JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 70, 546–562 (1979)

Optimal Constrained Selection of a Measurable Subset*

ROBERT J. T. MORRIS

Bell Laboratories, Holmdel, New Jersey 07733

Submitted by G. Leitmann

Necessary and sufficient conditions are given for a class of optimization problems involving optimal selection of a measurable subset from a given measure space subject to set function inequality constraints. Results are developed firstly for the case where the set functions involved possess a differentiability property and secondly where a type of convexity is present. These results are then used to develop numerical methods. It is shown that in a special case the optimal set can be obtained via solution of a fixed point problem in Euclidean space.

1. INTRODUCTION

A class of optimization problems involving optimal choice of a subset of a given space has been the subject of several recent papers [1–4]. This type of problem has been shown to arise in diverse applications including electrical insulator design [1], optimal plasma confinement [2] and fluid flow [3], yet the solution is not adequately catered for by existing optimization theory. This is partly because much of the existing theory is oriented towards the problem of optimal selection of a point rather than a subset from a vector space and the collection of subsets of a vector space fails to possess a linear space structure.

For the purposes of motivation we begin by presenting an extremely simple example.

EXAMPLE 1.1. Rainfall in a region $\mathscr{R}$ is distributed according to the function $r(x, y)$ where x is longitude, y is latitude. It is desired to plant a crop which has a per-acre yield of $p(r)$ where r is the rainfall. The cost per unit area of planting is K and the area to be planted must not exceed A. The return realized from total production p is $u(p)$. The optimization problem is to choose a subregion for planting, i.e., a set $\Omega \subseteq \mathscr{R}$, such that the net return $u(\int_\Omega p(r(x, y))\,dx\,dy) - K \cdot \text{measure}(\Omega)$ is maximized subject to the constraint: measure $(\Omega) \leqslant A$.

The general problem we will consider is now stated.

* This research was performed at the Dept. of System Science, University of California, Los Angeles, California, 90024, as part of the author's doctoral dissertation. The author acknowledges the partial support of a Gowrie Research Fellowship and Air Force Office of Scientific Research Grant No. 74-2662.

0022-247X/79/080546–17$02.00/0

PROBLEM 1.1. Given a measure space $(X, \mathcal{A}, m)$ and mappings $F, G_1, \ldots, G_n$: $\mathcal{A} \to \mathbf{R}$, find $\Omega^* \in \mathcal{A}$ which minimizes $F(\Omega)$ subject to the constraints $G_i(\Omega) \leqslant 0$, $i = 1, \ldots, n$.

The mappings $F, G_1, \ldots, G_n$ are set functions which are in general non-additive but will be assumed to possess various properties to obtain conditions of necessity, sufficiency, etc.

A problem of this type was considered in [1] where there were no constraints present and the function F was assumed to possess a "differentiability" property. A problem with a single constraint was studied in [2] but the set functions were additive. Different approaches have been presented in [3, 4] where the admissible sets are domains whose boundaries possess certain regularity properties. The aim of this paper is to show that for Problem 1.1 appropriate definitions can be made and theorems given which establish necessary and sufficient conditions for optimality analogous to the results of mathematical programming. These results enable several computational algorithms to be stated.

We preface the results with the following observation. If $(X, \mathcal{A}, m)$ is a finite measure space, then any set $\Omega \in \mathcal{A}$ can be identified with its characteristic function $\chi_\Omega \in L_1(X, \mathcal{A}, m)$. In this way Problem 1.1 can be regarded as constrained minimization of a nonlinear functional on $L_1(X, \mathcal{A}, m)$ over the set $\chi = \{\chi_\Omega : \Omega \in \mathcal{A}\}$. Thus Problem 1.1 can be phrased as a minimization over an infinite dimensional Banach space. We hasten to point out that the set χ is extremely poorly conditioned for the purposes of application of the standard theory, for χ is not convex, is not open and is actually nowhere dense. Thus we do not expect to get necessary conditions for optimality as strong as those which are available on, for example, open convex sets. Our results will be of the nature of being "minimum principles" (for discussion see [5, p. 162]), as is illustrated by the following example.

EXAMPLE 1.2. Let $g: X \to \mathbf{R}$ be integrable and suppose that it is desired to minimize $\int_\Omega g\, dm$ by choice of $\Omega \in \mathcal{A}$. A solution is clearly given by the set $\Omega^* = g^{-1}((-\infty, 0])$. Denoting an integral $\int_\Omega f\, dm$ by the customary notation for functionals $\langle f, \chi_\Omega \rangle$, the set Ω^* is seen to satisfy the necessary condition of the minimum principle type

$$\langle g, \chi_\Omega - \chi_{\Omega^*} \rangle \geqslant 0 \qquad \text{all} \quad \Omega \in \mathcal{A}.$$

2. LOCAL THEORY

Throughout the paper we will make the following

ASSUMPTION. The measure space $(X, \mathcal{A}, m)$ is

(i) finite, i.e., $mX < \infty$,

(ii) atomless, i.e., for any $\Omega \in \mathcal{A}$ with $m\Omega > 0$, there exists $\Omega' \subseteq \Omega$, $\Omega' \in \mathcal{A}$ with $0 < m\Omega' < m\Omega$.

The finiteness of $(X, \mathcal{A}, m)$ allows $\mathcal{A}$ to be made into a pseudometric space by way of the pseudometric

$$\rho(\Omega_1, \Omega_2) = m(\Omega_1 \triangle \Omega_2), \qquad \Omega_1, \Omega_2 \in \mathcal{A}.$$

This enables the following

DEFINITION 2.1. A set function $F: \mathcal{A} \to \mathbf{R}$ is said to be *differentiable* at $\Omega_0 \in \mathcal{A}$ if there exists $f_{\Omega_0} \in L_1(X, \mathcal{A}, m)$, the *derivative* at Ω_0, such that

$$F(\Omega) = F(\Omega_0) + \langle f_{\Omega_0}, \chi_\Omega - \chi_{\Omega_0} \rangle + E_F(\Omega_0, \Omega)$$

where $E_F(\Omega_0, \Omega)$ is $o[\rho(\Omega_0, \Omega)]$, i.e., $\lim_{\rho(\Omega_0,\Omega)\to 0} [E_F(\Omega_0, \Omega)/\rho(\Omega_0, \Omega)] = 0$.

We recall the following result.

LEMMA 2.1 (Liapunov). *Given an atomless measure space $(X, \mathcal{A}, m)$ and integrable functions $f_i: X \to \mathbf{R}$, $i = 1,..., n$ then the range of the vector measure $[\int_\Omega f_1\, dm,..., \int_\Omega f_n\, dm]$ is convex and compact.*

Proof. See [6, p. 113].

PROPOSITION 2.2. *If a set function is differentiable then its derivative is unique.*

Proof. Let f and f' both be derivatives of F at Ω_0 and set $g = f - f'$. Then $\langle g, \chi_\Omega - \chi_{\Omega_0} \rangle = o[\rho(\Omega_0, \Omega)]$ for $\Omega, \Omega_0 \in \mathcal{A}$. Fix $\epsilon > 0$ and let $\Lambda = \{x \in \Omega_0: g(x) > \epsilon\}$. Then by Lemma 2.1 there exist sets $\Lambda_n \subseteq \Lambda$ with $m\Lambda_n = (1/n)m\Lambda$, $\langle g, \chi_{\Lambda_n} \rangle = (1/n)\langle g, \chi_\Lambda \rangle$. Now unless $m\Lambda = 0$, $\langle g, \chi_{\Lambda_n} \rangle / m\Lambda_n = \langle g, \chi_{\Omega_0} - (\chi_{\Omega_0} - \chi_{\Lambda_n}) \rangle / m\Lambda_n \to 0$ as $n \to \infty$. But $\langle g, \chi_{\Lambda_n} \rangle / m\Lambda_n = \langle g, \chi_\Lambda \rangle / m\Lambda \geqslant \epsilon$, all n. Thus $m\Lambda = 0$ for arbitrary positive ϵ and it follows that $g \leqslant 0$ a.e. on Ω_0. Applying the same argument to $-g$ shows that $g = 0$ a.e. on Ω_0. Similarly $g = 0$ a.e. on $\Omega_0{}^c$ yielding the desired result. Q.E.D.

Remarks. (i) Set functions satisfying a property similar to, but somewhat stronger than, Definition 2.1 were considered in [1].

(ii) If F is countably additive and absolutely continuous with respect to m, then f_Ω is simply the Radon-Nikodym derivative dF/dm.

(iii) If $\tilde{F}$ is a Fréchet differentiable functional on $L_1(X, \mathcal{A}, m)$ then $F(\Omega) = \tilde{F}(\chi_\Omega)$ is a differentiable set function. In this case the Fréchet derivative of $\tilde{F}$ denoted $\tilde{F}'$ lies in $L_1^*(X, \mathcal{A}, m) = L_\infty(X, \mathcal{A}, m)$. According to Proposition 2.2, $\tilde{F}'$ coincides with f_Ω and consequently $f_\Omega \in L_\infty(X, \mathcal{A}, m)$.

EXAMPLE 2.1. An example of a differentiable set function is $F(\Omega) = u(\int_\Omega v_1\, dm,\dots, \int_\Omega v_n\, dm)$ where $u\colon \mathbf{R}^n \to \mathbf{R}$ is differentiable and $v_1,\dots, v_n$ are in $L_\infty(X, \mathcal{O}l, m)$. It has derivative $f_\Omega = \sum_{i=1}^n u_i(\int_\Omega v_1\, dm,\dots, \int_\Omega v_n\, dm)v_i$ where u_i denotes the i'th partial derivative of u.

Note that $F(\Omega) = \tilde{F}(\chi_\Omega)$, where $\tilde{F}(x) = u(\langle v_1, x\rangle,\dots, \langle v_n, x\rangle)$, $x \in L_1(X, \mathcal{O}l, m)$ defines a Fréchet differentiable functional $\tilde{F}\colon L_1(X, \mathcal{O}l, m) \to \mathbf{R}$.

EXAMPLE 2.2. Modifying Example 2.1 so that u is affine and $v_1,\dots, v_n$ are in $L_1(X, \mathcal{O}l, m)$ provides an example of a differentiable set function which may not be the restriction to χ of any Fréchet differentiable functional on $L_1(X, \mathcal{O}l, m)$.

DEFINITION 2.2. Given a function $f\colon X \to \mathbf{R}$ we say that f *separates* Ω_0 if $\langle f, \chi_\Omega - \chi_{\Omega_0}\rangle \geqslant 0$ for all $\Omega \in \mathcal{O}l$ or equivalently $f \leqslant 0$ a.e. on Ω_0, $f \geqslant 0$ a.e. on $\Omega_0{}^c$.

As usual we will say that $\Omega_0 \in \mathcal{O}l$ is a *local minimum* for Problem 1.1 if there exists $\epsilon > 0$ such that for Ω satisfying $\rho(\Omega_0, \Omega) < \epsilon$, $G_i(\Omega) \leqslant 0$, $i = 1,\dots, n$ it follows that $F(\Omega) \geqslant F(\Omega_0)$.

THEOREM 2.3 (Necessary condition for a constrained local minimum). *Let $(X, \mathcal{O}l, m)$ be a finite atomless measure space and $F, G_1,\dots, G_n$ be set functions differentiable at Ω^* with respective derivatives $f_{\Omega^*}, g^1_{\Omega^*},\dots, g^n_{\Omega^*}$.*

Suppose Ω^* is a local minimum of $F(\Omega)$ subject to $G_i(\Omega) \leqslant 0$, $i = 1,\dots, n$ and that Ω^* is regular, *i.e. there exists a set $\Omega_1 \in \mathcal{O}l$ with $G_i(\Omega^*) + \langle g^i_{\Omega^*}, \chi_{\Omega_1} - \chi_{\Omega^*}\rangle < 0$, $i = 1,\dots, n$.*

Then there exist nonnegative reals $\lambda_1^*,\dots, \lambda_n^*$ such that

$$f_{\Omega^*} + \sum_{i=1}^n \lambda_i^* g^i_{\Omega^*} \text{ separates } \Omega^*.$$

If $G_i(\Omega^*) < 0$ it follows that the corresponding $\lambda_i^* = 0$.

Proof. In the proof the derivative subscripts will be elided. Define

$$\begin{aligned} A &= \{(v_0, v_1,\dots, v_n)\colon \text{there exists } \Omega \in \mathcal{O}l \text{ with } v_0 \geqslant \langle f, \chi_\Omega - \chi_{\Omega^*}\rangle, \\ &\qquad v_i \geqslant G_i(\Omega^*) + \langle g^i, \chi_\Omega - \chi_{\Omega^*}\rangle, i = 1,\dots, n\} \\ B &= \{(v_0, v_1,\dots, v_n)\colon v_i < 0, i = 0,\dots, n\}. \end{aligned}$$

The set B is clearly convex and it follows from Lemma 2.1 that A is convex.

We now show that A and B are disjoint. For assume the contrary, that there exists $\Omega \in \mathcal{O}l$ with

$$\langle f, \chi_\Omega - \chi_{\Omega^*}\rangle < 0, \qquad G_i(\Omega^*) + \langle g^i, \chi_\Omega - \chi_{\Omega^*}\rangle < 0, \quad i = 1,\dots, n.$$

Now write $\Omega^+ = \Omega \sim \Omega^*$, $\Omega^- = \Omega^* \sim \Omega$ then $\chi_\Omega - \chi_{\Omega^*} = \chi_{\Omega^+} - \chi_{\Omega^-}$. By Lemma 2.1 there exist families $\Omega^+(\alpha) \subseteq \Omega^+$, $\Omega^-(\alpha) \subseteq \Omega^-$ satisfying

$$\int_{\Omega^\pm(\alpha)} [1, f, g^1, ..., g^n] \, dm = \alpha \int_{\Omega^\pm} [1, f, g^1, ..., g^n] \, dm, \qquad \alpha \in [0, 1].$$

Write $\Omega(\alpha) = \Omega^+(\alpha) \cup \Omega^* \sim \Omega^-(\alpha)$ then $\rho(\Omega^*, \Omega(\alpha)) = \alpha\rho(\Omega^*, \Omega)$. Thus

$$\begin{aligned} F(\Omega(\alpha)) &= F(\Omega^*) + \langle f, \chi_{\Omega^+(\alpha)} - \chi_{\Omega^-(\alpha)} \rangle + o(\alpha) \\ &= F(\Omega^*) + \alpha\langle f, \chi_{\Omega^+} - \chi_{\Omega^-} \rangle + o(\alpha) \end{aligned}$$

and similarly

$$G_i(\Omega(\alpha)) = G_i(\Omega^*) + \alpha\langle g^i, \chi_{\Omega^+} - \chi_{\Omega^-} \rangle + o(\alpha), \qquad i = 1, ..., n.$$

Letting $\alpha \to 0$, it follows that there exists arbitrarily small $\alpha' > 0$ such that

$$F(\Omega(\alpha')) < F(\Omega^*), \qquad G_i(\Omega(\alpha')) < 0, \qquad i = 1, ..., n.$$

This contradicts the local optimality of Ω^*.

Thus A and B are disjoint convex subsets of a finite dimensional space and can be separated by a hyperplane. (The remainder of the proof is quite standard (cf., for example, [7, p. 249]) but is given here for completeness.) Hence there exist reals $\lambda_0, \lambda_1, ..., \lambda_n$ not all zero and δ such that

$$\sum_{i=0}^{n} \lambda_i v_i \geqslant \delta \qquad \text{if} \qquad (v_0, ..., v_n) \in A$$

$$\sum_{i=0}^{n} \lambda_i v_i \leqslant \delta \qquad \text{if} \qquad (v_0, ..., v_n) \in B.$$

It follows from the nature of B that $\lambda_0, ..., \lambda_n, \delta$ are nonnegative and from the nature of A that δ is nonpositive, thus $\delta = 0$.

We now show $\lambda_0 > 0$. For assume that $\lambda_0 = 0$, and consider Ω_1 of the theorem data. Then

$$\sum_{i=1}^{n} \lambda_i (G_i(\Omega^*) + \langle g^i, \chi_{\Omega_1} - \chi_{\Omega^*} \rangle) \geqslant 0,$$

but $\lambda_1, ..., \lambda_n$ are nonnegative and $G_i(\Omega^*) + \langle g^i, \chi_{\Omega_1} - \chi_{\Omega^*} \rangle$, $i = 1, ..., n$ are strictly negative, thus $\lambda_i = 0$, $i = 0, ..., n$ contradicting the assumption that not all λ_i are zero. We conclude that $\lambda_0 > 0$.

Defining λ_i^* by λ_i/λ_0, $i = 1, ..., n$ yields

$$\left\langle f + \sum_{i=1}^{n} \lambda_i^* g^i, \chi_\Omega - \chi_{\Omega^*} \right\rangle + \sum_{i=1}^{n} \lambda_i^* G_i(\Omega^*) \geqslant 0$$

for all $\Omega \in \mathcal{O}\!\mathcal{l}$. Setting $\Omega = \Omega^*$ yields $\sum_{i=1}^{n} \lambda_i^* G_i(\Omega^*) \geqslant 0$, but by the nonnegativity of λ_i^* and nonpositivity of $G_i(\Omega^*)$ it follows that $\sum_{i=1}^{n} \lambda_i^* G_i(\Omega^*) = 0$.
Q.E.D.

Remark. Note that it is the convexity of the set A as provided by Lemma 2.1 rather than any property of χ that is crucial to the above proof.

In order to obtain sufficient conditions for a set to be a local minimum of Problem 1.1, we now introduce a notion of convexity. A definition of convexity for a general set function will be given in Section 3. However for differentiable set functions we can define a type of local convexity by analogy with the result for the differentiable function $h: \mathbf{R}^n \to \mathbf{R}$ that convexity of h is equivalent to the property

$$h(x) \geqslant h(y) + \nabla h(y)(x - y), \qquad \text{all} \quad x, y \in \mathbf{R}^n.$$

DEFINITION 2.3. A differentiable set function $F: \mathcal{O}\!\mathcal{l} \to \mathbf{R}$ is said to be *locally convex* at Ω_0 if there exists $\epsilon > 0$ such that $\rho(\Omega_0, \Omega) < \epsilon$ implies that $E_F(\Omega_0, \Omega) \geqslant 0$ or equivalently

$$F(\Omega) \geqslant F(\Omega_0) + \langle f_{\Omega_0}, \chi_\Omega - \chi_{\Omega_0} \rangle.$$

EXAMPLE 2.3. Requiring u to be locally convex in Example 2.1 makes the corresponding set function F locally convex.

THEOREM 2.4 (Sufficient conditions for a constrained local minimum.). *Suppose F, $G_1, \ldots, G_n$ are set functions which are differentiable and locally convex at Ω^* and that $G_i(\Omega^*) \leqslant 0$, $i = 1, \ldots, n$. Suppose further that there exist nonnegative reals $\lambda_1, \ldots, \lambda_n$ such that*

(i) *$f_{\Omega^*} + \sum_{i=1}^{n} \lambda_i g_{\Omega^*}^i$ separates Ω^**

(ii) *if $G_i(\Omega^*) < 0$ the corresponding $\lambda_i = 0$.*

Then Ω^* is a local minimum of $F(\Omega)$ subject to $G_i(\Omega) \leqslant 0$, $i = 1, \ldots, n$.

Proof. Let ϵ be the smallest of the constants associated with F, $G_1, \ldots, G_n$ in Definition 2.3. Let Ω satisfy $G_i(\Omega) \leqslant 0$, $i = 1, \ldots, n$, and $\rho(\Omega^*, \Omega) < \epsilon$. Then

$$\begin{aligned} F(\Omega) &\geqslant F(\Omega^*) + \langle f_{\Omega^*}, \chi_\Omega - \chi_{\Omega^*} \rangle, && \text{by local convexity of } F \\ &\geqslant F(\Omega^*) - \left\langle \sum_{i=1}^{n} \lambda_i g_{\Omega^*}^i, \chi_\Omega - \chi_{\Omega^*} \right\rangle, && \text{by (i)} \\ &\geqslant F(\Omega^*) + \sum_{i=1}^{n} \lambda_i (G_i(\Omega^*) - G_i(\Omega)), && \text{by local convexity of } G_i \\ &\geqslant F(\Omega^*), && \text{by (ii).} \end{aligned}$$

Q.E.D.

3. Global Theory

The first task in formulating a global theory is to give an appropriate definition for a convex set function.

The general problem of defining a convex function on a space $\mathscr{X}$ devoid of linear structure was considered by Von Neumann and Morgenstern in [8], for a recent survey we refer to [9]. In this approach the convex combination of two elements $x, y \in \mathscr{X}$ is given by $\langle \lambda, x, y \rangle$ where $\langle\ ,\ ,\ \rangle$: $[0, 1] \times \mathscr{X} \times \mathscr{X} \to \mathscr{X}$ is a mapping, sometimes referred to as a *mixture*, satisfying certain properties which are unimportant here. A function F: $\mathscr{X} \to \mathbf{R}$ is then said to be *convex* if $F[\langle \lambda, x, y \rangle] \leqslant \lambda F(x) + (1 - \lambda) F(y)$, all $x, y \in \mathscr{X}$, $\lambda \in [0, 1]$. We wish to ask whether it is possible to define a mixture on $\mathcal{Q}$. But we have the following result.

Proposition 3.1. *Given a non-null measure space* $(X, \mathcal{Q}, m)$, *there is no mixture which makes all additive set functions convex.*

Proof. Assume the contrary that there exists a mixture $\langle\ ,\ ,\ \rangle$. Consider the additive set function $F(\Omega) = \int_\Omega f\, dm$, where $f \in L_1(X, \mathcal{Q}, m)$. Then F and $-F$ are both additive and hence convex from which it follows that $F(\langle \frac{1}{2}, \varnothing, \Omega \rangle) = \frac{1}{2} F(\Omega)$ for all $\Omega \in \mathcal{Q}$. Now fix $\Omega \in \mathcal{Q}$ and let $f = \chi_{\langle 1/2, \varnothing, \Omega \rangle}$. Then

$$m\langle \tfrac{1}{2}, \varnothing, \Omega) = F(\langle \tfrac{1}{2}, \varnothing, \Omega \rangle) = \tfrac{1}{2} F(\Omega) = \tfrac{1}{2} m(\Omega \cap \langle \tfrac{1}{2}, \varnothing, \Omega \rangle).$$

But $\langle \frac{1}{2}, \varnothing, \Omega \rangle \supseteq \Omega \cap \langle \frac{1}{2}, \varnothing, \Omega \rangle$ thus $m(\frac{1}{2}, \varnothing, \Omega \rangle = 0$. Hence $0 = \int_{\langle 1/2, \varnothing, \Omega \rangle} f\, dm = \frac{1}{2} \int_\Omega f\, dm$ for any $\Omega \in \mathcal{Q}$, $f \in L_1(X, \mathcal{Q}, m)$. Choosing $f = 1$ shows that $mX = 0$, contradicting the assumption that $(X, \mathcal{Q}, m)$ is non-null. Q.E.D.

Since it is natural that at least the additive set functions be convex we conclude that there is no set which can effectively play the role of $\langle \lambda, \Omega, \Lambda \rangle$, the convex combination of the sets Ω and Λ.

Instead we associate with $\langle \lambda, \Omega, \Lambda \rangle$ the sequence $\Omega_n \cup \Lambda_n \cup (\Omega \cap \Lambda)$ where $\chi_{\Omega_n} \to^{w^*} \lambda \chi_{\Omega \sim \Lambda}$, $\chi_{\Lambda_n} \to^{w^*} (1 - \lambda) \chi_{\Lambda \sim \Omega}$ and $\to^{w^*}$ denotes weak* convergence of elements in $L_\infty(X, \mathcal{Q}, m)$. We will then show that the weak* limit of $\chi_{\Omega_n \cup \Lambda_n \cup (\Omega \cap \Lambda)}$ is $\lambda \chi_\Omega + (1 - \lambda) \chi_\Lambda$ which is a convex combination of χ_Ω and χ_Λ in the space $L_\infty(X, \mathcal{Q}, m)$.

Definition 3.1. A set function F: $\mathcal{Q} \to \mathbf{R}$ is said to be *convex* if given $\lambda \in [0, 1]$ and $\Omega, \Lambda \in \mathcal{Q}$ it follows that

$$\varlimsup_{n \to \infty} F(\Omega_n \cup \Lambda_n \cup (\Omega \cap \Lambda)) \leqslant \lambda F(\Omega) + (1 - \lambda) F(\Lambda)$$

for any sequences $\chi_{\Omega_n} \to^{w^*} \lambda \chi_{\Omega \sim \Lambda}$, $\chi_{\Lambda_n} \to^{w^*} (1 - \lambda) \chi_{\Lambda \sim \Omega}$.

PROPOSITION 3.2. *Given sets $\Omega, \Lambda \in \mathcal{A}$, real $\lambda \in [0, 1]$ and $L_\infty(X, \mathcal{A}, m)$ sequences*

$$\chi_{\Omega_n} \xrightarrow{w^*} \lambda \chi_{\Omega \sim \Lambda}, \qquad \chi_{\Lambda_n} \xrightarrow{w^*} (1-\lambda)\chi_{\Lambda \sim \Omega}$$

then

$$\chi_{\Omega_n \cup \Lambda_n \cup (\Omega \cap \Lambda)} \xrightarrow{w^*} \lambda \chi_\Omega + (1-\lambda)\chi_\Lambda .$$

Proof. Consider $f \in L_1(X, \mathcal{A}, m)$. Then

$$\begin{aligned}
&\langle f, \chi_{\Omega_n \cup \Lambda_n \cup (\Omega \cap \Lambda)} \rangle \\
&\quad = \langle f, \chi_{\Omega_n} + \chi_{\Lambda_n} + \chi_{\Omega \cap \Lambda} \rangle - \langle f, \chi_{\Omega_n \cap \Lambda_n} + \chi_{\Omega_n \cap \Omega \cap \Lambda} + \chi_{\Lambda_n \cap \Omega \cap \Lambda} \rangle \\
&\qquad + \langle f, \chi_{\Omega_n \cap \Lambda_n \cap \Omega \cap \Lambda} \rangle .
\end{aligned}$$

Now $\lim_{n\to\infty} \langle f, \chi_{\Omega_n \cap \Omega \cap \Lambda} \rangle = \lim_{n\to\infty} \langle f\chi_{\Omega \cap \Lambda}, \chi_{\Omega_n} \rangle = \lambda \langle f\chi_{\Omega \cap \Lambda}, \chi_{\Omega \sim \Lambda} \rangle = 0$ and similarly $\lim_{n\to\infty} \langle f, \chi_{\Lambda_n \cap \Omega \cap \Lambda} \rangle = 0$. It follows that $\lim_{n\to\infty} |\langle f, \chi_{\Omega_n \cap \Lambda_n \cap \Omega \cap \Lambda} \rangle| \leqslant \lim_{n\to\infty} \langle |f|, \chi_{\Omega_n \cap \Omega \cap \Lambda} \rangle = 0$. It remains to show that $\lim_{n\to\infty} \langle f, \chi_{\Omega_n \cap \Lambda_n} \rangle = 0$. To see this,

$$\begin{aligned}
\Omega_n \cap \Lambda_n &= (\Omega_n \cap \Lambda_n \cap \Omega) \cup (\Omega_n \cap \Lambda_n \cap \Omega^c) \\
&\subseteq (\Lambda_n \cap \Omega) \cup (\Omega_n \cap \Omega^c)
\end{aligned}$$

then

$$\begin{aligned}
\lim_{n\to\infty} |\langle f, \chi_{\Omega_n \cap \Lambda_n} \rangle| &\leqslant \lim_{n\to\infty} \langle |f|, \chi_{\Lambda_n \cap \Omega} \rangle + \lim_{n\to\infty} \langle |f|, \chi_{\Omega_n \cap \Omega^c} \rangle \\
&= (1-\lambda) \langle |f| \chi_\Omega, \chi_{\Lambda \sim \Omega} \rangle + \lambda \langle |f| \chi_{\Omega^c}, \chi_{\Omega \sim \Lambda} \rangle \\
&= 0.
\end{aligned}$$

Thus

$$\begin{aligned}
\lim_{n\to\infty} \langle f, \chi_{\Omega_n \cup \Lambda_n \cup (\Omega \cap \Lambda)} \rangle &= \lim_{n\to\infty} \langle f, \chi_{\Omega_n} + \chi_{\Lambda_n} + \chi_{\Omega \cap \Lambda} \rangle \\
&= \langle f, \lambda \chi_\Omega + (1-\lambda)\chi_\Lambda \rangle .
\end{aligned}$$

EXAMPLE 3.1. It follows from Proposition 3.2 that $F(\Omega) = u(\int_\Omega v_1\, dm, \ldots, \int_\Omega v_n\, dm)$ is a convex set function whenever $u\colon \mathbf{R}^n \to \mathbf{R}$ is convex and $v_1, \ldots, v_n \in L_1(X, \mathcal{A}, m)$.

LEMMA 3.3. *Let $(X, \mathcal{A}, m)$ be a finite atomless measure space with $L_1(X, \mathcal{A}, m)$ separable. Then for $\Omega \in \mathcal{A}$ and $\lambda \in [0, 1]$ it follows that $\lambda\chi_\Omega$ is in the weak* closure of $\chi = \{\chi_\Lambda : \Lambda \in \mathcal{A}\} \subseteq L_\infty(X, \mathcal{A}, m)$.*

Proof. Let $L_1(X, \mathcal{A}, m)$ have countable dense set $f_1, f_2, \ldots$. Applying Lemma 2.1 for each n, there exists $\Omega_n \subseteq \Omega$, $\Omega_n \in \mathcal{A}$ with $\int_{\Omega_n} f_i\, dm = \lambda \int_\Omega f_i\, dm$,

$i = 1,..., n$. Fix $f \in L_1(X, \mathcal{A}, m)$ and $\epsilon > 0$. Then there exists N such that $\|f_N - f\|_{L_1} \leqslant \epsilon/2$. For $n \geqslant N$,

$$\left| \int_{\Omega_n} f\, dm - \lambda \int_{\Omega} f\, dm \right| \leqslant \left| \int_{\Omega_n} f\, dm - \int_{\Omega_n} f_N\, dm \right| + \left| \int_{\Omega_n} f_N\, dm - \lambda \int_{\Omega} f_N\, dm \right|$$
$$+ \left| \lambda \int_{\Omega} f_N\, dm - \lambda \int_{\Omega} f\, dm \right| \leqslant \frac{\epsilon}{2} + 0 + \frac{\epsilon}{2} = \epsilon.$$

Thus χ_{Ω_n} is weak* convergent to $\lambda\chi_{\Omega}$. Q.E.D.

THEOREM 3.4 (Necessary conditions for a constrained minimum). *Let $(X, \mathcal{A}, m)$ be a finite atomless measure space with $L_1(X, \mathcal{A}, m)$ separable. Let $F, G_1, ..., G_n$ be convex set functions. Assume the* constraint qualification condition: *there exists $\Omega_1 \in \mathcal{A}$ with $G_i(\Omega_1) < 0$, $i = 1,..., n$.*

$$\textit{Let } \mu_0 = \inf F(\Omega) \textit{ subject to } \Omega \in \mathcal{A},\ G_i(\Omega) \leqslant 0,\ i = 1,..., n \tag{3.1}$$

and assume μ_0 is finite. Then there exist nonnegative reals $\lambda_1^*,..., \lambda_n^*$ such that

$$\mu_0 = \inf \left\{ F(\Omega) + \sum_{i=1}^{n} \lambda_i^* G_i(\Omega) : \Omega \in \mathcal{A} \right\}. \tag{3.2}$$

If the infimum is achieved in (3.2) *by an $\Omega^* \in \mathcal{A}$, it is achieved in* (3.1) *by Ω^* and $\sum_{i=1}^{n} \lambda_i^* G_i(\Omega^*) = 0$. Thus for any nonnegative $\lambda_1, ..., \lambda_n$ and $\Omega \in \mathcal{A}$ satisfying $G_i(\Omega) \leqslant 0$, $i = 1,..., n$ it follows that*

$$F(\Omega^*) + \sum_{i=1}^{n} \lambda_i G_i(\Omega^*) \leqslant F(\Omega^*) = F(\Omega^*) + \sum_{i=1}^{n} \lambda_i^* G_i(\Omega^*) \leqslant F(\Omega) + \sum_{i=1}^{n} \lambda_i^* G_i(\Omega).$$

Proof. Define

$$A = \{(v_0, v_1, ..., v_n): \text{there exists } \Omega \in \mathcal{A} \text{ with } v_0 \geqslant F(\Omega),\ v_i \geqslant G_i(\Omega),\ i = 1,..., n\}$$
$$B = \{(v_0, v_1, ..., v_n): v_0 < \mu_0,\ v_i < 0,\ i = 1,..., n\}.$$

We will show that $\bar{A}$, i.e. the $\mathbf{R}^{n+1}$ closure of A, is convex. Assume $\bar{A}$ is nonempty or else the result follows vacuously. Choose $\epsilon > 0$, $\lambda \in [0, 1]$ and $a_1, a_2 \in \bar{A}$. Then there exist $\Omega_1, \Omega_2 \in \mathcal{A}$ with

$$a_i \geqslant \left[F(\Omega_i) - \frac{\epsilon}{2}, G_1(\Omega_i) - \frac{\epsilon}{2}, ..., G_n(\Omega_i) - \frac{\epsilon}{2} \right], \qquad i = 1, 2.$$

By Lemma 3.3, there exist sequences $\chi_{\Omega_1^n} \xrightarrow{w^*} \lambda\chi_{\Omega_1 \sim \Omega_2}$, $\chi_{\Omega_2^n} \xrightarrow{w^*} (1 - \lambda)\chi_{\Omega_2 \sim \Omega_1}$.

Invoking the convexity of $F, G_1, \ldots, G_n$ shows that there exists n such that $A = \Omega_1{}^n \cup \Omega_2{}^n \cup (\Omega_1 \cap \Omega_2) \in \mathcal{Ol}$ satisfies

$$F(A) \leqslant \lambda F(\Omega_1) + (1-\lambda) F(\Omega_2) + \frac{\epsilon}{2}$$

$$G_i(A) \leqslant \lambda G_i(\Omega_1) + (1-\lambda)\, G_i(\Omega_2) + \frac{\epsilon}{2}, \qquad i = 1, \ldots, n.$$

Thus $\lambda a_1 + (1-\lambda) a_2 \geqslant [F(A) - \epsilon, G_1(A) - \epsilon, \ldots, G_n(A) - \epsilon]$. But ϵ is arbitrary positive hence $\lambda a_1 + (1-\lambda) a_2 \in \bar{A}$.

Now $\bar{A}$ and B are clearly disjoint, hence there exists a hyperplane separating $\bar{A}$ and B and consequently A and B. The statements of the theorem now follow in a manner similar to Theorem 2.3 and identical to the usual arguments in mathematical programming, see for example [7, p. 217]. Q.E.D.

THEOREM 3.5 (Lagrangian Duality). *Let everything be as in Theorem* 3.4. *Then*

$$\mu_0 = \sup_{\lambda_i \geqslant 0} \inf_{\Omega \in \mathcal{Ol}} \left[F(\Omega) + \sum_{i=1}^{n} \lambda_i G_i(\Omega) \right]$$

where the supremum is attained by $[\lambda_1^*, \ldots, \lambda_n^*] \geqslant 0$. *If* μ_0 *is attained in* (3.1) *by an* $\Omega^* \in \mathcal{Ol}$, *then* $\sum_{i=1}^{n} \lambda_i^* G_i(\Omega^*) = 0$ *and* Ω^* *minimizes* $F(\Omega) + \sum_{i=1}^{n} \lambda_i^* G_i(\Omega)$.

Proof. For any nonnegative $\lambda_1, \ldots, \lambda_n$

$$\begin{aligned}
&\inf \left\{ F(\Omega) + \sum_{i=1}^{n} \lambda_i G_i(\Omega) : \Omega \in \mathcal{Ol} \right\} \\
&\qquad \leqslant \inf \left\{ F(\Omega) + \sum_{i=1}^{n} \lambda_i G_i(\Omega) : \Omega \in \mathcal{Ol},\ G_i(\Omega) \leqslant 0,\ i = 1, \ldots, n \right\} \\
&\qquad \leqslant \inf\{ F(\Omega) : \Omega \in \mathcal{Ol},\ G_i(\Omega) \leqslant 0,\ i = 1, \ldots, n \} \\
&\qquad = \mu_0 .
\end{aligned}$$

Thus $\sup_{\lambda_i \geqslant 0} \inf\{F(\Omega) + \sum_{i=1}^{n} \lambda_i G_i(\Omega) : \Omega \in \mathcal{Ol}\} \leqslant \mu_0$, but this is attained for $\lambda_i = \lambda_i^*$ according to Theorem 3.4. The remaining statements of the theorem follow from Theorem 3.4. Q.E.D.

4. NUMERICAL METHODS

We now outline several numerical methods for the solution of Problem 1.1. It will be assumed that $(X, \mathcal{Ol}, m)$ is the Legesgue measure space over $X \subseteq \mathbf{R}^d$, $mX < \infty$, and that all set functions are differentiable.

Unconstrained optimization techniques will dominate our attention, it being assumed that any constraints that are present have been removed via the Langrangian techniques of Sections 2 and 3. Naturally, in the event of a constrained problem, an algorithm of the type given here would be embedded in an iterative scheme which develops the correct Lagrange multipliers, for example, a primal-dual method. We note that in the absence of convexity of both the objective and constraint set functions, the Lagrangian Duality of Theorem 3.5 may not apply and a primal-dual method can fail. For this reason, or perhaps to improve convergence, it may be necessary to employ an augmented Lagrangian instead of the usual Lagrangian. We refer to [10] for a survey of results on augmented Lagrangian methods.

This section is divided into two parts. The first employs a representation of a set in terms of finite elements whereas the second shows how an explicit representation can be avoided in a special case.

In minimizing the set function $F(\Omega)$ our basic strategy will be to seek a set satisfying the necessary conditions for optimality, as developed in Section 2.

4.1 *Finite Element Representation*

One approach which was introduced by Céa, Gioan and Michel in [1] is to partition X into a finite union of disjoint elementary sets or finite elements as follows:

Let $\{\Lambda_i^h\}_{i=1}^N$ be a family of disjoint measurable sets with $\rho(X, \bigcup_{i=1}^N \Lambda_i^h) = 0$ and $m\Lambda_i^h = h$, $i = 1,..., N$. Denote by $\mathcal{Cl}_{\Lambda^h}$ the power set $\mathscr{P}[\{\Lambda_i^h\}_{i=1}^N]$.

We can now state a simple algorithm to find an element of $\mathcal{Cl}_{\Lambda^h}$ which approximately satisfies the necessary conditions. It is a simplification of an algorithm from [1].

ALGORITHM 4.1.1. $M > 0$, $\Omega_{\text{init}} \in \mathcal{Cl}_{\Lambda^h}$.

(1) $\Omega \leftarrow \Omega_{\text{init}}$

(2) select (i) $\Lambda \in \{\Lambda_i^h\}_{i=1}^N$, $\Lambda \subseteq \Omega$ with $\int_\Lambda f_\Omega \, dm > Mh^2$ or (ii) $\Lambda \in \{\Lambda_i^h\}_{i=1}^N$, $\Lambda \subseteq \Omega^c$ with $\int_\Lambda f_\Omega \, dm < -Mh^2$ or if neither (i) or (ii) is satisfiable, stop.

(3) $\Omega \leftarrow \Omega \bigtriangleup \Lambda$

(4) go to 2.

PROPOSITION 4.1.1. *Suppose* $F\colon \mathcal{Cl} \to \mathbf{R}$ *is differentiable with derivative* f_Ω *and* $E_F(\Omega_1, \Omega_2) \leqslant M[\rho(\Omega_1, \Omega_2)]^2$, *all* $\Omega_1, \Omega_2 \in \mathcal{Cl}$. *Then Algorithm* 4.1 *stops with* Ω *satisfying*

$$\int_\Lambda f_\Omega \, dm \leqslant Mh^2, \qquad \textit{all} \qquad \Lambda \in \{\Lambda_i^h\}_{i=1}^N, \qquad \Lambda \subseteq \Omega$$

and

$$\int_{\Lambda} f_{\Omega}\, dm \geqslant -Mh^2, \qquad all \qquad \Lambda \in \{\Lambda_i^h\}_{i=1}^N, \qquad \Lambda \subseteq \Omega^c.$$

Proof. At step 2 of the algorithm, suppose (i) is satisfied. Then

$$\begin{aligned} F(\Omega \triangle \Lambda) &= F(\Omega) - \int_{\Lambda} f_{\Omega}\, dm + E_F(\Omega, \Omega \sim \Lambda) \\ &< F(\Omega) - Mh^2 + Mh^2 \\ &= F(\Omega). \end{aligned}$$

Similarly if (ii) is satisfied it also follows that $F(\Omega \triangle \Lambda) < F(\Omega)$. Thus if the algorithm does not halt at step 2, then $F(\Omega)$ decreases at each cycle. Since N is finite, such a decrease occurs only a finite number of times and the algorithm halts. At that point the statement of the proposition clearly holds. Q.E.D.

Remark. Depending on the comparative computational demands of finding f_{Ω} and performing integrals, Algorithm 4.1.1 may be improved in efficiency. We refer to [1] and [11] for other algorithms of this type for which Proposition 4.1.1 holds.

Since Algorithm 4.1.1 solves an approximated version of the problem it is important now to enquire whether as $\mathcal{O}_{\Lambda^h}$ is given a finer structure, the repeated application of Algorithm 4.1.1 will produce a sequence whose accumulation points satisfy the necessary conditions for optimality. To show this define a sequence of discretization schemes $\{\Lambda_i^h\}_{i=1}^{N(h)}$ where $h \in H = \{h_1, h_2, ...\}$ and the sequence H converges strictly monotonically to zero. Again define $\mathcal{O}_{\Lambda^h} = \mathscr{P}[\{\Lambda_i^h\}_{i=1}^{N(h)}]$. The discretization scheme is required to possess the following properties:

(1) for each $h \in H$, $\{\Lambda_i^h\}_{i=1}^{N(h)}$ is a family of disjoint measurable sets satisfying $m\Lambda_i^h = h$, $i = 1, ..., N(h)$ and $\rho(X, \bigcup_{i=1}^{N(h)} \Lambda_i^h) = 0$,

(2) given $\Omega \in \mathcal{O}$ and $\epsilon > 0$, there exists N such that $n \geqslant N$ implies there exists $A \in \mathcal{O}_{\Lambda^{h_n}}$ with $\rho(A, \Omega) < \epsilon$.

Note that property (2) provides that the discretization scheme does not deteriorate in its ability to approximate a given set as $n \to \infty$, and moreover becomes arbitrarily fine.

THEOREM 4.1.2. *Assume*

(i) $F\colon \mathcal{O} \to \mathbf{R}$ *is differentiable with derivative* f_{Ω}

(ii) $E_F(\Omega_1, \Omega_2) \leqslant M[\rho(\Omega_1, \Omega_2)]^2$, *all* $\Omega_1, \Omega_2 \in \mathcal{O}$

(iii) $\rho(\Omega, \Omega_n) \to 0$ *implies* $\| f_{\Omega} - f_{\Omega_n} \|_{L_1} \to 0$.

Suppose a discretization scheme satisfying properties (1) *and* (2) *above is employed, Algorithm* 4.1.1 *is executed successively with* $h = h_1$, $h = h_2$,... *and the resulting sets are denoted* $\Omega_1, \Omega_2, \ldots$. *Then if* $\{\Omega_n\}$ *has a* ρ*-accumulation point* Ω *it follows that* f_Ω *separates* Ω.

Proof. Rename the ρ-convergent subsequence of $\{\Omega_n\}_{n=1}^\infty$ as $\{\Omega_n\}_{n=1}^\infty$. Fix $\epsilon > 0$ and $A \subseteq \Omega$, $A \in \mathcal{O}\!\ell$. Now since $f_\Omega \in L_1(X, \mathcal{O}\!\ell, m)$, there exists $\delta > 0$ such that

$$mA < \delta \qquad \text{implies} \qquad \int_A |f_\Omega|\, dm < \frac{\epsilon}{5}. \tag{4.1.1}$$

Choose n so large that (4.1.2)–(4.1.5) hold

$$\|f_\Omega - f_{\Omega_n}\|_{L_1} < \frac{\epsilon}{5}, \tag{4.1.2}$$

there exists $A_n \in \mathcal{O}\!\ell_{A^{h_n}}$ with

$$\rho(A_n, A) \leqslant \frac{\delta}{2}, \tag{4.1.3}$$

$$\rho(\Omega, \Omega_n) \leqslant \frac{\delta}{2}, \tag{4.1.4}$$

$$Mh_n \cdot mX \leqslant \frac{\epsilon}{5}. \tag{4.1.5}$$

Then

$$\begin{aligned}
\int_A f_\Omega\, dm &\leqslant \left| \int_A f_\Omega\, dm - \int_{A_n} f_\Omega\, dm \right| + \left| \int_{A_n} f_\Omega\, dm - \int_{A_n} f_{\Omega_n}\, dm \right| \\
&\quad + \int_{A_n \sim \Omega_n} f_{\Omega_n}\, dm + \int_{A_n \cap \Omega_n} f_{\Omega_n}\, dm \\
&\leqslant \frac{\epsilon}{5} + \frac{\epsilon}{5} + \left| \int_{A_n \sim \Omega_n} f_{\Omega_n} - f_\Omega\, dm \right| + \left| \int_{A_n \sim \Omega_n} f_\Omega\, dm \right| + \int_{A_n \cap \Omega_n} f_{\Omega_n}\, dm, \\
&\qquad\qquad \text{using (4.1.3), (4.1.1), (4.1.2)} \\
&\leqslant \frac{2\epsilon}{5} + \frac{\epsilon}{5} + \frac{\epsilon}{5} + \int_{A_n \cap \Omega_n} f_{\Omega_n}\, dm, \\
&\qquad\qquad \text{using (4.1.2), (4.1.3), (4.1.4), (4.1.1).}
\end{aligned}$$

Now $A_n \cap \Omega_n \in \mathcal{O}\!\ell_{A^{h_n}}$ and $A_n \cap \Omega_n \subseteq \Omega_n$ thus by Proposition 4.1.1

$$\begin{aligned}
\int_{A_n \cap \Omega_n} f_{\Omega_n}\, dm &\leqslant Mh_n m(A_n \cap \Omega_n) \\
&\leqslant \frac{\epsilon}{5}, \qquad \text{using (4.1.5).}
\end{aligned}$$

Combining these results $\int_A f_\Omega\, dm \leqslant \epsilon$ for arbitrary positive ϵ and arbitrary $A \subseteq \Omega$, $A \in \mathcal{Cl}$. Thus $f_\Omega \leqslant 0$ a.e. on Ω. Similarly $f_\Omega \geqslant 0$ a.e. on Ω^c, i.e. f_Ω separates Ω. Q.E.D.

Theorem 4.1.2 allows the statement of an algorithm whose accumulation points are separated sets simply by repeatedly executing Algorithm 4.1.1 with $h = h_1$, $h = h_2$,... . However it is more sensible and efficient to state an algorithm in which the $n + 1$'th iteration uses the result of the n'th iteration. For this purpose it is convenient to require an additional property of the discretization scheme, namely:

$$\mathcal{Cl}_{A^{h_m}} \subseteq \mathcal{Cl}_{A^{h_n}} \qquad \text{whenever} \qquad m \leqslant n, \tag{3}$$

where (3) is understood with respect to the equivalence classes induced by the pseudometric ρ. This is a "compatibility" condition which provides that the result of one iteration can be used as an initial set for any later iteration.

4.2 *A Special Case*

We now develop a computational approach quite different from that of Section 4.1. This method has the advantage that it avoids entirely a representation in terms of elementary sets.

Attention is restricted to the following special case of the unconstrained optimization problem. It is desired to minimize $F(\Omega) = u(\int_\Omega v_1\, dm,..., \int_\Omega v_n\, dm)$ by choice of $\Omega \in \mathcal{Cl}$. F is assumed to be differentiable, with derivative $f_\Omega = \sum_{i=1}^n u_i(\int_\Omega v_1\, dm,..., \int_\Omega v_n\, dm) v_i$; according to Example 2.1 a sufficient condition for this is that u is differentiable and v_1,..., v_n are in $L_\infty(X, \mathcal{Cl}, m)$.

By Lemma 2.1,

$$\mathcal{R} = \text{Range}\left\{\left[\int_\Omega v_1\, dm,..., \int_\Omega v_n\, dm\right], \Omega \in \mathcal{Cl}\right\}$$

is a convex and compact subset of $\mathbf{R}^n$. Thus if u is lower semicontinuous a minimizing $\Omega \in \mathcal{Cl}$ necessarily exists. Furthermore if it were possible to find $\mathcal{R}$ explicitly then the problem would reduce to nonlinear minimization over a convex compact subset of Euclidean space. However in general exact determination of $\mathcal{R}$ is difficult and instead we will show that the minimization problem reduces to the solution of a nonlinear equation on $\mathbf{R}^n$.

DEFINITION 4.2.1. A function f is said to *strongly separate* $\Omega \in \mathcal{Cl}$ if $m(\Omega \triangle f^{-1}([-\infty, 0])) = 0$.

Remarks.

(i) Strong separation clearly implies separation.

(ii) If the set function derivative f_Ω has the property that $mf_\Omega^{-1}([0]) = 0$ for all $\Omega \in \mathcal{O}$ then strong separation and separation are equivalent.

Any set $\Omega \in \mathcal{O}$ induces a vector

$$\tilde{\omega}(\Omega) = \left[\int_\Omega v_1\, dm, \dots, \int_\Omega v_n\, dm\right] \in \mathbf{R}^n,$$

and any vector $\omega \in \mathbf{R}^n$ induces a set

$$\tilde{\Omega}(\omega) = \left\{x \in X\colon \sum_{i=1}^n u_i(\omega)\, v_i(x) \leqslant 0\right\} \in \mathcal{O}.$$

Define the mapping $g\colon \mathbf{R}^n \to \mathbf{R}^n$ by $g = \tilde{\omega} \circ \tilde{\Omega}$.

We now state the main result.

THEOREM 4.2.1. *If ω is a fixed point of g, then $f_{\tilde{\Omega}(\omega)}$ strongly separates $\tilde{\Omega}(\omega)$. If f_Ω strongly separates Ω then $\tilde{\omega}(\Omega)$ is a fixed point of g.*

Proof. Suppose ω is a fixed point of g. Then

$$f_{\tilde{\Omega}(\omega)} = \sum_{i=1}^n u_i(\tilde{\omega}(\tilde{\Omega}(\omega)))\, v_i = \sum_{i=1}^n u_i(g(\omega))\, v_i = \sum_{i=1}^n u_i(\omega)\, v_i\, .$$

Thus $f_{\tilde{\Omega}(\omega)}^{-1}([-\infty, 0]) = \tilde{\Omega}(\omega)$, i.e. $f_{\tilde{\Omega}(\omega)}$ strongly separates $\tilde{\Omega}(\omega)$.

Next suppose f_Ω strongly separates Ω. Then

$$g(\tilde{\omega}(\Omega)) = \tilde{\omega}(\tilde{\Omega}(\tilde{\omega}(\Omega))) = \tilde{\omega}(f_\Omega^{-1}([-\infty, 0])) = \tilde{\omega}(\Omega)$$

thus $\tilde{\omega}(\Omega)$ is a fixed point of g. Q.E.D.

It follows from Theorem 4.2.1 that if a fixed point of g were found, numerically or otherwise, it would yield a separated set and hence a candidate for a minimum of the optimization problem.

Since every solution to the optimization problem is a separated set, it would be nice to know that every separated set had a corresponding fixed point of g, but the theorem states this only for strongly separated sets. However as mentioned in an above remark, if the derivative f_Ω has the property that $mf_\Omega^{-1}([0]) = 0$ for all $\Omega \in \mathcal{O}$ then the strongly separated and separated sets coincide and the desired result is achieved. If a particular problem failed to enjoy this property it would appear that, from a numerical standpoint, a suitable perturbation of the functions $u, v_1, \dots, v_n$ should be possible to circumvent this degeneracy without significantly distorting the original problem. This is illustrated in the next example.

EXAMPLE 4.2.1. It is desired to minimize $F(\Omega) = (\int_\Omega x^2\, dm)^2$ by choice of measurable $\Omega \subseteq [0, 1]$. The solution is obviously any set of measure zero. Now $f_\Omega = 2(\int_\Omega x^2\, dm) \cdot x^2$ and

$$g(\omega) = 0, \quad \omega > 0; \qquad g(\omega) = \tfrac{1}{3}, \quad \omega \leqslant 0.$$

It is seen that g has no fixed points. This happens because any solution Ω to the problem is separated but not strongly separated by f_Ω.

Hence adopt a perturbation of the "u function." Let $F^\epsilon(\Omega) = (\epsilon^2 + \int_\Omega x^2\, dm)^2$. Then $f_\Omega{}^\epsilon = 2(\epsilon^2 + \int_\Omega x^2\, dm) \cdot x^2$ and

$$g^\epsilon(\omega) = 0, \quad \omega > -\epsilon^2; \qquad g^\epsilon(\omega) = \tfrac{1}{3}, \quad \omega \leqslant -\epsilon^2.$$

Now g^ϵ has the unique fixed point $\omega = 0$. This fixed point induces the set $\tilde{\Omega}(0) = \{0\}$ which satisfies the necessary conditions and differs negligibly from all solutions of the original problem.

Thus a problem of the type considered in this section can be solved by finding the fixed points of a mapping $g\colon \mathbf{R}^n \to \mathbf{R}^n$. Note that the evaluation of g requires multidimensional integration which may have to be performed numerically. There are many well known methods which can be employed to solve the fixed point problem ranging from numerical schemes for the solution of $g(\omega) - \omega = 0$ to the inefficient but robust methods of direct search for minimization of $\| g(\omega) - \omega \|^2$. The success or failure, and relative merits of such schemes will depend to a large extent on the particular problem. For further details and examples of the above technique we refer to [11].

5. CONCLUDING REMARKS

The extension of the results of Sections 2 and 3 to the case of a σ-finite measure space can be performed without major difficulties, see [11]. In this case it is desirable to allow extended real valued set functions in order to include, for example, the base measure of the space

The definition of convexity for a general set function was approached in Section 3 from the primal viewpoint using a concept akin to the "convex combination" of two sets. An alternative approach appealing to the dual and related to the generalized convexity of [12] is to define a convex set function F as one satisfying

$$F(\Omega) = \sup \left\{ c + \int_\Omega f\, dm.\ c \in \mathbf{R}, f \in L_1(X, \mathcal{O}\!\iota, m),\ c + \int_\Lambda f\, dm \leqslant F(\Lambda) \text{ for all } \Lambda \in \mathcal{O}\!\iota \right\}.$$

Using the techniques of [12] a result similar to Theorem 3.3 can be shown to hold.

Acknowledgment

The author is grateful for the useful suggestions and comments given to him by Professor Paul K. C. Wang of the University of California, Los Angeles.

References

1. J. Cea, A. Gioan, and J. Michel, Quelques résultats sur l'identification de domaines, *Calcolo* **10** (1973), 133–145.
2. P. K. C. Wang, On a class of optimization problems involving domain variations, *in* "International Symposium on New Trends in System Analysis, Versailles, France, Dec. 1976," Lecture Notes in Control and Information Sciences, No. 2, p. 49, Springer-Verlag, Berlin, 1977.
3. D. Begis and R. Glowinski, Application de la méthode des éléments finis à l'approximation d'un problème de domaine optimal. Méthodes de résolution des problèmes approchés, *Applied Math. Opt.* **2**, No. 2 (1975), 130–169.
4. F. Murat and J. Simon, Étude de problèmes d'optimal design, *in* "Proceedings of the Seventh IFIP Conference, Nice, France, Sept. 1975," pp. 54–62, Lecture Notes in Computer Science No. 41, Springer–Verlag, Berlin, 1976.
5. O. L. Mangasarian, "Nonlinear Programming," McGraw–Hill, New York, 1969.
6. W. Rudin, "Functional Analysis," McGraw–Hill, New York, 1973.
7. D. G. Luenberger, "Optimization by Vector Space Methods," Wiley, New York, 1969.
8. J. Von Neumann and O. Morgenstern, "Theory of Games and Economic Behavior," Princeton Univ. Press, Princeton, N.J., 1944.
9. S. P. Gudder, Convexity and Mixtures, *SIAM Review* **19**, No. 2 (1977), 221–240.
10. D. P. Bertsekas, Multiplier methods: A survey, *Automatica* **12** (1976), 133–145.
11. R. J. T. Morris, "Optimization Problems Involving Set Functions," Ph.D. dissertation, University of California, Los Angeles, March 1978; also available as UCLA Tech. Report UCLA-ENG-7815, March 1978.
12. S. Dolecki and S. Kurcyusz, Convexité généralisée et optimization, *C. R. Acad. Sci. Paris Sér. A* **283** (1976), 91–94.