DISCUSSION PAPER NO. 273

THE DUALITY BETWEEN SUBOPTIMIZATION AND PARAMETER DELETION

by

Elmor L. Peterson

February 22, 1977

The Duality between Suboptimization and Parameter Deletion

by

Elmor L. Peterson*

Abstract

The main tool used in studying the influence of given perturbation parameters on a given optimization problem is, of course, the corresponding Rockafellar dual problem (in which the dual variables are in a one-to-one correspondence with the parameters). However, even when the given optimization problem is defined in terms of simple formulas, there are many important cases where the corresponding dual problem can not be computed explicitly (in terms of simple formulas) unless certain additional (uninteresting) perturbation parameters are included. Under a very weak hypothesis, the main theorem to be given here asserts that a (sub)-optimization of the resulting dual problem over the additional (uninteresting) dual variables produces the desired dual problem (i.e., the dual problem that corresponds to the original optimization problem without the additional perturbation parameters). In addition to its uses in parametric analysis this theorem can be used to show that various decomposition principles are dual to one another and hence are essentially equivalent.

Key Words

Duality, suboptimization, parametric programming, decomposition.

CONTENTS

*Department of Industrial Engineering/Management Sciences and Department of Mathematics, Northwestern University, Evanston, Illinois 60201. Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. AFOSR-77-3134.

1. Introduction. A given optimization problem is seldom studied in isolation, but is usually embedded in a parameterized family of closely related optimization problems. There are generally many such families to choose from, but for practical reasons the family chosen usually includes as parameters at least all of the potentially variable "problem inputs". The other parameters included may not have any practical significance, but are chosen because their presence makes certain computations much more tractable. In defining the parameter scales it is advantageous to make zero coincide with the given problem, so that the input parameters can also be interpreted as perturbations of the given input data.

In engineering design, management science, and operations research there are at least two compelling practical reasons for studying such problem families: (1) a designer or decision-maker frequently needs to know how his optimal design or decision changes with changing input data; and (2) he also needs to know the sensitivity of his optimal design or decision to small perturbations of his given input data, so that he can establish the accuracy needed for such data. Moreover, in the prediction of scientific phenomena through the use of variational principles there are clearly analogous reasons for studying such problem families. Of course, the indispensable tool to be used in all such studies is the corresponding Rockafellar dual problem [13,14,15].

In "ordinary mathematical programming" the only perturbation parameters considered are those that perturb the given upper and/or lower bounds placed on the constraint functions. As demonstrated on page 320 of [14], the corresponding Rockafellar dual problem is then just the "ordinary dual problem" that evolved from the "Wolfe dual problem" [16] via the work of Falk [4]. Unlike duality in linear programming, the ordinary dual problem generally can not be computed explicitly (in terms of simple formulas), even when the objective and constraint functions in the given problem are expressed in terms of simple formulas.

According to Rockafellar (on page 322 of [14]), this major obstacle to working with the ordinary dual problem is, in a certain sense, due to the fact that the upper and/or lower bound perturbation parameters "are not enough to compensate for nonlinearities of the constraint functions".

In (generalized) "geometric programming" there are additional perturbation parameters that compensate for such nonlinearities. In fact, [12] (and, for a limited formulation, page 324 of [14]) shows that the corresponding Rockafellar dual problem is then just the "geometric dual problem" that evolved from Chapter VII of [3] via the work of Peterson [7,8,9]. Like duality in linear programming, the geometric dual problem generally can be computed explicitly (in terms of functions that in many important cases have very simple formulas). In fact, geometric programming duality is, in retrospect, a more natural extension of linear programming duality than is ordinary programming duality.

Some, though usually not all, of the additional perturbation parameters may actually perturb other problem input data and hence be of interest in their own right. In fact, this phenomena can be clearly illustrated by the following five classes of optimization problems (which, incidently, also indicate how geometric programming has more closely unified many of the major subfields of optimization theory).

In prototype geometric programming (described by examples 1 and 6 in sections 2.1 and 2.2 of [9]), nonconvex inseparable ordinary "posynomial" optimization problems are formulated as convex separable geometric programming problems in which the additional perturbation parameters perturb (the logarithm of) the given posynomial coefficients. Generally many, though usually not all, of the posynomial coefficients are potentially variable inputs, such as costs per unit quantity of materiél.

In the theory of "nonlinear networks" (described by example 5 in section 2.1 of [9]), the variational principles employed are in essence already formulated as separable geometric programming problems in which the perturbation parameters perturb the given input "flows" or input "potential differences" and certain other (physically imperturbable) quantities. In a recent economics extension [10] the perturbation parameters perturb the given input "flows" or input "costs" and "prices" and certain other (economically imperturbable) quantities.

In the theory of "dynamic programming" with linear "transition functions" (described by example 4 in section 2.1 of [9]), the multistage optimization problems considered are readily recognized as partially separable geometric programming problems in which the perturbation parameters translate and hence perturb the given "policy" and "state" sets as well as the given "initial state" and "final target" sets.

In facility location theory (described by example 3 in section 2.1 of [9]), the inseparable ordinary "generalized Weber problems" considered are easily formulated as (partially separable) geometric programming problems in which the additional perturbation parameters translate and hence perturb the given locations of the previously existing facilities.

In the theory of "ℓ_p programming" (described by example 2 in section 2.1 of [9]), ℓ_p constrained ℓ_p regression problems and inseparable ordinary quadratically constrained quadratic programming problems are formulated as separable geometric programming problems in which the additional perturbation parameters perturb the given vector that is being "optimally approximated".

In summary, the geometric programming approach to a given optimization problem is generally much more powerful than the ordinary programming approach. In fact, many inseparable ordinary programming problems can be formulated as separable geometric programming problems; and the corresponding geometric dual problem provides the means for analyzing the effect of a much broader class of input perturbations. Moreover, there are many important cases where the corresponding geometric dual problem can be derived in terms of very simple formulas even though the corresponding ordinary dual problem can be expressed only implicitly in terms of the solution of another parameterized family of optimization problems -- a dichotomy that is explained in [11].

Although such tractability provides a relatively easy analysis of the corresponding problem family, the preceding five examples indicate that in many important cases such problem families are larger than desired; that is, they have perturbation parameters that perturb given input data that is actually imperturable according to real-world considerations. Such excess perturbation parameters can be a dimensional curse because they can drastically increase the number of certain computations without any corresponding increase in the amount of usable information. However, the main theorem to be established here implies that such perturbation parameters can be effectively deleted (i.e. kept equal to zero) by suboptimizing the geometric dual problem over the corresponding dual variables.

In large-scale linear programming it is known that the "Benders decomposition" [1] of certain structured dual problems (by suboptimization) is "dual" to the "Dantzig-Wolfe decomposition" [2] of the corresponding primal problems (by Lagrangian optimization). That type of decomposition duality can be viewed as a corollary to the main theorem of this paper. In fact, the main theorem of this paper can also be used in large-scale geometric programming to dualize certain decomposition principles described in subsections 3.1.6 and 3.3.6 of [9].

The only prerequisite for this paper is some of the convexity theory in [14] -- especially the theory having to do with the "relative interior" (ri S) of an arbitrary convex set $S \subseteq E_N$ (N-dimensional Euclidean space).

2. Rockafellar duality. Suppose that $g:C$ is a (proper) function g with a nonempty (effective) domain $C \subseteq E_N$, and assume that the independent variable (d,p) in C is the Cartesian product of a "decision" (vector) variable d and a "perturbation" (vector) parameter p .

Consider the parameterized family $\mathcal{A}$ that consists of the following optimization problems $A(p)$.

PROBLEM $A(p)$. <u>Using the "feasible solution" set</u>

$$S(p) \triangleq \{d \mid (d,p) \in C\} \ ,$$

<u>calculate both the "problem infimum"</u>

$$\varphi(p) \triangleq \inf_{d \in S(p)} g(d,p)$$

<u>and the "optimal solution" set</u>

$$S^*(p) \triangleq \{d \in S(p) \mid g(d,p) = \varphi(p)\} \ .$$

For a given perturbation p , problem $A(p)$ is either "consistent" or "inconsistent", depending on whether the feasible solution set $S(p)$ is nonempty or empty. The (effective) domain of the infimum function φ is the "feasible perturbation" set

$$\mathcal{P} \triangleq \{p \mid \text{problem } A(p) \text{ is consistent}\} \ ,$$

which is obviously identical to $\{p \mid (d,p) \in C$ for at least one $d\}$ and hence is not empty. Unlike the function g, the function φ may assume the value $-\infty$. However, for our purposes, it is not advantageous to follow Rockafellar's custom of artificially defining g and φ to be $+\infty$ outside their respective domains C and $\mathcal{P}$.

Now, suppose that $g:C$ has a "conjugate transform" $h:D$; that is, suppose there is a function h with a nonempty domain

$$D \triangleq \{(q,e) \mid \sup_{(d,p) \in C} [\langle q,d\rangle + \langle e,p\rangle - g(d,p)] < +\infty\}$$

and function values

$$h(q,e) \triangleq \sup_{(d,p) \in C} [\langle q,d\rangle + \langle e,p\rangle - g(d,p)] \ .$$

The inner product $\langle q,d\rangle$ associates the "dual perturbation" parameter q with the "primal decision" variable d, and the inner product $\langle e,p\rangle$ associates the "dual decision" variable e with the "primal perturbation" parameter p .

Consider the parameterized family $\mathcal{B}$ that consists of the following optimization problems $B(q)$.

PROBLEM $B(q)$. Using the feasible solution set

$$T(q) \triangleq \{e \mid (q,e) \in D\} ,$$

calculate both the problem infimum

$$\psi(q) \triangleq \inf_{e \in T(q)} h(q,e)$$

and the optimal solution set

$$T^*(q) \triangleq \{e \in T(q) \mid h(q,e) = \psi(q)\} .$$

Needless to say, the domain of the infimum function ψ is the feasible perturbation set

$$\mathcal{D} \triangleq \{q \mid \text{problem } B(q) \text{ is consistent}\} ,$$

which is clearly not empty.

Due to the known symmetry [5,6,14] of the conjugate transformation on the class of all closed convex functions $g:C$ (as well as the obvious symmetry of the preceding association of perturbation parameters and decision variables), families $\mathcal{A}$ and $\mathcal{B}$ are termed Rockafellar dual families, and problems $A(0)$ and $B(0)$ are termed Rockafellar dual problems. Actually, Rockafellar [13,14,15] formulates $\mathcal{B}$ as a family of maximization problems by placing minus signs in front of the sup and e in the definition of $h:D$. Although that formulation facilitates specializations to (the standard formulations of) linear programming duality and ordinary programming duality, the preceding formulation facilitates a specialization [12] to geometric programming duality.

To (re)orient the reader toward the preceding formulation, we now summarize Rockafellar's most relevant results in terms of that formulation.

Theorem 0 . Suppose the function $g:C$ is convex and closed (and hence has a conjugate transform $h:D$). Then,

(I) either the infimum function φ is finite and convex on its domain $\mathcal{P}$ or $\varphi(p) \equiv -\infty$ for each $p \in (\mathrm{ri}\ \mathcal{P})$,

(II) the infimum function φ is finite on its domain $\wp$ if and only if the dual problem B(0) is consistent; in which case

(i) the dual objective function $h(0,\cdot):T(0)$ is the conjugate transform of $\varphi:\wp$,

(ii) the dual infimum $\psi(0)$ is finite if and only if 0 is in the domain $\wp^c$ of the closure $\varphi^c:\wp^c$ of $\varphi:\wp$, in which event

$$0 = \varphi^c(0) + \psi(0) \text{ and } \partial\varphi^c(0) = T^*(0),$$

(iii) if the primal problem A(0) is also consistent, then 0 is in the domain $\wp^c$ and

$$\varphi^c(0) \le \varphi(0), \text{ with equality only if } \partial\varphi^c(0) = \partial\varphi(0),$$

(iv) the infimum function φ can differ from its closure φ^c only at relative boundary points of $\wp$,

(v) if the dual problem B(0) has a (strongly) feasible solution e° for which $(0,e^\circ) \in (\text{ri } D)$, and if the dual infimum $\psi(0)$ is finite, then

(a) primal problem A(0) is also consistent,

(b) $\varphi^c(0) = \varphi(0)$ and hence $0 = \varphi(0) + \psi(0)$

(c) the primal optimal solution set $S^*(0) \neq \emptyset$.

Conclusion (v) is, of course, the relevant version of "Fenchel's theorem".

It is important to note that conclusions (II) and (i) to Theorem 0 provide a method for constructing, without the use of numerical optimization techniques, the closure $\varphi^c:\wp^c$ of $\varphi:\wp$ (which, by virtue of conclusion (iv) to Theorem 0, is essentially the desired infimum function $\varphi:\wp$). In particular, if the dual feasible solution set T(0) is not empty (which, as indicated by conclusions (I) and (II) to Theorem 0, is the only nontrivial convex case), T(0) can of course be covered with a mesh

$$M \triangleq \{e^1, e^2, \ldots, e^m, \ldots\} \subseteq T(0) \,,$$

which need only contain a finite number m of points e^i when T(0) is bounded. Each mesh M leads to both an approximating function that bounds φ^c from below and an approximating function that bounds φ^c from above. Moreover, it is a consequence of conjugate transform theory that these lower and upper approximations can be made with arbitrary accuracy, simply by choosing the mesh M to be sufficiently dense in T(0).

To obtain the lower approximating function, first note that conclusions (I) and (i) to Theorem 0 imply that $\varphi^c:\mathcal{P}^c$ is the conjugate transform of $h(0,\cdot):T(0)$. It is then clear that φ^c is bounded from below on $\mathcal{P}^c$ by the function φ_ℓ^c whose function values

$$\varphi_\ell^c(p) \triangleq \sup_{i\in\{1,2,\ldots\}} [\langle e^i,p\rangle - h(0,e^i)] \quad \text{for each} \quad p \in \mathcal{P}^c.$$

Being the supremum of affine functions makes φ_ℓ^c convex and, when M is finite, polyhedral.

To obtain the upper approximating function, first note that the (Young-Fenchel) "conjugate inequality" for $\varphi^c:\mathcal{P}^c$ and $h(0,\cdot):T(0)$ asserts that

$$\varphi^c(p) = \langle e^i,p\rangle - h(0,e^i) \quad \text{for each} \quad p \in \partial_e h(0,e^i), \qquad i=1,2,\ldots \quad .$$

It is then clear from the convexity of φ^c that affine interpolations between such function values bound φ^c from above. In fact, φ^c is obviously bounded from above by the function φ_u^c whose function values are simply the infimum of all corresponding interpolated values. Clearly, φ_u^c is convex and, when the number of interpolation values is finite, polyhedral.

For the preceding method to be practical, it is clear that the dual objective function $h(0,\cdot):T(0)$ must be computed explicitly (in terms of simple formulas) -- a computation that may be possible only if certain additional (uninteresting) perturbation parameters are included as components of p . However, the following section shows how to subsequently delete such additional parameters (i.e. keep them zero) by judiciously choosing each mesh M.

3. The main results. The duality between suboptimization and parameter deletion can be crystallized as the following theorem.

Theorem 1. *Suppose that* $g:C$ *is convex and closed, and let the component partitions*

$$(d,p) \triangleq (d,p',p^\circ)$$

and

$$(q,e) \triangleq (q,e',e^\circ)$$

be compatible in the sense that p' *and* e' *have the same dimension (which means of course that* p° *and* e° *also have the same dimension).* *If* $0 \in (\mathrm{ri}\, \mathcal{P}^\circ)$ *where*

$$\mathcal{P}^\circ \triangleq \{p^\circ \mid (d,p',p^\circ) \in C \text{ for at least one } (d,p')\},$$

then a deletion of p° *is the dual of a (sub)optimization over* e° *in the sense that*

(i) *the function* $g^\circ : C^\circ$ *with domain*

$$C^\circ \triangleq \{(d,p') \mid (d,p',0) \in C\}$$

and functional values

$$g^\circ(d,p') \triangleq g(d,p',0)$$

has as its conjugate transform the function $h^\circ : D^\circ$ *with domain*

$$D^\circ \triangleq \{(q,e') \mid (q,e',e^\circ) \in D \text{ for at least one } e^\circ\}$$

and function values

$$h^\circ(q,e') \triangleq \inf_{e^\circ \in T_s(q,e')} h(q,e',e^\circ)$$

where

$$T_s(q,e') \triangleq \{e^\circ \mid (q,e',e^\circ) \in D\}$$

(ii) *the set*

$$T_s^*(q,e') \triangleq \{e^\circ \in T_s(q,e') \mid h(q,e',e^\circ) = h^\circ(q,e')\}$$

is not empty for each $(q,e') \in D^\circ$.

Proof. For a fixed but arbitrary (q,e') consider the programming family whose infimum function $\varphi^{\circ}(\cdot;q,e'):\mathcal{P}^{\circ}$ has function values

$$\varphi^{\circ}(p^{\circ};q,e') \triangleq \inf_{(d,p')\in S^{\circ}(p^{\circ})} [g(d,p',p^{\circ}) - \langle q,d\rangle - \langle e',p'\rangle] ,$$

where

$$S^{\circ}(p^{\circ}) \triangleq \{(d,p') \mid (d,p',p^{\circ}) \in C\} .$$

The function $g(\cdot,\times,\circ) - \langle q,\cdot\rangle - \langle e',\times\rangle:C$ clearly inherits convexity and closedness from $g:C$, so Theorem 0 can be applied to this family and its dual. To compute its dual, simply observe that

$$\sup_{(d,p,p^{\circ})\in C} \{\langle r,d\rangle + \langle r',p'\rangle + \langle e^{\circ},p^{\circ}\rangle - [g(d,p',p^{\circ}) - \langle q,d\rangle - \langle e',p'\rangle]\}$$

is finite if and only if $(r,r',e^{\circ}) \in D - (q,e',0)$, in which case this supremum equals $h(r+q,r'+e',e^{\circ})$. Consequently, the dual family has an infimum function $\psi^{\circ}(\cdot,\times;q,e'):\mathcal{D}^{\circ}(q,e')$ with domain

$$\mathcal{D}^{\circ}(q,e') \triangleq \{(r,r') \mid (r,r',e^{\circ}) \in D - (q,e',0) \quad \text{for at least one } e^{\circ}\}$$

and function values

$$\psi^{\circ}(r,r';q,e') \triangleq \inf_{e^{\circ}\in T^{\circ}(r,r';q,e')} h(r+q,r'+e',e^{\circ}),$$

where

$$T^{\circ}(r,r';q,e') \triangleq \{e^{\circ} \mid (r,r',e^{\circ}) \in D - (q,e',0)\} .$$

Now, the hypothesis that $0 \in (\mathrm{ri}\ \mathcal{P}^{\circ})$ along with conclusions (I) through (i) of Theorem 0 imply that $\varphi^{\circ}(0;q,e')$ is finite if and only if $T^{\circ}(0,0;q,e') \neq \emptyset$. But this means that $\sup_{(d,p')\in C^{\circ}} \{\langle q,d\rangle + \langle e',p'\rangle - g^{\circ}(d,p')\}$ is finite if and only if $(q,e') \in D^{\circ}$; in which case (the dual of) conclusion (v) to Theorem 0 (relative to the infimum function $\varphi^{\circ}:\mathcal{P}^{\circ}$) asserts that

$$\sup_{(d,p')\in C^{\circ}} \{\langle q,d\rangle + \langle e',p'\rangle - g^{\circ}(d,p')\} = h^{\circ}(q,e')$$

and

$$T_s^{*}(q,e) \neq \emptyset ,$$

provided we can show that there exists $(d,p') \in S^{\circ}(0)$ for which $(d,p',0) \in (\text{ri } C)$.

Toward that end, observe that $\mathcal{P}^{\circ} = \mathfrak{J}\, C$ where $\mathfrak{J}$ is the orthogonal projector onto the p°-axes. From the hypothesis that C is convex and from Theorem 6.6 on page 48 of [14] we now infer that $(\text{ri } \mathcal{P}^{\circ}) = \mathfrak{J}(\text{ri } C)$. Consequently, the hypothesis that $0 \in (\text{ri } \mathcal{P}^{\circ})$ implies the existence of a point $(d,p',0) \in (\text{ri } C)$ for which $(d,p') \in S^{\circ}(0)$. q.e.d.

It is important to note that conclusion (i) to Theorem 1 is the key to extending our parametric programming method to <u>a method for constructing</u>, <u>with only a limited use of numerical optimization techniques</u>, <u>the restriction</u> $\varphi^{c}(\cdot,0)$ <u>of</u> $\varphi^{c}:\mathcal{P}^{c}$ <u>to the desired domain</u>

$$\mathcal{P}^{c'} \triangleq \{p' \mid (p',0) \in \mathcal{P}^{c}\} .$$

In particular, observe that the appropriate dual feasible solution set

$$T_{\circ}(0) \triangleq \{e' \mid (0,e') \in D^{\circ}\} = \{e' \mid (e',e^{\circ}) \in T(0) \text{ for at least one } e^{\circ}\} .$$

The latter representation shows that $T_{\circ}(0)$ is not empty if and only if $T(0)$ is not empty; in which case $T_{\circ}(0)$ can be computed explicitly in terms of $T(0)$, even though it is generally impossible to compute $h^{\circ}(0,e')$ explicitly in terms of $h(0,e',e^{\circ})$. Consequently, given that $T(0)$ is not empty and can be computed explicitly (in terms of simple formulas), the latter representation shows that $T_{\circ}(0)$ can also be computed explicitly, and hence can easily be covered with a mesh

$$M^{\circ} \triangleq \{e'^{1}, e'^{2}, \ldots, e'^{m}, \ldots\} \subseteq T_{\circ}(0) .$$

Each mesh M° leads to both an approximating function that bounds $\varphi^{c}(\cdot,0)$ from below and an approximating function that bounds $\varphi^{c}(\cdot,0)$ from above -- approximations that can be made with arbitrary accuracy, simply by choosing the mesh M° to be sufficiently dense in $T_{\circ}(0)$.

Now, the derivation at the end of section 2 (applied to $g^{\circ}: C^{\circ}$ and $h^{\circ}: D^{\circ}$, rather than to $g:C$ and $h:D$) produces the lower approximating function $\varphi_{\ell}^{c\circ}$ whose function values

$$\varphi_{\ell}^{co}(p',0) \triangleq \sup_{i\in\{1,2,\ldots\}} [\langle e'^{i},p'\rangle - h^{o}(0,e'^{i})] \quad \text{for each} \quad p'\in \mathcal{P}^{\infty} .$$

Note that the only nontrivial ingredients in this formula are the numbers $h^{o}(0,e'^{i}), i\in\{1,2,\ldots\}$ -- numbers that can be obtained from (sub)optimations of $h(0,e'^{i},e^{o})$ over $e^{o}\in T_{s}(0,e'^{i})$, hopefully "with only a limited use of numerical optimization techniques" (because of a hopefully small dimension for e^{o}).

The derivation at the end of section 2 also produces the upper approximating function φ_{u}^{co} whose function values are simply the infima of all corresponding interpolated values, computed from the interpolating values

$$\varphi^{c}(p',0) = \langle e'^{i},p'\rangle - h^{o}(0,e'^{i}) \quad \text{for each} \quad p'\in\partial_{e'}h^{o}(0,e'^{i}), \quad i=1,2,\ldots \quad .$$

However, the sets $\partial_{e'}h^{o}(0,e'^{i}), i\in\{1,2,\ldots\}$ in this formula are generally difficult (if not impossible) to compute, due to a lack of simple formulas for $h^{o}(0,\cdot):T_{o}(0)$.

The following theorem provides a remedy for the preceding difficulty.

<u>Theorem</u> 2. <u>If the hypotheses of Theorem</u> 1 <u>are satisfied and if</u> $e'\in T_{o}(0)$, <u>then</u>

(i) <u>there exists at least one vector</u> $p'\in\partial_{e'}h^{o}(0,e')$ <u>when</u> $e'\in(\mathrm{ri}\, T_{o}(0))$,

(ii) $(p',0)\in\partial_{e}h(0,e',e^{o*})$ <u>when</u> $p'\in\partial_{e'}h^{o}(0,e')$ <u>and</u> $e^{o*}\in T_{s}^{*}(0,e')$,

(iii) $\varphi^{c}(p',0) = \langle e',p'\rangle - h(0,e',e^{o*})$ <u>when</u> $(p',0)\in\partial_{e}h(0,e',e^{o*})$.

Proof. Conclusion (i) is an immediate consequence of convexity theory, because $h^{o}(0,\cdot):T_{o}(0)$ is the conjugate transform of $\varphi(\cdot,0)$ and hence is convex and closed.

To prove conclusion (ii), observe that

$$h^{o}(0,e'^{\dagger}) \geq h^{o}(0,e') + \langle p',e'^{\dagger} - e'\rangle \quad \text{for each} \quad e'^{\dagger}\in T_{o}(0).$$

But $h^{o}(0,e') = h(0,e',e^{o*})$, so

$$h^{o}(0,e'^{\dagger}) \geq h(0,e',e^{o*}) + \langle p',e'^{\dagger} - e'\rangle \quad \text{for each} \quad e'^{\dagger}\in T_{o}(0) .$$

Using the definition of $h^{\circ}(0,e'^{\dagger})$, we now infer that

$$h(0,e'^{\dagger},e^{\circ\dagger}) \geq h(0,e',e^{\circ *}) + \langle p', e'^{\dagger} - e' \rangle \quad \text{for each} \quad (e'^{\dagger},e^{\circ\dagger}) \in T(0),$$

which implies that

$$h(0,e'^{\dagger},e^{\circ\dagger}) \geq h(0,e',e^{\circ *}) + \langle (p',0),(e'^{\dagger},e^{\circ\dagger}) - (e',e^{\circ *}) \rangle \quad \text{for each} \quad (e'^{\dagger},e^{\circ\dagger}) \in T(0).$$

Hence $(p',0) \in \partial_e h(0,e',e^{\circ *})$.

Conclusion (iii) is immediate from conclusions (I) through (i) of Theorem 0 and conjugate transform theory. q.e.d.

It is important to note that conclusion (i) to Theorem 2 asserts that the sets $\partial_{e'} h^{\circ}(0,e'^{i})$, $i \in \{1,2,\ldots\}$, in the formula for interpolating values, are not empty when $M^{\circ} \subseteq (\mathrm{ri}\, T_{\circ}(0))$ -- a rather unrestrictive condition on M°. Moreover, conclusion (ii) to Theorem 2 and conclusion (ii) to Theorem 1 together imply that each $p' \in \partial_{e'} h^{\circ}(0,e'^{i})$ can be obtained from any set $\partial_e h(0,e'^{i},e^{\circ *})$ for which $e^{\circ *} \in T_s^{*}(0,e'^{i})$. Finally, conclusion (iii) to Theorem 2 asserts that our formula for interpolating values can actually be expressed in terms of the sets $\partial_e h(0,e'^{i},e^{\circ *})$, $i \in \{1,2,\ldots\}$ -- sets that tend to be (relatively) easy to compute, due to the frequent availability of (relatively) simple formulas for $h(0,\cdot):T(0)$.

4. Decomposition duality. This topic involves a minimax representation [14, part III] of the Rockafellar dual problem $B(0)$ in terms of a (generalized Lagrangian) saddle function $\mathcal{L}$.

To construct $\mathcal{L}$ from the given function $g:C$, let

$$\mathcal{D} \triangleq \{d \mid (d,p) \in C \quad \text{for at least one } p\},$$

and suppose that

$$P(d) \triangleq \{p \mid (d,p) \in C\} \quad \text{for} \quad d \in \mathcal{D}.$$

Also, suppose that

$$E(d) \triangleq \{e \mid \inf_{p \in P(d)} [g(d,p) - \langle e,p \rangle] \text{ is finite}\} \quad \text{for} \quad d \in \mathcal{D},$$

and let

$$\mathcal{E} \triangleq \bigcap_{d \in \mathcal{D}} E(d) \ .$$

Then, $\mathcal{L}$ has domain $\mathcal{D} \times \mathcal{E}$ and function values

$$\mathcal{L}(d;e) \triangleq \inf_{p \in P(d)} [g(d,p) - \langle e,p \rangle] \quad \text{for} \quad d \in \mathcal{D} \text{ and } e \in \mathcal{E} \ .$$

To obtain the minimax representation of problem $B(0)$, note that

$$\begin{aligned}
\inf_{d \in \mathcal{D}} \mathcal{L}(d;e) &= \inf_{d \in \mathcal{D}} \ \inf_{p \in P(d)} [g(d,p) - \langle 0,d \rangle - \langle e,p \rangle] \\
&= \inf_{(d,p) \in C} [g(d,p) - \langle 0,d \rangle - \langle e,p \rangle] \\
&= - \sup_{(d,p) \in C} [\langle 0,d \rangle + \langle e,p \rangle - g(d,p)] .
\end{aligned}$$

These equations clearly imply the objective function identity

$$\inf_{d \in \mathcal{D}} \mathcal{L}(d;e) + h(0,e) \equiv 0 \quad \text{for} \quad e \in T(0);$$

from which we readily infer the minimax (actually maximin) representation

$$\sup_{e \in \mathcal{E}} \ \inf_{d \in \mathcal{D}} \mathcal{L}(d;e) + \inf_{e \in T(0)} h(0,e) = 0 \ .$$

Now, under the hypotheses of Theorem 1, a saddle function $\mathcal{L}^{\circ} : \mathcal{D}^{\circ} \times \mathcal{E}^{\circ}$ can be constructed from the given function $g^{\circ} : C^{\circ}$ in the same way that $\mathcal{L} : \mathcal{D} \times \mathcal{E}$ has been constructed from $g:C$. Moreover, conclusion (i) to Theorem 1 then implies the minimax representation

$$\sup_{e' \in \mathcal{E}^{\circ}} \ \inf_{d \in \mathcal{D}^{\circ}} \mathcal{L}^{\circ}(d;e') + \inf_{e' \in T_{o}(0)} \ \inf_{e^{\circ} \in T_{s}(0,e')} h(0,e',e^{\circ}) = 0 \ .$$

This equation expresses the duality between a generalized Dantzig-Wolfe decomposition and a generalized Benders decomposition. The outer supremum and the outer infimum are the respective master problems; and the inner infima decompose into respective infima sums when there is sufficient "separability" and "sparsity" -- requirements on $g:C$ that can best be understood in the context of subsections 3.1.6 and 3.3.6 of [9], viewed from the vantage point of [12].

References

1. Benders, J. F., "Partitioning Procedures for Solving Mixed Variables Programming Problems", Num. Math., 4(1962).

2. Dantzig, G. B., and P. Wolfe, "The Decomposition Algorithm for Linear Programming", ORSA, 8, January, February, October (1960).

3. Duffin, R. J., and E. L. Peterson, and C. Zener, Geometric Programming - Theory and Applications, John Wiley, New York (1967).

4. Falk, J. E., "Lagrange Multipliers and Nonlinear Programming", Jour. Math. Anal. Appl., 19(1967).

5. Fenchel, W., "On Conjugate Convex Functions", Canadian Jour. Math., 1(1949).

6. __________, "Convex Cones, Sets and Functions", Mathematics Department mimeographed lecture notes, Princeton University (1951).

7. Peterson, E. L., "Symmetric Duality for Generalized Unconstrained Geometric Programming", SIAM J. Appl. Math., 19(1970).

8. ______________, "Mathematical Foundations of Geometric Programming", Appendix to "Geometric Programming and Some of Its Extensions", Optimization and Design (M. Avriel, M. J. Rijckaert, and D. J. Wilde, eds.), Prentice-Hall, Englewood Cliffs, N. J. (1972).

9. ______________, "Geometric Programming", SIAM Rev., 18(1976)

10. ______________, "The Conical Duality and Complementarity of Price and Quantity for Multicommodity Spatial and Temporal Network Allocation Problems", Ctr. for Math. Stud. in Econ. and Man. Sci. Discussion Paper, Northwestern University, April (1976).

11. ______________, "Ordinary Duality vis-a-vis Geometric Duality", Ctr. for Math. Stud. in Econ. and Man. Sci. Discussion Paper, Northwestern University, December (1976).

12. ______________, "Geometric Duality via Rockafellar Duality", Ctr. for Math. Stud. in Econ. and Man. Sci. Discussion Paper, Northwestern University, January (1977).

13. Rockafellar, R. T., "Duality in Nonlinear Programming", Amer. Math. Soc. Lects. Appl. Math., 11 (G. B. Dantzig and A. F. Veinott, eds.), Amer. Math. Soc., Providence, R.I. (1968).

14. ________________, Convex Analysis, Princeton University Press, Princeton, New Jersey (1970).

15. ________________, "Conjugate Duality and Optimization", CBMS Regional Conf. Ser. in Appl. Math., No. 16, SIAM, Philadelphia, Penna.,(1974).

16. Wolfe, P., "A Duality Theorem for Nonlinear Programming", Quart. Appl. Math., 19(1961).