AD656904

THE TECHNOLOGICAL INSTITUTE

31

NORTHWESTERN UNIVERSITY

EVANSTON, ILLINOIS

SYSTEMS RESEARCH MEMORANDUM No. .

The Technological Institute   The College of Arts and Sciences
Northwestern University

MATHEMATICAL PROGRAMMING
BY
GENERALIZED SEQUENTIAL
UNCONSTRAINED METHODS

by

Anthony V. Fiacco

Northwestern University
Evanston, Illinois
and
The Research Analysis Corporation
McLean, Virginia

June 1967

Part of the research underlying this report was undertaken for the Office of Naval Research, Contract NONR-1228(10), Project NR 047-021, and for the U. S. Army Research Office - Durham, Contract No. DA-31-124-ARO-D-322 at Northwestern University. Part of the research was also undertaken for the Research Analysis Corporation, McLean, Virginia.

SYSTEMS RESEARCH GROUP

A. Charnes, Director

Contents

1. INTRODUCTION[1]

Our objective is to provide a more general theoretical basis for those methods for solving constrained minimization problems that are based on successive unconstrained minimizations of a parametric auxiliary function. In order to do this, we first give basic defining properties of a general auxiliary function and obtain a proof of local convergence for the mildly regulated nonconvex problem. We proceed from this general result to develop a family of auxiliary functions by giving the general function more and more structure, eventually being led to the general form of the " penalty" functions that conventionally have been utilized. Similarly, the problem structure is increasingly specialized and we deal finally with the convex problem. As expected, stronger results are obtained as additional structure is assumed.

1] This paper is based on material developed in [1] and essentially constitutes a generalization of results originally obtained in [5].

The problem of interest is the general mathematical programming problem

$$\begin{array}{lll} \text{(A)} & \text{minimize } f(x) & \\ & \text{subject to } g_i(x) \geq 0 & , \ i=1,2,\ldots,p \ \text{ and} \\ & \qquad\qquad h_j(x) = 0 & , \ j=1,2,\ldots,q \ , \text{ where } x \in E^n. \end{array}$$

Since we wish to assume little more than the continuity of the problem functions for many of our results, we must account for local minima in problem (A). Even with local minima present, we could still take the point of view that only global minima are of interest, and concentrate on analyzing what is needed to bring about convergence to a global solution. However, our point of view is to attribute importance to the determination of any local minimum objective function value v^* of problem (A), and to determine at least one local minimum x^* such that $f(x^*) = v^*$. Global results follow as a byproduct of local results in this approach.

In view of this objective, we must guarantee that a sequence of points will be generated converging to any set of local minima yielding a given objective function value. For the techniques that we shall develop here, a key requirement is that, in the interior of a specified domain of interest, each such set contains a suitably prescribed closed subset. In the development given in [5], it is assumed that any such set is compact. We shall utilize this assumption, concentrating on generating bounded sequences converging to finite local solutions of (A).

We address ourselves to developing methods that generate a sequence of points converging to a local solution of (A) by means of successive unconstrained minimizations of a parametrized auxiliary function, over a specified sequence of values of the involved parameters. Also, we restrict our attention to auxiliary functions which, under suitable conditions, converge in value to the value of the objective function $f(x)$ at a local minimum of (A). In our development, we are less concerned with the form of the penalty function and concentrate on those properties the function must possess in order to define a convergent sequential unconstrained minimization scheme. Our results are considerably more general than those obtained previously.

Particular realizations of sequential unconstrained methods in the class we shall develop have been only recently validated, by Parisot [7] for the linear problem, by Pietrzykowski [8], Fiacco and McCormick [2], [3], [4], and Pomentale [9] for the convex problem, and by Stong [10], Zangwill [12] and Fiacco and McCormick [5] for nonconvex problems as well. Many of these results, and in particular the principal convergence theorems, will be largely subsumed in the development following.

In Section 2, we define a generalized auxiliary function and prove that it yields a sequence of local unconstrained minima whose limit points are local solutions of the problem, under mild regulatory conditions. The function is defined in such a manner that, if the set of points satisfying a specified constraint has a nonempty interior, then the constraint can be enforced for all points in the minimizing sequence.

In Section 3, this function is somewhat specialized to yield the general forms of "interior-feasible" and "exterior-feasible" auxiliary functions. The convergence theorems for these functions follow as immediate corollaries of the above general result.

Further specializations lead to the functions defined and developed in [5]. These subsume the class of interior point and exterior point "penalty" functions that have conventionally been utilized in this approach. A stronger characterization of convergence is given for these functions, in Section 4.

For the convex programming problem, the interior and exterior penalty functions can be defined to be convex, thereby leading to global and dual results. These were developed in [5] and are summarized in Section 5.

2. The Generalized Auxiliary Function for the Mixed Algorithm

For convenience, we shall write problem (A) in the form

(B) minimize $f(x)$ subject to $g_i(x) \geq 0$, $i = 1, 2, \ldots, q$.

The following definitions and results will be used. The first is a trivial extension of the fact that a continuous function attains its minimum on a compact set.

Lemma 2.1.

If $F(x)$ is continuous in the nonempty interior of a compact set V and such that possibly $F(y) = +\infty$ for $y \in (V - V^o)$, then there exists a finite value v and a point $\bar{x} \in V$ such that $F(\bar{x}) = v = \min_V F(x)$.

Definition.

A point $x^* \in G \equiv \{x | g_i(x) \geq 0,\ i = 1, \ldots, q\}$ is a <u>local minimum</u> of problem (B), if there exists an open set N such that $x^* \in N$ and $f(x^*) = \min_{G \cap N} f(x)$. If N can be selected such that $f(x^*) = \min_N f(x)$, x^* is referred to as an <u>unconstrained local minimum</u> of (B). Otherwise, x^* is called a <u>constrained local</u> minimum of (B).

The lemma following will be used in proving the convergence of a minimizing sequence to a prescribed compact set of minima yielding a given local minimum value of $f(x)$ in problem (B).

For significant additional generality, we require the definition of an "isolated" set.

Definition.

A nonempty set $A^* \subset A$ is called an isolated set of A if there exists a closed set E such that $E^o \supset A^*$ and such that if $x \in (E - A^*)$, then $x \notin A$.

We assume continuity of the functions of problem (B) and consider an isolated set A^* of local minima A yielding a particular local minimum value, say $f(x) = v^*$. The following result asserts that there exists a compact set S such that $S^o \supset A^*$, and such that the set of global minima of

(B*) minimize $f(x)$ subject to $x \in G \cap S$

is given by A^*.

Let $A \equiv \{x \mid f(x) = v^* \text{ and } x \text{ is a local solution of (B)}\}$.

Lemma 2.2. Existence of Compact Perturbation Set.

If the functions of (B) are continuous and if A^* is a nonempty compact isolated set of A, then there exists a compact set S such that $S^o \supset A^*$ and, if $\bar{x} \in G \cap S$ and $\min_{G \cap S} f(x) = f(\bar{x})$, then $\bar{x} \in A^*$.

Proof. By assumption there exists a closed set E such that $E^o \supset A^*$ and $x \in (E - A^*)$ implies $x \notin A$. Since A^* is compact, it follows that there exist compact sets S_k such that $A^* \subset S_k^o \subset E$ and such that $\{S_k\} \downarrow A^*$. If the conclusion of the lemma is false, then $\min_{G \cap S_k} f(x) = f(x^k) \le v^*$, with $x^k \in G \cap S_k$ and $x^k \notin A^*$ for every k. Since A^* is closed and since $\{x^k\}$ must contain a convergent subsequence $\{x^{k_j}\}$, it follows that $x^{k_j} \to \bar{x} \in A^*$.

If $f(x^{k_j}) < v^* = f(\bar{x})$ for all j, the above implies that $\bar{x}$ is not a local minimum of (B), a contradiction of the definition of A^*. If $f(x^{k_j}) = v^*$ for some $j = \bar{j}$ then, by construction of $\{S_k\}$, v^* must be the minimum value of $f(x)$ in $G \cap S_{k_j}$ for $j \ge \bar{j}$. But since $x^k \in S^o_{k_{\bar{j}}}$ for k large enough, this means x^{k_j} is a local solution of (B) with value v^* for j large enough; i. e., $x^{k_j} \in A \cap E$ and hence, $x^{k_j} \in A^*$ for all j large enough. Thus, the conclusion follows for this case, also.

Q. E. D.

The equivalent problems (A) and (B) can also be formulated as

$$\text{(M)} \qquad \text{minimize } f(x) \text{ s.t. } x \in R \cap Q,$$

where $R \equiv \{x \mid g_i(x) \geq 0,\ i = 1, \ldots, m\}$ and $Q \equiv \{x \mid g_j(x) \geq 0,\ j = m+1, \ldots, q\}$. If there are any equality constraints, we assume these appear among the last $q - m$ constraints (i. e., they are involved only in defining Q and not in defining R). However, Q may contain inequality constraints as well.

It is assumed that $R^o \neq \emptyset$ and $Q \neq \emptyset$. It is desired to solve (M) by generating a sequence of unconstrained minima of an auxiliary function in such a manner that the sequence must be restricted to R^o, but need not be restricted to Q.

As throughout, we tacitly assume that $x \in E^n$ and the functions of problem (M) are continuous. The generalized auxiliary function is defined as follows.

Defining Properties of $V(x, r, t)$.

i) $V(x, r, t)$ is continuous for $x \in R^o$, for any $r > 0$ and $t > 0$;

ii) If $\{x^k\} \subset R^o$ and $x^k \to y \in R - R^o$, where $\|y\| < \infty$, then

$$\lim_{k \to \infty} V(x^k, r, t) = +\infty \text{ for any } r > 0,\ t > 0;$$

iii) If $\{x^k\} \subset R^o$, $r_k > 0$ and $t_k > 0$ for every k, and $(x^k, r_k, t_k) \to (y, 0, +\infty)$, with $\|y\| < \infty$, then $\liminf_k V(x^k, r_k, t_k) \begin{cases} = +\infty & \text{if } y \notin Q, \\ \geq f(y) & \text{otherwise;} \end{cases}$

iv) If $y \in R^o \cap Q$, $r_k > 0$ and $t_k > 0$ for every k, and $(r_k, t_k) \to (0, +\infty)$, then $\lim_{k\to\infty} V(y, r_k, t_k) = f(y)$.

Realizations of $V(x, r, t)$ follow in the ensuing development.

With $V(x, r, t)$ so defined, a sequential technique having the desired properties can be validated under suitable conditions. Let $A \equiv \{x \mid f(x) = v^*$ and x is a local solution of (M)$\}$.

Theorem 2.1. Convergence of V-minima to Local Solutions of (M).

If the functions of (M) are continuous, $V(x, r, t)$ is as above, $R^o \cap Q \neq \emptyset$, $R \cap Q = \overline{R^o \cap Q}$, A^* is a nonempty compact isolated set of A, $r_k > 0$ and $t_k > 0$ for every k with $(r_k, t_k) \to (0, +\infty)$, then: (a) there exists a compact set S as given in Lemma 2.2, so that $S^o \supset A^*$ and, for all k large enough, the unconstrained minima $\{x^k\}$ of $\{V(x, r_k, t_k)\}$ in $R \cap S$ exist and every limit point of the uniformly bounded sequence $\{x^k\}$ is in A^*, and (b) $V(x^k, r_k, t_k) \to v^*$.

Proof. Let $V(x^k, r_k, t_k) = \inf_{R \cap S} V(x, r_k, t_k)$, where S is the compact set defined in Lemma 2.2. Recall that this also means that $R \cap S^o \supset A^*$. Since $A^* \subset R \cap Q$ and $R \cap Q = \overline{R^o \cap Q}$, it follows that $R^o \cap S \neq \emptyset$.

By property (i), $V(x, r, t)$ is continuous in $R^o \cap S$ for $r > 0$ and $t > 0$. By (ii), $V(x, r, t) \to +\infty$ as $x \to y$ if $x \in R^o$ and $y \in (R - R^o)$, for any $r > 0$, $t > 0$. These facts imply that $\{x^k\} \subset R^o \cap S$. The fact that, for any $r_k > 0$, $t_k > 0$, x^k exists minimizing $V(x, r_k, t_k)$ in the compact set $R \cap S$ follows from Lemma 2.1.

Since $R \cap S$ is compact, $\{x^k\}$ must have a convergent subsequence which, for convenience, shall still be denoted by $\{x^k\}$. Hence, we can assume $x^k \to y^o \in R \cap S$. In particular, this also means

$\|y^o\| < \infty$. We shall show that $y^o \in A^*$.

By hypothesis, $R \cap Q$ is the closure of $R^o \cap Q$ and hence, by the continuity of $f(x)$, there exists $x^o \in R^o \cap Q \cap S$ such that $f(x^o) \le v^* + \epsilon$, any $\epsilon > 0$. By definition of x^k,

$$(2.1) \qquad V(x^k, r_k, t_k) \le V(x^o, r_k, t_k)$$

and now property (iv) implies that

$$(2.2) \qquad \liminf_k V(x^k, r_k, t_k) \le \liminf_k V(x^o, r_k, t_k) = f(x^o) \le v^* + \epsilon$$

for any $\epsilon > 0$, by suitable choice of x^o.

If $y^o \notin Q$ then, by property (iii), $\lim_{k\to\infty} V(x^k, r_k, t_k) = +\infty$, which contradicts (2.2). Thus, $y^o \in R \cap S \cap Q$.

If $y^o \notin A^*$ then, by definition of S, it follows that $f(y^o) \ge v^* + \lambda$, for some $\lambda > 0$. By property (iii) this means that $\liminf_k V(x^k, r_k, t_k) \ge f(y^o) \ge v^* + \lambda$, which again contradicts (2.2), since we can select x^o above such that $\epsilon < \lambda$. Consequently, we must have $y^o \in A^*$.

Further, since $\{x^k\} \subset R^o$, $x^k \to y^o$ and $y^o \in A^* \subset R \cap S^o$, this also means that for k large enough, $x^k \in R^o \cap S^o$. Thus, the x^k are uniformly bounded unconstrained minima of $V(x, r_k, t_k)$ in $R^o \cap S^o$ for large k, and every limit point of $\{x^k\}$ is in A^*.

For part (b), since $y^o \in A^*$, (2.2) and property (iii) imply that

$$(2.3) \qquad v^* = f(y^o) \le \liminf_k V(x^k, r_k, t_k) \le v^* + \epsilon,$$

for any $\epsilon > 0$. Similarly, (2.1) and properties (iii) and (iv) yield

(2.4) $$v^* \le \lim\sup_{k\to\infty} V(x^k, r_k, t_k) \le v^* + \epsilon ,$$

for any $\epsilon > 0$. From (2.3) and (2.4), $\lim_{k\to\infty} V(x^k, r_k, t_k) = v^*$.

Q. E. D.

The following corollary gives the global result indicated in the previous discussion. It will be apparent that this result follows immediately from the theorem and the fact that the set of global minima of (M) must be closed under our assumptions on the problem functions. The second corollary is an obvious consequence.

Corollary 1. Convergence to Bounded Set of Global Minima.

Assuming the hypotheses of Theorem 2.1, if the set A^* of global minima of problem (M) is bounded, then the conclusions of the theorem hold and the sequence of unconstrained local minima $\{x^k\}$ of $\{V(x, r_k, t_k)\}$ in $R \cap S$ is such that every limit point is a global minimum of (M).

Corollary 2. Convergence to Isolated Local Minimum.

Assuming the hypotheses of Theorem 2.1, if $A^* = \{x^*\}$ and x^* is an isolated local minimum of (M), then the conclusions of the theorem hold and, furthermore, the sequence of unconstrained local minima $\{x^k\}$ of $\{V(x, r_k, t_k)\}$ in $R \cap S$ is such that $\{x^k\}$ itself converges to x^*.

Before proceeding with the general development, a few interesting facts may be noted about the general auxiliary function that has been defined. Suppose this function has the form $V[f(x), g_1(x), \ldots, g_q(x), r, t]$, V is once differentiable in $f, g_1, \ldots, g_q$, and these latter functions are differentiable in x, for $x \in R^o$ and any $r > 0$, $t > 0$. Then, if x^k is a local unconstrained minimum of V, denoting the corresponding value of V

by V^k, it follows that

$$\nabla V^k = \frac{\partial V}{\partial f(x^k)} \nabla f(x^k) + \sum_{i=1}^{q} \frac{\partial V}{\partial g_i(x^k)} \nabla g_i (x^k) = 0 .$$

This equation is the same as $\nabla L (x^k, u^k) = 0$, where L is the Lagrangian associated with Problem (M), providing we assume $\partial V/\partial f(x^k) \neq 0$ and we set

$$u^k = \frac{-\partial V/\partial g_i (x^k)}{\partial V/\partial f(x^k)} .$$

This shows how we can be led to establishing a direct connection between the conditions that hold at a local minimum of the penalty function and the conditions that hold at a local solution of (M).

The above also indicates the connection with duality in the convex programming problem that will be briefly summarized in Section 5. If V is convex increasing in the convex function f and decreasing in the concave functions $g_i(x)$, $i = 1, \ldots, q$, then it follows that $V(x, r, t)$ is a convex function in R^o. If we further assume $\partial V/\partial f > 0$, then it follows immediately that (x^k, u^k) is a dual feasible point.

Thus, the connections between the conditions associated with minimization of the penalty function and the optimality and duality conditions are apparent simply from the structure of this function, our defining properties having assumed very little concerning its particular form.

3. Generalized Interior and Exterior Auxiliary Functions

Returning to the general development, we wish to be able to solve Problem (M) when we require only that $x \in R$, or that $x \in Q$, i.e., we wish to solve the problems

M(R) minimize $f(x)$ subject to $x \in R$,

and

M(Q) minimize $f(x)$ subject to $x \in Q$,

using an auxiliary function technique analagous to that given above for Problem (M).

The Problems M(R) and M(Q) have the same basic structure, the essential difference being that we shall assume $R^o \neq \phi$ and $Q \neq \phi$ (so that we may have $Q^o = \phi$). The remaining distinction is procedural: we insist on restricting x to R^o in the course of solving M(R), whereas x need not be restricted to Q in the course of solving M(Q).

Note that M is the same as M(R) if $Q = E^n$ and M is the same as M(Q) if $R = E^n$. Since there were no restrictions on R and Q in the above development (other than $R^o \neq \phi$, $Q \neq \phi$), the above convergence theorem for Problem (M) utilizing $V(x, r, t)$ is valid for $R = E^n$ or $Q = E^n$. We shall modify the respective definitions of $V(x, r, t)$ in that we shall associate the parameter r with Problem M(R) and the parameter t with Problem M(Q).

We are thus led to the auxiliary functions for M(R) and M(Q) which we shall call $U(x, r)$ and $T(x, t)$, respectively.

We assume $R^o \neq \phi$ and essentially arrive at the defining properties of $U(x, r)$ by setting $Q = E^n$ and by suppressing t in the definition of $V(x, r, t)$.

Defining Properties of $U(x, r)$.

a) $U(x, r)$ is continuous for $x \in R^o$, for any $r > 0$;

b) If $\{x^k\} \subset R^o$ and $x^k \to y \in R - R^o$, where $\|y\| < \infty$, then $\lim_{k\to\infty} U(x^k, r) = +\infty$ for any $r > 0$;

c) If $\{x^k\} \subset R^o$, $r_k > 0$ for every k, and $(x^k, r_k) \to (y, 0)$, where $\|y\| < \infty$, then $\liminf_k U(x^k, r_k) \geq f(y)$;

d) If $y \in R^o$, $r_k > 0$ for every k, and $r_k \to 0$, then $\lim_{k\to\infty} U(y, r_k) = f(y)$.

Assuming $Q \neq \emptyset$, setting $R = E^n$ and suppressing r in the definition of $V(x, r, t)$, we obtain the defining properties of $T(x, t)$.

Defining Properties of $T(x, t)$.

A) $T(x, t)$ is continuous for $x \in E^n$, for any $t > 0$;

B) If $\{x^k\} \subset E^n$, $t_k > 0$ for every k, and $(x^k, t_k) \to (y, +\infty)$, where $\|y\| < \infty$, then $\liminf_k T(x^k, t_k) \begin{cases} = +\infty \text{ if } y \notin Q, \\ \geq f(y) \text{ otherwise;} \end{cases}$

C) If $y \in Q$, $t_k > 0$ for every k, and $t_k \to +\infty$, then $\lim_{k\to\infty} T(y, t_k) = f(y)$.

As direct consequences of these definitions and the previous results, we obtain methods for solving $M(R)$ and $M(Q)$.

Theorem 3.1. Convergence of U-Minima to Local Solutions of $M(R)$.

The convergence theorem for $V(x, r, t)$ applied to (M) is valid if $V(x, r, t)$ is replaced by $U(x, r)$, (M) is replaced by $M(R)$, $Q = E^n$, and t is suppressed, in the statement of that theorem.

Theorem 3.2. Convergence of T-Minima to Local Solutions of $M(Q)$.

The convergence theorem for $V(x, r, t)$ applied to (M) is valid if $V(x, r, t)$ is replaced by $T(x, t)$, (M) is replaced by $M(Q)$, $R = E^n$, and r is suppressed, in the statement of that theorem.

The steps in the proofs of these two theorems are precisely the same as those in the V-function theorem, with obvious modifications, on making the indicated substitutions.

4. Additional Results for General Interior and Exterior Penalty Functions

We obtain a realization of $V(x, r, t)$ in terms of $U(x, r)$ and $T(x, t)$ by defining $V_1(x, r, t) \equiv U(x, r) + T(x, t) - f(x)$. As required by the definition of $V(x, r, t)$, we assume the problem functions continuous, $R^o \neq \emptyset$, and $\Omega \neq \emptyset$. The following result is a direct consequence of the defining properties.

Lemma 4.1. $V_1(x, r, t)$ is a V-function (i.e., a function $V(x, r, t)$ satisfying (i) through (iv) above.)

"Penalty" functions used in this approach are usually obtained by adding to $f(x)$ a "penalty term" which absorbs the effects of the constraints of the given problem and the involved parameter. Towards developing these functions as particular realizations of the functions $U(x, r)$ and $T(x, t)$ defined above, we define the following. Assume $R^o \neq \emptyset$ and $\Omega \neq \emptyset$.

Defining Properties of $I(x, r)$.

a_1) $I(x, r)$ is continuous for $x \in R^o$, for any $r > 0$;

b_1) If $\{x^k\} \subset R^o$ and $x^k \to y \in R - R^o$, where $\|y\| < \infty$, then $\lim_{k\to\infty} I(x^k, r) = +\infty$ for any $r > 0$;

c_1) If $\{x^k\} \subset R^o$, $r_k > 0$ for every k, and $(x^k, r_k) \to (y, 0)$, where $\|y\| < \infty$, then $\liminf_k I(x^k, r_k) \geq 0$;

d_1) If $y \in R^o$, $r_k > 0$ for every k, and $r_k \to 0$, then $\lim_{k\to\infty} I(y, r_k) = 0$.

Defining Properties of $O(x, t)$.

A_1) $O(x, t)$ continuous for $x \in E^n$, for any $t > 0$;

E_1) If $\{x^k\} \subset E^n$, $t_k > 0$ for every k, and $(x^k, t_k) \rightarrow (y, +\infty)$, where $\|y\| < \infty$, then $\liminf_k \; O(x^k, t_k) \begin{cases} = +\infty & \text{if } y \notin \Omega \\ \geq 0 & \text{otherwise;} \end{cases}$

C_1) If $y \in \Omega$, $t_k > 0$ for every k, and $t_k \rightarrow +\infty$, then $\lim_{k \rightarrow \infty} O(y, t_k) = 0$.

With "U-function" and "T-function" meaning $U(x, r)$ and $T(x, t)$, respectively, as defined in the previous section, we obtain the following direct consequences.

Lemma 4.2. $U_1(x, r) \equiv f(x) + I(x, r)$ is a U-function.

Lemma r.3. $T_1(x, t) \equiv f(x) + O(x, t)$ is a T-function.

The proofs of the lemmas follow immediately from the continuity of $f(x)$ and the defining properties.

A further immediate consequence of the three lemmas above is the following.

Corollary 1. $V_2(x, r, t) \equiv U_1(x, r) + T_1(x, t) - f(x) = f(x) + I(x, r) + O(x, t)$ is a V-function.

This provides a realization of the V-function, associated with Problem (M), in terms of the objective function $f(x)$ of that problem, and the penalty functions associated with problems $M(R)$ and $M(Q)$ (through $U_1(x,r)$ and $T_1(x, t)$ above).

The penalty functions defined in [5] for the problems $M(R)$, $M(Q)$, and (M) are, respectively,

$$U_2(x, r) \equiv f(x) + s(r)\, I(x) ,$$

$$T_2(x, t) \equiv f(x) + p(t)\, O(x) ,$$

and

$$V_3(x, r, t) \equiv f(x) + s(r)\, I(x) + p(t)\, O(x) .$$

The penalty functions utilized to date generally subscribe to the above forms. We shall summarize the defining properties of

$s(r)$, $I(x)$, $p(t)$, and $0(x)$. It follows readily that $s(r)\ I(x)$ is an I-function (satisfies properties (a_1) through (d_1) above), and $p(t)\ 0(x)$ is an 0-function (satisfies (A_1) through (C_1) above).

Defining Properties of $I(x)$ and $s(r)$ [5].

1) $I(x)$ is continuous for $x \in R^o$;

2) If $\{x^k\} \subset R^o$ and $x^k \to y \in R - R^o$, then $\lim_{k\to\infty} I(x^k) = +\infty$;

3) $s(r)$ is a (scalar-valued) function of r, continuous for $r > 0$;

4) If $r_1 > r_2 > 0$, then $s(r_1) > s(r_2) > 0$;

5) If $r_k > 0$ for every k and $r_k \to 0$, then $\lim_{r_{k\to\infty}} s(r_k) = 0$.

Defining Properties of $0(x)$ and $p(t)$ [5].

1^1) $0(x)$ continuous for $x \in E^n$;

2^1) $0(x) \begin{cases} = 0 \text{ if } x \in C, \\ > 0 \text{ if } x \notin C; \end{cases}$

3^1 $p(t)$ is a (scalar-valued) function of t, continuous for $t > 0$;

4^1) If $t_2 > t_1 > 0$, then $p(t_2) > p(t_1) > 0$;

5^1) If $t_k > 0$ for every k and $t_k \to +\infty$, then $\lim_{k\to\infty} p(t_k) = +\infty$.

As straightforward consequences of these defining properties we obtain the following:

Lemma 4.4. $I_1(x, r) \equiv s(r)\ I(x)$ is an I-function.

Lemma 4.5. $0_1(x, t) \equiv p(t)\ 0(x)$ is an 0-function.

In view of the above results, this means that U_2 is a U-function (i.e., satisfies (a) through (d) above), T_2 is a T-function (i.e., satisfies (A) through (C) above) and V_3 is a V-function (i.e., satisfies (i) through (iv) above.)

From this point on, we shall be dealing only with the U_2, T_2, and V_3 functions defined above. For convenience, therefore, we shall revert to the notation, U, T, and V for these functions.

It is first noted from the above development that these penalty functions are special cases of the general auxiliary functions defined in Sections 2 and 3, so that the respective convergence theorems apply. Additional characterizations of convergence follow as corollaries of these theorems, from the particular structure of these functions.

In the following we assume the conditions of Theorems 3.1 and 3.2 are satisfied.

We know $\{x^k\} \subset R^o \cap S^o$ for k large, where

$$\min_{R \cap S} U(x, r_k) = U(x^k, r_k) \ ,$$

and S is the compact set defined in Lemma 2.2. Recall that $S^o \supset A^*$, a compact set of local minima associated with the value $f(x) = v^*$. Also, we know there exists at least one limit point x^* of $\{x^k\}$, and any such limit point is in A^*. The following results show that for $U(x, r)$ as presently defined, we can assure that the minimizing points yield values of f, U and I that converge monotonically.

Corollary 1 of Theorem 3.1.

If $U(x, r) = f(x) + s(r)\, I(x)$, where $I(x)$ and $s(r)$ satisfy properties (1) - (5) given above, $f(x)$ is continuous, $\{r_k\}$ is a positive strictly decreasing null sequence, and $\min_{R \cap S} U(x, r_k) = U(x^k, r_k)$, then $f(x^k) \downarrow v^*$, $I(x^k) \uparrow \gamma$, where $\gamma = +\infty$ if $x^* \in (R-R^o)$ and $|\gamma| < \infty$ if $x^* \in R^o$, and $s(r_k)\, I(x^k) \to 0$. Also, if $I(x) \geq 0$ in $R^o \cap S$, then $U(x^k, r_k) \downarrow v^*$.

Proof. By definition of x^k, $U(x^k, r_k) \leq U(x, r_k)$ for every $x \in S \cap R$, so that

(4.1) $$f(x^k) + s(r_k)\, I(x^k) \leq f(x^{k+1}) + s(r_k)\, I(x^{k+1}), \text{ and}$$

(4.2) $$f(x^{k+1}) + s(r_{k+1})\, I(x^{k+1}) \leq f(x^k) + s(r_{k+1})\, I(x^k).$$

Multiplying the first inequality by $s(r_{k+1})$, the second by $s(r_k)$, summing and rearranging yields

$$[s(r_k) - s(r_{k+1})] f(x^{k+1}) \le [s(r^k) - s(r^{k+1})] f(x^k).$$

Since $r_k > 0$ and $r_k \downarrow 0$ strictly, $s(r_k) > s(r_{k+1})$ by definition of $s(r)$, so that $f(x^{k+1}) \le f(x^k)$. This, together with (4.1), also implies that $I(x^{k+1}) \ge I(x^k)$.

Since we also have $f(x) \ge v^*$ for $x \in S \cap R$, we must have $f(x^k) \downarrow \bar{v} \ge v^*$. We know that at least one limit point of $\{x^k\}$ exists, and Theorem 3.2 assures us that limit points of $\{x^k\}$ must be in A^*, so that we must have $\bar{v} = v^*$; i.e., $f(x^k) \downarrow v^*$.

The conclusions regarding the convergence of $I(x^k)$ follow from the monotonicity shown above and from the properties (1) and (2) of $I(x)$.

By property (c) of the U function and by the result for $U(x, r)$ analagous to equation (2.1), it follows that

$$v^* \le \lim_{j\to\infty} U(x^{k_j}, r_{k_j}) \le v^* + \varepsilon$$

for <u>any</u> $\varepsilon > 0$ and <u>any</u> convergent subsequence $\{x^{k_j}\}$ of $\{x^k\}$. Hence, $U(x^{kj}, r_{k_j}) \to v^*$. It follows that $U(x^k, r_k) \to v^*$, by the compactness of $R \cap S$ and the fact that at least one convergent subsequence exists.

From the above results, we obtain immediately that $s(r_k)\, I(x^k) \to 0$.

Finally, if $I(x) \ge 0$ in $R^o \cap S^o$, then it follows that $U(x, r_{k+1}) \le U(x, r_k)$ for $r_k > r_{k+1} > 0$ and any $x \in R \cap S$. This implies $U(x^{k+1}, r_{k+1}) \le U(x^k, r_k)$ and using the result shown above, $U(x^k, r_k) \downarrow v^*$.

Q. E. D.

The analagous results for the $T(x, t)$ function can be proved in a similar manner, and are summarized as follows.

Corollary 1 of Theorem 3.2.

If $T(x, t) = f(x) + p(t) O(x)$, where $O(x)$ and $p(t)$ satisfy properties (1^1) - (5^1) given above, $f(x)$ is continuous, $\{t_k\}$ is a positive strictly increasing unbounded sequence, and $\min_S T(x, t_k) = T(x^k, t_k)$, then $T(x^k, t_k) \uparrow v^*$, $f(x^k) \uparrow v^*$, $O(x^k) \downarrow 0$, and $p(t_k) O(x^k) \to 0$.

5. Global and Dual Results for Convex Programming[1]

For the following development, we utilize the formulation

(C) minimize $f(x)$ subject to $g_i(x) \geq 0$, $i = 1, \ldots, m$.

We shall summarize in this section some of the principal results that hold when the interior and exterior penalty function methods are applied to solve (C), when this is a convex programming problem. These methods are specialized to take advantage of the convexity assumption, essentially by defining the penalty functions in such a way that these are themselves convex functions. The theorems given in the previous section are of course still valid, essentially with "local" replaced by "global." Furthermore, duality results are forthcoming that provide additional information.

We first discuss the interior point method. It is assumed that $R^o = \{x | g_i(x) > 0, i = 1, \ldots, m\} \neq \phi$. We shall further assume the penalty factor $I \equiv I[g(x)]$, where $g \equiv (g_1, \ldots, g_m)$, and that $h(r) = r$, so that the form of the interior penalty function is now $U(x, r) \equiv f(x) + rI[g(x)]$. Recall that $I[g(x)]$ is continuous in R^o and, with $\{x^k\} \subset R^o$ and $x^k \to y \in (R - R^o)$, $I[g(x^k)] \to +\infty$.

It is possible to assure the convexity of $U(x, r)$ in R^o by appropriately defining $I[g(x)]$. The following lemma is immediate and yields a suitable definition.

Lemma 5.1. Convexity of $I[g(x)]$.

If $I(g)$ is a convex decreasing function of g for $g > 0$, and the $g_i(x)$ are concave, then $I[g(x)]$ is a convex function of x in R^o.

---

[1]The material in this section is essentially a summary of some of the results obtained in [5] and is presented here for completeness. Most importantly, this development shows how the penalty function can be used to take advantage of the convexity assumption to yield global and dual results.

It is readily verified, for example, that the functions $-\Sigma \ln g_i$ and $\Sigma 1/g_i$ satisfy the conditions of the lemma. Assuming the hypothesis of the lemma holds, then $I[g(x)]$ is convex, and hence, $U(x, r) = f(x) + rI[g(x)]$ is convex in R^o, for $r > 0$.

The following theorem then holds for $U(x, r)$ as defined. Note that the conclusions are immediately implied by the interior-point auxiliary function Theorem 3.1, and Corollary 1 (Section 4) of that theorem, once the convexity assumptions are introduced.

Theorem 5.1. Interior Penalty Method Convergence to Solution of Convex Programming Problem [5].

If the solutions of the convex programming problem (C) are bounded, $R^o = \{x \mid g_i(x) > 0, i = 1, \ldots, m\} \neq 0$, $I(g)$ is defined, decreasing and convex in g when $g > 0$ and $\{r_k\} \downarrow 0$, then $U(x, r)$ is a convex function in R^o for $r > 0$, for r_k small there exists $x(r_k)$ that minimizes $U(x, r_k)$ in R^o, every local minimum of $U(x, r_k)$ in R^o is a global minimum in R^o, any limit point x^* of $\{x(r_k)\}$ solves (C), $U[x(r_k), r_k] \downarrow v^*$, $\{f[x(r_k)]\} \downarrow v^* \equiv \min_R f(x)$, and $I[x(r_k)] \uparrow \gamma$, where $\gamma = +\infty$ if $x^* \in (R - R^o)$, and $|\gamma| < \infty$ if $x^* \in R^o$, and $r_k I[x(r_k)] \to 0$.

As expected from the previous general theorem, Theorem 3.1, and the convexity assumption, the minimizing points are now global and all limit points are global solutions of (C). The monotone convergence results are also global. These results follow from the convexity of $U(x, r)$ and the well known fact that any local solution of the convex programming problem (C) is also a global solution. Hence, the compact set A^* of local minima of (C) is the <u>unique</u> set of <u>global</u> minima, and the compact set S, utilized for our previous local results, can be taken to be <u>any</u> compact set S such that $S^o \supset A^*$.

In effect, the role of convexity is essentially to make the previous theorem a "global" one. It should be mentioned that in [5] a condition is given that regularizes $I[g(x)]$ such that, for the convex problem the minimizing $x(r_k)$ exists for <u>all</u> $r_k > 0$, not merely for r_k small. A similar condition has not been obtained for the general problem.

A "dual" of problem (C) was formulated by Wolfe [11], assuming the problem functions are once differentiable. The Lagrangian function associated with problem (C) is defined as $L(x,u) \equiv f(x) - \sum_{i=1}^{m} u_i g_i(x)$. The dual of (C) is then defined as,

(D) maximize $L(x,u)$

subject to $\nabla_x L(x,u) = 0$, $u_i \geq 0$, $i = 1, \ldots, m$.

The dual relationship is defined in the sense given in the following two theorems.

Theorem 5.2. Primal-Dual Bounds.

If (C) is a convex programming problem, y is any primal feasible point, and (x,u) is any feasible point of (D), then $f(y) \geq L(x,u)$.

Theorem 5.3. Equality of Primal-Dual Values at Optimal Solutions.

If (C) is a convex programming problem and the Kuhn-Tucker constraint qualification (see Ref. [6]) holds at a solution x^* of (C), then there exists a solution (x^*, u^*) of (D) and $f(x^*) = L(x^*, u^*)$.

The nondifferentiable form of (D) is given by assuming only continuity and replacing $\nabla_x L(x,u) = 0$ by $L(x,u) = \inf_\xi L(\xi, u)$. The first dual theorem goes through easily with this change. By making this substitution and assuming $R^o \equiv \{x \mid g_i(x) > 0, \ i = 1, \ldots, m\} \neq \emptyset$, it is possible to dispense with the differentiability requirement and the Kuhn-Tucker constraint qualification, and arrive at a non-differentiable version of the second dual theorem given above, assuming only continuity.

In conjunction with the Kuhn-Tucker constraint qualification, it is relevant to point out the fact that this qualification is satisfied providing the constraints are differentiable and convex, and the interior of the constraint region defined by the nonlinear constraints is nonempty [13].

We now turn briefly to the significant duality results that follow from the $U(x, r)$ method. For this development, we give U additional structure, and now assume that the penalty factor has the form $\sum_{i=1}^{m} I_i[g_i(x)]$. The result applies to the nondifferentiable form of the dual problem,

$$(\overline{D}) \quad \text{maximize} \quad L(x,u) \text{ subject to } L(x,u) = \inf_{\xi} L(\xi,u), \ u_i \geq 0, \ i = 1, \ldots, m.$$

Theorem 5.4. Dual-Feasibility and Convergence for Interior-Point Methods [5].

Assume the conditions of Theorem 5.1. If, in addition, each $I_i(g_i)$ is differentiable in g_i when $g_i > 0$, then for r_k small enough $[x(r_k), u(r_k)]$ satisfies the constraints of $(\overline{D})$ where

$$u_i(r_k) \equiv -r_k \frac{\partial I_i\{g_i[x(r_k)]\}}{\partial g_i}, \ i = 1, \ldots, m. \text{ Furthermore,}$$

$\lim_{k \to \infty} L[x(r_k), u(r_k)] = v^* = \min_R f(x)$ so that all limit points of $\{[x(r_k), u(r_k)]\}$ solve $(\overline{D})$.

The definition of the quantities $u_i(r_k)$ and the result given above is most readily appreciated by assuming the functions differentiable, differentiating $U(x, r)$ directly, and comparing the result to $\nabla_x L(x, u)$, which would appear in the differentiable form of $(\overline{D})$.

The above result is a significant one, not generally shared by other methods. One very important application of it is that upper and lower bounds on the optimal value v^* are generated with the determination of each minimizing $x(r_k)$. This is a direct consequence of the first dual theorem given above. It provides a natural convergence criterion. Also, the Lagrange multipliers are critically involved in providing optimality conditions. In effect, the dual theorem provides an explicit relationship between $U(x,r)$ and the Lagrangian $L(x,u)$ associated with problem (C).

A parallel development for the exterior-point method for the convex problem is given in [5]. Again, similar conditions are invoked on the form of $T(x,t)$, assuring its convexity. The convergence theorem is essentially a "global" version of the general exterior-point theorem given in Section 3. Analagous to the development for the $U(x,r)$ interior method, a condition is given in [5] that guarantees the existence of the minimum of $T(x,t)$ for <u>all</u> $t > 0$, for the convex problem. Finally, a comparable dual theorem is proved. In view of the fact that the optimal value is approached from below in this method, the dual result provides only a lower bound on the optimal value.

Recalling that $R^o \neq \emptyset$ is <u>not</u> required for the feasible-exterior $T(x,t)$ - function method, the above indicated results imply, in particular, that the Kuhn-Tucker constraint qualification need not be satisfied for problem (C). This means, when conditions assure that a minimizing sequence of $T(x,t_k)$ leads to a solution of the problem, then we have a way of characterizing such a solution although the constraint qualification may not hold.

REFERENCES

[1] Fiacco, A. V., Sequential Unconstrained Minimization Methods for Nonlinear Programming, Doctoral Dissertation, Northwestern University, Evanston, Illinois, June 1967.

[2] Fiacco, A. V. and G. P. McCormick, "The Sequential Unconstrained Minimization Technique for Nonlinear Programming, A Primal-Dual Method," Management Science, 10 (2): 360-66 (1964).

[3] Fiacco, A. V. and G. P. McCormick, "The Sequential Unconstrained Minimization Technique for Convex Programming with Equality Constraints," Research Analysis Corp., McLean, Virginia, RAC-TP-155, April 1965.

[4] Fiacco, A. V. and G. P. McCormick, "The Slacked Unconstrained Minimization Technique for Convex Programming," paper presented to the 1966 IFORS Meeting, MIT, Cambridge, Mass., 29 Aug. - 2 Sept. 66, paper to be published.

[5] Fiacco, A. V. and G. P. McCormick, Monograph Draft, unpublished manuscript.

[6] Kuhn, H. W. and A. W. Tucker, "Non-Linear Programming," Proceedings of the Second Symposium on Mathematical Statistics and Probability, J. Neyman (ed), University of California Press, Berkeley, 1951, pp 481-93.

[7] Parisot, G. R., Resolution Numerique Approchee du Probleme de Programmation Lineaire par Application de la Programmation Logarithmique, Ph. D. Thesis, Universite de Lille, Lille, France, 11 March 61.

[8] Pietrzykowski, T., "Application of the Steepest Descent Method to Concave Programming," Proceedings of the IFIPS Congress, Munich, 1962, North Holland Company, Amsterdam, Holland, 1962, pp 185-89.

[9] Pomentale, T., "A New Method for Solving Conditioned Maxima Problems," J. Math. Analysis and Application, 10: 216-20 (1965).

[10] Stong, Robert E., "A Note on the Sequential Unconstrained Minimization Technique for Non-Linear Programming," Management Science, 12 (1): 142-44 (1965).

[11] Wolfe, P., "A Duality Theory for Nonlinear Programming," Quart. Appl. Math., XIX (3): (1961).

[12] Zangwill, W. I., "Non-Linear Programming via Penalty Functions," Management Science, 13 (5): 344-58 (Jan 67).

[13] Zoutendijk, G., Methods of Feasible Directions, Elsevier Publishing Company, Amsterdam and New York, 1960.

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Northwestern University	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP Not Applicable

3. REPORT TITLE

MATHEMATICAL PROGRAMMING BY GENERALIZED SEQUENTIAL UNCONSTRAINED METHODS

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report

5. AUTHOR(S) (Last name, first name, initial)

Fiacco, Anthony V.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
June 1967	28	13

8a. CONTRACT OR GRANT NO. NONR-1228(10) b. PROJECT NO. NR 047-021 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Systems Research Memorandum No. 178 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. AVAILABILITY/LIMITATION NOTICES

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Logistics and Mathematical Statistics Branch, Office of Naval Research

13. ABSTRACT

Our objective is to provide a more general theoretical basis for those methods for solving constrained minimization problems that are based on successive unconstrained minimizations of a parametric auxiliary function. In order to do this, we first give basic defining properties of a general auxiliary function and obtain a proof of local convergence for the mildly regulated nonconvex problem. We proceed from this general result to develop a family of auxiliary functions by giving the general function more and more structure, eventually being led to the general form of the "penalty" functions that conventionally have been utilized. Similarly, the problem structure is increasingly specialized and we deal finally with the convex problem. As expected, stronger results are obtained as additional structure is assumed.

DD FORM 1 JAN 64 1473

Unclassified

Security Classification

Unclassified

Security Classification

14. KEY WORDS	LINK A ROLE	LINK A WT	LINK B ROLE	LINK B WT	LINK C ROLE	LINK C WT
Nonlinear Programming (constrained minimization in E^n) Generalized Parametric Auxiliary Function Sequential Unconstrained Minimization Hierarchy of Auxiliary Functions Nonconvexity, Convexity, and Dual Results						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2*a*. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2*b*. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7*a*. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7*b*. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8*a*. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8*b*, 8*c*, & 8*d*. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9*a*. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9*b*. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through ________________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through ________________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through ________________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.

DD FORM 1 JAN 64 1473 (BACK)

Unclassified

Security Classification