DUALITY RELATIONSHIPS FOR ENTROPY-LIKE MINIMIZATION PROBLEMS*

J. M. BORWEIN† AND A. S. LEWIS‡

Abstract. This paper considers the minimization of a convex integral functional over the positive cone of an L_p space, subject to a finite number of linear equality constraints. Such problems arise in spectral estimation, where the objective function is often entropy-like, and in constrained approximation. The Lagrangian dual problem is finite-dimensional and unconstrained. Under a quasi-interior constraint qualification, the primal and dual values are equal, with dual attainment. Examples show the primal value may not be attained. Conditions are given that ensure that the primal optimal solution can be calculated directly from a dual optimum. These conditions are satisfied in many examples.

Key words. convex programming, duality, spectral estimation, entropy, moment problem

AMS(MOS) subject classifications. primary 90C25, 49A27; secondary 41A46, 42A70

1. Introduction. A number of authors have recently considered dual approaches to the solution of optimization problems of the form

$$\text{(EP)} \qquad \begin{aligned} \inf \quad & \int_T \phi(x(t))\, dt \\ \text{subject to} \quad & Ax = b, \\ & x \geqq 0, \qquad x \in L_p(T). \end{aligned}$$

Here, (T, dt) is a σ-finite measure space, $1 \leqq p \leqq \infty$, $\phi : \mathbb{R} \to (-\infty, \infty]$ is convex, and $A : L_p \to \mathbb{R}^n$ is continuous. Such problems arise in a number of areas. When (for $p < \infty$), $\phi(x) = (1/p)x^p$ we obtain the constrained L_p approximation problem (see [Micchelli, Smith, Swetits, and Ward, 1985]). These problems appear in the theory of constrained interpolation [Irvine, Marin, and Smith, 1986], and also as a result of using the L_p norm as the objective function in spectral estimation problems (see, for example, [Goodrich and Steinhardt, 1986] and [Ben-Tal, Borwein, and Teboulle, 1988a]). A number of different objective functions can be used. Typically, they are "entropic" in form, for instance $\phi(x) = -\log x$ [Burg, 1975] and $\phi(x) = x \log x$ [Johnson and Shore, 1984]. A survey of common objective functions may be found in [Ben-Tal, Borwein, and Teboulle, 1988b]. For simplicity of exposition we will only consider the autonomous case, where the function ϕ does not depend explicitly on the variable t. The non-autonomous case is a simple extension.

Let us denote the function $\phi + \delta(\cdot \mid \mathbb{R}_+)$ by ϕ_+ (where δ denotes an indicator function [Rockafellar, 1970]). Assuming ϕ_+ is a *normal convex integrand* in the sense of [Rockafellar, 1968], we can write the Lagrangian dual problem as

$$\text{(DEP)} \qquad \begin{aligned} \sup \quad & b^T \lambda - \int_T (\phi_+)^*((A^T \lambda)(t))\, dt \\ \text{subject to} \quad & \lambda \in \mathbb{R}^n, \end{aligned}$$

---

† Department of Mathematics, Statistics, and Computing Science, Dalhousie University, Halifax, Nova Scotia, Canada B3H 3J5.

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

where $*$ denotes convex conjugation and $A^T : \mathbb{R}^n \to L_q(T)$ $((1/p)+(1/q)=1)$ is the adjoint map. It is known (see, for example, [Ben-Tal, Borwein, and Teboulle, 1988b]) that under a suitable constraint qualification the values of (EP) and (DEP) are equal, with dual attainment.

Suppose now that ϕ_+ is closed and essentially strictly convex, so, by [Rockafellar, Thm. 26.3, 1970], $(\phi_+)^*$ is essentially smooth. (We defer precise definitions to a later section.) Suppose that $\bar{\lambda}$ is optimal for (DEP). *Assuming that* $(A^T\bar{\lambda})(t) \in$ int $(\text{dom} (\phi_+)^*)$, *and that we can differentiate through the integral*, we then obtain

(1.1) $b - A(((\phi_+)^*)'((A^T\bar{\lambda})(\cdot))) = 0.$

If we now set $\bar{x}(t) = ((\phi_+)^*)'((A^T\bar{\lambda})(t))$, $t \in T$, then by [Rockafellar, Thm. 23.5, 1970],

$$(A^T\bar{\lambda})(t) \in \partial\phi_+(\bar{x}(t)), \quad \text{a.e.},$$

so $\phi_+(\bar{x}(t)) + (\phi_+)^*((A^T\bar{\lambda})(t)) = \bar{x}(t)(A^T\bar{\lambda})(t)$, almost everywhere and, since $(\phi_+)^*$ is increasing, $\bar{x} \geqq 0$. Thus, by (1.1) $\bar{x}(t)$ is feasible for (EP) and, integrating over T, has the same objective value as the dual and hence is optimal (by weak duality).

The question of when the above assumptions are justified has, to the authors' knowledge, never been addressed in the published literature on the subject. Since the derivation of primal solutions is of paramount importance, this question is clearly extremely significant. The aim of this paper is therefore to give these matters a rigorous treatment. We begin by deriving the basic duality result from the theory developed in [Borwein and Lewis, 1988]. We then give some examples to show how the above assumptions can fail. This will motivate our rigorous treatment of the derivation of primal solutions.

2. Duality. We will begin by deriving the fundamental duality result.

DEFINITION 2.1 [Borwein and Lewis, 1988]. Let (X, τ) be a topological vector space, with convex $C \subseteq X$. The *quasi-relative interior* of C ($\tau - \text{qri } C$) is the set of those $x \in C$ for which cl $\mathbb{P}(C - x)$ is a subspace.

(Here, cl $\mathbb{P}B$ denotes the closed cone generated by B.) Note that if X is normed, the weak and norm quasi-relative interiors coincide.

Example 2.2 [Borwein and Lewis, 1988]. Let T be a σ-finite measure space, $1 \leqq p \leqq \infty$, $(1/p)+(1/q)=1$ and let $X = L_p(T)$, $Y = L_q(T)$. Then

$$\sigma(X, Y) - \text{qri } X_+ = \{x \,|\, x(t) > 0 \text{ a.e.}\}.$$

LEMMA 2.3. *Let* X *be a topological vector space with convex* C_1, $C_2 \subseteq X$ *satisfying* cl $C_1 = $ cl C_2. *For any* $x \in C_1 \cap C_2$, $x \in \text{qri } C_1$ *if and only if* $x \in \text{qri } C_2$.

Proof. The result follows immediately from the fact that cl $\mathbb{P}(C_1 - x) = $ cl $\mathbb{P}(C_2 - x)$.   □

The following duality result may be found in [Borwein and Lewis, Cor. 4.10, 1988].

THEOREM 2.4. *Let* X *be locally convex,* $f: X \to (-\infty, \infty]$ *convex,* $A: X \to \mathbb{R}^n$ *continuous and linear,* $b \in \mathbb{R}^n$, *and* $P \subset \mathbb{R}^n$ *a polyhedral cone. Consider the following dual pair of problems:*

$$\inf \quad f(x)$$

(CM)   *subject to* $Ax \in b + P,$

$$x \in C,$$

$$\max \quad b^T\lambda - (f + \delta(\cdot\,|\,C))^*(A^T\lambda)$$

(DCM)

subject to $\lambda \in P^+.$

If there exists a feasible $\hat{x} \in \text{qri} ((\text{dom} f) \cap C)$ *for* (CM), *then the values of* (CM) *and* (DCM) *are equal (with attainment in* (DCM)).

To use the above result in our case we need to compute the conjugate of certain integral functionals. We will need the following result concerning normal convex integrands.

THEOREM 2.5. *Let T be a finite measure space, $1 \leq p \leq \infty$, and suppose $\psi : \mathbb{R} \to (-\infty, \infty]$ is closed, convex, and proper. Define $I_\psi : L_p(T) \to (-\infty, \infty]$ by $I_\psi(x) = \int_T \psi(x(t)) \, dt$. Then $(I_\psi)^* : L_q(T) \to (-\infty, \infty]$ is given by $(I_\psi)^* = \int_T \psi^*(y(t)) \, dt$.*

Proof. For the proof see the corollary to Theorem 2 in [Rockafellar, 1968]. □
We can now derive the required duality result.

COROLLARY 2.6. *Suppose T is a finite measure space, $1 \leq p \leq \infty$, $\phi : \mathbb{R} \to (-\infty, \infty]$ with ϕ_+ closed and convex, $a_i \in L_q(T)$ for $i = 1, \cdots, n$, and $b \in \mathbb{R}^n$. Consider the following dual pair of problems:*

$$\inf \quad \int_T \phi(x(t)) \, dt$$

(EP$_p$) $\qquad$ *subject to* $\quad \int_T x(t) a_i(t) \, dt = b_i, \qquad i = 1, \cdots, n,$

$$x \geq 0, \qquad x \in L_p(T),$$

$$\max \quad b^T \lambda - \int_T (\phi_+)^* \left[\sum_{i=1}^n \lambda_i a_i(t) \right] dt$$

(DEP$_p$)
$\qquad$ *subject to* $\quad \lambda \in \mathbb{R}^n.$

Suppose further that the following constraint qualification holds:
(CQ) $\qquad$ *There exists a feasible $\hat{x}$ for* (EP$_p$) *with $\hat{x} \in \sigma(L_p, L_q) - \text{qri} (\text{dom} I_\phi \cap (L_p)_+)$.*
Then the values of (EP$_p$) *and* (DEP$_p$) *are equal (with dual attainment).*

Suppose furthermore that $(0, \infty) \subset \text{dom} \phi$. Then (CQ) *is equivalent to requiring the existence of a feasible $\hat{x} \in \text{dom} I_\phi$ with $\hat{x}(t) > 0$ almost everywhere. In particular, if $\text{ess sup} \, x < +\infty$ and $\text{ess inf} \, x > 0$ then $x \in \text{dom} I_\phi$.*

Proof. In Theorem 2.4, set $X = L_p(T)$ with the $\sigma(L_p, L_q)$ topology, $f = I_\phi$, $C = (L_p)_+$, A defined by $(Ax)_i = \langle x, a_i \rangle = \int_T x(t) a_i(t) \, dt$ for $i = 1, \cdots, n$, and $P = \{0\}$. Then $f + \delta(\cdot \mid C) = I_{\phi_+}$, so by Theorem 2.5, $(f + \delta(\cdot \mid C))^* = I_{\phi_+^*}$. Also, $P^+ = \mathbb{R}^n$, and $A^T \lambda = \sum_{i=1}^n \lambda_i a_i$, so the duality result follows.

If $(0, \infty) \subset \text{dom} \phi$ then the set $K := \{x \in L_\infty \mid \text{ess inf} \, x > 0\}$ is contained in $\text{dom} I_\phi \cap (L_p)_+$. To see this, observe that if $x \in L_\infty$ with $\text{ess inf} \, x > 0$, then there exist $M, \varepsilon > 0$ with $\varepsilon \leq x(t) \leq M$, almost everywhere. Since $(0, \infty) \subset \text{dom} \phi$, ϕ is continuous on $(0, \infty)$, by [Rockafellar, Thm. 10.1, 1970], and so the set $\phi[\varepsilon, M] \subset \mathbb{R}$ is compact. Thus, $\phi(x(\cdot))$ is bounded on T, and so $x \in \text{dom} I_\phi$.

Now K is dense in $(L_p)_+$ (the proof essentially follows [Rudin, Thm. 3.13, 1966]). Thus, we have

$$\text{cl} \, K \subset \text{cl} (\text{dom} I_\phi \cap (L_p)_+)$$

$$\subset \text{cl} ((L_p)_+)$$

$$= (L_p)_+$$

$$= \text{cl} \, K,$$

so $\text{cl} (\text{dom} I_\phi \cap (L_p)_+) = (L_p)_+$. Applying Lemma 2.3 it follows that if $\hat{x} \in (\text{dom} I_\phi) \cap \text{qri} (L_p)_+$ then $\hat{x} \in \text{qri} ((\text{dom} I_\phi) \cap (L_p)_+)$, and the result follows by Example 2.2. □

In the case where the measure space is nonatomic and totally finite, we can handle nonconvex objectives (as observed by the referee). Suppose that ϕ_+ is lower semicontinuous (not necessarily convex), but that ϕ_+^* and ϕ_+^{**} are proper. Then, with A as before, if we define

$$h(b) := \inf \left\{ \int_T \phi_+(x(t)) \, dt \,|\, Ax = b, \, x \in L_1 \right\},$$

the domain of h is convex, and h is a convex function. This is a consequence of the Liapunov convexity theorem [Holmes, p. 108, 1975]. In fact if we define the value function of the regularized problem,

$$\bar{h}(b) := \inf \left\{ \int_T \phi_+^{**}(x(t)) \, dt \,|\, Ax = b, \, x \in L_1 \right\},$$

then whenever the constraint qualification is satisfied at b, in other words $b \in A(\text{qri}\,(\text{dom}\, I_{\phi_+}))$, we have $h(b) = \bar{h}(b)$. This follows from [Rockafellar, Thm. 3H, 1976], and the fact that the constraint qualification forces h to be lower semicontinuous at b. In conclusion, our results apply equally well to ϕ as to ϕ^{**}. See also [Ioffe and Tihomorov, 1968] for similar ideas.

In some common cases the constraint qualification is particularly easy to check. The next result gives conditions under which the constraint qualification is no more arduous than primal consistency. Note that in most applications the a_i's are actually continuous (at least piecewise).

DEFINITION 2.7 [Borwein and Lewis, 1988]. A set of measurable functions $a_i : T \to \mathbb{R}$, $i = 1, \cdots, n$, on a measure space T are called *pseudo-Haar* if they are linearly independent on every nonnull subset of T.

PROPOSITION 2.8. *Consider the measure space* $[\alpha, \beta] \subset \mathbb{R}$, *with Lebesgue measure, and suppose* $a_i : [\alpha, \beta] \to \mathbb{R}$ *are analytic and linearly independent. Then they are pseudo-Haar on* $[\alpha, \beta]$.

(A real function f defined on an open interval in $\mathbb{R}$ is called *analytic* if it is represented locally by an absolutely convergent power series at any point in the interval, cf. [Bochner and Martin, 1964].)

Proof. We will show that if a real function f is analytic on an open interval $T \supset [\alpha, \beta]$, and is not identically zero on $[\alpha, \beta]$, then it has at most finitely many zeros on $[\alpha, \beta]$. The result then follows immediately.

Suppose f has infinitely many zeros on $[\alpha, \beta]$. By compactness, there exists a sequence $x_i \to x_0 \in [\alpha, \beta]$ as $i \to \infty$, with $f(x_i) = 0$, $x_i \neq x_0$, for $i = 1, 2, \cdots$. Since f is analytic, for some $\delta > 0$ there is an absolutely convergent power series $f(x) = \sum_{n=0}^{\infty} c_n(x - x_0)^n$, for $x \in N_\delta(x_0)$. Suppose the c_n's are not all 0, and let c_m be the first nonzero coefficient. Then

$$f(x) = (x - x_0)^m \left\{ c_m + (x - x_0) \sum_{n=m+1}^{\infty} c_n(x - x_0)^{n-m-1} \right\},$$

for $x \in N_\delta(x_0)$. Since the series $\sum_{n=m+1}^{\infty} c_n(x - x_0)^{n-m-1}$ converges absolutely on $N_\delta(x_0)$, by a standard power series result it converges uniformly on $N_{\delta/2}(x_0)$, so in particular it is continuous at x_0. We then have

$$0 = (x_i - x_0)^{-m} f(x_i)$$

$$= c_m + (x_i - x_0) \sum_{n=m+1}^{\infty} c_n(x_i - x_0)^{n-m-1},$$

and letting $i \to \infty$ implies $c_m = 0$, which is a contradiction.

Thus, $c_n = 0$ for all n, so $f(x) = 0$ on $N_\delta(x_0)$. But 0 is an analytic function on the domain T which agrees with f on $N_\delta(x_0)$, so by [Bochner and Martin, Thm. 4, Chap. II, 1964], $f \equiv 0$ on T. This is a contradiction.   □

The following result shows essentially that if the constraint functions a_i are pseudo-Haar and (EP$_p$) is consistent for any $1 \leq p \leq \infty$ then there is in fact a feasible $\hat{x}$ in the norm interior of $(L_\infty)_+$. In the case when $(0, \infty) \subset \text{dom } \phi$, $\hat{x}$ clearly satisfies the constraint qualification, so the result shows that the constraint qualification is always satisfied for a consistent primal (providing 0 is not the only primal feasible solution).

THEOREM 2.9. *Suppose (T, μ) is a finite measure space and $0 \leq x \in L_p(T)$ is nonzero $(1 \leq p \leq \infty)$. Suppose further that $a_i \in L_q(T)$, $i = 1, \cdots, n$, are pseudo-Haar. Then there exists a $y \in L_\infty(T)$ and $\varepsilon > 0$ with $y(t) \geq \varepsilon$ almost everywhere and $\langle x, a_i \rangle = \langle y, a_i \rangle$, each i.*

Proof. Since $0 \neq x \geq 0$, there exists $T_1 \subset T$ with $\mu(T_1) > 0$ and a $\delta > 0$ such that $x(t) \geq \delta$ almost everywhere on T_1. We claim that

$$(2.10) \qquad \left\{ \left(\int_{T_1} u a_i \, d\mu \right)_{i=1}^n \,\middle|\, u \in L_\infty(T_1) \right\} = \mathbb{R}^n.$$

Suppose not. Since the left-hand side is clearly a subspace, there exists $\lambda \neq 0$ such that $\sum_{i=1}^n \lambda_i \int_{T_1} u a_i \, d\mu = 0$, for all $u \in L_\infty(T_1)$. This implies $\sum_{i=1}^n \lambda_i a_i(t) = 0$ almost everywhere on T_1, which contradicts the fact that the a_i's are pseudo-Haar. Thus, (2.10) holds.

Now define

$$C = \left\{ \left(\int_{T_1} u a_i \, d\mu \right)_{i=1}^n \,\middle|\, u \in L_\infty(T_1), \|u\|_\infty < \delta/2 \right\}.$$

Since, by (2.10), $\mathbb{P}C = \mathbb{R}^n$, it follows that $0 \in \text{int } C$ [Rockafellar, Cor. 6.4.1, 1970].

Let us now define a sequence of functions $(x_m) \subset L_\infty(T)$ by

$$x_m(t) = \begin{cases} m, & \text{if } x(t) > m, \\ x(t), & \text{if } \dfrac{1}{m} \leq x(t) \leq m, \\ \dfrac{1}{m}, & \text{if } x(t) < \dfrac{1}{m}. \end{cases}$$

For $p < \infty$ we have

$$\|x_m - x\|_p^p \leq \int_{\{t \mid x(t) > m\}} x(t)^p \, d\mu + \int_{\{t \mid x(t) < 1/m\}} \left(\frac{1}{m}\right)^p d\mu \to 0, \quad \text{as } m \to \infty,$$

and in the case $p = \infty$, clearly $\|x_m - x\|_\infty \leq 1/m$, for m large. Thus, $(\langle x_m - x, a_i \rangle)_{i=1}^n \to 0$ as $m \to \infty$, and so, since $0 \in \text{int } C$, for $m > 1/\delta$ sufficiently large, $(\langle x_m - x, a_i \rangle)_{i=1}^n \in C$. It then follows that $\langle x_m - x, a_i \rangle = \langle v, a_i \rangle$, each $i = 1, \cdots, n$, for some $v \in L_\infty(T)$ with $\|v\|_\infty < \delta/2$ and $v(t) = 0$ almost everywhere on T_1^c.

Finally, set $y = x_m - v$. Then $y \in L_\infty(T)$, and almost everywhere on T_1^c, $y(t) = x_m(t) \geq 1/m$. On T_1 we know $x(t) \geq \delta > 1/m$ almost everywhere, so $x_m(t) \geq \min\{x(t), m\} \geq \delta$, almost everywhere. Since $v(t) \leq \delta/2$ almost everywhere it follows that $y(t) \geq \delta/2$ almost everywhere on T_1. The result now follows.   □

For a given set of constraint functions a_i, $i = 1, \cdots, n$, the question of for what b's the problem (EP$_p$) is consistent is known as the *extendibility problem*. Simple conditions are known, for example, when $T = [0, 1]$ with Lebesgue measure and $a_i(t) = t^{i-1}$ (one of the "classical" moment problems), and for the case when $T = [-\pi, \pi]$ with Lebesgue measure and $\{a_1(\theta), \cdots, a_{2k+1}(\theta)\} = \{1, \cos \theta, \sin \theta, \cdots, \cos k\theta, \sin k\theta\}$ (the trigonometric moment problem). See, for example, [Karlin and Studden, 1966] and [Ben-Tal, Borwein, and Teboulle, 1988b] for more information.

PROPOSITION 2.11. *If ϕ_+ is strictly convex on* dom ϕ_+, *then any optimal solution to* (EP_p) *is unique.*

Proof. If ϕ_+ is strictly convex on dom ϕ_+ then so is I_{ϕ_+} on dom I_{ϕ_+}. To see this, suppose $x_1, x_2 \in$ dom I_{ϕ_+}, with $x_1 \neq x_2$, and $0 < \gamma < 1$. By strict convexity of ϕ_+, $\gamma\phi_+(x_1(t)) + (1 - \gamma)\phi_+(x_2(t)) \leqq \phi_+(\gamma x_1(t) + (1 - \gamma)x_2(t))$, with strict inequality on a nonnull set. It follows that $\gamma I_{\phi_+}(x_1) + (1 - \gamma)I_{\phi_+}(x_2) < I_{\phi_+}(\gamma x_1 + (1 - \gamma)x_2)$. The result now follows.   □

3. Examples. In the course of the derivation of a primal solution described in § 1, the crucial step was differentiating through the integral in the dual objective function to obtain (1.1). The following example shows clearly the sort of difficulties that can occur.

For $1 \leqq p < 2$, consider the convex integral functional $I : L_p[0, 1] \to (-\infty, \infty]$ defined by

$$I(x) = \begin{cases} -\displaystyle\int_0^1 \log x(t)\, dt, & x(t) > 0 \text{ a.e.,} \\ +\infty, & \text{otherwise.} \end{cases}$$

By [Rockafellar, Thm. 22, 1974], $y \in \partial I(x)$ if and only if $y(t) \in \partial(-\log)(x(t)) = \{-1/x(t)\}$, almost everywhere, for $x(t) > 0$ almost everywhere. With a slight abuse of notation, let us denote the function that is identically equal to one by 1. Then $\partial I(1) = \{-1\}$, so I has a unique subgradient at 1.

Now define $h \in L_p[0, 1]$ by $h(t) = 1/\sqrt{t}$. For any $\varepsilon > 0$, $1 - \varepsilon/\sqrt{t} \leqq 0$ on a set of nonzero measure, so $I(1 - (\varepsilon/\sqrt{t})) = +\infty$. Thus, the directional derivative $I'(1; -h) = +\infty$.

On the other hand, we have

$$I'(1; h) = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon}\left\{-\int_0^1 \log\left(1 + \frac{\varepsilon}{\sqrt{t}}\right) dt\right\}.$$

It is straightforward to check that as $\varepsilon \downarrow 0$, $(1/\varepsilon)\log(1 + (\varepsilon/\sqrt{t})) \uparrow (1/\sqrt{t})$, so by the monotone convergence theorem, $I'(1; h) = -\int_0^1 (1/\sqrt{t})\, dt = -2$.

Thus, I has no Gateaux derivative at 1 although it has a unique subgradient there. The choice of space is clearly important here. In the case $p = \infty$, we have $1 \in \|\cdot\|_\infty - \text{int (dom } I)$, and it is easy to see that I is $\|\cdot\|_\infty$-continuous at 1. Thus, in this case the Gateaux derivative does exist: $\nabla I(1) = -1$ (see, for example, [Holmes, 1975]).

Simple examples show that even when the constraint qualification is satisfied we cannot necessarily expect primal attainment in (EP_p).

Example 3.1. Consider the following semi-infinite linear program:

$$\inf \qquad \int_0^1 x(t)\, dt$$

$$\text{subject to} \quad \int_0^1 tx(t)\, dt = 1, \qquad x \geqq 0, \quad x \in L_p[0, 1].$$

In this problem $\phi(x) = x$, so $(\phi_+)^*(y) = \sup_{x \geqq 0}\{xy - x\} = \delta(y|(-\infty, 1])$. Thus, the dual problem is

$$\max \qquad \lambda - \int_0^1 \delta(\lambda t|(-\infty, 1])\, dt$$

$$\text{subject to} \quad \lambda \in \mathbb{R}.$$

The optimal dual solution is clearly $\bar{\lambda} = 1$, giving value 1. The constraint qualification is clearly satisfied by $\hat{x}(t) = 2$, so the primal value is also 1, by Corollary 2.6. However,

since $x(t) > tx(t)$ almost everywhere, $\int_0^1 x(t)\, dt > \int_0^1 tx(t)\, dt = 1$ for any feasible x, so the optimal value is not attained (for any $1 \leq p \leq \infty$).

The above behavior can be observed even when the function ϕ is essentially strictly convex.

Example 3.2. Consider the following problem:

$$\text{inf} \qquad \int_0^{2\pi} \left(\frac{1}{x(t)} \right) dt$$

$$\text{subject to} \qquad \int_0^{2\pi} x(t) \sin t\, dt = 1, \qquad x \geq 0, \quad x \in L_p[0, 2\pi].$$

In this case $\phi(x) = 1/x$, $x > 0$, and ∞ otherwise. Thus, $(\phi_+)^*(y) = \sup_{x>0} \{xy - (1/x)\}$, so by differentiation,

$$(\phi_+)^*(y) = \begin{cases} -2\sqrt{-y}, & y \leq 0, \\ \infty, & \text{otherwise.} \end{cases}$$

The dual problem therefore becomes

$$\text{max} \qquad \lambda - \int_0^{2\pi} (-2\sqrt{-\lambda \sin t})\, dt$$

$$\text{subject to} \quad -\lambda \sin t \geq 0 \text{ a.e. on } [0, 2\pi], \qquad \lambda \in \mathbb{R}.$$

The only dual feasible solution is $\bar{\lambda} = 0$, so the dual value is 0. The constraint qualification is satisfied by

$$\hat{x}(t) = \begin{cases} 1, & 0 \leq t < \pi \\ \frac{1}{2}, & \pi \leq t \leq 2\pi, \end{cases}$$

since

$$\int_0^{2\pi} \hat{x}(t) \sin t\, dt = [-\cos \theta]_0^\pi + \frac{1}{2}[-\cos \theta]_\pi^{2\pi} = 1.$$

Thus, the primal value is also 0, by Corollary 2.6, or direct computation. However, this value clearly is not attained.

In both of the above examples the function $(\phi_+)^*$ is differentiable on the interior of its domain, but the manipulation described in the introduction is manifestly invalid since there exists no primal optimal solution. Our last example shows how attainment in the primal problem can depend on the choice of space for the primal variables.

Example 3.3. Consider the following problem:

$$\text{inf} \qquad \int_0^1 \left(\frac{1}{x(t)} \right) dt$$

$$\text{subject to} \int_0^1 tx(t)\, dt = 1, \qquad x \geq 0, \quad x \in L_p[0, 1].$$

In a similar fashion to Example 3.2, the dual problem is

$$\text{max} \qquad \lambda - \int_0^1 (-2\sqrt{-\lambda t})\, dt$$

$$\text{subject to} \quad -\lambda t \geq 0, \text{ a.e. on } [0, 1], \qquad \lambda \in \mathbb{R}.$$

This is equivalent to max $\{\lambda + \frac{4}{3}\sqrt{-\lambda} \mid \lambda \leq 0\}$. By differentiation, the optimum occurs at $\bar{\lambda} = -\frac{4}{9}$, giving a dual value of $\frac{4}{9}$.

The constraint qualification is again satisfied by, for instance, $\hat{x}(t) = 2$, almost everywhere on $[0, 1]$. The primal value is therefore also $\frac{4}{9}$. The manipulation described in the introduction (differentiating the dual objective function) may now be performed, at least formally, to obtain $\bar{x}(t) = 3/2\sqrt{t}$. It may be checked that $\bar{x}$ satisfies the primal constraint and has value $\frac{4}{9}$. Thus, it is optimal, at least for $1 \leq p < 2$, and by Proposition 2.11 it is the unique primal optimal solution. However, $\bar{x} \notin L_p[0, 1]$ for $2 \leq p \leq \infty$, so for these values of p there exists no primal optimal solution (otherwise we would obtain a contradiction to Proposition 2.11). Note finally that the question of primal attainment for this problem is not affected by the addition of the extra constraint $\int_0^1 x(t)\, dt = 3$, since $\bar{x}$ remains feasible. This provides a counterexample whose constraints are in the form of a standard moment problem.

4. Primal solutions. We shall now give a rigorous treatment of conditions ensuring the existence of a primal optimal solution. We begin with some simple results aimed at identifying the domain of the functional $I_{\phi_+^*}$.

PROPOSITION 4.1. *The function* $(\phi_+)^* : \mathbb{R} \to (-\infty, \infty]$ *is monotone increasing.*

Proof. The result follows immediately from the definition:

$$(\phi_+)^*(y) = \sup_{x \geq 0} \{yx - \phi(x)\}. \qquad \square$$

LEMMA 4.2. *Suppose* $\phi : \mathbb{R} \to (-\infty, \infty]$ *with* ϕ_+ *closed, convex, and proper. Define* $d := \lim_{x \to \infty} (\phi(x)/x)$ *(this limit exists). Then* cl $(\mathrm{dom}\, (\phi_+)^*) = (-\infty, d]$.

Proof. Define $\psi(x) = \phi_+(x) - y(x)$. Then by [Rockafellar, Cor. 13.3.4, 1970], $y \in$ cl $(\mathrm{dom}\, (\phi_+)^*)$ if and only if $(\psi 0^+)(z) \geq 0$, for all z. By [Rockafellar, Thm. 8.5, 1970],

$$(\psi 0^+)(z) = \lim_{\lambda \to \infty} \frac{\psi(v + \lambda z) - \psi(v)}{\lambda}$$

$$= \lim_{\lambda \to \infty} \frac{\phi_+(v + \lambda z) - \phi_+(v) - y\lambda z}{\lambda}$$

$$= \lim_{\lambda \to \infty} \left(\frac{\phi_+(v + \lambda z)}{v + \lambda z} \frac{v + \lambda z}{\lambda} \right) - yz$$

$$= (d - y)z,$$

for any $v \in \mathrm{dom}\, \phi_+$ and for $z > 0$. For $z < 0$, $(\psi 0^+)(z) = +\infty$, and $(\psi 0^+)(0) = 0$. Thus, $(\psi 0^+)(z) \geq 0$, for all z, if and only if $y \leq d$. The result now follows by Proposition 4.1. $\square$

LEMMA 4.3. *Suppose* ϕ *is as in Lemma 4.2. The functional* $I_{(\phi_+)^*} : L_\infty(T) \to (-\infty, \infty]$ *is* $\|\cdot\|_\infty$*-continuous at any* $y \in L_\infty(T)$ *for which* ess sup $y < d$.

Proof. There exist $M > 0$ and $\varepsilon > 0$ such that $-M \leq y(t) \leq d - \varepsilon$, almost everywhere on T. By Lemma 4.2, int $(\mathrm{dom}\, (\phi_+)^*) = (-\infty, d)$, so $(\phi_+)^*$ is continuous on this set, by [Rockafellar, Thm. 10.1, 1970]. It follows that $(\phi_+)^*$ is uniformly continuous on $[-2M, d - (\varepsilon/2)]$. Thus, if $y_n \to y$ uniformly on T, $(\phi_+)^*(y_n(t)) \to (\phi_+)^* y(t)$ uniformly, and so $I_{\phi_+^*}(y_n) \to I_{\phi_+^*}(y)$, as required. Alternatively, it is sufficient to observe that $I_{\phi_+^*}$ is bounded above on some neighborhood of y [Holmes, 1975]. $\square$

We will use the following ideas from [Rockafellar, 1970].

DEFINITION 4.4. *A proper convex function* $f : \mathbb{R}^n \to (-\infty, \infty]$ *is essentially strictly convex if* f *is strictly convex on every convex subset of* $\{x \mid \partial f(x) \neq \phi\}$.

Note that if $n = 1$, f is essentially strictly convex if and only if f is strictly convex on dom f. This follows from the fact that if $f: \mathbb{R}^n \to (-\infty, \infty]$ is convex, the only way in which it can fail to be strictly convex on dom f is if it is actually affine on some line segment. To see this, suppose f is not strictly convex on dom f. Without loss of generality suppose $f(0) = f(x_0) = 0$ for some $x_0 \neq 0$, and that for some $0 < \lambda_0 < 1$, $f(\lambda_0 x_0) = 0$. Convexity implies $f(\lambda x_0) \leq 0$ for all $0 \leq \lambda \leq 1$, so suppose we have strict inequality for some λ, and without loss of generality suppose $0 < \lambda < \lambda_0$. We then have

$$0 = f(\lambda_0 x_0)$$

$$= f\left(\left(\frac{1 - \lambda_0}{1 - \lambda}\right)(\lambda x_0) + \left(\frac{\lambda_0 - \lambda}{1 - \lambda}\right)(x_0)\right)$$

$$\leq \left(\frac{1 - \lambda_0}{1 - \lambda}\right)f(\lambda x_0) + \left(\frac{\lambda_0 - \lambda}{1 - \lambda}\right)f(x_0)$$

$$< 0,$$

which is a contradiction. Thus, $f(x) = 0$ on $[0, x_0]$.

However, if f is essentially strictly convex then it is certainly strictly convex on ri $(\text{dom} f) \subset \{x \mid \partial f(x) \neq \phi\}$, and the above argument then shows it must be strictly convex on dom f (provided $n = 1$).

DEFINITION 4.5. A proper convex function $f: \mathbb{R}^n \to (-\infty, \infty]$ is *essentially smooth* if f is differentiable on int $(\text{dom} f) \neq \phi$, and $\|\nabla f(x_i)\| \to \infty$ whenever $(x_i) \subset \text{int}(\text{dom} f)$ with $x_i \to$ some x in the boundary of dom f.

THEOREM 4.6. *If $f: \mathbb{R}^n \to (-\infty, \infty]$ is closed, proper, and convex, then f is essentially strictly convex if and only if f^* is essentially smooth.*

Proof. For the proof see [Rockafellar, Thm. 26.3, 1970]. □

The following result gives conditions for dual solutions to be unique.

THEOREM 4.7. *Consider the problem (DEP_p) of Corollary 2.6. Suppose $\{a_1, \cdots, a_n\}$ is linearly independent and ϕ_+ is essentially smooth. Then any optimal solution is unique.*

Proof. Suppose $\lambda^1 \neq \lambda^2$ are both optimal solutions of (DEP_p). The two functions $\sum_{i=1}^n \lambda_i^j a_i(t) \in \text{dom}(\phi_+)^*$ almost everywhere, $j = 1, 2$. Since the a_i's are linearly independent, they differ on a nonnull subset of T, T_1 say. By Theorem 4.5, $(\phi_+)^*$ is strictly convex on dom $(\phi_+)^*$, so if we set $\lambda^3 = \frac{1}{2}\lambda^1 + \frac{1}{2}\lambda^2$, then

$$(\phi_+)^*\left(\sum_{i=1}^n \lambda_i^3 a_i(t)\right) \leq \frac{1}{2}(\phi_+)^*\left(\sum_{i=1}^n \lambda_i^1 a_i(t)\right) + \frac{1}{2}(\phi_+)^*\left(\sum_{i=1}^n \lambda_i^2 a_i(t)\right),$$

almost everywhere, with strict inequality on T_1. It follows that λ^3 is an improvement on λ^1 and λ^2, which is a contradiction. □

We are now ready to prove our main result. We shall be concerned specifically with the dual pair (EP_1) and (DEP_1), so the primal variable will lie in $L_1(T)$. In this result we shall give conditions allowing us to obtain the primal optimal solution by differentiating the dual objective function at the optimum, as described in the introduction. This function involves the convex functional $I_{\phi_+^*}: L_\infty(T) \to (-\infty, \infty]$, whose subgradients lie in $L_\infty^*(T)$. We shall use the results in [Rockafellar, 1971] to decompose such subgradients into their singular and continuous (lying in $L_1(T)$) parts. Finally, we shall find a condition ensuring the singular part vanishes, which will lead to the desired conclusion. The result could alternatively be proved by a direct differentiation argument (see [Ben-Tal, Borwein, and Teboulle, 1988b]), but we shall use the ideas from our proof again in the last section.

THEOREM 4.8. *Consider the dual pair of problems* (EP$_1$) *and* (DEP$_1$). *Suppose that all the assumptions of Corollary 2.6 are met, and that ϕ_+ is strictly convex on* dom ϕ_+. *Let $\bar{\lambda}$ be dual optimal. Suppose finally that the following assumption is satisfied*:

$$(4.9) \qquad d := \lim_{x \to \infty} \frac{\phi(x)}{x} > \text{ess sup} \sum_{i=1}^{n} \bar{\lambda}_i a_i.$$

Then the unique primal optimal solution is given by

$$\bar{x}(t) = ((\phi_+)^*)' \left(\sum_{i=1}^{n} \bar{\lambda}_i a_i(t) \right).$$

Proof. By Corollary 2.6 we know that the primal and dual values are equal, with dual attainment. Let us denote the constraint map by $A : L_1(T) \to \mathbb{R}^n$, so $(Ax)_i = \langle x, a_i \rangle$, each i. The adjoint map $A^T : \mathbb{R}^n \to L_\infty(T)$ is therefore defined by $A^T \lambda = \sum_{i=1}^{n} \lambda_i a_i$. It is thus $\mathbb{R}^n - \| \cdot \|_\infty$ continuous, so we can define $A^{TT} : L_\infty^*(T) \to \mathbb{R}^n$, and we will have $A^{TT}|_{L_1(T)} = A$.

The dual objective function is $-g(\lambda)$, where $g : \mathbb{R}^n \to (-\infty, \infty]$ is defined by $g(\lambda) = -b^T \lambda + I_{\phi_+^*}(A^T \lambda)$, and since $\bar{\lambda}$ is dual optimal, $0 \in \partial g(\bar{\lambda})$. However, $\sum_{i=1}^{n} \bar{\lambda}_i a_i \in$ range (A^T) and $I_{\phi_+^*}$ is continuous at this point, by (4.9) and Lemma 4.3. It follows by [Borwein, Thm. 4.1, 1981] or [Rockafellar, Thm. 19, 1974] that $\partial g(\bar{\lambda}) = -b + A^{TT} \partial I_{\phi_+^*}(A^T \bar{\lambda})$, and so there exists $\bar{\mu} \in \partial I_{\phi_+^*}(A^T \bar{\lambda}) \subset L_\infty^*(T)$ with $A^{TT} \bar{\mu} = b$.

By definition of the subgradient, for all $y \in L_\infty(T)$, $\bar{\mu}(y - A^T \bar{\lambda}) \le I_{\phi_+^*}(y) - I_{\phi_+^*}(A^T \bar{\lambda})$. Thus, $\bar{\mu} \ge 0$, since if not, we could find a $y \le A^T \bar{\lambda}$ with $\bar{\mu}(y - A^T \bar{\lambda}) > 0$, implying $I_{\phi_+^*}(A^T \bar{\lambda}) < I_{\phi_+^*}(y)$, which contradicts the fact that $(\phi_+)^*$ is monotone increasing by Proposition 4.1.

We can now apply [Rockafellar, Cor. 1B, 1971] to deduce the existence of a $z \in L_1(T)$ and a singular component $\nu \in L_\infty^*(T)$ such that $\bar{\mu} = z + \nu$, where $z(t) \in \partial(\phi_+)^*((A^T \bar{\lambda})(t))$, almost everywhere and ν attains its maximum over dom $I_{\phi_+^*}$ at $A^T \bar{\lambda}$. However, we know $A^T \bar{\lambda} \in$ int (dom $I_{\phi_+^*}$), from which it follows that $\nu = 0$ (alternatively, see [Rockafellar, Cor. 2C, 1971]). Furthermore, we know $(A^T \bar{\lambda})(t) \in$ int (dom $(\phi_+)^*$), almost everywhere, and from Theorem 4.6 we know $(\phi_+)^*$ is essentially smooth. Thus, $z \in L_1(T)$, $z \ge 0$, $Az = b$, and $z(t) = ((\phi_+)^*)'((A^T \bar{\lambda})(t))$, almost everywhere. In particular, z is primal feasible.

Finally, since ϕ_+ is closed,

$$\phi_+(z(t)) + (\phi_+)^*((A^T \bar{\lambda})(t)) = z(t)(A^T \bar{\lambda})(t), \quad \text{a.e.,}$$

by [Rockafellar, Thm. 23.5, 1970]. Integrating over T gives

$$I_{\phi_+}(z) + I_{\phi_+^*}(A^T \bar{\lambda}) = \langle z, A^T \bar{\lambda} \rangle$$
$$= (Az)^T \bar{\lambda}$$
$$= b^T \bar{\lambda},$$

so z has the same objective value as the dual value, so it is optimal by weak duality. Uniqueness follows by Proposition 2.11.   □

5. Special cases. The last result of the previous section (Theorem 4.8) showed that under suitable conditions, if the dual optimal solution $\bar{\lambda}$ satisfied the condition that ess sup $\sum_{i=1}^{n} \bar{\lambda}_i a_i < d$, where d was defined to be $\lim_{x \to \infty} (\phi(x)/x)$, we could obtain the unique primal solution by differentiating the dual objective function. If $d = \infty$ then this condition is clearly no restriction. However, if $d < \infty$ then condition (4.9) may fail; this is the case, for instance, in Example 3.2. In this final section we shall consider conditions on ϕ that ensure a priori that condition (4.9) will hold.

We begin by defining another constant associated with ϕ. Assuming ϕ_+ is essentially strictly convex, we know by Lemma 4.2 that $\mathrm{cl}\,(\mathrm{dom}\,(\phi_+)^*) = (-\infty, d]$, and that $(\phi_+)^*$ is essentially smooth by Theorem 4.6. Let us define (assuming $d < \infty$)

$$(5.1) \qquad c := \lim_{y \uparrow d} (d - y)((\phi_+)^*)'(y).$$

LEMMA 5.2. *Suppose ϕ_+ is essentially strictly convex and essentially smooth. Then*

$$(5.3) \qquad c = \lim_{x \to \infty} (d - \phi_+'(x))x.$$

Proof. By [Rockafellar, Thm. 26.5, 1970], ϕ_+' is one-to-one from $\mathrm{int}\,(\mathrm{dom}\,\phi_+)$ to $\mathrm{int}\,(\mathrm{dom}\,(\phi_+)^*) = (-\infty, d)$, continuous in both directions, and $((\phi_+)^*)' = (\phi_+')^{-1}$. Note that since $d < \infty$, $(k, \infty) \subset \mathrm{dom}\,\phi_+$, for k sufficiently large, so expression (5.3) is well defined. Furthermore, $\phi_+'(x)$ increases to d as x increases to ∞, continuously, using l'Hôpital's rule and the convexity of ϕ_+. Thus, we have

$$
\begin{aligned}
c &= \lim_{y \uparrow d} (d - y)((\phi_+)^*)'(y) \\
&= \lim_{x \uparrow \infty} (d - \phi_+'(x))((\phi_+)^*)'(\phi_+'(x)) \\
&= \lim_{x \to \infty} (d - \phi_+'(x))x,
\end{aligned}
$$

as required. $\square$

We will now restrict our attention to the case where the underlying measure space T is a compact real interval with Lebesgue measure. We will need the following lemma.

LEMMA 5.4. *Suppose $h \in C[\alpha, \beta]$, $h(t_0) = 0$ for some $\alpha \leq t_0 \leq \beta$, and h is Lipschitz (or in particular, continuously differentiable) at t_0. Then $1/h \notin L_1[\alpha, \beta]$.*

Proof. Suppose first that h is continuously differentiable at t_0. Define a function $g: [\alpha, \beta] \to \mathbb{R}$ by

$$
g(t) = \begin{cases} h(t)/(t - t_0), & t \neq t_0, \\ h'(t_0), & t = t_0. \end{cases}
$$

By L'Hôpital's rule, g is continuous, so by compactness, $|g(t)| \leq M$ for all $t \in [\alpha, \beta]$, for some M. Since $h(t) = (t - t_0)g(t)$ for all $t \in [\alpha, \beta]$, $|h(t)| \leq M|t - t_0|$, so h is Lipschitz at t_0.

Now assume h is Lipschitz at t_0, so $|h(t)| \leq M|t - t_0|$, for all $t \in [\alpha, \beta]$. Thus,

$$\frac{1}{|h(t)|} \geq \frac{1}{M|t - t_0|} \quad \forall t \in [\alpha, \beta], \quad t \neq t_0.$$

Since clearly $1/|t - t_0| \notin L_1[\alpha, \beta]$, the result follows. $\square$

For the last step in the above argument, it is critical that the underlying measure is Lebesgue measure (or at least is greater than some positive multiple of Lebesgue measure). The underlying space is important, too. Consider, for example, $T = [-1, 1] \times [-1, 1] \subset \mathbb{R}^2$, with Lebesgue measure. Then $1/\|t\| \in L_1(T)$.

THEOREM 5.5. *Let $T = [\alpha, \beta]$, with Lebesgue measure, and each a_i be locally Lipschitz (or in particular, continuously differentiable), $i = 1, \cdots, n$. Consider the dual pair of problems (EP_1) and (DEP_1). Suppose that all of the assumptions of Corollary 2.6 are met, and that ϕ_+ is strictly convex on $\mathrm{dom}\,\phi_+$. Suppose that*

$$(5.6) \qquad \textit{there exists } \mu \in \mathbb{R}^n \textit{ with } \sum_{i=1}^n \mu_i a_i(s) < d \quad \forall s \in [\alpha, \beta],$$

where $d := \lim_{x \to \infty} (\phi(x)/x)$, and if $d < \infty$ define $c := \lim_{y \uparrow d} (d - y)((\phi_+)^*)'(y)$. Suppose that either $d = \infty$ or $d < \infty$ and $c > 0$. Then the unique primal optimal solution is given by $\bar{x}(t) = ((\phi_+)^*)'(\sum_{i=1}^n \bar{\lambda}_i a_i(t))$, where $\bar{\lambda}$ is a dual optimal solution.

Proof. Just as in the proof of Theorem 4.8, we know that $0 \in \partial g(\bar{\lambda})$, where $g(\lambda) = -b^T \lambda + I_{\phi_+^*}(A^T \lambda)$. We can write (5.6) as

$$\text{range}(A^T) \cap \text{cont}(I_{\phi_+^*}) \neq \phi,$$

by Lemma 4.3, so it follows by [Rockafellar, Thm. 19, 1974] or [Borwein, Thm. 4.1, 1981] that $\partial g(\lambda) = -b + A^{TT} \partial I_{\phi_+^*}(A^T \lambda)$. Thus, there exists $\bar{\mu} \in \partial I_{\phi_+^*}(A^T \bar{\lambda})$. Applying [Rockafellar, Cor. 1B, 1971] just as in the proof of Theorem 4.8, it follows that there exists $z \in L_1(T)$ with $z(t) \in \partial(\phi_+)^*((A^T\bar{\lambda})(t))$, almost everywhere.

The next step is to observe that since $(\phi_+)^*$ is essentially smooth (by Theorem 4.6), $\partial(\phi_+)^*(d) = \phi$. To see this, recall that as $y \uparrow d$, $((\phi_+)^*)'(y) \uparrow \infty$, by the definition of essentially smooth, and the fact that, by Lemma 4.2, $\text{cl}(\text{dom}(\phi_+)^*) = (-\infty, d]$. If $u \in \partial(\phi_+)^*(d)$ then $u(y - d) \leq (\phi_+)^*(y) - (\phi_+)^*(d)$, for all y. By the mean value theorem, given any $M > 0$ there exists $y < d$ with $(\phi_+)^*(d) - (\phi_+)^*(y) = ((\phi_+)^*)'(z)(d - y) \geq M(d - y)$, for some $y < z < d$. It follows that $u \geq M$, which is a contradiction as M was arbitrary.

Since $z(t) \in \partial(\phi_+)^*((A^T\bar{\lambda})(t))$ almost everywhere, it follows that $(A^T\bar{\lambda})(t) < d$, almost everywhere, and so $z(t) = ((\phi_+)^*)'((A^T\bar{\lambda})(t))$, almost everywhere. We know from Theorem 4.8 that if ess sup $A^T\bar{\lambda} < d$ then $\bar{x} = z$ is the unique primal optimal solution, as required. Since the a_i's are continuous on $[\alpha, \beta]$, the only alternative is $(A^T\bar{\lambda})(t_0) = d$, for some $\alpha \leq t_0 \leq \beta$. We will show that, in this case, z cannot possibly lie in $L_1[\alpha, \beta]$, giving a contradiction. This will complete the proof.

Assume therefore that for some $\alpha \leq t_0 \leq \beta$, $\sum_{i=1}^n \bar{\lambda}_i a_i(t_0) = d < \infty$. By the definition of c, there exists $\varepsilon > 0$ such that $(d - y)((\phi_+)^*)'(y) \geq c/2 > 0$, for all $d > y \geq d - \varepsilon$. By the continuity of the a_i's, there exists $\delta > 0$ such that $\sum_{i=1}^n \bar{\lambda}_i a_i(t) \geq d - \varepsilon$, for all $|t - t_0| \leq \delta$. Thus,

$$z(t) = ((\phi_+)^*)'\left(\sum_{i=1}^n \bar{\lambda}_i a_i(t)\right)$$

$$\geq \frac{c}{2(d - \sum_{i=1}^n \bar{\lambda}_i a_i(t))},$$

for all $|t - t_0| \leq \delta$. But by Lemma 5.4, the right-hand side is not integrable (on either $[t_0 - \delta, t_0]$ or $[t_0, t_0 + \delta]$), so $z \notin L_1[\alpha, \beta]$. This completes the proof. $\square$

Note in particular that condition (5.6) will always hold if one of the a_i's is a nonzero constant function.

We will conclude with a number of examples of typical objective functions ϕ, taken from [Ben-Tal, Borwein, and Teboulle, 1988b]. Some of these objectives are taken from the literature, others are new. For each function we give the numbers d and c of Theorem 5.5. It is easy to check that ϕ_+ is closed and essentially strictly convex in each case.

Examples 5.6.

(i) $\qquad \phi_+(x) = \begin{cases} -\log x, & x > 0, \\ \infty, & x \leq 0, \end{cases}$

$\qquad (\phi_+)^*(y) = \begin{cases} -1 - \log(-y), & y < 0, \\ \infty, & y \geq 0, \end{cases} \qquad\qquad d = 0, \quad c = 1.$

(ii)
$$\phi_+(x) = \begin{cases} x \log x - x, & x > 0, \\ 0, & x = 0, \\ \infty, & x < 0, \end{cases}$$

$$(\phi_+)^*(y) = e^y, \qquad\qquad d = \infty.$$

(iii)
$$\phi_+ + (x) = \begin{cases} \dfrac{1}{p} x^p, & x \geqq 0, \quad (\text{with } 1 < p < \infty), \\ \infty, & x < 0, \end{cases}$$

$$(\phi_+)^*(y) = \frac{1}{q} (y^+)^q, \qquad\qquad d = \infty.$$

(iv)
$$\phi_+(x) = \begin{cases} e^x - 1, & x \geqq 0, \\ \infty, & x < 0, \end{cases}$$

$$(\phi_+)^*(y) = \begin{cases} y \log y - y + 1, & y \geqq 1, \\ 0, & y < 1, \end{cases} \qquad d = \infty.$$

(v)
$$\phi_+(x) = \begin{cases} \dfrac{\gamma x - x^\gamma}{1 - \gamma}, & x \geqq 0, \\ \infty, & x < 0, \end{cases} \left(\text{with } 0 < \gamma < 1, \ \frac{1}{\gamma} - \frac{1}{\psi} = 1 \right),$$

$$(\phi_+)^*(y) = \begin{cases} \left(1 - \dfrac{1}{\psi} y\right)^{-\psi}, & y \leqq \psi, \\ \infty, & \text{otherwise}, \end{cases} \qquad d = \psi, \quad c = \infty.$$

(vi)
$$\phi_+(x) = \begin{cases} k - \sqrt{k^2 - x^2}, & 0 \leqq x \leqq k, \\ \infty, & \text{otherwise}, \end{cases}$$

$$(\phi_+)^*(y) = \begin{cases} k(\sqrt{1 + y^2} - 1), & y \geqq 0, \\ 0, & y < 0, \end{cases} \qquad d = \infty.$$

(vii)
$$\phi_+(x) = \begin{cases} x \log x - (1 + x) \log (1 + x), & x > 0, \\ 0, & x = 0, \\ \infty, & x < 0, \end{cases}$$

$$(\phi_+)^*(y) = \begin{cases} -\log (1 - e^y), & y < 0, \\ \infty, & y \geqq 0, \end{cases} \qquad d = 0, \qquad c = \infty.$$

(viii)
$$\phi_+(x) = \begin{cases} \dfrac{1}{x}, & x > 0, \\ \infty, & x \leqq 0, \end{cases}$$

$$(\phi_+)^*(y) = \begin{cases} -2\sqrt{-y}, & y \leqq 0, \\ \infty, & y > 0, \end{cases} \qquad d = 0, \qquad c = 0.$$

In all of the above examples, ϕ_+ is strictly convex on its domain, and except in cases (iii), (iv), and (vi) it is also essentially smooth. With the exception of case (vi), $(0, \infty) \subset \text{dom } \phi_+$, so the constraint qualification will simply require the existence of a feasible $\hat{x} \in \text{dom } I_{\phi_+}$ with $\hat{x}(t) > 0$, almost everywhere. In the case of (vi) it can be

338 J. M. BORWEIN AND A. S. LEWIS

checked using the results of [Borwein and Lewis, 1988] that the constraint qualification requires the existence of a feasible $\hat{x}$ with $0 < \hat{x}(t) < k$, almost everywhere. Example (viii) shows, as we would expect, that the assumptions of Theorem 5.5 fail for Example 3.2.

Suppose finally that $(\phi_+)^*$ is actually continuously differentiable on (∞, d), as is the case in all of the above examples. With the assumptions of Theorem 5.5 it then follows that the unique optimal solution of (EP_1), $\bar{x}$, is actually continuous on T, so it certainly lies in L_p for any $1 \leq p \leq \infty$. Thus, $\bar{x}$ will in fact be the unique optimal solution for (EP_p), for any $1 \leq p \leq \infty$, and also for the analogous problem posed in $C[\alpha, \beta]$.

REFERENCES

A. BEN-TAL, J. M. BORWEIN, AND M. TEBOULLE (1988a), *A dual approach to multidimensional L_p spectral estimation problems*, SIAM J. Control Optim., 26, pp. 985–996.
——— (1988b), *Spectral estimation via convex programming*, to appear.
S. BOCHNER AND W. T. MARTIN (1964), *Several Complex Variables*, Princeton University Press, Princeton, NJ.
J. M. BORWEIN (1981), *A Lagrange multiplier theorem and a sandwich theorem for convex relations*, Math. Scand., 48, pp. 188–204.
J. M. BORWEIN AND A. S. LEWIS (1988), *Partially finite convex programming, Parts I and II*, Math. Programming, to appear.
J. P. BURG (1975), *Maximum entropy spectral analysis*, Ph.D. dissertation, Stanford University, Stanford, CA.
B. K. GOODRICH AND A. STEINHARDT (1986), *L_2 spectral estimation*, SIAM J. Appl. Math., 46, pp. 417–428.
B. B. HOLMES (1975), *Geometric Functional Analysis and its Applications*, Springer-Verlag, New York.
A. D. IOFFE AND V. M. TIHOMOROV (1968), *Extension of variational problems*, Trans. Moscow Math. Soc., 18, pp. 207–273.
L. D. IRVINE, S. P. MARIN, AND P. W. SMITH (1986), *Constrained interpolation and smoothing*, Constructive Approximation, 2, pp. 129–151.
R. W. JOHNSON AND J. E. SHORE (1984), *Which is the better entropy expression for speech processing: slogs or logs?*, IEEE Trans. on Acoustics, Speech and Signal Processing, ASSP 32, pp. 129–136.
S. KARLIN AND W. STUDDEN (1966), *Tchebycheff systems with applications in analysis and statistics*, Interscience, New York.
C. A. MICCHELLI, P. W. SMITH, J. SWETITS, AND J. D. WARD (1985), *Constrained L_p approximation*, Constructive Approximation, 1, pp. 93–102.
R. T. ROCKAFELLAR (1968), *Integrals which are convex functionals*, Pacific J. Math., 24, pp. 525–539.
——— (1970), *Convex Analysis*, Princeton University Press, Princeton, NJ.
——— (1971), *Integrals which are convex functionals II*, Pacific J. Math., 39, pp. 439–469.
——— (1974), *Conjugate Duality and Optimization*, CBMS-NSF Regional Conference Series in Applied Mathematics 16, Society for Industrial and Applied Mathematics, Philadelphia, PA.
——— (1976), *Integral functions, normal integrands and measurable selections*, in Nonlinear Operators and the Calculus of Variations, L. Waelbroeck, ed., Lecture Notes in Math. 543, Springer-Verlag, New York, Berlin, pp. 157–207.
W. RUDIN (1966), *Real and Complex Analysis*, McGraw-Hill, New York.