k = 0 \; \forall k \; ,$$

the formula for $\partial(z \circ g)$ suggests that the pairs (x, y) in graph $\partial f \cap$ graph $\partial(z \circ g)$ might approach, in some sense, the solutions of (C_0) as $z \to 0$. We are thus led to consider the three problems

(P_z) $\qquad\qquad\qquad \min_x \{ f(x) - (z \circ g)(x) \} \; ,$

(P_z^*) $\qquad\qquad\qquad \min_y \{ f^*(y) - (z \circ g)(y) \} \; ,$

and

(C_z) $\qquad$ find $(x, y) \in (Q \times Q) \cap$ graph ∂f such that $x_k y_k = z_k \forall k$.

Our analysis starts out in §2 with an existence, duality, and characterization result (Theorem 1) which follows easily from Fenchel's Duality Theorem [2] (see also [6, § 31]) together with the cited properties of $z \circ g$. Assuming throughout the condition

$$\varnothing \neq P \cap \operatorname{dom} f \quad \text{and} \quad \varnothing \neq P \cap \operatorname{dom} f^* \; ,$$

we find for each $z \in P$ that: (i) problems (P_z), (P_z^*), (C_z) each have at least one solution; (ii) (x, y) solves (C_z) if and only if x solves (P_z) and y solves (P_z^*); and (iii) $\inf(P_z) + \inf(P_z^*) = \sum (z_k - z_k \log z_k)$. Since $z \circ g$ is strictly concave, the solutions to (P_z) and (P_z^*) are necessarily unique and hence for convenience will be denoted by x^z and y^z. It is shown later (Corollary 5A), after substantially more work, that the mapping

$$z \longrightarrow (x^z, y^z)$$

is continuous on P and in fact differentiable there almost everywhere. Taken together, the preceding facts constitute an analogue of Moreau's Proximation Theorem [5] (see also [6, Theorem 31.5 and ff.]). Two new general facts follow concerning the geometric nature of graph ∂f. The above existence result together with the arithmetic-geometric mean inequality imply that $(P \times P) \cap$ graph ∂f is a maximal monotone set with respect to $P \times P$ (Theorem 3), and the continuous dependence result implies that $(P \times P) \cap$ graph ∂f is homeomorphic to P via the map $(x, y) \to z$, where $z_k = x_k y_k \; \forall k$ (Theorem 4). These are nonobvious analogues of results known to hold for the entire set graph ∂f with respect to $R^n \times R^n$ and R^n (see [6, Corollaries 31.5.1 and 31.5.2]).

Another general fact established in §2 is that the functions f and f^* may each be replaced in (P_z) and (P_z^*) by their "coordinate-

wise nondecreasing hulls" without affecting either the optimal values
or the optimal solutions (Theorem 2). These two nondecreasing
hull functions are shown to be the monotone conjugates of each
other (Proposition 3), and the corresponding "subdifferential" mapp-
ing from Q to Q is characterized in terms of the original ∂f
(Corollary 2A). Thus, our analogue of Moreau's theorem may be
viewed alternatively as pertaining to monotone conjugate pairs of
coordinatewise nondecreasing convex functions on Q.

The major part of the paper, §§ 3 and 4, is devoted to uncover-
ing the relationships between the parametrized trio (P_z), (P_z^*), (C_z)
and the original problem trio (P_0), (P_0^*), (C_0). It is here that the
nice properties of the special penalty terms we have chosen come
particularly into play. For instance, one strong property which is
not obvious at first glance, but upon which most of the later results
depend, is that the function

$$(x, z) \longrightarrow \sum z_k \log x_k + \sum (z_k - z_k \log z_k)$$

is concave *jointly* in $(x, z) \in P \times P$. Indeed, this function admits
an extension to all of $R^n \times R^n$ which is closed proper concave, and
the associated conjugate function and subdifferential mapping can
be found explicitly (Proposition 4). Using this and techniques of
convex analysis, we show (Theorem 5) that the parametrized optimal
values,

$$z \longrightarrow \inf (P_z) \,,$$

form a continuous concave function of $z \in Q$, actually continuously
differentiable on $z \in P$, and having certain subdifferential and con-
jugate formulas. We also show (Theorem 7) that the parametrized
optimal solutions,

$$z \longrightarrow X^z \quad \text{and} \quad z \longrightarrow Y^z \,,$$

where X^z and Y^z denote the solution sets of (P_z) and (P_z^*), form
upper-semicontinuous mappings on Q. In particular, the distance
between x^z and the solution set of (P_0) tends to zero as z tends to
0 along any path (or sequence) in P.

This last point raises the prospect that stronger convergence
facts for x^z might be possible if the passage of z to 0 is controlled
in some fashion. The rest of §§ 3 and 4 is devoted to exploring
this in detail, and a number of results are obtained for the case in
which z tends to 0 along (any sequence from) a fixed ray. A
sampling of these results will now be outlined. Together with all
the other information, they justify a solution procedure which is
described in § 5.

A little notation is needed.   For any $z \in Q$, let $\sigma(z) = \{k \, | \, z_k > 0\}$ be called the *support* of z, and let I and J denote the (well-defined) maximal supports from among those of all the solutions to (P_0) and (P_0^*) respectively.   For any fixed $a \in P$ satisfying $\sum a_k = 1$, let the solutions x^z and y^z corresponding to z's of the form $z = \zeta a$ (for $\zeta > 0$) be denoted simply by x^ζ and y^ζ.   In what follows it is understood that a is as just described and fixed.

The most striking result, and the one which most reflects the special properties of our penalty term, is the following (Theorem 8). If (x°, y°) is any accumulation point of (x^ζ, y^ζ) as $\zeta \downarrow 0$, then

$$\sigma(x^\circ) = I \quad \text{and} \quad \sigma(y^\circ) = J \, ,$$

and one has the inequality

$$\sum_I a_k(x_k/x_k^\circ) + \sum_J a_k(y_k/y_k^\circ) \leqq 1$$

for any x solving (P_0) and any y solving (P_0^*).   From the inequality follow *a posteriori* estimates on the solution sets of (P_0) and (P_0^*) (Corollary 8A).   Using both parts of the above result together with the arithmetic-geometric mean inequality, we obtain the following surprising convergence result (Theorem 9).   If (C_0) has a strictly complementary solution (i.e., if $I \cup J = \{1, \cdots, n\}$),   then (x^ζ, y^ζ) converges to such a solution as $\zeta \downarrow 0$; moreover,

$$\lim_{\zeta \downarrow 0} (x^\zeta, y^\zeta) = (\tilde{x}, \tilde{y}) \, ,$$

where $\tilde{x}$ and $\tilde{y}$ are the unique solutions of (P_0) and (P_0^*) which maximize the concave functions $x \to \prod_I x_k^{a_k}$ and $y \to \prod_J y_k^{a_k}$, respectively.   This same conclusion holds even in the absence of a strictly complementary solution to (C_0), provided that f is polyhedral (Theorem 10).   An example in R^2 shows that this behavior may fail in general.   It still seems very likely, however, that convergence of (x^ζ, y^ζ) to some particular solution of (C_0) must occur as $\zeta \downarrow 0$ (i.e., that there is at most one accumulation point as $\zeta \downarrow 0$).   It would be nice to resolve this.

Other facts are also derived in §§ 3 and 4 concerning (x^ζ, y^ζ) as $\zeta \downarrow 0$.   These include, for instance, a stopping criterion and a result guaranteeing that some improvement must continually occur, either in the x-variable or in the y-variable, until optimality is achieved. These facts are discussed further in § 5, where the solution method is formalized.   Using it, one can locate the very special solutions $\tilde{x}$ and $\tilde{y}$ defined above even in the general case.

In § 6 the analogy between the results in § 2 and Moreau's Proximation Theorem is described quite explicitly.   It is shown that in the trivial case $n = 1$ the present analogue (consisting of

Theorem 1 and Corollary 5A) is actually equivalent to Moreau's theorem. This follows from a one-to-one correspondence between the closed proper convex functions on R and certain closed proper convex functions on the nonnegative halfline (Theorem 14). Simple examples are given in the case of $n = 2$, however, which show that in general neither of the two "theorems" follows from the other. Also in §6 the Proximal Point Algorithm is reviewed in the context of Moreau's theorem and then transformed into the present context. Likewise, the present solution procedure is transformed into the context of Moreau's theorem. The differences between the two procedures are thereby clarified. This suggests some questions for further research.

In an appendix it is shown how the present framework can be made to handle problems similar to (P_0) and (C_0) but involving non-negativity constraints in only some of the variables.

The ideas of this paper admit extensions to complementarity problems (C_0) and (C_z) in which ∂f is replaced by a general maximal monotone (multivalued) operator $T: R^n \to R^n$ satisfying the condition

$$\varnothing \neq (P \times P) \cap \operatorname{graph} T .$$

This will be taken up elsewhere [4].

Throughout, we freely use terminology, notation, and facts from finite-dimensional convex analysis as presented in Rockafellar [6]. In particular, the symbols $*, \partial$, and 0^+ applied to a function signify respectively the conjugate function, subdifferential mapping, and recession function. Of course, each of these is to be interpreted in either the convex or the concave sense, according to whether the given function itself is convex or concave. We do depart slightly from [6], however, by denoting the indicator function of a set C by ψ_C (rather than $\delta(\cdot \,|\, C)$); thus, $\psi_C(x)$ is 0 if $x \in C$ and is $+\infty$ otherwise. Throughout, we use the convention that $0 \log 0 = 0$. Also, unless otherwise indicated, the summation symbol $\sum$ signifies $\sum_{k=1}^{n}$ and the product symbol $\prod$ signifies $\prod_{k=1}^{n}$.

2. **An analogue of Moreau's Proximation Theorem.** We begin by extending to all $z \in Q$ the definition of the penalty function $z \circ g$ appearing in problems (P_z) and (P_z^*). Let $h: R \to [-\infty, +\infty)$ be the function

$$(2.1) \qquad h(\xi) = \begin{cases} \log \xi & \text{if } \xi > 0 \\ -\infty & \text{if } \xi \leq 0 , \end{cases}$$

and define functions $\zeta h: R \to [-\infty, +\infty)$ for $\zeta \geq 0$ by setting

$$(2.2) \qquad (\zeta h)(\xi) = \begin{cases} \zeta(h(\xi)) & \text{if } \zeta > 0, \ \xi \in R \\ -\psi_{[0, +\infty)}(\xi) & \text{if } \zeta = 0, \ \xi \in R . \end{cases}$$

It is elementary to check that, for each fixed $\zeta \geqq 0$, the function ζh is closed proper concave and satisfies

$$(2.3) \qquad ((\zeta h)0^+)(\xi) = -\psi_{[0,+\infty)}(\xi) ,$$

$$(2.4) \qquad (\xi, \eta) \in \text{graph } \partial(\zeta h) \Longleftrightarrow \xi \geqq 0, \eta \geqq 0, \xi\eta = \zeta ,$$

$$(2.5) \qquad (\zeta h)^*(\eta) = (\zeta h)(\eta) + \zeta - \zeta \log \zeta .$$

Now let $g: R^n \to [-\infty, +\infty)$ be the function

$$(2.6) \qquad g(x) = \sum h(x_k) ,$$

and define the functions $z \circ g: R^n \to [-\infty, +\infty)$ for $z \in Q$ by setting

$$(2.7) \qquad (z \circ g)(x) = \sum (z_k h)(x_k) .$$

PROPOSITION 1. *For each $z \in Q$ the function $z \circ g$ is closed proper concave and satisfies*

$$(2.8) \qquad ((z \circ g)0^+)(x) = -\psi_Q(x) ,$$

$$(2.9) \qquad (x, y) \in \text{graph } \partial(z \circ g) \Longleftrightarrow x \in Q, y \in Q, x_k y_k = z_k \; \forall k ,$$

$$(2.10) \qquad (z \circ g)^*(y) = (z \circ g)(y) + \sum (z_k - z_k \log z_k) .$$

Proof. The additive separability structure of $z \circ g$ implies that

$$((z \circ g)0^+)(x) = \sum ((z_k h)0^+)(x_k) ,$$

$$(x, y) \in \text{graph } \partial(z \circ g) \Longleftrightarrow (x_k, y_k) \in \text{graph } \partial(z_k h) \; \forall k ,$$

$$(z \circ g)^*(y) = \sum (z_k h)^*(y_k) .$$

From these facts it is clear that (2.8), (2.9), (2.10) follow from (2.3), (2.4), (2.5), respectively.

For general $z \in Q$, problems (P_z), (P_z^*), (C_z) are defined just as for the special case $z \in P$ in §1. Notice that when $z = 0$ these problems are indeed the same as problems (P_0), (P_0^*), (C_0) from §1. In general, the sets of optimal solutions to (P_z) and (P_z^*) will be denoted by X^z and Y^z, respectively.

The next result provides several characterizations of the condition which will shortly become our blanket hypothesis. Notice that characterization (a) says that the constraints in (P_0) are satisfiable strictly (Slater condition) and that the function $f + \psi_Q$ minimized in (P_0) has bounded level sets.

PROPOSITION 2. *For any $z \in P$ and $z' \in P$ the following six conditions are pairwise equivalent:*

(a) $\varnothing \neq P \cap \operatorname{dom} f$ *and* $(f0^+)(x) > 0$ *for all nonzero* $x \in Q$;

(b) $\varnothing \neq P \cap \operatorname{dom} f^*$ *and* $(f^*0^+)(y) > 0$ *for all nonzero* $y \in Q$;

(c) $\varnothing \neq P \cap \operatorname{dom} f$ *and* $\varnothing \neq P \cap \operatorname{dom} f^*$;

(d) $+\infty > \inf \{f - z \circ g\} > -\infty$;

(e) $+\infty > \inf \{f^* - z' \circ g\} > -\infty$;

(f) $+\infty > \inf \{f - z \circ g\}$ *and* $+\infty > \inf \{f^* - z' \circ g\}$.

Proof. The proposition follows from the following lemma together with its dualized version.

LEMMA 1. *Let* $z \in P$ *be fixed. Among the conditions*

(i) $+\infty > \inf \{f - z \circ g\}$

(ii) $\varnothing \neq P \cap \operatorname{dom} f$

(iii) $(f^*0^+)(y) > 0$ *for all nonzero* $y \in Q$

(iv) $\inf \{f^* - z \circ g\} > -\infty$

one has the implications

$$(\text{i}) \Longleftrightarrow (\text{ii}) \Longrightarrow (\text{iii}) \Longrightarrow (\text{iv}) \ ,$$

and if $\varnothing \neq P \cap \operatorname{dom} f^*$ *then all four conditions are pairwise equivalent.*

Proof. Since $P = \operatorname{dom}(z \circ g)$, (i) $\Leftrightarrow$ (ii). Since $\operatorname{dom} f$ and P are convex and P is open, it is easy to check that (ii) is equivalent to

$$\varnothing \neq P \cap \operatorname{ri}(\operatorname{dom} f) \ .$$

By separation theory [6, Theorems 11.3 and 11.1], the above condition fails if and only if there exists a nonzero y such that

$$\inf_P \langle \cdot, y \rangle \geqq \sup_{\operatorname{dom} f} \langle \cdot, y \rangle \quad \text{and} \quad \sup_P \langle \cdot, y \rangle > \inf_{\operatorname{dom} f} \langle \cdot, y \rangle \ .$$

Since

$$\sup_{\operatorname{dom} f} \langle \cdot, y \rangle = (f^*0^+)(y) \quad \text{and} \quad \sup_P \langle \cdot, y \rangle = \psi_Q(-y)$$

[6, Theorem 13.3], the situation just described occurs if and only if there exists a nonzero y such that

$$-\psi_Q(y) \geqq f^*0^+(y) \quad \text{and} \quad \psi_Q(-y) > - f^*0^+(-y) \ .$$

Since f^*0^+ is never $-\infty$ [6, Theorem 8.5], ψ_Q takes only the values 0 and $+\infty$, and $Q \cap (-Q) = \{0\}$, the latter situation occurs if and only if there exists a nonzero $y \in Q$ satisfying $(f^*0^+)(y) \leqq 0$. Taking contrapositives, we obtain (ii) $\Leftrightarrow$ (iii). To see (i) $\Rightarrow$ (iv), note first that the definitions of f^* and $(z \circ g)^*$ yield

$$f(x) - (z \circ g)(x) \geqq (z \circ g)^*(y) - f^*(y) \ , \quad \forall x \forall y \ .$$

Using identity (2.10), we conclude that

$$\inf \{f - z \circ g\} \geqq \sum (z_k - z_k \log z_k) - \inf \{f^* - z \circ g\} \ .$$

From this, (i) $\Rightarrow$ (iv) is clear. Finally, notice that

$$\inf \{f^* - z \circ g\} = -(f^* - z \circ g)^*(0) \ ,$$

so that (iv) $\Leftrightarrow 0 \in \text{dom} \, (f^* - z \circ g)^*$. Now assume that $\varnothing \neq P \cap \text{dom} \, f^*$. Then [6, Theorem 16.4] implies

$$(f^* - z \circ g)^* = f \, \square \, (-z \circ g)^* \ ,$$

where $\square$ denotes infimal convolution, so that

$$\begin{aligned}
\text{dom} \, (f^* - z \circ g)^* &= \text{dom} \, f + \text{dom} \, (-z \circ g)^* \\
&= \text{dom} \, f - P \ .
\end{aligned}$$

It follows that (iv) $\Rightarrow$ (ii) in the presence of $\varnothing \neq P \cap \text{dom} \, f^*$.

For the remainder of §§ 2 through 5, unless otherwise specified we assume that the mutually equivalent conditions of Proposition 2 are satisfied.

THEOREM 1. *For each $z \in Q$,*
(i) *(P_z), (P_z^*) and (C_z) each have at least one solution;*
(ii) *$x \in X^z$ and $y \in Y^z$ if and only if (x, y) solves (C_z);*
(iii) *$\inf (P_z) + \inf (P_z^*) = \sum (z_k - z_k \log z_k)$;*
(iv) *X^z and Y^z are compact convex;*
(v) *the solutions are unique when $z \in P$, so that we can write $X^z = \{x^z\}$ and $Y^z = \{y^z\}$ if $z \in P$.*

Proof. Let $z \in Q$ be fixed. Since $\text{dom} \, f$ and $\text{dom} \, f^*$ are each convex and since $\text{ri} (\text{dom} \, (z \circ g)) = P = \text{ri} (\text{dom} \, (z \circ g)^*)$, condition (c) of Proposition 2 is equivalent to

$$\varnothing \neq \text{ri} \, (\text{dom} \, f) \cap \text{ri} \, (\text{dom} \, (z \circ g)) \ \text{and} \ \varnothing \neq \text{ri} \, (\text{dom} \, (z \circ g)^*) \cap \text{ri} \, (\text{dom} \, f^*) \ .$$

Hence Fenchel's Duality Theorem [2], [6, Theorem 31.1] implies that

$$\inf \{f - z \circ g\} = \sup \{(z \circ g)^* - f^*\} \ ,$$

where both of these extrema are attained finitely. Also, by [6, Theorem 31.3] the pairs (x, y) such that x yields the infimum and y yields the supremum are exactly the elements in graph $\partial f \cap$ graph $\partial(z \circ g)$. In view of formulas (2.9) and (2.10), assertions (i), (ii), (iii) follow. To see (iv), observe that by [6, Theorems 8.4 and 8.7] the set X^z is compact if and only if

$$((f - z \circ g)0^+)(x) > 0, \ \forall x \neq 0 \ .$$

Since

$$((f - z \circ g)0^+)(x) = f0^+(x) - ((z \circ g)0^+)(x) = f0^+(x) + \psi_Q(x)$$

by [6, Theorem 9.3] and formula (2.8), this means that

$$X^z \text{ is compact} \iff f0^+(x) > 0 \text{ whenever } 0 \neq x \in Q .$$

Similarly,

$$Y^z \text{ is compact} \iff f^*0^+(y) > 0 \text{ whenever } 0 \neq y \in Q .$$

Since each of these recession conditions is satisfied, according to our blanket hypothesis (see characterizations (a) and (b) in Proposition 2), assertion (iv) is proved. Assertion (v) is immediate from (i) and the strict concavity of $z \circ g$ for $z \in P$.

COROLLARY 1A. *For any* $z \in Q$, $x \in X^z$, $y \in Y^z$ *one has*

$$f(x) \leq \inf (P_0) + \sum z_k$$

and

$$f^*(y) \leq \inf (P_0^*) + \sum z_k .$$

Proof. The vectors x, y, z satisfy

$$f(x) + f^*(y) = \langle x, y \rangle$$

and

$$x \in Q, \; y \in Q, \; x_k y_k = z_k \; \forall k .$$

Also,

$$\inf (P_0) + \inf (P_0^*) = 0 .$$

Hence

$$\inf (P_0) \leq f(x) = \sum x_k y_k - f^*(y) ,$$

which implies

$$f^*(y) \leq - \inf (P_0) + \sum z_k = \inf (P_0^*) + \sum z_k .$$

The first inequality is proved similarly.

COROLLARY 1B. *Let* $z \in P$, *and let* x^z *and* y^z *be as in Theorem* 1(v). *Write* $a_k = z_k / \sum_{l=1}^n z_l$, $\forall k$. *Then*

$$\{x \mid f(x) \leq f(x^z)\} \subset \{x \mid \sum x_k (a_k / x_k^z) \leq 1\} ,$$

$$\{y \mid f^*(y) \leq f^*(y^z)\} \subset \{y \mid \sum y_k (a_k / y_k^z) \leq 1\} .$$

Proof. It suffices to prove the first inclusion. Consider

$$S = \{x \mid f(x) \leqq f(x^z)\} , \quad T = \{x \mid (z \circ g)(x) \geqq (z \circ g)(x^z)\} .$$

Since int $T = \{x \in P \mid (z \circ g)(x) > (z \circ g)(x^z)\}$, the fact that x^z solves inf $\{f - z \circ g\}$ implies $\varnothing = S \cap \text{int } T$. Hence, there exists a hyperplane separating S from T, i.e., there exist $0 \neq b \in R^n$ and $\beta \in R$ such that

$$\langle x, b \rangle \leqq \beta \;\; \forall x \in S \quad \text{and} \quad \beta \leqq \langle x, b \rangle \;\; \forall x \in T .$$

Since $x^z \in S \cap T$ yields $\langle x^z, b \rangle = \beta$, it follows that

$$\langle x - x^z, -b \rangle \leqq 0 \;\; \forall x \in T .$$

By [6, Corollary 23.7.1], this means there exists some $\lambda > 0$ (recall $b \neq 0$) such that $b \in \lambda \partial(z \circ g)(x^z)$. By (2.9) this means $b_k = \lambda(z_k/x_k^z)$, $\forall k$. Hence, each $x \in S$ satisfies

$$\lambda \sum x_k(z_k/x_k^z) = \langle x, b \rangle \leqq \beta = \lambda \sum z_k ,$$

which implies the first inclusion.

Theorem 1 can be regarded as a rather special nonlinear "perfect duality" result analogous to the Gale-Kuhn-Tucker Duality Theorem in linear programming (see also [6, discussion on p. 337 concerning Corollary 31.4.1]). This is expressed in the following corollary, for which we make an exception and do *not* assume the blanket hypothesis.

COROLLARY 1C. *Let* $z \in P$. *If either*

$$\inf \{f - z \circ g\} \quad or \quad \inf \{f^* - z \circ g\}$$

is finite, then both are finite, in which case

$$\inf \{f - z \circ g\} + \inf \{f^* - z \circ g\} = \sum (z_k - z_k \log z_k)$$

and the two problems have unique solutions x *and* y *characterized by the conditions*

$$(x, y) \in (P \times P) \cap \text{graph } \partial f \quad and \quad x_k y_k = z_k \; \forall k .$$

Proof. Immediate from Proposition 2 and the theorem.

Many of our subsequent results can best be stated using the concept of support. Recall from §1 that for any $z \in Q$ we define the *support* of z to be

$$\sigma(z) = \{k \mid z_k > 0\} .$$

In particular, the support of the zero vector is the empty set $\varnothing$, and to this set we assign cardinality zero in what follows.

It will also be convenient to have notation to describe the various faces of Q. If K is any nonempty subset of $\{1, \cdots, n\}$ let us agree to write

$$Q^K = \{z \in Q \,|\, z_k = 0 \,\forall k \notin K\} \,.$$

For the degenerate case $K = \varnothing$ it is natural in what follows to adopt the convention that Q^ϕ is the singleton consisting of the zero vector of R^n.

Now for any fixed $b \in Q$, the nonnegative integer

$$\max \{\text{card } \sigma(x) \,|\, x \in X^z\}$$

is clearly attained by at least one element of X^z. Let $I(z)$ denote the support corresponding to any such element, and to eliminate the trivial case $X^z = \{0\}$ assume that $I(z) \neq \varnothing$. For any $x \in X^z$ it is easy to see, using the convexity of X^z and the maximality of card $I(z)$, that $\sigma(x) \subset I(z)$. It follows that $I(z)$ is independent of which element of X^z was chosen to define it, and also that X^z is contained in the particular face $Q^{I(z)}$. Similarly, if $Y^z \neq \{0\}$ then there is a well-defined minimal index set, call it $J(z)$, such that Y^z lies entirely within the face $Q^{J(z)}$ of Q. Finally, note that our earlier definitions yield $I(z) = \varnothing$ if $X^z = \{0\}$ (respectively $J(z) = \varnothing$ if $Y^z = \{0\}$), and in either case the convention $Q^\phi = \{0\}$ is just what's needed. Notice that, for all $z \in Q$, the "complementarity" conditions $x_k y_k = z_k$ in (C_z) imply that

$$I(z) \cap J(z) = \sigma(z) \,.$$

The next several results deal with a variant of Theorem 1 which involves replacing f and f^* by certain related functions. We define the (*coordinatewise*) *nondecreasing hull* of f to be the function

$$(\text{ndh } f)(x) = \inf \{f(x') \,|\, x' \in x + Q\}, \quad \forall x \in R^n \,.$$

It is easy to verify that ndh f is the greatest (coordinatewise) nondecreasing minorant of f. Similarly, the function

$$(\text{ndh } f^*)(y) = \inf \{f^*(y') \,|\, y' \in y + Q\}, \quad \forall y \in R^n \,,$$

called the nondecreasing hull of f^*, is the greatest nondecreasing minorant of f^*. Note that from now on the blanket hypothesis is again in force.

PROPOSITION 3. *The nondecreasing hulls of f and f^* are closed proper convex functions on R^n satisfying:*
  (i) $(\text{ndh } f)(x) = \min \{f(x') \,|\, x' \in x + Q\}, \; \forall x \in R^n,$
       $(\text{ndh } f^*)(y) = \min \{f^*(y') \,|\, y' \in y + Q\}, \; \forall y \in R^n;$

(ii) $(\text{ndh } f)(x) = \sup_{y \in Q} \{\langle x, y \rangle - (\text{ndh } f^*)(y)\}, \; \forall x \in Q,$

$(\text{ndh } f^*)(y) = \sup_{x \in Q} \{\langle x, y \rangle - (\text{ndh } f)(x)\}, \; \forall y \in Q;$

(iii) *for any pair* $(x, y) \in P \times P$,

$$f(x) + f^*(y) = \langle x, y \rangle \Longleftrightarrow (\text{ndh } f)(x) + (\text{ndh } f^*)(y) = \langle x, y \rangle,$$

and in this event

$$(\text{ndh } f)(x) = f(x) \; and \; (\text{ndh } f^*)(y) = f^*(y).$$

Proof. We can rewrite ndh f as

$$(\text{ndh } f)(x) = \inf \{f(x') + \psi_{-Q}(x'') \,|\, x = x' + x''\}$$
$$= (f \,\square\, \psi_{-Q})(x),$$

where $\square$ denotes infimal convolution. By the hypothesis $\varnothing \neq P \cap \text{dom } f^*$ and [6, Theorem 16.4],

$$(f^* + \psi_Q)^*(x) = (f \,\square\, \psi_{-Q})(x)$$

where moreover the infimal convolution is attained for each x. Hence

(2.11) $\text{ndh } f = f \,\square\, \psi_{-Q} = (f^* + \psi_Q)^*,$

where the infimal convolution is always attained. This shows in particular that ndh f, being the conjugate of a closed proper convex function, is itself closed proper convex. Similarly, the hypothesis $\varnothing \neq P \cap \text{dom } f$ implies

(2.12) $\text{ndh } f^* = f^* \,\square\, \psi_{-Q} = (f + \psi_Q)^*,$

where the infimal convolution is always attained. In particular, ndh f^* is also closed proper convex. Assertion (i) is now established. Now notice from the inf-convolution term in (2.12) that

$$\text{dom}\,(\text{ndh } f^*) = \text{dom } f^* - Q.$$

Hence

$$\varnothing \neq \text{ri}\,(\text{dom}\,(\text{ndh } f^*)) \cap \text{ri}\,(\text{dom } \psi_Q)$$

holds by virtue of our hypothesis $\varnothing \neq P \cap \text{dom } f^*$. Therefore we can apply the other part of (2.12) together with [6, Theorem 16.4] to obtain

$$\sup_{y \in Q} \{\langle x, y \rangle - (\text{ndh } f^*)(y)\} = (\text{ndh } f^* + \psi_Q)^*(x)$$
$$= ((f + \psi_Q) \,\square\, \psi_{-Q})(x)$$
$$= \inf \{(f + \psi_Q)(x') + \psi_{-Q}(x'') \,|\, x = x' + x''\}$$
$$= \inf \{f(x') \,|\, x' \in Q \cap (x + Q)\}.$$

It follows that

$$\sup_{y \in Q} \{\langle x, y \rangle - (\text{ndh } f^*)(y)\} \geqq (\text{ndh } f)(x), \quad \forall x \in R^n ,$$

with equality whenever $x \in Q$. In particular the first identity in (ii) is now established. The second identity in (ii) follows similarly from (2.11) and the hypothesis $\varnothing \neq P \cap \text{dom } f^*$. Let us now establish (iii). Note first that (ii) implies the general inequality

$$(2.13) \qquad \langle x, y \rangle \leqq (\text{ndh } f)(x) + (\text{ndh } f^*)(y), \quad \forall (x, y) \in Q \times Q .$$

Since $\text{ndh } f \leqq f$ and $\text{ndh } f^* \leqq f^*$, one implication in (iii) follows trivially. Now let $(x, y) \in P \times P$ satisfy

$$(2.14) \qquad\qquad (\text{ndh } f)(x) + (\text{ndh } f^*)(y) = \langle x, y \rangle .$$

By (i) we have

$$(2.15) \qquad\qquad (\text{ndh } f)(x) = f(x + x') \text{ for some } x' \in Q$$

and

$$(2.16) \qquad (\text{ndh } f*)(y) = f^*(y + y') \text{ for some } y' \in Q .$$

Hence

$$\langle x, y \rangle = f(x + x') + f^*(y + y')$$
$$\geqq \langle x + x', y + y' \rangle ,$$

where the inequality is just Fenchel's inequality. Therefore

$$0 \geqq \langle x, y' \rangle + \langle x', y \rangle + \langle x', y' \rangle .$$

Since each term on the right is nonnegative, this yields

$$\langle x, y' \rangle = 0 = \langle x', y \rangle .$$

Since x and y are in P, this forces

$$x' = 0 = y' .$$

This together with (2.14), (2.15), (2.16) yields

$$f(x) + f^*(y) = \langle x, y \rangle .$$

Assertion (iii) is now established.

Proposition 3(ii) says that the restrictions of $\text{ndh } f$ and $\text{ndh } f^*$ to Q are the monotone conjugates of each other, in the sense of Rockafellar [6, Theorem 12.4]. The "subdifferential" characterization given in Proposition 3(iii) will be extended to all pairs $(x, y) \in Q \times Q$ in Corollary 2A below. Notice that from Proposition 3(iii), together with parts (v) and (ii) of Theorem 1, it follows that

$$f(x^z) = (\mathrm{ndh}\, f)(x^z) \quad \text{and} \quad f^*(y^z) = (\mathrm{ndh}\, f^*)(y^z) \,, \quad \forall z \in P \,.$$

For each $z \in Q$ let $(\bar{P}_z)$ and $\bar{X}^z$ denote the problem and corresponding solution set obtained by replacing f in (P_z) by $\mathrm{ndh}\, f$. Similarly, let $(\bar{P}_z^*)$ and $\bar{Y}^z$ denote the problem and corresponding solution set obtained by replacing f^* in (P_z^*) by $\mathrm{ndh}\, f^*$. Finally, the present counterpart of the complementarity problem (C_z) is the problem

$$(\bar{C}_z) \quad \begin{cases} \text{find } (x, y) \in Q \times Q \text{ such that} \\ (\mathrm{ndh}\, f)(x) + (\mathrm{ndh}\, f^*)(y) = \langle x, y \rangle \text{ and } x_k y_k = z_k\ \forall k \,. \end{cases}$$

THEOREM 2. *For each $z \in Q$,*

(i) $(\bar{P}_z)$, $(\bar{P}_z^*)$ *and* $(\bar{C}_z)$ *each have at least one solution;*

(ii) $x \in \bar{X}^z$ *and* $y \in \bar{Y}^z$ *if and only if* (x, y) *solves* $(\bar{C}_z)$;

(iii) $\inf (\bar{P}_z) = \inf (P_z)$ *and* $\inf (\bar{P}_z^*) = \inf (P_z^*)$, *so that*

$$\inf (\bar{P}_z) + \inf (\bar{P}_z^*) = \sum (z_k - z_k \log z_k);$$

(iv) $\bar{X}^z$ *and* $\bar{Y}^z$ *are compact convex, and in fact,*

(2.17) $$Q^{I(z)} \supset \bar{X}^z = Q \cap (X^z - Q^{I(z) \setminus \sigma(z)}) \supset X^z \,,$$

(2.18) $$Q^{J(z)} \supset \bar{Y}^z = Q \cap (Y^z - Q^{J(z) \setminus \sigma(z)}) \supset Y^z \,;$$

(v) $\bar{X}^z = \{x^z\}$ *and* $\bar{Y}^z = \{y^z\}$ *if* $z \in P$.

Proof. First, notice that (2.11) implies

$$\mathrm{dom}\, (\mathrm{ndh}\, f) = \mathrm{dom}\, f - Q \,,$$

$$\mathrm{dom}\, (\mathrm{ndh}\, f)^* = \mathrm{dom}\, f^* \cap Q \,,$$

and from these it follows that

$$\varnothing \neq P \cap \mathrm{dom}\, (\mathrm{ndh}\, f) \Longleftarrow \varnothing \neq P \cap \mathrm{dom}\, f \,,$$

$$\varnothing \neq P \cap \mathrm{dom}\, (\mathrm{ndh}\, f)^* \Longleftarrow \varnothing \neq P \cap \mathrm{dom}\, f^* \,.$$

Thus, our blanket hypothesis is met by $\mathrm{ndh}\, f$, and so Theorem 1 applies to the modified problems

(2.19) $$\inf \{\mathrm{ndh}\, f - z \circ g\}$$

and

(2.20) $$\inf \{(\mathrm{ndh}\, f)^* - z \circ g\}$$

obtained by replacing f by $\mathrm{ndh}\, f$ in both (P_z) and (P_z^*). Since problem (2.19) is $(\bar{P}_z)$ and problem (2.20) is (P_z^*) (since $(\mathrm{ndh}\, f)^* = f^* + \psi_Q$), Theorem 1(i) applied twice yields

$$\inf(\bar{P}_z) = -\inf(P_z^*) + \sum(z_k - z_k \log z_k) = \inf(P_z) \, .$$

A similar argument based on (2.12) yields

$$\inf(\bar{P}_z^*) = -\inf(P_z) + \sum(z_k - z_k \log z_k) = \inf(P_z^*) \, .$$

This establishes assertion (iii). Now by (2.13) and (2.10), for each $x, y \in Q$ we have

(2.19)
$$(\mathrm{ndh}\, f)(x) + (\mathrm{ndh}\, f^*)(y) \geqq \langle x, y \rangle$$
$$\geqq (z \circ g)(x) + (z \circ g)(y) + \sum(z_k - z_k \log z_k) \, .$$

Hence the condition that $x \in \bar{X}^z$ and $y \in \bar{Y}^z$, which by (iii) is equivalent to

$$(\mathrm{ndh}\, f)(x) - (z \circ g)(x) + (\mathrm{ndh}\, f^*)(y) - (z \circ g)(y) = \sum(z_k - z_k \log z_k) \, ,$$

is in turn equivalent to the two equations

$$(\mathrm{ndh}\, f)(x) + (\mathrm{ndh}\, f*)(y) = \langle x, y \rangle \, ,$$

$$(z \circ g)(x) + (z \circ g)(y) + \sum(z_k - z_k \log z_k) = \langle x, y \rangle \, .$$

Since by (2.9) and (2.10) the latter equation is equivalent to the condition

$$x \in Q, \, y \in Q, \, x_k y_k = z_k \, \forall k \, ,$$

assertion (ii) follows. Now suppose for the moment that (2.17) and (2.18) have already been established. We claim that the rest of the theorem follows easily. Indeed, since X^z is nonempty and bounded (Theorem 1), (2.17) implies that $\bar{X}^z$ is also. Since $\bar{X}^z$ is automatically closed and convex (recall $\mathrm{ndh}\, f$ and $-z \circ g$ are each closed convex), this shows that $\bar{X}^z$ is nonempty compact convex. Similarly, (2.18) implies $\bar{Y}^z$ is nonempty compact convex. This establishes the rest of assertion (iv) and also (by (ii)) assertion (i). Assertion (v) follows trivially from the set equalities in (2.17) and (2.18), because $z \in P$ implies $\sigma(z) = \{1, \cdots, n\}$ and hence

$$I(z) \backslash \sigma(z) = \varnothing = J(z) \backslash \sigma(z) \, .$$

It remains, then, to establish (2.17) and (2.18). Of these, we shall only prove (2.17), as the proof of (2.18) is analogous. Let y be any solution to problem (2.20) having maximal support, i.e.,

(2.22)
$$y \in Y^z \text{ and } \sigma(y) = J(z) \, .$$

By Theorem 1(ii), $\bar{x} \in \bar{X}^z$ if and only if $(\bar{x}, y)$ solves the complementarity problem associated with problems (2.19), (2.20):

(2.23)
$$\bar{x} \in Q, \, y \in Q, \, \bar{x}_k y_k = z_k \, \forall k$$

and

(2.24) $(\text{ndh } f)(\bar{x}) + (\text{ndh } f)^*(y) = \langle \bar{x}, y \rangle$.

Using (2.11) together with [6, Theorem 23.8], one sees that (2.24) is equivalent to

$$\bar{x} = x + x' \text{ for some } x \in \partial f^*(y) \text{ and } x' \in \partial \psi_Q(y) .$$

Since $x' \in \partial \psi_Q(y)$ is equivalent to $y \in Q$, $-x' \in Q$ and $-x_k' y_k = 0 \ \forall k$, it follows (in view of (2.22)) that conditions (2.23), (2.24) occur if and only if

$$\bar{x} \in Q \cap (x - Q^{J'(z)})$$

for some $x \in Q$ satisfying $x \in \partial f^*(y)$ and $x_k y_k = z_k \ \forall k$, where $J'(z) = \{1, \cdots, n\} \backslash J(z)$. But the conditions on x just listed mean precisely that $x \in X^z$, in view of (2.22) and Theorem 1(ii). Hence

$$\bar{X}^z = Q \cap (X^z - Q^{J'(z)}) .$$

Since $X^z \subset Q$ and $\sigma(z) = I(z) \cap J(z)$, it can be checked that

$$Q \cap (X^z - Q^{J'(z)}) = Q \cap (X^z - Q^{I(z) \backslash \sigma(z)}) .$$

This establishes the equality part of (2.17). Since $X^z \subset Q^{I(z)}$, the equality implies $\bar{X}^z \subset Q^{I(z)}$, and since $0 \in Q^{I(z) \backslash \sigma(z)}$, it trivially yields $\bar{X}^z \supset X^z$. This concludes the proof of (2.17) and the theorem.

COROLLARY 2A. *For each* $(\bar{x}, \bar{y}) \in Q \times Q$,

$$(\text{ndh } f)(\bar{x}) + (\text{ndh } f^*)(\bar{y}) = \langle \bar{x}, \bar{y} \rangle$$

if and only if there exists $(x, y) \in Q \times Q$ *such that*

$$f(x) + f^*(y) = \langle x, y \rangle ,$$

$$x_k y_k = \bar{x}_k \bar{y}_k \ \forall k ,$$

$$\left. \begin{array}{l} \bar{x}_k \leqq x_k \text{ if } \bar{x}_k \bar{y}_k = 0 \text{ and } x_k > 0 \\ \bar{x}_k = x_k \text{ otherwise} \end{array} \right\} \forall k ,$$

$$\left. \begin{array}{l} \bar{y}_k \leqq y_k \text{ if } \bar{x}_k \bar{y}_k = 0 \text{ and } y_k > 0 \\ \bar{y}_k = y_k \text{ otherwise} \end{array} \right\} \forall k .$$

Proof. Let $(\bar{x}, \bar{y}) \in Q \times Q$ be given, and define $z \in Q$ by setting $z_k = \bar{x}_k \bar{y}_k \ \forall k$. Then

$$(\text{ndh } f)(\bar{x}) + (\text{ndh } f^*)(\bar{y}) = \langle \bar{x}, \bar{y} \rangle$$

is equivalent to $(\bar{x}, \bar{y})$ solving $(\bar{C}_z)$, which by parts (ii) and (iv) of Theorem 2 is equivalent to the existence of vectors $x \in X^z$ and

$y \in Y^z$ such that

$$\bar{x}_k = x_k \ \forall k \in I'(z) \cup \sigma(z)$$

$$\bar{x}_k \leqq x_k \ \forall k \in I(z) \backslash \sigma(z)$$

and

$$\bar{y}_k = y_k \ \forall k \in J'(z) \cup \sigma(z)$$

$$\bar{y}_k \leqq y_k \ \forall k \in J(z) \backslash \sigma(z) \ ,$$

where we write

$$I'(z) = \{1, \cdots, n\} \backslash I(z) \ \text{and} \ J'(z) = \{1, \cdots, n\} \backslash J(z) \ .$$

Using

$$\sigma(z) \subset \sigma(x) \subset I(z) \ \text{and} \ \sigma(z) \subset \sigma(y) \subset J(z)$$

together with Theorem 1(ii), one can check that these conditions are equivalent to the ones asserted in the corollary.

This section concludes with two new general facts concerning the geometric nature of subdifferential mappings in R^n.

THEOREM 3. *Let f be any closed proper convex function on R^n satisfying*

$$\varnothing \neq P \cap \text{dom} \, f \quad and \quad \varnothing \neq P \cap \text{dom} \, f^* \ .$$

Then $\Gamma = (P \times P) \cap \text{graph} \, \partial f$ satisfies

$$(2.25) \qquad \langle x' - x, y' - y \rangle \geqq 0, \quad \forall (x, y), (x', y') \in \Gamma \ ,$$

and for each $(x, y) \in (P \times P) \backslash \Gamma$ there exists a pair $(x', y') \in \Gamma$ such that

$$(2.26) \qquad \langle x' - x, y' - y \rangle < 0 \ .$$

That is, Γ is a maximal monotone subset of $P \times P$.

Proof. The monotonicity property (2.25) is easy to verify. Now let $(x, y) \in (P \times P) \backslash \Gamma$. We shall exhibit an $(x', y') \in \Gamma$ satisfying (2.26). Define a vector $z \in P$ by setting $z_k = x_k y_k \ \forall k$. By Theorem 1 there exists a pair $(x', y') \in \Gamma$ such that $x'_k y'_k = z_k \ \forall k$. Then

$$\langle x' - x, y' - y \rangle = \langle x', y' \rangle - \langle x', y \rangle - \langle x, y' \rangle + \langle x, y \rangle$$
$$= \sum z_k - \sum z_k(x'_k/x_k) - \sum z_k(y'_k/y_k) + \sum z_k$$

yields

$$(2.27) \qquad 2 - \zeta^{-1} \langle x' - x, y' - y \rangle = \sum a_k(x'_k/x_k) + \sum a_k(y'_k/y_k) \ ,$$

where

$$\zeta = \sum z_k \quad \text{and} \quad a_k = \zeta^{-1} z_k \; \forall k \; .$$

Now consider two cases.

Case 1: either there is no $\sigma > 0$ such that $x' = \sigma x$ or else there is no $\tau > 0$ such that $y' = \tau y$. Then we can apply the arithmetic-geometric mean inequality to each summation on the right-hand side of (2.27), obtaining

$$2 - \zeta^{-1} \langle x' - x, \, y' - y \rangle > \prod (x'_k/x_k)^{a_k} + \prod (y'_k/y_k)^{a_k} \; .$$

This together with another application of the arithmetic-geometric mean inequality yields

$$
\begin{aligned}
1 - (2\zeta)^{-1} \langle x' - x, \, y' - y \rangle &> \frac{1}{2} \prod (x'_k/x_k)^{a_k} + \frac{1}{2} \prod (y'_k/y_k)^{a_k} \\
&\geqq \{ \prod (x'_k/x_k)^{a_k} \}^{1/2} \cdot \{ \prod (y'_k/y_k)^{a_k} \}^{1/2} \\
&= \{ \prod (x'_k y'_k / x_k y_k)^{a_k} \}^{1/2} \\
&= 1 \; ,
\end{aligned}
$$

where the last equality uses $x'_k y'_k = z_k = x_k y_k \; \forall k$. Hence (2.26) holds.

Case 2: there exists $\sigma > 0$ such that $x' = \sigma x$ and there exists $\tau > 0$ such that $y' = \tau y$. Then

$$x'_k y'_k = \sigma \tau x_k y_k = \sigma \tau z_k = \sigma \tau x'_k y'_k \; \forall k$$

implies that $\sigma \tau = 1$. Now we cannot have $(\sigma, \tau) = (1, 1)$, since that would imply that $(x', y') = (x, y)$ is both in and out of graph ∂f. Hence $(\sigma, \tau) \neq (1, 1)$, which implies that $\sigma + \tau > 2$ (since $\sigma \tau = 1$). This together with (2.27) and $\sum a_k = 1$ yields

$$2 - \zeta^{-1} \langle x' - x, \, y' - y \rangle = \sum a_k \sigma + \sum a_k \tau = \sigma + \tau > 2 \; .$$

Hence (2.26) holds in this case also, and the proof is complete.

We remark, incidentally, that the uniqueness result Theorem 1(v) can be given an alternate, operator-theoretic proof based on the above argument. Indeed, for fixed $z \in P$ suppose that

$$(x, \, y), \, (x', \, y') \in \text{graph } \partial f \cap \text{graph } \partial (z \circ g) \; .$$

Then $(x, y), (x', y') \in \Gamma$, with $x_k y_k = z_k = x'_k y'_k \; \forall k$. Hence,

$$0 \leqq \langle x' - x, \, y' - y \rangle < 0$$

unless we find ourselves in Case 2 above with $\sigma = \tau = 1$, and in that situation $(x, y) = (x', y')$.

THEOREM 4. *Let f be any closed proper convex function on R^n satisfying*

$$\varnothing \neq P \cap \operatorname{dom} f \quad and \quad \varnothing \neq P \cap \operatorname{dom} f^* .$$

Then the mapping $h(z) = (x^z, y^z)$ is one-to-one from P onto $\Gamma = (P \times P) \cap \operatorname{graph} \partial f$, and it is continuous with continuous inverse $h^{-1}(x, y) = z$, where $z_k = x_k y_k \, \forall k$. That is, h is a homeomorphism between P and Γ.

Proof. Consider the function $l: \Gamma \to P$ given by $l(x, y) = z$, where $z_k = x_k y_k \, \forall k$. This is one-to-one, by the uniqueness part of Theorem 1. By the existence part of Theorem 1, for each $z \in P$ there is a pair $(x^z, y^z) \in \Gamma$ satisfying $l(x^z, y^z) = z$. Hence l is onto P and (since it is also one-to-one) its inverse is the function h. Clearly l is continuous. The continuity of h is established in Corollary 5A below, which does not depend on the present result.

In Corollary 11B, results are established which are similar to Theorems 3 and 4 but which involve the positive reals instead of the positive orthant.

3. Parametric analysis of optimal values. Central to the analysis of both $\inf (P_z)$ and X^z as functions of $z \in Q$ is certain joint convexity structure related to our penalty terms. This is presented in Proposition 4, which depends on the following.

LEMMA 2. *The function $\sigma: R^2 \to [-\infty, +\infty)$ defined by*

$$(3.1) \qquad \sigma(\xi, \zeta) = \begin{cases} \zeta \log \xi + \zeta - \zeta \log \zeta & if \ \xi > 0 \ and \ \zeta > 0 \\ 0 & if \ \xi \geqq 0 \ and \ \zeta = 0 \\ -\infty & otherwise \end{cases}$$

is positively homogeneous closed proper concave with conjugate

$$(3.2) \qquad \sigma^*(\eta, \omega) = \begin{cases} 0 & if \ \eta > 0 \ and \ 0 \leqq \omega + \log \eta \\ -\infty & otherwise , \end{cases}$$

and $(\eta, \omega) \in \partial \sigma(\xi, \zeta)$ if and only if

$$(3.3) \qquad \begin{cases} either \ \ \zeta > 0, \ \xi > 0, \ \eta > 0, \ \xi\eta = \zeta, \ 0 = \omega + \log \eta \\ or \ \ \ \ \ \zeta = 0, \ \xi = 0, \ \eta > 0, \ \ \ \ \ \ \ \ \ \ \ 0 \leqq \omega + \log \eta . \end{cases}$$

Proof. One could argue this directly. However, for brevity we indicate an alternate proof. It is easy to check that

$$(3.4) \qquad \sigma(\xi, \zeta) = \begin{cases} (h^*\zeta)(\xi) & \text{if } \zeta > 0 \\ (h^*0^+)(\xi) & \text{if } \zeta = 0 \\ -\infty & \text{if } \zeta < 0 \,, \end{cases}$$

where h is the function defined in (2.1) and we have used $h^*(\eta) = h(\eta) + 1$ from (2.5). The lemma now follows from (3.4) and a one-dimensional application of [3, Proposition 1], which provides formulas for the conjugate and subdifferential of the indicator function of a general convex epigraph set.

PROPOSITION 4. *Let* $\gamma \colon R^{2n} \to [-\infty, +\infty)$ *be the function*

$$(3.5) \qquad \gamma(x, z) = \begin{cases} (z \circ g)(x) + \sum (z_k - z_k \log z_k) & \text{if } z \in Q \\ -\infty & \text{if } z \notin Q \,, \end{cases}$$

where $z \circ g$ *is as defined in* (2.7), (2.2), (2.1). *Then* γ *is a positively homogeneous closed proper concave function with conjugate*

$$(3.6) \qquad \gamma^*(y, w) = \begin{cases} 0 & \text{if } y \in P \text{ and } 0 \le w_k + \log y_k \ \forall k \\ -\infty & \text{otherwise} \,, \end{cases}$$

and $(y, w) \in \partial\gamma(x, z)$ *if and only if*

$$(3.7) \qquad \begin{cases} x \in Q, z \in Q, y \in P, x_k y_k = z_k \ \forall k \,, \\ 0 = w_k + \log y_k \text{ if } z_k > 0, 0 \le w_k + \log y_k \text{ if } z_k = 0 \,. \end{cases}$$

Proof. For each $z \in Q$, (2.5) implies that

$$\gamma(x, z) = \sum (z_k h)^*(x_k) \,.$$

Using (2.2) and the fact that $h^* = h + 1$ has recession function $-\psi_{[0,+\infty)}$, one can check that

$$(z_k h)^*(x_k) = \begin{cases} (h^* z_k)(x_k) & \text{if } z_k > 0 \\ (h^* 0^+)(x_k) & \text{if } z_k = 0 \,. \end{cases}$$

Together with (3.4), this information yields

$$(3.8) \qquad \gamma(x, z) = \sum \sigma(x_k, z_k) \,.$$

From this additive separability structure it follows that

$$(3.9) \qquad \gamma^*(y, w) = \sum \sigma^*(y_k, w_k)$$

and

$$(3.10) \qquad (y, w) \in \partial\gamma(x, z) \Longleftrightarrow (y_k, w_k) \in \partial\sigma(x_k, z_k) \ \forall k \,.$$

Combining (3.9) with (3.2) yields (3.6), and combining (3.10) with

(3.3) yields (3.7). This concludes the proof.

Define functions ν_f and ν_{f*} on R^n by setting

$$(3.11) \qquad \nu_f(z) = \begin{cases} \inf (P_z) & \text{if } z \in Q \\ -\infty & \text{if } z \notin Q , \end{cases}$$

$$(3.12) \qquad \nu_{f*}(z) = \begin{cases} \inf (P_z^*) & \text{if } z \in Q \\ -\infty & \text{if } z \notin Q . \end{cases}$$

By Theorem 1(iii) we already know that

$$(3.13) \qquad \nu_f(z) + \nu_{f*}(z) = \sum (z_k - z_k \log z_k), \quad \forall z \in Q .$$

For a further study of ν_f and ν_{f*}, the following auxiliary functions are useful:

$$(3.14) \qquad \mu_f(z) = \begin{cases} \inf (P_z) - \sum (z_k - z_k \log z_k) & \text{if } z \in Q \\ +\infty & \text{if } z \notin Q , \end{cases}$$

$$(3.15) \qquad \mu_{f*}(z) = \begin{cases} \inf (P_z^*) - \sum (z_k - z_k \log z_k) & \text{if } z \in Q \\ +\infty & \text{if } z \notin Q . \end{cases}$$

From (3.13) it is immediate that

$$(3.16) \qquad \nu_f = -\mu_{f*}, \quad \nu_{f*} = -\mu_f .$$

THEOREM 5. *The functions ν_f and ν_{f*} given by (3.11) and (3.12) are closed proper concave with effective domain Q. They are continuous relative to all of Q, continuously differentiable on P, and twice differentiable almost everywhere on P. Furthermore, their conjugates are given by*

$$(3.17) \qquad \nu_f^*(w) = -\min \{f(x) \,|\, x_k \geqq \exp (-w_k) \,\forall k\} ,$$

$$(3.18) \qquad \nu_{f*}^*(w) = -\min \{f^*(y) \,|\, y_k \geqq \exp (-w_k) \,\forall k\} ,$$

and their subdifferentials are given by

$$(3.19) \quad w \in \partial \nu_f(z) \Longrightarrow \begin{cases} z \in Q \text{ and } \exists (x, y) \in (P \times Q) \cap \text{graph } \partial f \\ \text{such that } x_k y_k = z_k \,\forall k \text{ and} \\ -w_k = \log x_k \text{ if } z_k > 0, \ -w_k \leqq \log x_k \text{ if } z_k = 0 , \end{cases}$$

$$(3.20) \quad w \in \partial \nu_{f*}(z) \Longrightarrow \begin{cases} z \in Q \text{ and } \exists (x, y) \in (Q \times P) \cap \text{graph } \partial f \\ \text{such that } x_k y_k = z_k \,\forall k \text{ and} \\ -w_k = \log y_k \text{ if } z_k > 0, \ -w_k \leqq \log y_k \text{ if } z_k = 0 . \end{cases}$$

In particular, for each $z \in P$,

$$\nabla \nu_f(z) = -L(x^z) \text{ and } \nabla \nu_{f*}(z) = -L(y^z)$$

where $L: P \to R^n$ is the mapping

$$L(z_1, \cdots, z_n) = (\log z_1, \cdots, \log z_n) .$$

Proof. We shall prove only the results concerning ν_{f*}, as those for ν_f are proved similarly. By (3.16) it suffices to prove corresponding results for μ_f, since in particular

(3.21) $$\nu_{f*}^*(w) = -\mu_f^*(-w) ,$$

(3.22) $$w \in \partial \nu_{f*}(z) \Longleftrightarrow -w \in \partial \mu_f(z) .$$

We begin by observing that

(3.23) $$\begin{cases} \mu_f(z) = \inf_x \{f(x) - (z \circ g)(x)\} - \sum (z_k - z_k \log z_k) \\ \qquad = \inf_x \{\varphi(x, z) - \gamma(x, z)\} , \end{cases}$$

where γ is as in (3.5) and φ is the function on R^{2n} defined by

(3.24) $$\varphi(x, z) = f(x), \quad \forall (x, z) \in R^{2n} .$$

Let $A: R^{2n} \to R^n$ be the linear transformation $A(x, z) = z$. Then (3.23) can be expressed as

(3.25) $$\mu_f(z) = (A(\varphi - \gamma))(z) .$$

(See [6] for the operation of forming the image of a convex function under a linear transformation.) The remainder of the proof consists of establishing various facts about μ_f, working from the representation (3.25). Let A^* denote the adjoint linear transformation, i.e., $A^*(w) = (0, w)$. We shall show that $\varphi - \gamma$ and $(\varphi - \gamma)^* A^*$ are closed proper convex and also derive their conjugates and subdifferentials. Observe first from (3.24) that

(3.26) $$\varphi^*(y, w) = \begin{cases} f^*(y) & \text{if } w = 0 \\ +\infty & \text{if } w \neq 0 , \end{cases}$$

(3.27) $$(y, w) \in \partial \varphi(x, z) \Longleftrightarrow z \in R^n, w = 0, y \in \partial f(x) .$$

It is easily seen that

$$\varnothing \neq \mathrm{ri}\,(\mathrm{dom}\,\varphi) \cap \mathrm{ri}\,(\mathrm{dom}\,(-\gamma)) \Longleftrightarrow \varnothing \neq P \cap \mathrm{dom}\,f.$$

Hence, by the blanket hypothesis and [6, Theorems 16.4 and 23.8], it follows that $\varphi - \gamma$ is closed proper convex with conjugate

(3.28) $$\begin{aligned} (\varphi - \gamma)^*(y, w) = \min \{\varphi^*(y_1, w_1) - \gamma^*(-y_2, -w_2)| \\ y = y_1 + y_2, w = w_1 + w_2\} \end{aligned}$$

and subdifferential satisfying

$$(3.29) \qquad \partial(\varphi - \gamma)(x, z) = \partial\varphi(x, z) - \partial\gamma(x, z) \, .$$

Now (3.28) implies, in particular, that

$$\text{dom}\,(\varphi - \gamma)^* = \text{dom}\,\varphi^* - \text{dom}\,\gamma^* \, .$$

Hence, $\varnothing \neq \text{range}\,A^* \cap \text{ri}\,(\text{dom}\,(\varphi - \gamma)^*)$

$$\begin{aligned}
&\Longleftrightarrow \exists w,\, (0,\, w) \in \text{ri}\,(\text{dom}\,\varphi^*) - \text{ri}\,(\text{dom}\,\gamma^*) \\
&\Longleftrightarrow \exists w \exists y,\, y \in \text{ri}\,(\text{dom}\,f^*) \text{ and } (y,\, -w) \in \text{ri}\,(\text{dom}\,\gamma^*) \\
&\Longleftrightarrow \exists y \in P \cap \text{ri}\,(\text{dom}\,f^*) \quad \exists w,\, w_k < \log y_k \,\forall k \\
&\Longleftrightarrow \exists y \in P \cap \text{ri}\,(\text{dom}\,f^*) \\
&\Longleftrightarrow \varnothing \neq P \cap \text{dom}\,f^* \, ,
\end{aligned}$$

where we have used (3.26) and (3.6) to find $\text{dom}\,\varphi^*$ and $\text{dom}\,\gamma^*$ and then [6, Theorem 6.8] to find $\text{ri}\,(\text{dom}\,\gamma^*)$. Therefore, by the blanket hypothesis and [6, Theorems 16.3 and 23.9], it follows that $(\varphi - \gamma)^* A^*$ is closed proper convex with conjugate

$$(3.30) \qquad \begin{aligned}
((\varphi - \gamma)^* A^*)^*(z) &= (A^{**}(\varphi - \gamma)^{**})(z) \\
&= (A(\varphi - \gamma))(z) \\
&= \min_x \{\varphi(x,\, z) - \gamma(x,\, z)\}
\end{aligned}$$

and subdifferential satisfying

$$(3.31) \qquad \partial((\varphi - \gamma)^* A^*)(w) = A\partial(\varphi - \gamma)(A^*(w)) \, .$$

By (3.25) and (3.30), μ_f is the conjugate of a closed proper convex function; hence μ_f itself is such a function. From the finiteness of $\inf\,(P_z)$ for all $z \in Q$ (Theorem 1(iii)), it is clear that $\text{dom}\,\mu_f = Q$. Since Q is polyhedral convex, hence locally simplicial, [6, Theorem 10.2] implies that μ_f is continuous relative to all of Q. Combining (3.23) with (3.28) and (3.26) yields

$$(3.32) \qquad \begin{aligned}
\mu_f^*(w) &= ((\varphi - \gamma)^* A^*)(w) \\
&= \min_y \{\varphi^*(y,\, 0) - \gamma^*(y,\, -w)\} \\
&= \min_y \{f^*(y)\,|\,y \in P \text{ and } w_k \leqq \log y_k \forall k\} \\
&= \min_y \{f^*(y)\,|\,y_k \geqq \exp w_k \,\forall k\} \, .
\end{aligned}$$

By (3.21), (3.32) implies (3.18). Now observe that [6, Theorem 23.5], (3.25) and (3.30), together with (3.31), (3.29), (3.27) and (3.7), imply that

$$w \in \partial \mu_f(z) \iff z \in \partial((\varphi - \gamma)^* A^*)(w)$$

$$\iff \exists x, \ (x, z) \in \partial(\varphi - \gamma)^*(0, w)$$

$$\iff \exists x, \ (0, w) \in \partial\varphi(x, z) - \partial\gamma(x, z)$$

$$\iff \exists x \exists y, \ (y, 0) \in \partial\varphi(x, z) \text{ and } (y, -w) \in \partial\gamma(x, z)$$

(3.33) $\iff \begin{cases} \exists (x, y) \in \text{graph } \partial f \text{ such that} \\ x \in Q, \ z \in Q, \ y \in P, \ x_k y_k = z_k \ \forall k \ . \\ w_k = \log y_k \text{ if } z_k > 0, \ w_k \leqq \log y_k \text{ if } z_k = 0 \ . \end{cases}$

By (3.22), (3.33) implies (3.20). Finally, notice that (3.33) implies for each $z \in P$ that

$$w \in \partial \mu_f(z) \iff \begin{cases} \exists (x, y) \in (P \times P) \cap \text{graph } \partial f \text{ such that} \\ x_k y_k = z_k \text{ and } w_k = \log y_k \ \forall k \ . \end{cases}$$

Now by Theorem 1, the only pair (x, y) which can possibly satisfy the conditions on the right is (x^z, y^z). It follows that for each $z \in P$,

$$w \in \partial \mu_f(z) \iff w_k = \log y_k^z \ \forall k \ .$$

Hence [6, Theorem 25.1], μ_f is differentiable on P with

(3.34) $\qquad \nabla \mu_f(z) = (\log y_1^z, \cdots, \log y_n^z), \quad \forall z \in P \ .$

Since μ_f is finite convex on P, [6, Theorem 25.5] implies that μ_f is continuously differentiable on P, and Alexandroff's Theorem [1] implies that μ_f is twice differentiable almost everywhere on P. This concludes the proof.

COROLLARY 5A. *The mapping* $z \to (x^z, y^z)$ *is continuous on* P, *and in fact, differentiable almost everywhere on* P.

Proof. It suffices to establish the asserted properties for each of the two vector coordinates separately. So consider, for example, the mapping $z \to y^z$, and let $E \colon R^n \to P$ be the mapping

$$E(w_1, \cdots, w_n) = (\exp w_1, \cdots, \exp w_n) \ .$$

By (3.34),

(3.35) $\qquad E(\nabla \mu_f(z)) = y^z, \quad \forall z \in P \ .$

Since it was shown in the theorem that $\nabla \mu_f$ has the properties in question relative to P, and since these properties are preserved under composition with E, the proof is complete by (3.35).

We conclude § 3 by establishing a variant of Theorem 5 to deal

with a situation treated extensively in the sequel, namely, the situation in which z is made to approach 0 along a given ray in P. Let there be given a fixed vector

(3.36) $a \in P$ such that $\sum a_k = 1$,

and consider the z's of the form $z = \zeta a$. Write (P_ζ), inf (P_ζ), x^ζ, $\nu_f(\zeta)$, etc. to denote the previously treated objects when regarded as restricted to such z's. Thus, for example,

$$\nu_{f^*}(\zeta) = \begin{cases} \inf (P_\zeta^*) & \text{if } \zeta \geqq 0 \\ -\infty & \text{if } \zeta < 0 , \end{cases}$$

where inf $(P_\zeta^*) = \inf \{f^* - \zeta a \circ g\}$ for each $\zeta \geqq 0$. *In all discussions involving the one-dimensional parameter ζ, as opposed to the n-dimensional parameter z, we assume that a is a given fixed vector satisfying* (3.36).

THEOREM 6. *The functions $\nu_f(\zeta)$ and $\nu_{f^*}(\zeta)$ described above are closed proper concave with effective domain $[0, +\infty)$. They are continuous relative to $[0, +\infty)$, continuously differentiable on $(0, +\infty)$, and twice differentiable almost everywhere on $(0, +\infty)$. Furthermore, their conjugates are given by*

(3.37) $\nu_f^*(\omega) = -\min \{f(x) \mid \prod x_k^{a_k} \geqq \exp(-\omega)\}$,

(3.38) $\nu_{f^*}^*(\omega) = -\min \{f^*(y) \mid \prod y_k^{a_k} \geqq \exp(-\omega)\}$,

and their subdifferentials are given by

(3.39) $\partial \nu_f(\zeta) = \begin{cases} \{-\sum a_k \log x_k^\zeta\} & \text{if } \zeta > 0 \\ \{\omega \mid \exists x \in P \cap \partial f^*(0), \ \omega \geqq -\sum a_k \log x_k\} & \text{if } \zeta = 0 \\ \varnothing & \text{if } \zeta < 0 , \end{cases}$

(3.40) $\partial \nu_{f^*}(\zeta) = \begin{cases} \{-\sum a_k \log y_k^\zeta\} & \text{if } \zeta > 0 \\ \{\omega \mid \exists y \in P \cap \partial f(0), \ \omega \geqq -\sum a_k \log y_k\} & \text{if } \zeta = 0 \\ \varnothing & \text{if } \zeta < 0 . \end{cases}$

Proof. We shall prove only the assertions concerning $\nu_{f^*}(\zeta)$, as those for $\nu_f(\zeta)$ are proved similarly. Let $B: R \to R^n$ be the linear transformation $B(\zeta) = \zeta a$, with adjoint $B^*(w) = \langle a, w \rangle$. Since (3.16) implies

(3.41) $\nu_{f^*}(\zeta) = -(\mu_f B)(\zeta), \quad \forall \zeta \in R$,

it suffices to prove the corresponding assertions about $\mu_f B$. Using the facts about μ_f established in the course of proving Theorem 5,

we obtain by [6, Theorems 9.5, 16.3 and 23.9] that $\mu_f B$ is a closed proper convex function on R having effective domain $[0, +\infty)$, its conjugate is

$$(3.42) \qquad (\mu_f B)^*(\omega) = (B^* \mu_f^*)(\omega) = \min \{\mu_f^*(w) \,|\, B^* w = \omega\} \,,$$

and its subdifferential satisfies

$$(3.43) \qquad \partial(\mu_f B)(\zeta) = B^* \partial \mu_f(B(\zeta)) \,.$$

Combining (3.43) with (3.33), we obtain that $\omega \in \partial(\mu_f B)(\zeta)$ if and only if either $\zeta > 0$ and $\omega = \sum a_k \log y_k^\zeta$ or else $\zeta = 0$ and $\omega \leqq \sum a_k \log y_k$ for some $y \in P \cap \partial f(0)$. In view of (3.41) this proves (3.40). Combining (3.42) with (3.32), we obtain that

$$(\mu_f B)^*(\omega) = \min_{\langle a, w \rangle = \omega} \min \{f^*(y) \,|\, y_k \geqq \exp w_k \; \forall k\}$$
$$= \min_S \{f^*(y)\} \,,$$

where $S = \{(w, y) \,|\, \omega = \langle a, w \rangle \text{ and } w_k \leqq h(y_k) \; \forall k\}$ (and h is given by (2.1)). Hence,

$$(\mu_f B)^*(\omega) \geqq \inf_T \{f^*(y)\} \,,$$

where $T = \{(w, y) \,|\, \omega = \langle a, w \rangle \leqq \sum a_k h(y_k)\}$. On the other hand, for any y satisfying $\omega \leqq \sum a_k h(y_k)$ the w defined via

$$w_1 = h(y_1) - a_1^{-1}(\sum a_k h(y_k) - \omega), \quad w_k = h(y_k) \text{ for } k = 2, \cdots, n$$

satisfies $\omega = \langle a, w \rangle$ (as well as $w_k \leqq h(y_k) \; \forall k$). It follows that

$$(\mu_f B)^*(\omega) = \min_T \{f^*(y)\}$$
$$= \min \{f^*(y) \,|\, \omega \leqq \sum a_k h(y_k)\}$$
$$= \min \{f^*(y) \,|\, \prod y_k^{a_k} \geqq \exp \omega\} \,.$$

In view of (3.41), this proves (3.38). The assertions about continuity, continuous differentiability, and almost everywhere twice differentiability follow from what has already been proved, much as in the proof of Theorem 5.

COROLLARY 6A. *For each $\zeta \in (0, +\infty)$ write*

$$\xi(\zeta) = \prod (x_k^\zeta)^{a_k} \,, \qquad \eta(\zeta) = \prod (y_k^\zeta)^{a_k} \,.$$

Then the mapping $\zeta \to (\xi(\zeta), \eta(\zeta))$ is continuous and coordinatewise nondecreasing from $(0, +\infty)$ into $(0, +\infty) \times (0, +\infty)$, and in fact is differentiable almost everywhere on $(0, +\infty)$.

Proof. It suffices to establish the asserted properties for each of the two coordinate functions separately. By the theorem, $\nu_f(\zeta)$

is concave, continuously differentiable on $(0, +\infty)$, and in fact twice differentiable almost everywhere there. Since

$$\nabla \nu_f(\zeta) = -\log \xi(\zeta) \qquad \forall \zeta \in (0, +\infty)$$

by (3.39), the asserted properties for $\zeta \to \xi(\zeta)$ on $(0, +\infty)$ now follow.

The mapping described in Corollary 6A has other properties also, as explained below in Theorem 11 and Corollary 11B.

4. Parametric analysis of optimal solutions. We begin by establishing some general facts concerning the behavior of the solution sets of certain linearly perturbed versions of (P_z) and (P_z^*).

For each $y \in R^n$, write $f_y = f - \langle \cdot, y \rangle$. Consider the problems

$$(P_{z,y}) \qquad\qquad \min \{f_y - (z \circ g)\},$$

which are perturbed versions of (P_z). Clearly $\varnothing \neq P \cap \mathrm{dom}\, f_y$, and since $(f_y)^* = f^*(\cdot + y)$, we have $\varnothing \neq P \cap \mathrm{dom}\, (f_y)^*$ if and only if $y \in D = \mathrm{dom}\, f^* - P$. Hence, for each $(z, y) \in Q \times D$ the blanket hypothesis is met for $(P_{z,y})$. Denote the set of solutions to $(P_{z,y})$ by $X^{z,y}$, and let $I(z, y)$ denote the maximal support of these solutions (in the sense defined following Corollary 1C).

Similarly, for each $x \in R^n$ write $f_x^* = f^* - \langle x, \cdot \rangle$ and consider the class of problems

$$(P_{z,x}^*) \qquad\qquad \min \{f_x^* - (z \circ g)\},$$

which are perturbed versions of (P_z^*). Clearly $\varnothing \neq P \cap \mathrm{dom}\, f_x^*$, and since $(f_x^*)^* = f(\cdot + x)$, we have $\varnothing \neq P \cap \mathrm{dom}\, (f_x^*)^*$ if and only if $x \in C = \mathrm{dom}\, f - P$. Hence, for each $(z, x) \in Q \times C$ the blanket hypothesis is met for $(P_{z,x}^*)$. Denote the set of solutions to $(P_{z,x}^*)$ by $Y^{z,x}$, and let $J(z, x)$ denote the maximal support of these solutions.

Problems $(P_{z,y})$ and $(P_{z,x}^*)$ are not necessarily the duals of each other in the sense of Theorem 1 unless $y = 0$ and $x = 0$, in which case they coincide with (P_z) and (P_z^*), respectively. (Note that the blanket hypothesis $\varnothing \neq P \cap \mathrm{dom}\, f$ and $\varnothing \neq P \cap \mathrm{dom}\, f^*$ implies that the choices $x = 0$ and $y = 0$ do in fact belong to C and D, respectively.) For general $(z, y) \in Q \times D$, for instance, the dual of $(P_{z,y})$ is the problem

$$\min \{f^*(\cdot + y) - (z \circ g)(\cdot)\},$$

i.e., (P_z^*) with f^* subjected to a "horizontal" translation in the amount of y. The solutions to this latter problem are, as we know, closely related to the solutions $X^{z,y}$ to $(P_{z,y})$. The upper semicontinuity result in Theorem 7(a) below, which for y's near the origin is a kind of stability result for the problems (P_z) with respect to a

certain class of perturbations, can therefore also be interpreted in terms of a dual type of stability for the problems (P_z^*). The roles of these two types of stability are interchanged in Theorem 7(b).

THEOREM 7. (a) *The mapping* $(z, y) \to X^{z,y}$ *is an upper semicontinuous point-to-set mapping on* $Q \times D$, *and it is actually a continuous singlevalued mapping on* $P \times D$. *Moreover, whenever a sequence* $\{(z^m, y^m)\} \subset P \times D$ *converges to* $(z^\infty, y^\infty) \in (Q \backslash P) \times D$, *one has*

(i) $\inf \{\|x^{z^m,y^m} - x\| \mid x \in X^{z^\infty,y^\infty}\} \to 0$;

(ii) $x_k^{z^m,y^m} \to 0$, $\forall k \notin I(z^\infty, y^\infty)$;

(iii) $\{x^{z^m,y^m}\}$ *has at least one cluster point, and all such cluster points belong to* X^{z^∞,y^∞}.

(b) *The mapping* $(z, x) \to Y^{z,x}$ *is an upper semicontinuous point-to-set mapping on* $Q \times C$, *and it is actually a continuous singlevalued mapping on* $P \times C$. *Moreover whenever a sequence* $\{(z^m, x^m)\} \subset P \times C$ *converges to* $(z^\infty, x^\infty) \in (Q \backslash P) \times C$, *one has*

(i) $\inf \{\|y^{z^m,x^m} - y\| \mid y \in Y^{z^\infty,x^\infty}\} \to 0$;

(ii) $y_k^{z^m,x^m} \to 0$, $\forall k \notin J(z^\infty, x^\infty)$;

(iii) $\{y^{z^m,x^m}\}$ *has at least one cluster point, and all such cluster points belong to* Y^{z^∞,x^∞}.

Proof. We shall only prove (a), as the proof of (b) is similar. Let $\nu_{f,y}$ be defined in a manner analogous to (3.11), namely

$$\nu_{f,y}(z) = \begin{cases} \inf (P_{z,y}) & \text{if } z \in Q \\ -\infty & \text{if } z \notin Q, \end{cases}$$

and define closed proper convex functions $h_{f,z}$ for each $z \in Q$ by setting

$$h_{f,z}(x) = f(x) - (z \circ g)(x), \quad \forall x \in R^n.$$

It follows from the definitions that

(4.1) $$h_{f,z}^*(y) = -\nu_{f,y}(z), \quad \forall y \in R^n \ \forall z \in Q.$$

Hence,

(4.2)
$$\begin{aligned} x \in \partial h_{f,z}^*(y) &\Longleftrightarrow h_{f,z}(x) + h_{f,z}^*(y) = \langle x, y \rangle \\ &\Longleftrightarrow h_{f,z}(x) - \langle x, y \rangle = \nu_{f,y}(z) \\ &\Longleftrightarrow f_y(x) - (z \circ g)(x) = \nu_{f,y}(z) \\ &\Longleftrightarrow x \in X^{z,y} \end{aligned}$$

for every $z \in Q$. By (4.1) and the continuity assertion of Theorem 5 applied to $\nu_{f,y}$, the quantity $h_{f,z}^*(y)$ is (finite and) continuous in $z \in Q$ for each fixed $y \in D$. Hence, for each fixed sequence $\{z^m\} \subset Q$ converging to $z^\infty \in Q$, the sequence of convex functions $\{h_{f,z^m}^*\}$ is

finite and converges pointwise to the finite convex function h^*_{f,z^∞} everywhere on the open convex set D. Therefore [6, Theorem 24.5] implies that, for each $\varepsilon > 0$ and each sequence $\{y^m\} \subset D$ converging to $y^\infty \in D$, there exists an integer $\bar{m}$ such that

$$\partial h^*_{f,z^m}(y^m) \subset \partial h^*_{f,z^\infty}(y^\infty) + \varepsilon B, \quad \forall m \geqq \bar{m} \, ,$$

where B here denotes the Euclidean unit ball in R^n. In view of (4.2) we can restate what has just been proved as follows: for each sequence $\{(z^m, y^m)\} \subset Q \times D$ converging to $(z^\infty, y^\infty) \in Q \times D$ and each $\varepsilon > 0$, there exists an $\bar{m}$ such that

$$(4.3) \qquad\qquad X^{z^m, y^m} \subset X^{z^\infty, y^\infty} + \varepsilon B, \quad \forall m \geqq \bar{m} \, .$$

By Theorem 1, the set $X^{z,y}$ is nonempty and compact for all $(z, y) \in Q \times D$ and is a singleton, denoted by $x^{z,y}$, if actually $(z, y) \in P \times D$. In view of (4.3), the first assertion of (a) follows. Now suppose $\{(z^m, y^m)\} \subset P \times D$ converges to $(z^\infty, y^\infty) \in (Q\backslash P) \times D$. It follows from (4.3) that for each $\varepsilon > 0$ there exists an integer $\bar{m}$ such that

$$\inf \{\|x^{z^m, y^m} - x\| \mid x \in X^{z^\infty, y^\infty}\} \leqq \varepsilon, \quad \forall m \geqq \bar{m} \, .$$

This establishes (i). Now clearly (i) implies that

$$\inf \{\|x^{z^m, y^m} - x\|^2 \mid x \in X^{z^\infty, y^\infty}\} \longrightarrow 0 \, ,$$

where in fact the infimum is attained for each m. Since for the index set

$$I = I(z^\infty, y^\infty) = \max \{\operatorname{card} \sigma(x) \mid x \in X^{z^\infty, y^\infty}\}$$

and for each $x \in X^{z^\infty, y^\infty}$ we have

$$\|x^{z^m, y^m} - x\|^2 \geqq \sum_{k \notin I} |x_k^{z^m, y^m} - x_k|^2 = \sum_{k \notin I} |x_k^{z^m, y^m}|^2 \, ,$$

it follows that

$$\sum_{k \notin I} |x_k^{z^m, y^m}|^2 \longrightarrow 0 \, .$$

This establishes (ii). Assertion (iii) follows from (i) and the fact that X^{z^∞, y^∞} is nonempty and compact. This concludes the proof.

The idea of the additional parameters involved in Theorem 7 can be combined with the proof technique used in Theorem 5 to obtain an extension of Corollary 5A involving the additional parameters. That is, not only are the solution mappings

$$(z, y) \longrightarrow x^{z,y} \quad \text{and} \quad (z, x) \longrightarrow y^{z,x}$$

continuous on the open domains $P \times D$ and $P \times C$, respectively, (we already know this by Theorem 7) but they are also differenti-

able there almost everywhere. For instance, to prove the assertion concerning $y^{z,x}$, one considers the function

$$\mu_f(z, x) = \begin{cases} \inf_{x'} \{f(x' + x) - (z \circ g)(x')\} - \sum (z_k - z_k \log z_k) & \text{if } z \in Q, \\ +\infty & \text{if } z \notin Q, \end{cases}$$

and establishes by means of a representation analogous to (3.25) that μ_f is closed proper convex with known conjugate and sub-differential formulas. In particular, one obtains that this μ_f is differentiable everywhere on $\text{int} (\text{dom } \mu_f) = P \times C$ with gradient

$$\nabla \mu_f(z, x) = (y^{z,x}, L(y^{z,x})), \qquad L(y^{z,x}) = (\log y_1^{z,x}, \cdots, \log y_n^{z,x}).$$

The continuity and almost everywhere differentiability of $(z, x) \rightarrow y^{z,x}$ now follow by [6, Theorem 25.5] and Alexandroff's theorem [1], respectively.

The proof just sketched also yields, of course, an extension of all of Theorem 5 itself to include the additional parameters. We leave the statement of this to the reader.

For the remainder of this section we put aside the additional parameters just discussed and concentrate on the behavior of X^z and Y^z as z approaches 0 in a specified manner. Throughout this section and the next, let a be any given, fixed vector satisfying (3.36). Also, let the index sets $I(z)$ and $J(z)$ corresponding to $z = 0$ be denoted simply by I and J.

THEOREM 8. *Let the sequence* $\{z^m\} \subset P$ *converge to* 0 *in such a way that*

$$\lim_{m \to \infty} a_k^m = a_k \ \forall k, \quad \text{where } a_k^m = z_k^m / \sum_{l=1}^n z_l^m,$$

and assume that the corresponding solution sequence $\{(x^{z^m}, y^{z^m})\}$ *converges to* (x^0, y^0). *Then*

(4.4) $$\sigma(x^0) = I \quad \text{and} \quad \sigma(y^0) = J,$$

and for each $(x, y) \in X^0 \times Y^0$ *one has*

(4.5) $$\sum_I a_k(x_k/x_k^0) + \sum_J a_k(y_k/y_k^0) \leqq 1.$$

Proof. Let (x, y) be any pair in $X^0 \times Y^0$. Then

(4.6) $$\sigma(x) \subset I \quad \text{and} \quad \sigma(y) \subset J.$$

Since ∂f is monotone and (x^{z^m}, y^{z^m}), $(x, y) \in \text{graph } \partial f$, we have

$$0 \leqq \langle x^{z^m} - x, y^{z^m} - y \rangle$$
$$= \sum z_k^m - \langle x^{z^m}, y \rangle - \langle x, y^{z^m} \rangle + 0.$$

Since $x_k^{z^m} \cdot y_k^{z^m} = z_k^m \; \forall k$ together with (4.6) implies

$$\langle x^{z^m}, y \rangle = \sum_J z_k^m (y_k / y_k^{z^m}) \text{ and } \langle x, y^{z^m} \rangle = \sum_I z_k^m (x_k / x_k^{z^m}) \,,$$

it follows that

$$(4.7) \qquad \sum_I a_k^m (x_k / x_k^{z^m}) + \sum_J a_k^m (y_k / y_k^{z^m}) \leqq 1 \,, \quad \forall m \,.$$

We now show that $I \subset \sigma(x^0)$. Indeed, for each $k \in I$, (4.7) implies

$$(4.8) \qquad a_k^m (x_k / x_k^{z^m}) \leqq 1 \,, \quad \forall m \,.$$

If we had $x_k^0 = 0$ for some $k \in I$, then $x_k^{z^m} \to x_k^0$ would imply $(x_k / x_k^{z^m}) \to +\infty$, and since $a_k^m \to a_k > 0$ this would yield $a_k^m (x_k / x_k^{z^m}) \to +\infty$, in violation of (4.8). Therefore $x_k^0 > 0 \; \forall k \in I$, that is, $I \subset \sigma(x^0)$. Similarly, (4.7) implies $J \subset \sigma(y^0)$. Since $(x^0, y^0) \in X^0 \times Y^0$ (e.g., by Theorem 7(iii)), the maximal character of I and J implies that actually (4.4) holds. Taking that limit in (4.7) as $m \to \infty$ now yields (4.5).

COROLLARY 8A. *Under the assumptions of Theorem 8 one has*

$$(4.9) \qquad X^0 \subset \{ x \in Q^I \,|\, \sum_I x_k (a_k / x_k^0) \leqq 1 - \sum_J a_k \} \,,$$

$$(4.10) \qquad Y^0 \subset \{ y \in Q^J \,|\, \sum_J y_k (a_k / y_k^0) \leqq 1 - \sum_I a_k \} \,,$$

$$X^0 \times Y^0 \subset \{ (x, y) \in Q^I \times Q^J \,|\, \sum_I x_k (a_k / x_k^0) + \sum_J y_k (a_k / y_k^0) \leqq 1 \} \,.$$

In addition, (x^0, y^0) *satisfies*

$$(4.11) \qquad \mu \cdot \beta^\beta \leqq \prod_I (x_k^0)^{a_k} \cdot \prod_J (y_k^0)^{a_k} \leqq \mu \,,$$

where $\beta = \sum_{I \cup J} a_k$ *and* μ *is defined by*

$$\mu = \sup \{ \prod_I (x_k)^{a_k} \cdot \prod_J (y_k)^{a_k} \,|\, (x, y) \in X^0 \times Y^0, \, \sigma(x) = I, \, \sigma(y) = J \} \,.$$

(*Recall that* Q^I, Q^J *are faces of* Q *as defined following Corollary* 1C.)

Proof. To obtain (4.9), apply inequality (4.5) with x an abitrary element of X^0 and with $y = y^0$. Inclusion (4.10) follows analogously. Dividing (4.5) through by β and then applying the arithmetic-geometric mean inequality, we obtain

$$\beta^{-1} \geqq \prod_I (x_k / x_k^0)^{b_k} \cdot \prod_J (y_k / y_k^0)^{b_k} \,,$$

where $b_k = \beta^{-1} a_k$, $\forall k$. Rearranging this, then raising to the βth power and taking the supremum, we obtain (4.11).

The convergence permitted for z in Theorem 8 and its corollary includes, of course, that in which z approaches 0 along the ray $\{\zeta a \,|\, \zeta \geqq 0\}$. This is the situation treated for the remainder of § 4. Thus, we consider z's of the form $z = \zeta a$, $\zeta > 0$, and write simply (x^ζ, y^ζ) to denote the corresponding solutions.

We already know (by Theorem 7) that when $\zeta \downarrow 0$ the solutions (x^ζ, y^ζ) have at least one cluster point (x^0, y^0) and that each such cluster point belongs to $X^0 \times Y^0$. The next two theorems give conditions under which the pairs (x^ζ, y^ζ) must actually converge as $\zeta \downarrow 0$, and moreover to a special pair $(\tilde{x}, \tilde{y})$ in $X^0 \times Y^0$. This special solution pair will now be described.

PROPOSITION 5. *The problem*

$$(4.12) \qquad \sup \{ \textstyle\prod_I (x_k)^{a_k} \,|\, x \in X^0, \quad \sigma(x) = I \}$$

has a unique solution, call it $\tilde{x}$, *and* $\tilde{x}$ *is a strong Pareto optimal element of* X^0. *Similarly, the problem*

$$(4.13) \qquad \sup \{ \textstyle\prod_J (y_k)^{a_k} \,|\, y \in Y^0, \; \sigma(y) = J \}$$

has a unique solution, call it $\tilde{y}$, *and* $\tilde{y}$ *is a strong Pareto optimal element of* Y^0.

Proof. Consider the problem

$$(4.14) \qquad \sup \{ \theta(x) \,|\, x \in S \} \,,$$

where θ is the concave function on R^n given by

$$\theta(x) = \begin{cases} \log \prod_I (x_k)^{a_k} & \text{if } x_k > 0 \; \forall k \in I \\ -\infty & \text{otherwise} \end{cases}$$

and S is the nonempty convex set

$$S = \{ x \,|\, x \in X^0, \; \sigma(x) = I \} \,.$$

Pick any $x^0 \in S$. Observe that: (i) the level set $\{ x \,|\, \theta(x) \geqq \theta(x^0) \}$ is closed since θ is upper semicontinuous; (ii) $x \in Q$ and $\theta(x) \geqq \theta(x^0)$ imply $I \subset \sigma(x)$; and (iii) $x \in X^0$ implies $\sigma(x) \subset I$. Together with the compactness of X^0, these facts imply that

$$T = X^0 \cap \{ x \,|\, \theta(x) \geqq \theta(x^0) \}$$

is a nonempty compact subset of S. Since $x \in S \backslash T$ implies $\theta(x) < \theta(x^0)$ and $x^0 \in T$, it follows that the solution set of problem (4.14) coincides with that of the problem

$$\sup \{\theta(x) \mid x \in T\} \ .$$

Now this problem has at least one solution $\tilde{x}$, since T is nonempty compact and θ is upper semicontinuous. Also, it has at most one solution, since T is convex and θ is strictly concave on the subspace $\{x \mid x_k = 0 \ \forall k \notin I\}$, which contains T. Since the exponential function is strictly increasing, it follows that $\tilde{x}$ is the unique solution of problem (4.12). Finally, suppose that there were another element $x^0 \in X^0$ satisfying $x_k^0 \geqq \tilde{x}_k \ \forall k$, with at least one of these inequalities strict. Then we would have both $\sigma(x^0) = I$ and $\theta(x^0) > \theta(\tilde{x})$, contradicting the fact that $\tilde{x}$ solves (4.12). The assertions concerning problem (4.13) and $\tilde{y}$ are proved similarly.

THEOREM 9. *Assume that $I \cup J = \{1, \cdots, n\}$, that is, assume (C_0) has at least one strictly complementary solution. Then*

$$\lim_{\zeta \downarrow 0} (x^\zeta, y^\zeta) = (\tilde{x}, \tilde{y}) \ ,$$

where $\tilde{x}$ and $\tilde{y}$ are as in Proposition 5.

Proof. Let (x^0, y^0) be any cluster point of (x^ζ, y^ζ) as ζ decreases to zero through discrete values. Then by Theorem 7(iii), Theorem 8 and Proposition 5 we have

(4.15) $$(x^0, y^0) \in X^0 \times Y^0, \ \sigma(x^0) = I, \ \sigma(y^0) = J$$

and

(4.16) $$1 \geqq \sum_I a_k (\tilde{x}_k / x_k^0) + \sum_J a_k (\tilde{y}_k / y_k^0) \ .$$

Since $I \cup J = \{1, \cdots, n\}$ implies $\sum_{I \cup J} a_k = 1$, we can apply the arithmetic-geometric mean inequality to (4.16), obtaining

$$1 \geqq \prod_I (\tilde{x}_k / x_k^0)^{a_k} \cdot \prod_J (\tilde{y}_k / y_k^0)^{a_k} \ .$$

Rearranging this yields

(4.17) $$\prod_I (x_k^0)^{a_k} \cdot \prod_J (y_k^0)^{a_k} \geqq \prod_I (\tilde{x}_k)^{a_k} \cdot \prod_J (\tilde{y}_k)^{a_k} \ .$$

But in view of (4.15) and the maximality properties of $\tilde{x}$ and $\tilde{y}$ (Proposition 5), we also have

(4.18) $$\prod_I (\tilde{x}_k)^{a_k} \geqq \prod_I (x_k^0)^{a_k}$$

and

(4.19) $$\prod_J (\tilde{y}_k)^{a_k} \geqq \prod_J (y_k^0)^{a_k} \ .$$

By (4.17) it follows that the weak inequalities in both (4.18) and (4.19) must in fact be equalities. By the uniqueness properties of $\tilde{x}$ and $\tilde{y}$ (Proposition 5), this means that $x^0 = \tilde{x}$ and $y^0 = \tilde{y}$, which concludes the proof.

The proof of Theorem 9 actually shows that, assuming $I \cup J = \{1, \cdots, n\}$, the same limiting behavior holds for z's approaching 0 along the more general paths treated in Theorem 8.

THEOREM 10. *Assume that f is polyhedral. Then*

$$\lim_{\zeta \downarrow 0} (x^\zeta, y^\zeta) = (\tilde{x}, \tilde{y}) ,$$

where $\tilde{x}$ and $\tilde{y}$ are as in Proposition 5.

Proof. Let (x^0, y^0) be any cluster point of (x^ζ, y^ζ) as ζ decreases to zero through discrete values. We must show that $(x^0, y^0) = (\tilde{x}, \tilde{y})$. Since f polyhedral implies f^* polyhedral [6, Theorem 19.2], the argument we shall present showing $x^0 = \tilde{x}$ can be mimicked to obtain $y^0 = \tilde{y}$ also. Thus, it will suffice to assume that

$$(4.20) \qquad\qquad x^0 \neq \tilde{x}$$

and deduce a contradiction. From Theorem 7(a)(iii) and Theorem 8 we have

$$(4.21) \qquad\qquad x^0 \in X^0, \ \sigma(x^0) = I .$$

Passing to a subsequence if necessary, we can suppose that actually

$$(4.22) \qquad x^\zeta \longrightarrow x^0 \text{ as } \zeta \downarrow 0 \text{ through discrete values.}$$

For the function θ used in the proof of Proposition 5, it follows from (4.20), (4.21) and Proposition 5 that $\theta(x^0) < \theta(\tilde{x})$. Hence, the equation

$$3\varepsilon = \theta(\tilde{x}) - \theta(x^0)$$

defines a positive ε. Since θ is continuous on $\{x \,|\, x_k > 0 \ \forall k \in I\}$, there exists a $\delta > 0$ such that, for each $x \in Q$ satisfying $\sigma(x) \supset I$,

$$\|x - x^0\| \leqq \delta \Longrightarrow |\theta(x) - \theta(x^0)| \leqq \varepsilon$$

and

$$\|x - \tilde{x}\| \leqq \delta \Longrightarrow |\theta(x) - \theta(\tilde{x})| \leqq \varepsilon .$$

By (4.22), there exists a $\zeta_1 > 0$ such that

$$0 < \zeta \leqq \zeta_1 \Longrightarrow \|x^\zeta - x^0\| \leqq \delta \text{ and } \|s^\zeta - \tilde{x}\| \leqq \delta ,$$

where s^ζ is defined by

$$s^{\zeta} = x^{\zeta} + (\tilde{x} - x^0) \ .$$

Hence,

$$0 < \zeta \leqq \zeta_1 \Longrightarrow |\theta(x^{\zeta}) - \theta(x^0)| \leqq \varepsilon \text{ and } |\theta(s^{\zeta}) - \theta(\tilde{x})| \leqq \varepsilon \ .$$

In view of the definition of ε, this yields

$$0 < \zeta \leqq \zeta_1 \Longrightarrow \theta(s^{\zeta}) - \theta(x^{\zeta}) \geqq \varepsilon > 0 \ .$$

Since θ is concave, this in turn yields

(4.23) $\qquad 0 < \zeta \leqq \zeta_1 \text{ and } x \in (x^{\zeta}, s^{\zeta}] \Longrightarrow \theta(x) > \theta(x^{\zeta}) \ .$

(We use the familiar notations $[p, q]$ and $(p, q]$ to denote the straight line segments between vectors p and q which include and exclude p, respectively.)

Next, since f is polyhedral, it can be represented in the form

$$f(x) = \begin{cases} \max \{h_{\lambda}(x) \,|\, \lambda \in \varLambda\} & \text{if } x \in C \\ + \infty & \text{if } x \notin C \ , \end{cases}$$

where C is a nonempty polyhedral convex subset of R^n and h_{λ} is affine for each index λ in some nonempty finite set $\varLambda$. For fixed $0 < \rho < 1$ to be specified in a moment, let vectors r^{ζ} be defined for each $\zeta > 0$ by setting

$$r^{\zeta} = x^{\zeta} + \rho(\tilde{x} - x^0) \ .$$

Using (4.22) and $x^{\zeta}, x^0, \tilde{x} \in C$, it can be deduced from the fact that C is polyhedral convex that there exist $0 < \rho < 1$ and $\zeta_2 > 0$ such that

$$0 < \zeta \leqq \zeta_2 \Longrightarrow [x^{\zeta}, r^{\zeta}] \subset C \ .$$

Observe that

$$C = \bigcup_{\lambda \in \varLambda} C(\lambda) \ ,$$

where for each $\lambda \in \varLambda$ the set

$$C(\lambda) = \{x \in C \,|\, f(x) = h_{\lambda}(x)\}$$

can be seen to be convex and closed (although possibly empty). Suppose $0 < \zeta \leqq \zeta_2$. For each $\lambda \in \varLambda$ the set $[x^{\zeta}, r^{\zeta}] \cap C(\lambda)$ is either empty or else a closed line segment (possibly a singleton). Since

$$[x^{\zeta}, r^{\zeta}] = \bigcup_{\lambda \in \varLambda} [x^{\zeta}, r^{\zeta}] \cap C(\lambda), \quad \forall 0 < \zeta \leqq \zeta_2 \ ,$$

and the length of $[x^{\zeta}, r^{\zeta}]$ is $\rho \|\tilde{x} - x^0\| > 0$ (this uses (4.20)), it follows that there exists at least one $\lambda(\zeta) \in \varLambda$ such that the length of $[x^{\zeta}, r^{\zeta}] \cap C(\lambda(\zeta))$ is at least

$$l = \rho ||\tilde{x} - x^0||/\text{card } \Lambda > 0 .$$

Since Λ is finite, there exists at least one $\bar{\lambda} \in \Lambda$ such that $\lambda(\zeta) = \bar{\lambda}$ for infinitely many ζ's in our sequence satisfying $0 < \zeta \leq \zeta_2$. By passing again to a subsequence if necessary, we can suppose that $\lambda(\zeta) = \bar{\lambda}$ for each ζ in our sequence satisfying $0 < \zeta \leq \zeta_2$. Now write

$$[p^\zeta, q^\zeta] = [x^\zeta, r^\zeta] \cap C(\bar{\lambda}), \quad \forall 0 < \zeta \leq \zeta_2 ,$$

where p^ζ denotes the endpoint of the segment nearer x^ζ, and put

$$\zeta_0 = \min \{\zeta_1, \zeta_2\} .$$

We conclude the proof by obtaining a contradiction in each of two possible cases.

Case 1: $f(p^\zeta) \geq f(q^\zeta)$ for some $0 < \zeta \leq \zeta_0$. By the convexity of f, it follows that

$$(4.24) \qquad\qquad f(x^\zeta) \geq f(q^\zeta) .$$

Since $[p^\zeta, q^\zeta]$ has positive length, $q^\zeta \in (x^\zeta, r^\zeta]$. Hence (4.23) implies $\theta(q^\zeta) > \theta(x^\zeta)$, that is,

$$(4.25) \qquad\qquad \sum_I a_k \log q_k^\zeta > \sum_I a_k \log x_k^\zeta .$$

In view of the source of q^ζ, we also have

$$(4.26) \qquad\qquad q_k^\zeta = x_k^\zeta \quad \forall k \notin I .$$

From (4.25), (4.26) and the definition of $z \circ g$, it follows that

$$(4.27) \qquad\qquad -(z \circ g)(x^\zeta) > -(z \circ g)(q^\zeta), \text{ for } z = \zeta a .$$

Adding (4.24) to (4.27), we obtain the contradiction

$$\inf (P_z) = f(x^\zeta) - (z \circ g)(x^\zeta) > f(q^\zeta) - (z \circ g)(q^\zeta)$$

for $z = \zeta a$.

Case 2: $f(p^\zeta) < f(q^\zeta)$ for each $0 < \zeta \leq \zeta_0$. Since the segments $[p^\zeta, q^\zeta] \subset [x^\zeta, s^\zeta]$ are all parallel with positive length at least l for $0 < \zeta \leq \zeta_0$, it can be shown using (4.22) that there exists a segment $[p^0, q^0] \subset [x^0, \tilde{x}]$ having length at least l, and such that some subsequence of the p^ζ's converges to p^0 and the corresponding subsequence of the q^ζ's converges to q^0. Passing to this subsequence if necessary, we can suppose that $p^\zeta \to p^0$ and $q^\zeta \to q^0$ as $\zeta \downarrow 0$ through discrete values. Since $C(\bar{\lambda})$ is closed convex and contains $[p^\zeta, q^\zeta]$ for each $0 < \zeta \leq \zeta_0$, it follows that

(4.28) $[p^0, q^0] \subset C(\bar{\lambda})$.

Now the Case 2 hypothesis implies that the affine function $h_{\bar{\lambda}}$ strictly increases in the direction $\tilde{x} - x^0$. Since $[p^0, q^0]$ has positive length and direction $\tilde{x} - x^0$, it follows from this and (4.28) that

$$f(p^0) = h_{\bar{\lambda}}(p^0) < h_{\bar{\lambda}}(q^0) = f(q^0) .$$

But since $[p^0, q^0] \subset [x^0, \tilde{x}] \subset X^0$, we also have

$$f(p^0) = \inf (P_0) = f(q^0) .$$

This contradiction concludes the proof of the theorem.

The following shows that the sharp convergence established in Theorems 9 and 10 can fail in general.

EXAMPLE 1.   Let $n = 2$, and define a nonpolyhedral f by setting

$$f(x_1, x_2) = \begin{cases} x_1 & \text{if } 2x_1 \geqq x_2^2 \\ +\infty & \text{otherwise} . \end{cases}$$

It can be shown that

$$f^*(y_1, y_2) = \begin{cases} y_2^2/2(1 - y_1) & \text{if } y_1 < 1 \\ 0 & \text{if } y_1 = 1 \text{ and } y_2 = 0 \\ +\infty & \text{otherwise} . \end{cases}$$

With these functions the blanket hypothesis is met, and one clearly has for (P_0) and (P_0^*) that

$$X^0 = \{(0, 0)\}, \ I = \varnothing, \ \tilde{x} = (0, 0) ,$$

and

$$Y^0 = \{(y_1, 0) \,|\, 0 \leqq y_1 \leqq 1\}, \ J = \{1\}, \ \tilde{y} = (1, 0) .$$

For each a satisfying (3.36) and each $\zeta > 0$ one can solve (P_ζ^*), that is,

$$\inf \{y_2^2/2(1 - y_1) - \zeta a_1 \log y_1 - \zeta a_2 \log y_2 \,|\, 0 < y_1 < 1, 0 < y_2 < +\infty\} ,$$

analytically for y^ζ and then trivially obtain x^ζ via $x_k^\zeta y_k^\zeta = \zeta a_k \ \forall k$. Writing $\beta = a_2 + 2a_1$, one finds that

$$x^\zeta = \left(\frac{\zeta\beta}{2}, \ \sqrt{\zeta\beta}\right) \longrightarrow (0, 0) = \tilde{x}$$

but

$$y^\zeta = \left(\frac{2a_1}{\beta}, \ a_2\sqrt{\frac{\zeta}{\beta}}\right) \longrightarrow \left(\frac{2a_1}{\beta}, \ 0\right) \neq \tilde{y} .$$

The next theorem and its corollaries present various monotonicity

properties of (x^ζ, y^ζ) as $\zeta \downarrow 0$. These rely in part on the following technical preliminary.

PROPOSITION 6. *For each z and z' in P, the corresponding solutions (x^z, y^z) and $(x^{z'}, y^{z'})$ satisfy*

$$0 \leqq \langle x^z - x^{z'}, y^z - y^{z'} \rangle$$
$$\leqq \min \{ \sum (z_k - z'_k) \log (x_k^z / x_k^{z'}), \sum (z_k - z'_k) \log (y_k^z / y_k^{z'}) \} .$$

Proof. Since (x^z, y^z) and $(x^{z'}, y^{z'})$ belong to graph ∂f and ∂f is monotone,

(4.29) $$0 \leqq \langle x^z - x^{z'}, y^z - y^{z'} \rangle .$$

Since $x^z, y^z \in P$ and $x_k^z y_k^z = z_k > 0 \; \forall k$, Proposition 4 implies that

(4.30) $$(y^z, -L(y^z)) \in \partial \gamma(x^z, z) ,$$

where $L : P \to R^n$ is the mapping

$$L(z_1, \cdots, z_n) = (\log z_1, \cdots, \log z_n) .$$

Similarly, Proposition 4 implies that

(4.31) $$(y^{z'}, -L(y^{z'})) \in \partial \gamma(x^{z'}, z') .$$

Since $\partial \gamma$ is antimonotone (i.e., its negative is monotone), (4.30) and (4.31) yield

$$0 \geqq \langle (x^z, z) - (x^{z'}, z'), (y^z, -L(y^z)) - (y^{z'}, -L(y^{z'})) \rangle ,$$
that is,

(4.32) $$\langle x^z - x^{z'}, y^z - y^{z'} \rangle \leqq \langle z - z', L(y^z) - L(y^{z'}) \rangle .$$

Now notice that Proposition 4 can also be applied with the roles of the x- and the y-variables interchanged. It follows that

$$(x^z, -L(x^z)) \in \partial \gamma(y^z, z)$$
and

$$(x^{z'}, -L(x^{z'})) \in \partial \gamma(y^{z'}, z') ,$$

which implies similarly by the antimonotonicity of $\partial \gamma$ that

(4.33) $$\langle y^z - y^{z'}, x^z - x^{z'} \rangle \leqq \langle z - z', L(x^z) - L(x^{z'}) \rangle .$$

Combining (4.29) with (4.32) and (4.33) concludes the proof.

It will be convenient to have available the (positively homogeneous closed proper concave) function $\rho : R^n \to [-\infty, +\infty)$ defined by

(4.34)
$$\rho(x) = \begin{cases} \prod (x_n)^{a_k} & \text{if } x \in Q \\ -\infty & \text{if } x \notin Q . \end{cases}$$

For the z's of the form $z = \zeta a$, $\zeta > 0$, notice that

$$(z \circ g)(x) = \zeta \log \rho(x) \quad \forall x ,$$

provided we interpret $\log 0$ and $\log (-\infty)$ as $-\infty$.

THEOREM 11. *Let $\zeta \geqq \zeta' > 0$. Then*

(4.35) $\rho(x^\zeta) \geqq \rho(x^{\zeta'}) > 0$, *with equality only if $x^\zeta = x^{\zeta'}$* ,

(4.36) $\rho(y^\zeta) \geqq \rho(y^{\zeta'}) > 0$, *with equality only if $y^\zeta = y^{\zeta'}$* ,

and

(4.37) $\rho(x^\zeta) = \rho(x^{\zeta'})$ *and* $\rho(y^\zeta) = \rho(y^{\zeta'})$ *only if* $\zeta = \zeta'$.

Proof. We first prove (4.35). Without loss of generality assume $\zeta > \zeta'$. Choosing $z = \zeta a$ and $z' = \zeta' a$ in Proposition 6, we obtain

$$0 \leqq \sum (\zeta - \zeta') a_k \log (x_k^\zeta / x_k^{\zeta'}) .$$

This is equivalent to

$$1 \leqq \{\prod (x_k^\zeta / x_k^{\zeta'})^{a_k}\}^{\zeta - \zeta'}$$

that is, $\rho(x^\zeta) \geqq \rho(x^{\zeta'})$. Now suppose that

(4.38) $\rho(x^\zeta) = \rho(x^{\zeta'})$.

Using the first inequality in Proposition 6, we have

$$0 \leqq \langle x^\zeta - x^{\zeta'}, y^\zeta - y^{\zeta'} \rangle$$
$$= \zeta - \sum \zeta' a_k (x_k^\zeta / x_k^{\zeta'}) - \sum \zeta a_k (x_k^{\zeta'} / x_k^\zeta) + \zeta' .$$

Writing

$$\beta = \zeta(\zeta + \zeta')^{-1} \text{ and } \beta' = \zeta'(\zeta + \zeta')^{-1} ,$$

we deduce that

$$1 \geqq \beta' \sum a_k (x_k^\zeta / x_k^{\zeta'}) + \beta \sum a_k (x_k^{\zeta'} / x_k^\zeta)$$
$$\geqq \beta' \prod (x_k^\zeta / x_k^{\zeta'})^{a_k} + \beta \prod (x_k^{\zeta'} / x_k^\zeta)^{a_k}$$
$$= 1 .$$

Here the second inequality follows by applying the arithmetic-geometric mean inequality to each of the two sums separately, and the equality follows by (4.38). Thus, equality holds in both applications of the arithmetic-geometric mean inequality. Therefore there exists a positive number σ such that

$$x_k^\zeta / x_k^{\zeta'} = \sigma \quad \forall k \,.$$

This implies (using (4.38) again) that

$$0 < \rho(x^{\zeta'}) = \rho(x^\zeta) = \prod (\sigma x_k^{\zeta'})^{a_k} = \sigma \rho(x^{\zeta'}) \,.$$

Hence $\sigma = 1$, so that $x^\zeta = x^{\zeta'}$. This establishes (4.35). The proof of (4.36) is similar. Finally, suppose that

$$\zeta \geqq \zeta' > 0, \ \rho(x^\zeta) = \rho(x^{\zeta'}), \ \rho(y^\zeta) = \rho(y^{\zeta'}) \,.$$

By (4.35) and (4.36) this implies $x^\zeta = x^{\zeta'}$ and $y^\zeta = y^{\zeta'}$. But then

$$\zeta = \langle x^\zeta, y^\zeta \rangle = \langle x^{\zeta'}, y^{\zeta'} \rangle = \zeta' \,.$$

This concludes the proof.

COROLLARY 11A. *If* $x^\zeta \in X^0$ *for some* $\zeta > 0$, *then* $x^\zeta = x^{\zeta'}$ *for all* $\zeta \geqq \zeta' > 0$. *If* $y^\zeta \in Y^0$ *for some* $\zeta > 0$, *then* $y^\zeta = y^{\zeta'}$ *for all* $\zeta \geqq \zeta' > 0$.

Proof. Let $x^\zeta \in X^0$ and $\zeta > \zeta' > 0$, and assume for the moment that $\rho(x^\zeta) > \rho(x^{\zeta'})$. Then

$$-(\zeta' a \circ g)(x^{\zeta'}) = -\zeta' \log \rho(x^{\zeta'})$$
$$> -\zeta' \log \rho(x^\zeta)$$
$$= -(\zeta' a \circ g)(x^\zeta) \,.$$

Since $x^\zeta \in X^0$, we also have

$$f(x^{\zeta'}) \geqq f(x^\zeta) \,.$$

Invoking the optimality of $x^{\zeta'}$, we obtain the contradiction

$$\inf \{f - (\zeta' a \circ g)\} = f(x^{\zeta'}) - (\zeta' a \circ g)(x^{\zeta'})$$
$$> f(x^\zeta) - (\zeta' a \circ g)(x^\zeta) \,.$$

Therefore $\rho(x^\zeta) \leqq \rho(x^{\zeta'})$. Hence, (4.35) implies $x^\zeta = x^{\zeta'}$. The other assertion is proved similarly.

COROLLARY 11B. *For each* $\zeta \in (0, +\infty)$ *write* $\xi(\zeta) = \rho(x^\zeta)$ *and* $\eta(\zeta) = \rho(y^\zeta)$. *Then*

$$\Gamma_a = \{(\xi(\zeta), \, \eta(\zeta)) \,|\, 0 < \zeta < +\infty\}$$

is a maximal monotone subset of $(0, +\infty) \times (0, +\infty)$. *Moreover, the mapping*

$$h_a(\zeta) = (\xi(\zeta), \, \eta(\zeta))$$

is a homeomorphism from $(0, +\infty)$ *onto* Γ_a *with inverse given by*

(4.39)　　　　　　$h_a^{-1}(\xi, \eta) = \alpha^{-1}\xi\eta$, *where* $\alpha = \prod (a_k)^{a_k}$.

(*Recall that a and* ρ *are as in* (3.36) *and* (4.34), *respectively.*)

Proof. Clearly h_a is onto, and by (4.37) it is also one-to-one. By Corollary 5A the mapping $\zeta \to (x^\zeta, y^\zeta)$ is continuous from $(0, +\infty)$ into $P \times P$. Since ρ is continuous on P, it follows that h_a is continuous. For any given $(\xi, \eta) \in \Gamma_a$ let ζ be the unique element of $(0, +\infty)$ such that $h_a(\zeta) = (\xi, \eta)$. Then

$$
\begin{aligned}
\xi\eta &= \rho(x^\zeta) \cdot \rho(y^\zeta) \\
&= \prod (x_k^\zeta y_k^\zeta)^{a_k} \\
&= \prod (\zeta a_k)^{a_k} \\
&= \zeta\alpha
\end{aligned}
$$

where α is defined above. This establishes (4.39), and from it the continuity of h_a^{-1} is clear. This completes the proof that Γ_a is homeomorphic to $(0, +\infty)$. To see the monotonicity, let

$$(\xi_i, \eta_i) = (\xi(\zeta_i), \eta(\zeta_i)) \in \Gamma_a \text{ for } i = 1, 2 ,$$

and suppose without loss of generality that $\zeta_1 \geq \zeta_2 > 0$. By (4.35) and (4.36) we have

$$\xi_1 \geq \xi_2 \text{ and } \eta_1 \geq \eta_2 ,$$

and hence

$$\langle \xi_1 - \xi_2, \eta_1 - \eta_2 \rangle \geq 0 .$$

Finally, we establish the maximal monotonicity. For any given $\xi', \eta' \in (0, +\infty)$ such that $(\xi', \eta') \notin \Gamma_a$, we must produce a pair $(\xi, \eta) \in \Gamma_a$ for which

(4.40)　　　　　　$\langle \xi' - \xi, \eta' - \eta \rangle < 0$

holds. For the value $\zeta = \alpha^{-1}\xi'\eta'$, let $(\xi, \eta) = h_a(\zeta)$. Then (4.39) yields $\alpha^{-1}\xi\eta = \zeta$. Hence $\xi'\eta' = \xi\eta$, or equivalently,

$$\log \xi' + \log \eta' = \log \xi + \log \eta .$$

Therefore

$$
\begin{aligned}
\langle \log \xi' - \log \xi, \log \eta' - \log \eta \rangle &= -|\log \xi' - \log \xi|^2 \\
&= -|\log \eta' - \log \eta|^2 .
\end{aligned}
$$

Since $(\xi, \eta) \in \Gamma_a$ but $(\xi', \eta') \notin \Gamma_a$, it follows that

$$\langle \log \xi' - \log \xi, \log \eta' - \log \eta \rangle < 0 .$$

Thus, either

$$\log \xi' > \log \xi \quad \text{and} \quad \log \eta > \log \eta'$$

or else

$$\log \xi > \log \xi' \quad \text{and} \quad \log \eta' > \log \eta \ .$$

Upon exponentiating, each of these implies (4.40). This concludes the proof.

The above proof of maximality, together with (4.35) and (4.36), shows in addition that

$$\{(\nabla \nu_f(\zeta), \nabla \nu_{f*}(\zeta), \,|\, 0 < \zeta < +\infty\}$$

is a maximal monotone subset of $R \times R$. Here ν_f and ν_{f*} are the parametric optimal value functions treated in Theorem 6. (Cf. (3.39), (3.40), and also Proposition 8 in §6.)

PROPOSITION 7. *One and only one of the alternatives*

$$(4.41) \qquad \lim_{\zeta \downarrow 0} \rho(x^\zeta) = 0 \quad or \quad \lim_{\zeta \downarrow 0} x^\zeta = \tilde{x} \in P$$

holds. Similarly, one and only one of the alternatives

$$(4.42) \qquad \lim_{\zeta \downarrow 0} \rho(y^\zeta) = 0 \quad or \quad \lim_{\zeta \downarrow 0} y^\zeta = \tilde{y} \in P$$

holds. (Here $\tilde{x}$ and $\tilde{y}$ are as in Proposition 5.)

Proof. It suffices to show (4.41). Both alternatives in (4.41) cannot occur simultaneously, since that would imply

$$0 = \lim_{\zeta \downarrow 0} \rho(x^\zeta) = \rho(\tilde{x}) > 0$$

by the continuity of ρ on Q. Now suppose that

$$\lim_{\zeta \downarrow 0} x^\zeta = \tilde{x} \in P$$

fails. If it fails because $\tilde{x} \notin P$, then

$$(4.43) \qquad \exists k \in \{1, \cdots, n\}\backslash I$$

holds by Proposition 5, while if it fails because $\lim_{\zeta \downarrow 0} x^\zeta = \tilde{x}$ fails, then (4.43) holds by Theorem 9. So in any case, (4.43) holds. By Theorem 7(a)(iii), this entails $x_k^\zeta \to 0$ for some k, which implies $\rho(x^\zeta) \to 0$ and concludes the proof.

The following result partly involves the parametric optimal value functions ν_f and ν_{f*} restricted to the line $\{\zeta a \,|\, \zeta \in R\}$, as described following Corollary 5A and treated in Theorem 6. The right directional derivative of $\nu_f(\zeta)$ at $\zeta = 0$ will be denoted by $\nu_f'(0; 1)$. Thus,

$$\nu'_f(0; 1) = \lim_{\zeta \downarrow 0} \zeta^{-1}[\inf (P_\zeta) - \inf (P_0)]$$

$$= \lim_{\zeta \downarrow 0} \zeta^{-1}[\inf \{f - (\zeta a \circ g)\} - \inf_Q \{f\}] .$$

Similarly for $\nu'_{f^*}(0; 1)$.

THEOREM 12. *One and only one of the following three situations or "modes" is the case:*
   (A) $\lim_{\zeta \downarrow 0} x^\zeta = \tilde{x} \in P$ *and* $\lim_{\zeta \downarrow 0} y^\zeta = \tilde{y} = 0$;
   (B) $\lim_{\zeta \downarrow 0} x^\zeta = \tilde{x} = 0$ *and* $\lim_{\zeta \downarrow 0} y^\zeta = \tilde{y} \in P$;
   (C) $\lim_{\zeta \downarrow 0} \rho(x^\zeta) = 0$ *and* $\lim_{\zeta \downarrow 0} \rho(y^\zeta) = 0$.
Moreover,
   (A) $\Rightarrow \nu'_f(0; 1) = -\log \rho(\tilde{x})$ *and* $\nu'_{f^*}(0; 1) = + \infty$,
   (B) $\Rightarrow \nu'_f(0; 1) = + \infty$ *and* $\nu'_{f^*}(0; 1) = - \log \rho(\tilde{y})$,
   (C) $\Rightarrow \nu'_f(0; 1) = + \infty$ *and* $\nu'_{f^*}(0; 1) = + \infty$.
(*Here* $\tilde{x}$ *and* $\tilde{y}$ *are as in Proposition* 5.)

Proof. Obviously both (A) and (B) cannot hold, and by Proposition 7 each of (A) and (B) is incompatible with (C). Now suppose, for instance, that

(4.44) $$\lim_{\zeta \downarrow 0} x^\zeta = \tilde{x} \in P$$

holds. Then $I = \{1, \cdots, n\}$, so that the complementary slackness of solutions to (C_0) yields $J = \varnothing$, that is, $Y^0 = \{0\}$. Hence

$$\lim_{\zeta \downarrow 0} y^\zeta = 0 = \tilde{y}$$

by Theorem 7(b)(i) and Proposition 5. Similarly,

(4.45) $$\lim_{\zeta \downarrow 0} y^\zeta = \tilde{y} \in P$$

implies

$$\lim_{\zeta \downarrow 0} x^\zeta = 0 = \tilde{x} .$$

If neither (4.44) nor (4.45) holds, then Proposition 7 implies that situation (C) of the theorem holds. This concludes the proof of the first part. Now suppose that (A) is the case. Since we always have

(4.46) $$x \in P \cap X^0 \Longleftrightarrow x \in P \cap \partial f^*(0)$$

and

(4.47) $$y \in P \cap Y^0 \Longleftrightarrow y \in P \cap \partial f(0) ,$$

it follows by subdifferential formulas (3.39) and (3.40) in Theorem

6 that

(4.48) $$\varnothing \neq \partial \nu_f(0) = \{\omega \mid \exists x \in P \cap X^\circ, \ \omega \geqq - \log \rho(x)\}$$

and

(4.49) $$\varnothing = \partial \nu_{f^*}(0) \ .$$

From (4.49) and the fact (Theorem 6) that $\varnothing \neq \partial \nu_{f^*}(\zeta) \ \forall \zeta > 0$, it follows by the theory of concave functions of one variable that $\nu'_{f^*}(0; 1) = + \infty$. From (4.48) it follows that

$$
\begin{aligned}
\nu'_f(0; 1) &= \inf \{\langle 1, \omega \rangle \mid \omega \in \partial \nu_f(0)\} \\
&= \inf \{\omega \mid \exists x \in P \cap X^\circ, \ \omega \geqq - \log \rho(x)\} \\
&= -\sup \{\log \rho(x) \mid x \in P \cap X^\circ\} \\
&= -\log \rho(\tilde{x}) \ ,
\end{aligned}
$$

where the last equality is by Proposition 5 and the strictly increasing nature of the logarithm. The remaining assertions follow similarly.

THEOREM 13. *Assume that situation* (C) *in Theorem 12 occurs, and suppose* $\overline{\zeta} > 0$ *is such that*

(4.50) $$\rho(x^{\overline{\zeta}}) \leqq 1 \ and \ \rho(y^{\overline{\zeta}}) \leqq 1 \ .$$

Then for all $\overline{\zeta} \geqq \zeta > 0$, *the quantities*

$$\inf (P_\zeta) - \inf (P_0)$$

and

$$\inf (P_\zeta^*) - \inf (P_0^*)$$

are each positive and sum to

(4.51) $$\zeta - \zeta \log \zeta \alpha, \ where \ \alpha = \prod (a_k)^{a_k} \ .$$

Condition (4.50) *can occur only if* $\overline{\zeta} < \alpha^{-1} e$, *and* $\alpha^{-1} \leqq n$ *for all vectors* $a \in P$ *satisfying* $\sum a_k = 1$. (*Note that the quantity in* (4.51) *is differentiable concave for* $\zeta > 0$, *is 0 at 0, and has derivative tending to* $+\infty$ *as* $\zeta \downarrow 0$.)

Proof. For each $\zeta \geqq 0$, (3.13) yields

$$\nu_f(\zeta) + \nu_{f^*}(\zeta) = \sum (\zeta a_k - \zeta a_k \log \zeta a_k)$$

(where, as always, $0 \log 0 = 0$). Hence

$$[\nu_f(\zeta) - \nu_f(0)] + [\nu_{f^*}(\zeta) - \nu_{f^*}(0)] = \zeta - \zeta \log \zeta \alpha, \ \forall \zeta > 0 \ ,$$

which establishes the sum assertion. By the situation (C) hypothesis together with Theorems 12 and 6, both of the concave functions $\nu_f(\zeta)$ and $\nu_{f^*}(\zeta)$ are differentiable for $\zeta > 0$ and are finite with right derivative $+ \infty$ at $\zeta = 0$. It follows that

$$\nu_f(\zeta) > \nu_f(0) > - \infty \quad \text{whenever} \quad \nabla\nu_f(\zeta) \geqq 0 \ ,$$

and similarly

$$\nu_{f^*}(\zeta) > \nu_{f^*}(0) > - \infty \quad \text{whenever} \quad \nabla\nu_{f^*}(\zeta) \geqq 0 \ .$$

But from the subdifferential formulas (3.39) and (3.40),

$$\nabla\nu_f(\zeta) \geqq 0 \Longleftrightarrow \zeta > 0 \ \text{and} \ \rho(x^\zeta) \leqq 1$$

and

$$\nabla\nu_{f^*}(\zeta) \geqq 0 \Longleftrightarrow \zeta > 0 \ \text{and} \ \rho(y^\zeta) \leqq 1 \ ,$$

and by Theorem 11 and (4.50),

$$\rho(x^\zeta) \leqq 1 \ \text{and} \ \rho(y^\zeta) \leqq 1, \ \forall\zeta \in (0, \bar{\zeta}] \ .$$

Combining all this yields

$$\nu_f(\zeta) - \nu_f(0) > 0 \ \text{and} \ \nu_{f^*}(\zeta) - \nu_{f^*}(0) > 0, \ \forall\zeta \in (0, \bar{\zeta}] \ .$$

Now suppose $\bar{\zeta} > 0$ satisfies (4.50). What has already been shown yields in particular that

$$\bar{\zeta} - \bar{\zeta} \log \bar{\zeta}\alpha > 0 \ ,$$

that is, $\bar{\zeta} < \alpha^{-1}e$. Finally, it is routine to show that the supremum of the quantity α^{-1} over all vectors $a \in P$ satisfying $\sum a_k = 1$ is attained uniquely by $a = (1/n, \cdots, 1/n)$.

5. **A method for locating prescribed Pareto optimal solutions.** In this section we sketch a general (and somewhat idealized) procedure for solving (C_0), or equivalently, (P_0). After describing the method, we discuss some of the more prominent aspects of it. The proof of its various properties will be seen to draw on many of the results established above in §§ 2 through 4.

THE BASIC METHOD. Choose any $a \in P$ satisfying $\sum a_k = 1$, and let $\{\zeta^m\}$ be any sequence of positive real numbers decreasing to zero. For each $m = 1, 2, \cdots$, calculate the unique pair $(x^{\zeta^m}, y^{\zeta^m})$ which solves the complementarity problem (C_z) for the parameter value $z = \zeta^m a$. Then pick any cluster pair (x^0, y^0) of the sequence $\{(x^{\zeta^m}, y^{\zeta^m})\}$.

The justification of this is as follows. For each m, Theorem 1 guarantees the existence of a unique solution $(x^{\zeta^m}, y^{\zeta^m})$, and furthermore it can be computed by solving either one of two equivalent, relatively nice, convex minimization problems. By Theorems 7 and 8, at least one cluster pair (x^0, y^0) is assured, and any one selected is guaranteed to be a maximal strictly complementary solution of (C_0), that is, to have the maximum number of coordinate indices k such that either $x_k^0 > 0$ or $y_k^0 > 0$. Of course, by Theorems 7 and 8 these same conclusions hold even if the approach to 0 by the parameter z is from the direction of a only in an asymptotic sense.

Suppose now that one actually wants to locate the particular solution pair $(\tilde{x}, \tilde{y})$, which has the strong Pareto optimal properties with respect to a described in Proposition 5. There are several cases in which any cluster pair (x^0, y^0) selected through the above method will necessarily coincide with $(\tilde{x}, \tilde{y})$. For instance, according to Theorem 12 the above method operates in exactly one of the three possible "modes" (A), (B), or (C). If either of the relatively trivial modes (A) or (B) is the case, then $(x^0, y^0) = (\tilde{x}, \tilde{y})$ is automatic. Alternatively, according to Theorems 10 and 9, $(x^0, y^0) = (\tilde{x}, \tilde{y})$ is guaranteed if either f is polyhedral or else (C_0) has a strictly complementary solution (which, incidentally, can be detected by inspection of (x^0, y^0)).

If none of these criteria applies to ensure $(x^0, y^0) = (\tilde{x}, \tilde{y})$, then by Theorem 12 the method must be in Mode (C), and by Theorem 8 any cluster pair (x^0, y^0) selected must at least determine the minimal faces of Q, Q^I and Q^J, containing $\tilde{x}$ and $\tilde{y}$ respectively. With this information, it is possible to find each of $\tilde{x}$ and $\tilde{y}$ by applying the method to a certain subsidiary problem having lower dimensionality.

To FIND $\tilde{x}$: Let f_I be the restriction of f to the subspace $R^I = \{x \in R^n \,|\, x_k = 0, \ \forall k \notin I\}$ of dimension card $I < n$. Apply the method to the problem min $\{f_I(x) \,|\, x \in Q^I\}$, where $Q^I = \{x \in R^I \,|\, x_k \geq 0, \ \forall k \in I\}$. The x-component of any resulting cluster pair will be the $\tilde{x}$ corresponding to the original problem (P_0).

To FIND $\tilde{y}$: Let f_J^* be the restriction of f^* to the subspace $R^J = \{y \in R^n \,|\, y_k = 0, \ \forall k \notin J\}$ of dimension card $J < n$. Apply the method to the problem min $\{f_J^*(y) \,|\, y \in Q^J\}$, where $Q^J = \{y \in R^J \,|\, y_k \geq 0, \ \forall k \in J\}$. The y-component of any resulting cluster pair will be the $\tilde{y}$ corresponding to the original problem (P_0^*).

The justification of the procedure for locating $\tilde{x}$ is as follows. The subsidiary problem, involving essentially card I real variables, has solution set X^0 coinciding with that of the original, full-dimen-

sional problem (P_0). Also, it can be seen (e.g., by the results in the Appendix) that the subsidiary problem also satisfies the blanket hypothesis used throughout the paper (i.e., the conditions in Lemma 2) with respect to the space R^I. Finally, X^0 clearly meets the positive orthant $P^I = \{x \in R^I \mid x_k > 0, \; \forall k \in I\}$ of R^I. Hence, Theorems 8 and 12 imply that the method, when applied to the subsidiary problem, will necessarily be in Mode (A). This means that the x-component of any cluster pair is the strong Pareto optimum of X^0 with respect to a.

The justification of the procedure for locating $\tilde{y}$ is similar, except that the method, when applied to that subsidiary problem, will necessarily be in Mode (B) of Theorem 12.

Of course, in solving either of the subsidiary problems for $\tilde{x}$ or $\tilde{y}$ one is free to revise the choice of the "direction" vector a. This vector can also be thought of as describing the relative weights assigned to the various coordinate axes.

The function ρ defined in (4.34) provides a precise measure for gauging the improvement in x^{ζ^m} and y^{ζ^m} made during the course of the iterations. See Theorem 11 and its corollaries for results along these lines.

As mentioned above, the "typical" mode for the method to be in is Mode (C). In this event, Theorem 13 provides a stopping criterion for the method. It also assures that, once the iterations have brought one "close enough" (as measured in terms of ρ), the convergence (as measured in the differences $\inf (P_{\zeta^m}) - \inf (P_0)$ and $\inf (P^*_{\zeta^m}) - \inf (P_0^*))$ is very fast.

Another, simpler measure of convergence is provided by Corollary 1A, which implies for instance that

$$f(x^{\zeta^m}) - \zeta^m \leqq \inf (P_0) \leqq \zeta^m - f^*(y^{\zeta^m}) \,,$$

where the interval thus represented has total length ζ^m.

Corollary 1B provides bounds for X^0 and Y^0 when the iterations in the method are stopped short of finding a cluster pair (x^0, y^0).

By Corollary 8A, any cluster pair obtained by the method also provides bounds for X^0 and Y^0.

Finally, Theorem 7 presents a great deal of information relevant to implementing the method under only approximate, rather than "exact" calculations. The various estimates available from other results can be combined fruitfully with this.

6. **Similarities with Moreau's theorem and differences with the Proximal Point Algorithm.** The present paper treats the parametrized class of problems (C_z), $z \in Q$. These can clearly be reformulated as

(C_z) find $(x, y) \in$ graph $T \cap$ graph S_z ,

where

(6.1) $T = \partial f$ and $S_z = \partial(z \circ g)$.

Consider now the homeomorphism $L: P \to R^n$ given by

$$L(z_1, \cdots, z_n) = (\log z_1, \cdots, \log z_n) ,$$

which has inverse $E: R^n \to P$ given by

$$E(w_1, \cdots, w_n) = (\exp w_1, \cdots, \exp w_n) .$$

By means of the product mapping $L \times L$, the results of the present paper can be transformed from the context of $P \times P$ into the context of $R^n \times R^n$. This is the key to seeing the analogy with Moreau's theorem.

In particular, those portions of multivalued operators from R^n to R^n which map P into P can be transformed into multivalued operators which themselves map R^n into R^n. Thus, new operators T' and S'_w from R^n to R^n are induced by the operators in (6.1) by setting

(6.2)   graph $T' = \{(u, v) \mid \exists (x, y) \in (P \times P) \cap$ graph T,
                    $(L(x), L(y)) = (u, v)\}$

and

(6.3)   graph $S'_w = \{(u, v) \mid \exists (x, y) \in$ graph S_z, $(L(x), L(y)) = (u, v)\}$,

where $w = L(z)$. Of course, (6.3) makes sense only for $z \in P$. In terms of these induced operators, problem (C_z) can be reexpressed equivalently as

(D_w) find $(u, v) \in$ graph $T' \cap$ graph S'_w .

Here $w = L(z)$ and $z \in P$. Since S_z for $z \in P$ is characterized by

$$(x, y) \in \text{graph } S_z \Longleftrightarrow x \in P, y \in P, x_k y_k = z_k \quad \forall k ,$$

the induced S'_w, $w = L(z)$, is characterized by

$$(u, v) \in \text{graph } S'_\omega \Longleftrightarrow u \in R^n, v \in R^n, u + v = w .$$

It follows that

(6.4) $S'_w = \partial \left(-\dfrac{1}{2} \| \cdot - w \|^2 \right) ,$

where $\| \cdot \|$ denotes the Euclidean norm on R^n. Hence, *if* (see below) the T' induced by $T = \partial f$ via (6.2) is of the form

(6.5) $T' = \partial \varphi$,

for some closed proper convex function $\varphi: R^n \to (-\infty, +\infty]$, then Moreau's theorem (see formulation in [6, Theorem 31.5]) states the following. For each $w \in R^n$ there exists a unique solution (u^w, v^w) to (D_w), where u^w and v^w are characterized as the unique solutions to the dual optimization problems

(Q_w) $\min\limits_{u} \left\{ \varphi(u) + \dfrac{1}{2} \| u - w \|^2 \right\}$

and

(Q_w^*) $\min\limits_{v} \left\{ \varphi^*(v) + \dfrac{1}{2} \| v - w \|^2 \right\}$,

and one has

$$\inf (Q_w) + \inf (Q_w^*) = \frac{1}{2} \| w \|^2 \ .$$

Moreover, the mappings $w \to u^w$ and $w \to v^w$ are each gradient mappings of certain differentiable convex functions, so that $w \to (u^w, v^w)$ is continuous on R^n. All of this is in direct analogy with the conclusions of the present Theorem 1 and Corollary 5A concerning problems (P_z), (P_z^*) and (C_z) for the parameter values $z \in P$.

From Proposition 8 and Theorem 14 below, it follows that in the trivial case $n = 1$ the T' induced by $T = \partial f$ via (6.2) does in fact satisfy (6.5). It then follows from (6.4) and (6.5) that, in the case $n = 1$, Theorem 1 and Corollary 5A are logically equivalent to Moreau's theorem. Two simple examples given at the end of the section, however, show that for $n \geqq 2$ our treatment and Moreau's are in general distinct.

In the framework of Moreau's theorem and problems (Q_w), (Q_w^*) and (D_w), there is a closely associated iterative method for globally minimizing φ. Known as the Proximal Point Algorithm, it generates a sequence $\{u^m\} \subset R^n$ as follows: any $u^0 \in R^n$ is chosen to initiate the algorithm; for $m = 0, 1, 2, \cdots, u^{m+1}$ is taken to be the unique global minimizer of the function

$$u \longrightarrow \varphi(u) + \frac{1}{2} \| u - u^m \|^2 \ ,$$

that is, the unique element u satisfying the relation

$$0 \in (T' - S'_{u^m})(u) \ ,$$

where T' and S'_{u^m} are as in (6.5) and (6.4). It is known that if φ achieves its minimum at all, the sequence $\{u^m\}$ thus generated converges to some particular u minimizing φ, that is, satisfying the

relation $0 \in T'(u)$ for $T' = \partial\varphi$.

What becomes of this algorithm when transformed via $E \times E$ from the $R^n \times R^n$ setting to the $P \times P$ setting of the present paper? A sequence $\{x^m\} \subset P$ is generated as follows: any $x^0 \in P$ is chosen to initiate the algorithm; for $m = 0, 1, 2, \cdots, x^{m+1}$ is taken to be the unique element x satisfying the relation

$$0 \in (T - S_{x^m})(x) .$$

Here T can be any multivalued operator induced, through setting

$$\text{graph } T = \{(x, y) \mid \exists(u, v) \in \text{graph } T', (E(u), E(v)) = (x, y)\} ,$$

by means of a multivalued operator T' of the form (6.5). Assuming that the relation

$$e \in T(x), e = (1, \cdots, 1) ,$$

has any solution x at all in P, the sequence $\{x^m\}$ thus generated converges to a particular such solution. It would be interesting to know whether this convergence behavior also holds for T an arbitrary maximal monotone operator from R^n to R^n satisfying

$$\varnothing \neq (P \times P) \cap \text{graph } T .$$

A more restrictive form of this is whether such convergence behavior holds for every T of the form $T = \partial f$, for any closed proper convex function $f: R^n \to (-\infty, +\infty]$ satisfying

(6.6) $\varnothing \neq P \cap \text{dom } f$ and $\varnothing \neq P \cap \text{dom } f^* .$

We have not investigated either of these questions.

The preceding shows that the Proximal Point Algorithm is quite different in character from the method described in § 5. We conclude this discussion by transforming the method of § 5, by means of $L \times L$, from the $P \times P$ setting into the $R^n \times R^n$ setting. It is easy to see that the transformed method entails solving a sequence of problems (D_w), where the w's are of the form $w = \omega e + b$ for a sequence of real ω's tending to $-\infty$. Here $e = (1, \cdots, 1)$, $b = L(a)$ for the fixed vector a in § 5, and T' is of the form (6.2) for some $T = \partial f$, where f is closed proper convex on R^n satisfying (6.6). The corresponding solutions (u^ω, v^ω) to these problems have coordinates tending to either some finite limit or else toward $-\infty$, and the idealized accumulation points (u, v) of these sequences satisfy transformed complementary slackness conditions, that is,

(6.7) $\forall k = 1, \cdots, n$, either $u_k = -\infty$ or $v_k = -\infty .$

The transformed procedure thus develops certain "asymptotic" in-

formation concerning the way in which graph T' is situated in $R^n \times R^n$. In particular, if (6.5) happens to hold, the data in (6.7) corresponds to recession information about φ and/or information about the rate of growth of φ at relative boundary points of dom φ. It would be interesting to know whether the transformed method just described behaves similarly for every T' of the form (6.5) or, more generally, for an arbitrary maximal monotone operator T'. We have not investigated these questions.

Next, we establish the two results already cited, pertaining to the case $n = 1$, and then we conclude the section with the two examples dealing with the case $n \geqq 2$.

Note that since $n = 1$ in the next two results, P is now simply the open half-line $(0, +\infty)$.

PROPOSITION 8.   *The relations*

$$(6.9) \qquad \text{graph } \varPhi = \{(\log x, \log y) \mid (x, y) \in \text{graph } F\} \, ,$$

$$(6.10) \qquad \text{graph } F = \{(\exp u, \exp v) \mid (u, v) \in \text{graph } \varPhi\}$$

express a one-to-one correspondence between the multivalued operators $\varPhi: R \to R$ whose graphs are maximal monotone with respect to $R \times R$ and the multivalued operators $F: P \to P$ whose graphs are maximal monotone with respect to $P \times P$. This correspondence includes in particular a one-to-one correspondence between the singlevalued selections $\varPhi_0$ of such $\varPhi$'s and the singlevalued selections F_0 of such F's.

Proof. Relations (6.9) and (6.10) clearly provide a one-to-one correspondence between all the multivalued operators $\varPhi: R \to R$ and all the multivalued operators $F: P \to P$. We must show that graph $\varPhi$ maximal monotone with respect to $R \times R$ implies that the corresponding set graph F defined by (6.10) is maximal monotone with respect to $P \times P$, and vice versa. We check only the first of these implications, as the converse follows similarly upon interchanging the roles of the exponential and the logarithm. To see that graph F is monotone, for $i = 1, 2$ suppose that

$$(x_i, y_i) = (\exp u_i, \exp v_i), \text{ where } (u_i, v_i) \in \text{graph } \varPhi \, .$$

The assumed monotonicity of graph $\varPhi$ yields

$$\langle u_2 - u_1, v_2 - v_1 \rangle \geqq 0 \, ,$$

so that

$$\text{either } [u_2 \geqq u_1 \text{ and } v_2 \geqq v_1] \text{ or } [u_1 \geqq u_2 \text{ and } v_1 \geqq v_2] \, .$$

Exponentiating, we obtain

either $[x_2 \geqq x_1$ and $y_2 \geqq y_1]$ or $[x_1 \geqq x_2$ and $y_1 \geqq y_2]$,

and hence

$$\langle x_2 - x_1, y_2 - y_1 \rangle \geqq 0 .$$

To see that graph F is maximal monotone with respect to $P \times P$ we must show that, given any $(x_1, y_1) \in (P \times P) \backslash \text{graph } F$, there exists an $(x_2, y_2) \in \text{graph } F$ such that

(6.11) $\langle x_2 - x_1, y_2 - y_1 \rangle < 0 .$

Since $(u_1, v_1) = (\log x_1, \log y_1) \in (R \times R) \backslash \text{graph } \Phi$ and graph Φ is maximal monotone with respect to $R \times R$, there exists some $(u_2, v_2) \in$ graph Φ such that

$$\langle u_2 - u_1, v_2 - v_1 \rangle < 0 ,$$

and hence

either $[u_2 > u_1$ and $v_1 > v_2]$ or $[u_1 > u_2$ and $v_2 > v_1]$.

For $(x_2, y_2) = (\exp u_2, \exp v_2) \in \text{graph } F$, it follows that

either $[x_2 > x_1$ and $y_1 > y_2]$ or $[x_1 > x_2$ and $y_2 > y_1]$,

so that (6.11) holds. The assertion concerning singlevalued selections now follows, due to the strictly increasing nature of the logarithm and exponential functions.

THEOREM 14. *The functions $\varphi \colon R \to (-\infty, +\infty]$ which are lower semicontinuous, convex and somewhere finite are in one-to-one correspondence, up to additive real constants, with the functions $f \colon P \to (-\infty, +\infty]$ which are lower semicontinuous, nondecreasing, convex and somewhere finite. This correspondence can be given explicit form in terms of arbitrary singlevalued selections Φ_0 from Φ and F_0 from F, where $\Phi \colon R \to R$ and $F \colon P \to P$ are multivalued operators given by*

(6.12)   graph $\Phi = \{(u, v) \in R \times R \mid \varphi(u') \geqq \varphi(u) + \langle u'-u, v \rangle, \ \forall u' \in R\}$,

(6.13)   graph $F = \{(x, y) \in P \times P \mid f(x') \geqq f(x) + \langle x'-x, y \rangle, \ \forall x' \in P\}$.

Specifically, any such selections $\Phi_0 \colon \text{dom } \Phi \to R$ and $F_0 \colon \text{dom } F \to P$ are extended to functions $\bar{\Phi}_0 \colon R \to [-\infty, +\infty]$ and $\bar{F}_0 \colon P \to [0, +\infty]$ by setting

(6.14) $\bar{\Phi}_0(u) = \begin{cases} -\infty & \textit{if } u \textit{ lies to the left of } \text{dom } \Phi \\ \Phi_0(u) & \textit{if } u \in \text{dom } \Phi \\ +\infty & \textit{if } u \textit{ lies to the right of } \text{dom } \Phi , \end{cases}$

$$(6.15) \qquad \bar{F}_0(x) = \begin{cases} 0 & \textit{if } x \textit{ lies to the left of } \mathrm{dom}\, F, \\ F_0(x) & \textit{if } x \in \mathrm{dom}\, F \\ +\infty & \textit{if } x \textit{ lies to the right of } \mathrm{dom}\, F. \end{cases}$$

The correspondence between φ and f is then expressible in terms of $\bar{\Phi}_0$, $\bar{F}_0$ and arbitrary numbers $u_0 \in \mathrm{dom}\, \Phi$, $x_0 \in \mathrm{dom}\, F$ by means of the relations

$$(6.16) \qquad \varphi(u) = \int_{u_0}^{u} \log\left(\bar{F}_0(\exp \mu)\right) d\mu, \quad \forall u \in R,$$

$$(6.17) \qquad f(x) = \int_{x_0}^{x} \exp\left(\bar{\Phi}_0(\log \xi)\right) d\xi, \quad \forall x \in P,$$

where we use the conventions

$$\log 0 = -\infty, \ \log(+\infty) = +\infty,$$

$$\exp(-\infty) = 0, \ \exp(+\infty) = +\infty.$$

Proof. It is known (e.g., [6, § 24]) that the functions φ (as described above) are, up to additive real constants, in one-to-one correspondence with the multivalued operators Φ (as described above). This correspondence is given explicitly by means of (6.12) and the relation

$$(6.18) \qquad \varphi(u) = \int_{u_0}^{u} \bar{\Phi}_0(\mu) d\mu, \quad \forall u \in R,$$

where $\bar{\Phi}_0$ is as in (6.14) for any singlevalued selection Φ_0 from Φ and u_0 is any element of $\mathrm{dom}\, \Phi$. Furthermore, it is known also that such Φ's are precisely the multivalued operators from R to R whose graphs are maximal monotone subsets of $R \times R$. One can apply the same ideas and techniques used in the proofs of these results to establish analogous results with $P = (0, +\infty)$ playing the role of R. We omit the lengthy details and simply state the findings. The functions f (as described above) are, up to additive real constants, in one-to-one correspondence with the multivalued operators F (as described above). This correspondence is given explicitly by means of (6.13) and the relation

$$(6.19) \qquad f(x) = \int_{x_0}^{x} \bar{F}_0(\xi) d\xi, \quad \forall x \in P,$$

where $\bar{F}_0$ is as in (6.15) for any singlevalued selection F_0 from F and x_0 is any element of $\mathrm{dom}\, F$. Furthermore, these F's are precisely the multivalued operators from P to P whose graphs are maximal monotone subsets of $P \times P$. Finally, it follows from Pro-

position 8 that the extended selections $\bar{\varPhi}_0$ and $\bar{F}_0$ given by (6.14) and (6.15) are in one-to-one correspondence by means of

$$(6.20) \qquad \bar{\varPhi}_0(\mu) = \log\left(\bar{F}_0(\exp\mu)\right), \ \forall\mu \in R \ ,$$

$$(6.21) \qquad \bar{F}_0(\xi) = \exp\left(\bar{\varPhi}_0(\log\xi)\right), \ \forall\xi \in P \ ,$$

where we adopt the conventions described above. The theorem follows by combining all these one-to-one correspondences. In particular, (6.18) and (6.20) give (6.16), while (6.19) and (6.21) give (6.17). This concludes the proof.

We remark that the correspondence just noted between functions f on P and multivalued operators F from P to P extends trivially and in the obvious way to the situation in which $P = (0, +\infty)$ is replaced by $Q = [0, +\infty)$. The resulting correspondence is between the Young's functions f on Q and the multivalued operators F from Q to Q which are maximal monotone with respect to $Q \times Q$. Also associated with such f's, of course, are the conjugate Young's functions $f^+ : Q \to (-\infty, +\infty]$ given by

$$f^+(y) = \sup_{x \in Q} \{\langle x, y\rangle - f(x)\}, \ \forall y \in Q \ .$$

All this is the case $n = 1$ of monotone conjugacy, which is treated for general n by [6, Theorem 12.4] and Corollary 2A above (see also Proposition 3(ii)).

EXAMPLE 2.   Let $n = 2$, and take φ to be

$$\varphi(u) = \psi_C(u), \ C = \{(u_1, u_2) \mid u_1 + u_2 = 0\} \ .$$

Then $T' = \partial\varphi$ is given by

$$\text{graph } T' = \{(u, v) \in R^2 \times R^2 \mid u_1 + u_2 = 0, \ v_1 = v_2\} \ .$$

The T induced by this T' via (6.6) is given by

$$\text{graph } T = \{(x, y) \in P \times P \mid x_1 x_2 = 1, \ y_1 = y_2\} \ .$$

For any (x, y), $(x', y') \in \text{graph } T$ one finds that

$$\langle x' - x, y' - y\rangle = \xi\eta \ ,$$

where

$$\xi = (x_1'x_1 - 1)(x_1' - x_1)(x_1'x_1)^{-1} \in R \ ,$$
$$\eta = y_1' - y_1 \in R \ .$$

By appropriate choice of $x_1, x_1', y_1, y_1' > 0$ the expression $\langle x' - x, y' - y\rangle$ can therefore be made to assume any real value. This shows T fails even to be monotone, much less a subdifferential. (This example

extends in the obvious way to all $n \geq 2$.)

EXAMPLE 3. Let $n = 2$, and take f to be

$$f(x) = \psi_C(x), \quad C = \{(x_1, x_2) \mid x_1 + x_2 = 4 \text{ and } 1 \leq x_k \leq 3 \; \forall k\} \; .$$

Then $T = \partial f$ is given by

$$T(x) = \begin{cases} \{y \mid y_1 = y_2\} & \text{if } x \in \mathrm{ri}\, C \\ \{y \mid y_1 \leq y_2\} & \text{if } x = (1, 3) \\ \{y \mid y_1 \geq y_2\} & \text{if } x = (3, 1) \\ \varnothing & \text{if } x \notin C \; . \end{cases}$$

The T' induced by this T via (6.2) is given by

$$T'(u) = \begin{cases} \{v \mid v_1 = v_2\} & \text{if } u = L(x) \text{ for some } x \in \mathrm{ri}\, C \\ \{v \mid v_1 \leq v_2\} & \text{if } u = (0, \log 3) \\ \{v \mid v_1 \geq v_2\} & \text{if } u = (\log 3, 0) \\ \varnothing & \text{otherwise} \; . \end{cases}$$

We now show that this T' fails to be monotone. Consider (u, v), $(u', v') \in \mathrm{graph}\; T'$ such that

$$u = (\log x_1, \log x_2) \text{ for } x_1 + x_2 = 4 \text{ and } 1 < x_k < 3 \; \forall k \; ,$$
$$u' = (\log x_1', \log x_2') \text{ for } x_1' + x_2' = 4 \text{ and } 1 < x_k' < 3 \; \forall k \; ,$$
$$v = (\nu, \nu) \text{ for } \nu \in R \; ,$$
$$v' = (\nu', \nu') \text{ for } \nu' \in R \; .$$

One finds that

$$\langle u' - u, v' - v \rangle = \mu(\nu' - \nu), \quad \mu = \log\, (x_1' x_2' / x_1 x_2) \; .$$

It follows that $\langle u' - u, v' - v \rangle$ can be made to assume any real value. Hence, T' fails to be monotone, and so *a fortiori* it cannot be a subdifferential. (This example also extends in the obvious way to all $n \geq 2$.)

Appendix. Here we show how the framework developed above can be made to handle problems similar to (P_0) or (C_0) but in which only some of the variables are constrained to be nonnegative. More precisely, the minimal general hypothesis is identified under which such problems are fully equivalent to those treated in this paper, and the formulas essential to make use of this equivalence are provided. Except for differences in notation, the situation treated here is exactly that encountered in §5, where certain lower-dimensional problems were introduced in connection with finding $\tilde{x}$

and $\tilde{y}$ in the general case.

Let f be a closed proper convex function on R^n, but do *not* assume the earlier blanket hypothesis

$$\varnothing \neq P \cap \operatorname{dom} f \quad \text{and} \quad \varnothing \neq P \cap \operatorname{dom} f^* .$$

Let m be a fixed integer satisfying $1 \leqq m < n$, and let C and C^* be the dual cones

$$C = \{x \in R^n \mid x_k \geqq 0 \ \forall k \leqq m \text{ and } x_k \in R \ \forall k > m\} ,$$

$$C^* = \{y \in R^n \mid y_k \geqq 0 \ \forall k \leqq m \text{ and } y_k = 0 \ \forall k > m\} .$$

We consider the three problems

(A.1) $\qquad\qquad \min \{f(x) \mid x \in C\} ,$

(A.2) $\qquad\qquad \min \{f^*(y) \mid y \in C^*\} ,$

and

(A.3) $\quad$ find $(x, y) \in (C \times C^*) \cap \operatorname{graph} \partial f$ such that $\langle x, y \rangle = 0 .$

Note that the extreme case $m = n$ corresponds to the problems (P_0), (P_0^*) and (C_0).

It is known [6, Theorem 31.4] that the hypothesis

$$\varnothing \neq C \cap \operatorname{ri}(\operatorname{dom} f) \text{ and } \varnothing \neq C^* \cap \operatorname{ri}(\operatorname{dom} f^*)$$

is the weakest general condition for ensuring that all three of these problems have solutions. Moreover, it also implies that

$$\inf_C \{f\} + \inf_{C^*} \{f\} = 0$$

and that a pair (x, y) solves (A.3) if and only if x solves (A.1) and y solves (A.2). Propositions 9 and 10 to follow show that the slightly stronger condition

(A.4) $\qquad \varnothing \neq \operatorname{ri} C \cap \operatorname{ri}(\operatorname{dom} f) \text{ and } \varnothing \neq \operatorname{ri} C^* \cap \operatorname{ri}(\operatorname{dom} f^*)$

is precisely what is needed to reduce problems (A.1), (A.2), (A.3) to the form of (P_0), (P_0^*), (C_0) and moreover have the blanket hypothesis met. The reduction involves the function $f_m \colon R^m \to [-\infty, +\infty]$ defined by

$$f_m(x_1, \cdots, x_m) = \inf \{f(x_1, \cdots, x_n) \mid x_k \in R \ \forall k > m\} .$$

PROPOSITION 9. *Assume that f satisfies the condition*

(A.5) $\qquad \begin{cases} \textit{there is no } x \in R^n \textit{ such that} \\ x_k = 0 \ \forall k \leqq m, \ f0^+(x) \leqq 0, \ f0^+(-x) > 0 . \end{cases}$

Then f_m is a closed proper convex function satisfying

$$f_m(x_1, \cdots, x_m) = \min \{f(x_1, \cdots, x_n) \mid x_k \in R, \ \forall k > m\} \ ,$$
$$f_m 0^+(x_1, \cdots, x_m) = \min \{f 0^+(x_1, \cdots, x_n) \mid x_k \in R, \ \forall k > m\} \ ,$$
$$f_m^*(y_1, \cdots, y_m) = f^*(y_1, \cdots, y_m, 0, \cdots, 0) \ ,$$
$$f_m^* 0^+(y_1, \cdots, y_m) = f^* 0^+(y_1, \cdots, y_m, 0, \cdots, 0) \ ,$$

and also

$$(y_1, \cdots, y_m) \in \partial f_m(x_1, \cdots, x_m)$$

if and only if there exist $x_{m+1}, \cdots, x_m \in R$ such that

$$(y_1, \cdots, y_m, 0, \cdots, 0) \in \partial f(x_1, \cdots, x_m, x_{m+1}, \cdots, x_n) \ .$$

Proof. Observe that $f_m = Af$, where $A \colon R^n \to R^m$ is the linear transformation

$$A(x_1, \cdots, x_m, x_{m+1}, \cdots, x_n) = (x_1, \cdots, x_m) \ .$$

Clearly, the adjoint $A^* \colon R^m \to R^n$ is given by

$$A^*(y_1, \cdots, y_m) = (y_1, \cdots, y_m, 0, \cdots, 0) \ .$$

By [6, Corollary 16.2.1], condition (A.5) is equivalent to the existence of a vector y such that $A^* y \in \mathrm{ri}\,(\mathrm{dom}\, f^*)$. Hence, [6, Theorems 9.5 and 16.3] imply that

$$(f^* A^*)(y_1, \cdots, y_m) = f^*(y_1, \cdots, y_m, 0, \cdots, 0)$$

is a closed proper convex function with conjugate

$$(Af)(x_1, \cdots, x_m) = \inf \{f(x_1, \cdots, x_m, x_{m+1}, \cdots, x_n) \mid x_k \in R, \ \forall k > m\} \ ,$$

where the infimum is actually attained. This establishes that f_m is closed proper convex (being the conjugate of such a function) and also the formulas asserted for f_m and f_m^*. The formula for $f_m^* 0^+$ follows by [6, Theorem 9.5], and that for $f_m 0^+$ follows by a closer examination of the proof of [6, Theorem 9.2]. The characterization of ∂f_m in terms of ∂f follows from [6, Theorem 23.9]. This concludes the proof.

On the basis of Proposition 9, we can consider seriously in place of (A.1), (A.2), (A.3) the "reduced" problems

$$(A.6) \qquad\qquad \min_{Q_m} \{f_m\} \ ,$$

$$(A.7) \qquad\qquad \min_{Q_m} \{f_m^*\} \ ,$$

and

(A.8)   find $(x, y) \in (Q_m \times Q_m) \cap \operatorname{graph} \partial f_m$ such that $\langle x, y \rangle = 0$

where

$$Q_m = \{x \in R^m \mid x_k \geqq 0, \ \forall k\} \ .$$

These latter three problems are simply (P_0), (P_0^*), (C_0) posed in the space R^m for the function f_m. Assuming condition (A.5), it is clear from Proposition 9 that solving (A.6), (A.7) or (A.8) is equivalent to solving (A.1), (A.2) or (A.3), respectively. What remains is to characterize, in terms of the original function f, the blanket hypothesis required to apply Theorem 1 and all the other results to problems (A.6), (A.7), (A.8). This is done in our last proposition, which uses the following technical facts.

LEMMA 3. *Let K be a nonempty convex cone in R^n, and let f be a proper convex function on R^n. Then*

$$\varnothing \neq \operatorname{ri} K \cap \operatorname{ri} (\operatorname{dom} f)$$

is equivalent to the condition

$$[y \in K^* \ and \ f^*0^+(y) \leqq 0] \Longrightarrow [-y \in K^* \ and \ f^*0^+(-y) \leqq 0] \ ,$$

where $K^* = \{y \mid \langle x, y \rangle \geqq 0, \ \forall x \in K\}$.

Proof. We modify the proof of [6, Lemma 16.2], in which K is actually a subspace. By [6, Theorems 11.3 and 11.1], $\varnothing = \operatorname{ri} K \cap \operatorname{ri} (\operatorname{dom} f)$ occurs if and only if

$$\exists y, \ \sup_{\operatorname{dom} f} \langle \cdot, y \rangle \leqq \inf_{riK} \langle \cdot, y \rangle \quad \text{and} \quad \inf_{\operatorname{dom} f} \langle \cdot, y \rangle < \sup_{riK} \langle \cdot, y \rangle \ .$$

By [6, Theorem 13.3] and the fact that K is a cone, the latter condition occurs if and only if

$$\exists y, \ f^*0^+(y) \leqq - \psi_{K^*}(y) \ \text{and} \ -f^*0^+(-y) < \psi_{K^*}(-y) \ .$$

Since f^*0^+ is never $-\infty$ [6, Theorem 8.5] and ψ_{K^*} assumes only the values 0 and $+\infty$, the latter is equivalent to

$$\exists y \in K^*, \ f^*0^+(y) \leqq 0 \ \text{and} \ [0 < f^*0^+(-y) \ \text{if} \ -y \in K^*] \ .$$

The lemma follows by taking contrapositives.

In the following result we write $P_m = \{x \in R^m \mid x_k > 0, \ \forall k\}$.

PROPOSITION 10. *Then following three conditions are pairwise equivalent:*

(a) $\varnothing \neq P_m \cap \operatorname{dom} f_m$

(b) $\varnothing \neq \operatorname{ri} C \cap \operatorname{ri} (\operatorname{dom} f)$

(c) $f^*0^+(y) > 0$ *whenever* $y \in C^*$ *and* $y_k > 0$ *for some* $k \leq m$.
If condition (A.5) *is satisfied, the following three conditions are pairwise equivalent:*

(a*) $\varnothing \neq P_m \cap \operatorname{dom} f_m^*$

(b*) $\varnothing \neq \operatorname{ri} C^* \cap \operatorname{ri}(\operatorname{dom} f^*)$

(c*) $f0^+(x) > 0$ *whenever* $x \in C$ *and* $x_k > 0$ *for some* $k \leq m$.
Moreover, even without assuming condition (A.5), *condition* (b*) *is equivalent to the conjunction of* (c*) *and* (A.5).

Proof. Since P_m is open and $\operatorname{dom} f_m$ is convex, (a) is equivalent to

$$\varnothing \neq P_m \cap \operatorname{ri}(\operatorname{dom} f_m).$$

Since $f_m = Af$ implies $\operatorname{dom} f_m = A \operatorname{dom} f$, by [6, Theorem 6.6] the latter condition is equivalent to (b). By Lemma 3 and the fact that C^* contains no lines, (b) is equivalent to (c). To prove the corresponding equivalences involving f_m^*, note first that, by [6, Corollary 16.2.1],

$$(A.5) \text{ is satisfied} \iff \varnothing \neq A^{*-1}(\operatorname{ri}(\operatorname{dom} f^*)).$$

Also, from the formula $f_m^* = f^* A^*$ in Proposition 9 we have

$$\operatorname{dom} f_m^* = A^{*-1} \operatorname{dom} f^*.$$

From [6, Theorem 6.7], it follows that condition (A.5) implies

$$(y_1, \cdots, y_m) \in \operatorname{ri}(\operatorname{dom} f_m^*) \iff (y_1, \cdots, y_m, 0, \cdots, 0) \in \operatorname{ri}(\operatorname{dom} f^*).$$

Hence,

$$\varnothing \neq P_m \cap \operatorname{dom} f_m^* \iff \varnothing \neq P_m \cap \operatorname{ri}(\operatorname{dom} f_m^*)$$
$$\iff \varnothing \neq \operatorname{ri} C^* \cap \operatorname{ri}(\operatorname{dom} f^*)$$
$$\iff (C \cap \operatorname{rec} \operatorname{cone} f) \subset -(C \cap \operatorname{rec} \operatorname{cone} f),$$

where the last equivalence is by Lemma 3. By using the form of C, the last condition can be seen equivalent to the conjunction (c*) and (A.5). This concludes the proof.

REFERENCES

1. A. D. Alexandroff, *Almost everywhere existence of the second differential of a convex function and some properties of convex surfaces connected with it*, Leningrad State Univ. Ann., Math. Ser., **6** (1939), 3-35 (Russian).

2. W. Fenchel, *Convex Cones, Sets and Functions*, mimeographed lecture notes, Princeton University, 1951.

3. L. McLinden, *Symmetrized separable convex programming*, Trans. Amer. Math. Soc., **247** (1979), 1-44.

4. L. McLinden, *The complementarity problem for maximal monotone multifunctions*, in *Variational Inequalities and Complementarity Problems: Theory and Applications*, ed. by R. W. Cottle, F. Giannessi and J.-L. Lions, Wiley, Chichester, 1980, 20 pp.

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

6. R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, N.J., 1970.

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
URBANA, IL 61801