AD616411

AVAILABLE ONLY FOR REFERENCE USE
AT DDC FIELD SERVICES. COPY IS
NOT AVAILABLE FOR PUBLIC SALE.

QUASI-CONCAVE PROGRAMMING

Kenneth J. Arrow
Alain C. Enthoven

Economics Division
The RAND Corporation

P-1[illegible]

December 1[illegible], 19[illegible]

Approved for [illegible]

COPY ___ OF ___
HARD COPY $.
MICROFICHE $.

DDC
JUN 22 1965
TISIA B

Reproduced by

The RAND Corporation • Santa Monica • [illegible]

The views expressed in this paper are not necessarily those of the Corporation

EVALUATION COPY

PROCESSING COPY

ARCHIVE COPY

CLEARINGHOUSE FOR FEDERAL SCIENTIFIC AND TECHNICAL INFORMATION, CFSTI
INPUT SECTION 410.11

LIMITATIONS IN REPRODUCTION QUALITY OF TECHNICAL ABSTRACT BULLETIN DOCUMENTS, DEFENSE DOCUMENTATION CENTER (DDC)

[ ] 1. AVAILABLE ONLY FOR REFERENCE USE AT DDC FIELD SERVICES. COPY IS NOT AVAILABLE FOR PUBLIC SALE.

[ ] 2. AVAILABLE COPY WILL NOT PERMIT FULLY LEGIBLE REPRODUCTION. REPRODUCTION WILL BE MADE IF REQUESTED BY USERS OF DDC.

[ ] A. COPY IS AVAILABLE FOR PUBLIC SALE.

[ ] B. COPY IS NOT AVAILABLE FOR PUBLIC SALE.

[ ] 3. LIMITED NUMBER OF COPIES CONTAINING COLOR OTHER THAN BLACK AND WHITE ARE AVAILABLE UNTIL STOCK IS EXHAUSTED. REPRODUCTIONS WILL BE MADE IN BLACK AND WHITE ONLY.

TSL-121-2 65

DATE PROCESSED:

PROCESSOR:

QUASI-CONCAVE PROGRAMMING

Kenneth J. Arrow
Alain C. Enthoven

I. INTRODUCTION

Our problem is to maximize a differentiable function, $f(x)$, over an $\underline{n}$ dimensional vector $\underline{x} = (x_1, \ldots, x_n)$, subject to the constraints $g(x) \geq 0$, where $g(x)$ is a differentiable $\underline{m}$ dimensional vector function, $g^1(x), \ldots, g^m(x)$, and $\underline{x} \geq 0$. H. W. Kuhn and A. W. Tucker, in their frequently quoted paper on Nonlinear Programming [4], proved that if $g(x)$ satisfies their <u>Constraint Qualification</u>,[1] the necessary conditions for x^0 to maximize $f(x)$ subject to $g(x) \geq 0$ and $\underline{x} \geq 0$ (the Kuhn-Tucker-Lagrange conditions, or KTL) are

$$
\text{(KTL)} \qquad
\begin{aligned}
f_x^0 + \lambda^0 g_x^0 &\leq 0 \\
x^0(f_x^0 + \lambda^0 g_x^0) &= 0 \\
\lambda^0 g(x) &= 0 \\
\lambda^0 &\geq 0,
\end{aligned}
$$

where, for example, f_x^0 is the vector of partial derivatives of $f(x)$ evaluated at the point $\underline{x}^0$.[2] Kuhn and Tucker also proved that if $f(x)$ and $g(x)$ are <u>concave</u> functions, (KTL) are sufficient conditions for a constrained maximum.

---

1. See [4], pp. 483-484, also 1].

2. In general, we denote by subscript partial differentiation with respect to the indicated arguments.

A function is <u>concave</u> if the chord joining any two points on any plane profile of its graph lies everywhere on or below the function. That is, $f(x)$ is concave if

$$(1.1) \qquad f[\theta x + (1 - \theta)x^0] \geq \theta f(x) + (1 - \theta)f(x^0) \qquad (0 \leq \theta \leq 1)$$

for all points x and x^0 in the region of definition of $f(x)$. Write (1.1) in the form

$$(1.2) \qquad \frac{f[x^0 + \theta(x - x^0)] - f(x^0)}{\theta} \geq f(x) - f(x^0) \qquad (0 < \theta \leq 1)$$

and take the limit of the left-hand side as $\theta \to 0$,[3] to obtain

$$(1.3) \qquad f_x^0(x - x^0) + f(x^0) \geq f(x)$$

which is an alternative definition of concavity for differentiable functions. The inequality (1.3) states that if $f(x)$ is concave, it lies everywhere on or below its tangent planes.

A function is <u>quasi-concave</u> if, for each real number c, the set X defined by the inequality

$$(1.4) \qquad f(x) \geq c$$

is convex. That is, $f(x)$ is quasi-concave if

$$(1.5) \qquad f(x) \geq f(x^0) \text{ implies } f[\theta x + (1 - \theta)x^0] \geq f(x^0)$$

for $0 \leq \theta \leq 1$. Now, for any x satisfying (1.4), let

---

3. See [4], pp. 485-486. A function, $f(x)$, of several variables is differentiable if $f(x+h) = f(x) + ch + eh$, where c is a vector depending on x but not on h and e is a vector which goes to zero with h. If a function is differentiable, then its partial derivatives exist, and $f_x = c$, but the existence of the partial derivatives does not necessarily imply differentiability (see [2], pp. 59-61. In particular, $f[x^0 + \theta(x - x^0)] = f(x^0) + f_x^0\theta(x - x^0) + \quad \theta$, where $\quad = e(x - x^0)$ goes to zero with θ. Then,

$$\frac{f[x^0 + \theta(x - x^0)] - f(x^0)}{\theta} = f_x^0(x - x^0) + \varepsilon \quad \to f_x^0(x - x^0) \text{ as } \theta \to 0.$$

(1.6) $$F(\theta) = f[\theta x + (1 - \theta)x^0] \geq f(x^0) = f(0).$$

Therefore, $F'(0) \geq 0$. Thus, differentiating $F(\theta)$ and setting θ equal to zero, we have

(1.7) $$f(x) \geq f(x^0) \text{ implies } f^0_x(x - x^0) \geq 0 \text{ for}$$

differentiable quasi-concave functions.[4]

It is clear from (1.1) that all concave functions are quasi-concave. It also can be shown that any monotonic non-decreasing function of a quasi-concave function -- and therefore of a concave function -- is quasi-concave.[5] However, not every quasi-concave function can be expressed as a monotonic non-decreasing function of a concave function.[6] Thus quasi-concavity is a generalization of the notion of concavity.

In terms of traditional economic theory, a concave function is one which satisfies the second order condition for a maximum, that is,

(1.8) $$d^2 f = \sum_{i=1}^{n} \sum_{j=1}^{n} f_{x_i x_j} dx_i dx_j \leq 0.$$

Quasi-concavity is a weaker condition (1.8) does not have to hold for quasi-concave functions.[7] A quasi-concave function is one that has a diminishing

4. The differentiation with respect to θ is, in effect, taking the directional derivative of $f(x)$ at x^0 in the direction of the point x. It is clear from the definition of quasi-concavity that this derivative, $f^0_x \frac{x-x^0}{d}$, where the terms $\frac{x-x^0}{d}$ are the direction cosines ($d = [\Sigma(x-x^0)^2]$), must be non-negative. For a definition of directional derivatives, see [2], pp. [illegible]. Wold [7] defines a function to be <u>convex towards the origin</u> if (1.7) holds. (His terminology is geometrically valid in the case he considers, where $x^0 = 0$); since (1.7) can be shown to imply quasi-concavity, the two definitions are equivalent.

5. (1.5) can be written $f[\theta x + (1-\theta)x^0] \geq \min[f(x), f(x^0)]$. Let ϕ be a monotonic non-decreasing transformation. Then ϕ does not reverse rankings. That is, $f(x) > f(x^0)$ implies $\phi[f(x)] \geq \phi[f(x^0)]$. Therefore,
$\phi[f(\theta x + (1-\theta)x^0] \geq \phi[\min[f(x), f(x^0)]]$. Since ϕ does not reverse rankings, $\phi[\min[f(x), f(x^0)]] = \min[\phi[f(x)], \phi[f(x^0)]]$, whence $\phi[f(x)]$ is quasi-concave.

6. For example, $f(x, y) = [(x-1) + (1-x) + (x+y)]^2$ is quasi-concave. Its contour lines are straight lines that are not parallel. See also the example at the end of Part II below. Fenchel [3], p. 154, has proved that such a function cannot be transformed into a concave function by a monotonic non-increasing transformation.

7. For example, $x_1 x_2$ is quasi-concave.

marginal rate of substitution if $f_x > 0$, or an increasing marginal rate of transformation if $f_x < 0$, between any pair of variables, or between any distinct composite variables. Let $\underline{x}^0$ and $\underline{x}^1$ be any two non-negative vectors not zero and not proportional to each other. Then, if we let

$$g(u, v) = f(u\underline{x}^0 + v\underline{x}^1), \qquad u \geq 0, \; v \geq 0, \tag{1.9}$$

$$g_u^2 \, g_{vv} - 2g_u \, g_v \, g_{uv} + g_v^2 \, g_{uu} \geq 0, \tag{1.10}$$

if $f(x)$ is quasi-concave and twice differentiable. It can also be shown that if (1.10) holds everywhere, $f(x)$ is quasi-concave.

Alternatively, if $f(x)$ is quasi-concave, $(-1)^r D_r \geq 0$, for $r = 1, \ldots, n$ and for all $\underline{x}$, where D_r is the bordered determinant

$$D_r = \begin{vmatrix} 0 & f_{x_1} & \cdots & f_{x_r} \\ f_{x_1} & f_{x_1 x_1} & \cdots & f_{x_1 x_r} \\ \vdots & \vdots & & \vdots \\ f_{x_r} & f_{x_r x_1} & \cdots & f_{x_r x_r} \end{vmatrix} . \tag{1.11}$$

Moreover, a sufficient condition for $f(x)$ to be quasi-concave for $\underline{x} \geq 0$ is that D_r have the sign $(-1)^r$ for all x and all $r = 1, \ldots, n$.[8]

We seek sufficient conditions for $\underline{x}^0 \geq 0$ to maximize $f(x)$ subject to the constraints $g(x) \geq 0$ when $f(x)$ and $g(x)$ are differentiable quasi-concave functions. It is not true that (KTL) alone are sufficient conditions for a constrained maximum, as the following examples illustrate.

Any monotonic function of one variable is clearly quasi-concave. Let

8. These propositions are frequently used in the literature on utility functions, but rigorous proofs starting from the concept of quasi-concavity seem to be lacking. We give such proofs in Part VI below.

(1.12) $$f(x) = (x - 1)^3, \quad x \geqq 0$$

and maximize it subject to the constraint

(1.13) $$g(x) = 2 - x \geqq 0.$$

If $\underline{x}^0 = 1$, $\lambda^0 = 0$, (KTL) is satisfied, yet clearly the constrained maximum occurs at $\underline{x} = 2$, not $\underline{x} = 1$.

More generally, let $F(x)$ be any quasi-concave function and $\underline{x}^0$ any point, and let

(1.14) $$f(x) = [F(x) - F(x^0)]^3.$$

Then $f(x)$ is quasi-concave and has the same maxima as $F(x)$. But $f^0_x = 0$, although $\underline{x}^0$ was chosen arbitrarily. Moreover, if $g(x)$ is any vector function for which $g(x^0) \geqq 0$, (KTL) is satisfied if $\underline{x} = \underline{x}^0$ and $\lambda^0 = 0$, although $\underline{x}^0$ certainly need not be the constrained maximum for $f(x)$ subject to $g(x) \geqq 0$.

We also seek conditions under which (KTL) will be necessary for a constrained maximum, when the constraints are quasi-concave. The following example makes clear the fact that (KTL) are not always necessary conditions, and that some additional condition must be satisfied. Maximize $x_1 x_2$ subject to the constraints $x_1 \geqq 0$, $x_2 \geqq 0$ and

(1.15) $$g(x) = (1 - x_1 - x_2)^3 \geqq 0.$$

The constrained maximum occurs at $x_1^0 = x_2^0 = \frac{1}{2}$, but there is no value of λ for which (KTL) can be satisfied at that point. This example also illustrates the fact that it is the constraint <u>function</u>, and not the constraint <u>set</u> which must satisfy the additional condition. For (1.15) and

(1.16) $$1 - x_1 - x_2 \geqq 0$$

define the same convex set. Yet (KTL) are satisfied at x^0 with $\lambda^0 =$

when the constraint is (1.16), and, in fact, in this case (KTL) is a necessary condition for a maximum. The Kuhn-Tucker Constraint Qualification is designed to meet the problem. Since it is rather complicated to apply, in Part III below, we present a simpler condition on quasi-concave constraints which, when satisfied, implies that the Constraint Qualification must be satisfied, and therefore that (KTL) are necessary for a constrained maximum.

II. SUFFICIENT CONDITIONS FOR A CONSTRAINED MAXIMUM

Let a relevant variable be one which can take on a positive value without necessarily violating the constraints. Or, more formally, x_{i_o} is a relevant variable if there is some point in the constraint set, say x^*, at which $x^*_{i_o} > 0$. Then we shall prove the following theorem:

Theorem 1: Let $f(x)$ be a differentiable quasi-concave function of the n-dimensional vector x, and let $g(x)$ be an m-dimensional differentiable quasi-concave vector function, both defined for $x \geq 0$. Let x^o and λ^o satisfy (KTL), and let one of the following conditions be satisfied:

(a) $f^o_{x_{i_o}} < 0$ for at least one variable x_{i_o}

(b) $f^o_{x_{i_o}} > 0$ for some relevant variable x_{i_o}

(c) $f^o_x \neq 0$ and $f(x)$ is twice differentiable in the neighborhood of x^o [9]

(d) $f(x)$ is concave.

Then x^o maximizes $f(x)$ subject to the constraints $g(x) \geq 0$, $x \geq 0$.

Only one of these four conditions -- and there may be others -- need be satisfied for x^o to maximize $f(x)$ subject to the constraints, if (KTL) is satisfied at x^o.[10] Condition (b) will be satisfied if $x^o_{i_o}$ $f^o_{x_{i_o}} > 0$, if any $f^o_{x_{i_o}} < 0$ and all x_{i_o} are relevant (the usual case in economic theory), or if $f^o_x = 0$ and any x_{i_o} is relevant. If no x_{i_o} is relevant, the problem is trivial.

---

9. That is, all of the second order partial derivatives of $f(x)$ exist at x^o. However, they may be equal to zero.

10. In fact, we developed a series of conditions on $g(x)$ analogous to conditions (a) through (d), only to discover that the cases in which they added anything to conditions (a) through (d) were vacuous.

From (a) and (b) it follows that $f^o_x \neq 0$ is sufficient if all x_i are relevant.

Perhaps these conditions [illegible] conditions $f(x)$ must satisfy if the theorem does not apply. Thus, from (a), $f(x)$ must be a quasi-concave function that is not concave; from (c), $f^o_x = 0$; from (b), $f^o_{x_i} = 0$ for all relevant variables. Then from (c), either $f^o_x = 0$ or $f^o_{x_i} = 0$ for all relevant variables and $f(x)$ is not twice differentiable. Thus, (KTL) fails to be sufficient in the case of the cubic transforms shown in Part I because $f^o_x = 0$. An example in which (KTL) fails but $f^o_x \neq 0$ follows this proof.

<u>Proof</u>: We use the following identity:

$$(1.1) \qquad f^o_x(x^1 - x^o) = (x^1 - x^o)(g^o_x + \lambda^o g^o_x) - \lambda^o g^o_x(x^1 - x^o).$$

If $\underline{x}^o$ satisfies (KTL) and $\underline{x}^1$ is in the constraint set, the first term on the right-hand side is non-positive. The second term on the right-hand side is also non-positive under these conditions. If $\lambda^o_j = 0$, the jth component of the term vanishes. If $\lambda^o_j > 0$, $g^j(x^o) = 0$, and the fact that $\underline{x}^1$ is in the constraint set, that is $g^j(x^1) \geq 0$, implies $g^j(x^1) \geq g^j(x^o)$ or, by (1.7), $g^j_x(x^1 - x^o) \geq 0$. Therefore, for $f(x)$ and $g(x)$ quasi-concave,

$$(1.2) \qquad g(x^1) \geq 0,\ x^1 \neq x^o \text{ implies } f^o_x(x^1 - x^o) \geq 0$$

if $\underline{x}^o$ satisfies (KTL).

(a) <u>$f^o_{x_{i_o}} \neq 0$ for at least one variable x_{i_o}.</u>

Suppose the theorem is false, and that for some $\underline{x}^1$ [illegible] $g(x^1) \geq 0$, we have $f(x^1) = f(x^o)$. We shall show that in this case, $f^o_x(x^1 - x^o) = 0$ which contradicts (1.2).

Let $\underline{h}$ be the unit vector in the i_o^{th} direction.[11] Let $x = x^1 + \epsilon \underline{h}$.

11. $\underline{h}$ is the unit vector in, e.g., the direction of x_1, $\underline{h} = (0, 1, 0, \ldots, 0)$.

Then, for $\varepsilon > 0$ sufficiently small

(2.3) $$f(x^1) \geq f(x^0) \quad .$$

From (1.7) and (2.3),

(2.4) $$f_x^0(x^1 - x^0) \geq 0.$$

Since $f^0_{x_{i_0}} < 0$, it follows that $\underline{h} f^0_x < 0$, or

(2.5) $$-\varepsilon \underline{h} f^0_x < 0.$$

Adding (2.4) and (2.5), and recalling the definition of x^1, we have

(2.6) $$f_x^0(x^1 - x^0) < 0.$$

This contradicts (2.2) so the theorem must be true.

(b) $f_{x_{i_0}} > 0$ for some relevant x_{i_0}.[12]

Suppose the theorem is false, and that for some x^1 [illegible] $g(x^1)$ [illegible] we have $f(x^1) > f(x^0)$. Again we shall show (2.6) holds and that the supposition leads to a contradiction.

Let $\underline{h}$ be the negative of the unit vector in the $\underline{i}_0^{th}$ direction and define $x^2 = x^1 + \varepsilon \underline{h}$ as before. There are two cases to be considered: $x^1_{i_0} > 0$ and $x^1_{i_0} = 0$. Consider the former case first. For $\varepsilon > 0$ sufficiently small, $x^2 \geq 0$ and (2.3) is satisfied. Again (2.3) implies (2.4). Since $f^0_{x_{i_0}} > 0$ and $\underline{h}$ is now the negative of the unit vector, again (2.5) and (2.6) and therefore the contradiction follow. Thus we have shown that

---

12. It can be shown that if we exclude case (a), $x^0 f^0_x = 0$ and (b) are equivalent. Clearly $x^0 f^0_x = 0$ implies that (b) is satisfied. But also, (b) implies $x^0 f^0_x = 0$. By (2.2), $f^0_x(x^1 - x^0) \geq 0$ or $f^0_x x^1 \geq f^0_x x^0$ for all x^1 in the constraint set. Excluding (a), $x^0 f^0_x \geq 0$, so if $x^0 f^0_x$ is not positive, $f^0_x x^1 \leq 0$. But (b) and not (a) imply $f^0_x \geq 0$ and for some x^* in the constraint set and for some i_0, $f^0_{x_{i_0}} > 0$, and $x^*_{i_0} > 0$, which implies $x^* f^0_x > 0$. Therefore $x^0 f^0_x = 0$ implies that (b) is not satisfied, whence (b) implies $x^0 f^0_x > 0$.

(2.7) $$f(x^0) \leq f(x^1) \quad \text{for } g(x^1) \geq 0,\ x^1 \geq 0,\ x^1_{i_0} > 0.$$

The inequality remains to be proved for $x^1_{i_0} = 0$. Since x_{i_0} is a relevant variable, there will exist some point in the constraint set, say x^*, for which $x^*_{i_0} > 0$. Let x^1 be any point in the constraint set with $x^1_{i_0} = 0$. Let $x(t) = (1 - t)\, x^1 + tx^*$. Since the constraint set is convex, $x(t)$ satisfies the constraints for $0 \leq t \leq 1$. Moreover $x_{i_0}(t) > 0$ if $t > 0$. Then, from (2.7)

(2.8) $$f(x^0) \leq f[x(t)] \quad \text{for } t > 0.$$

Now, $x(t)$ approaches x^1 as t approaches 0. If (2.8) must hold for t arbitrarily close to zero, it must hold for $t = 0$, and we have

(2.9) $$f(x^0) \leq f(x^1).$$

(c) <u>$f^0_x \neq 0$ and $f(x)$ is twice differentiable in the neighborhood of x^0.</u>

Partition the vector x^0 into two sub-vectors, y^0 and z^0 corresponding to the relevant and irrelevant variables respectively. Then if we exclude the two cases already covered, but assume $f^0_y \neq 0$, we have

(2.10) $$f^0_y = 0,\quad f^0_z \geq 0,\quad z^0_{i_0} > 0 \quad \text{for some } i_0.$$

By the definition of an irrelevant variable, $z^0 = 0$ and $z^1 = 0$ for all $x^1 = (y^1, z^1)$ in the constraint set. Therefore, to prove the theorem it is sufficient to prove that $f(y^0, 0) \leq f(y^1, 0)$ for all $y^1 \geq 0$.

Define the function

(2.11) $$\phi(u, v) = f[(1 - u)y^0 + uy^1, v\bar{z}] - f(y^0, 0)$$

for $0 \leq u \leq 1$, and $v \geq 0$. For any $y^1 \geq 0$ and for any $\bar{z} \geq 0$ such that $\bar{z}_{i_0} > 0$

Because it is essentially $f(x)$ with the range of variation of x restricted to a convex subset of the non-negative orthant, $\phi(u, v)$ is quasi-concave. Then we have

$$\phi(0, 0) = 0, \tag{2.12}$$

$$\phi_u(0, 0) = f_y^0(y^1 - y^0) = 0, \tag{2.13}$$

and

$$\phi_v(0, 0) = f_y^0 \bar{y} = 0. \tag{2.14}$$

We want to prove $\phi(1, 0) \geq 0$, or to disprove $\phi(1, 0) < 0$. To do so, first we shall establish the fact that within a sufficiently small neighborhood of zero, $\phi(u, 0)$ is either positive, zero, or negative (but not more than one of the three). Then we shall show that $\phi(u, 0) = 0$ and $\phi(u, 0) < 0$ in a neighborhood of zero are incompatible with $\phi(1, 0) < 0$ while $\phi(u, 0) > 0$ contradicts the hypotheses of the theorem.

First, if for some $\bar{u} > 0$, $\phi(\bar{u}, 0) > 0$, then by quasi-concavity, (1.5) and (2.12), $\phi(u, 0) \geq 0$ for all u such that $0 \leq u \leq \bar{u}$. Thus, either $\phi(u, 0) \geq 0$ or $\phi(u, 0) < 0$ for all u in the interval. If $\phi(u, 0) \geq 0$, either there exists some sequence of points u_n approaching zero on which $\phi(u, 0) = 0$, or there does not. If there does not, $\phi(u, 0) > 0$ for $u > 0$ sufficiently small. If there does, then, by quasi-concavity and (1.5), $\phi(u, 0) > 0$ in the intervals between the points in the sequence, and therefore $\phi(u, 0) > 0$ for $u > 0$ sufficiently small. Therefore, either $\phi(u, 0) > 0$, or $\phi(u, 0) = 0$, or $\phi(u, 0) > 0$ in a neighborhood of $u = 0$.

Clearly, if $\phi(1, 0) > 0$, by (1.5), $\phi(u, 0) > 0$ for all u in the interval $0 \leq u \leq 1$, and $\phi(u, 0)$ cannot be negative.

Now, suppose $\phi(u, 0) = 0$ in an neighborhood of 0. If $\phi(1, 0) \neq 0$, we must have

$$(2.15) \qquad \begin{aligned} \phi(u, 0) &= 0 \qquad 0 \leq u \leq u^* \\ \phi(u, 0) &\neq 0 \qquad u^* < u \leq 1 \end{aligned}$$

where $u^* > 0$. Since, by (2.14), $\phi_v(u^*, 0) \neq 0$, there is a solution $u(v)$ to the equation

$$(2.16) \qquad \phi[u(v), 0] = \phi(0, v),$$

with $u(v) \geq u^*$, for v sufficiently small. The solution may not be unique, but this does not matter. In any case,

$$(2.17) \qquad \lim_{v \to 0} u(v) = u^*.$$

Let $\epsilon = 1 - \frac{u^*}{u(v)}$, and form a combination of the points $[u(v), 0]$ and $(0, v)$ with the weights $1 - \epsilon$ and ϵ respectively. Then, (2.16) and (2.17) imply

$$(2.18) \qquad \phi[(1 - \epsilon)u(v), \epsilon v] = \phi(u^*, \epsilon v) \geq \phi(0, v).$$

By Rolle's Theorem (the law of the mean),

$$(2.19) \qquad \phi_v(u^*, v^*) = \frac{\phi(u^*, \epsilon v) - \phi(u^*, 0)}{\epsilon v}$$

for some v^* in the interval $0 < v^* < \epsilon v$. But $\phi(u^*, 0) = 0$, by (2.15), so that (2.18) and (2.19) imply

$$(2.20) \qquad \phi_v(u^*, v^*) \geq \frac{1}{\epsilon} \frac{\phi(0, v)}{v}.$$

Now take the limits of both sides as v approaches zero. By (2.17), ϵ approaches zero as v does. $\frac{\phi(0, v)}{v}$ approaches $\phi_v(0, 0)$ which is positive. Therefore, the right-hand side approaches infinity. Since ϕ_v is differentiable by hypothesis, it is continuous, so that the left-hand side approaches $\phi_v(u^*, 0)$ which is finite. Therefore, the hypotheses lead to a contradiction, and $\phi(u, 0) = 0$ for $u \geq u^*$ and $\phi(1, 0) \neq 0$ are incompatible.

Finally, suppose $\phi(u, 0) > 0$ for $u > 0$ sufficiently small. Define $u(v)$ as in (2.16). Now

$$\lim_{v \to 0} u(v) = 0 \tag{2.21}$$

Consider $\phi(u, v)$ on the line connecting the points $(0, v)$ and $[u(v), 0]$. Since $\phi(u, v)$ is quasi-concave, its value along this line must be greater than or equal to its value at the end points. Therefore, the directional derivative of $\phi(u, v)$ at $(0, v)$ in the direction of $[u(v), 0]$ must be non-negative. That is

$$u(v)\, \phi_u(0, v) - v\, \phi_v(0, v) \geq 0.\text{[13]} \tag{2.22}$$

This can be written

$$u(v) \frac{\phi_u(0, v)}{v} \geq \phi_v(0, v). \tag{2.23}$$

Taking limits as v, and therefore $u(v)$ approach zero, we obtain $\phi_v(0, 0) > 0$ on the right-hand side. On the left-hand side, the limit of $\frac{\phi_u(0, v)}{v}$ as v approaches zero is $\phi_{uv}(0, 0)$. The existence of this derivative is, of course, one of the hypotheses of the theorem. The limit of the left-hand side is zero, which is a contradiction. Therefore, $\phi(u, 0) > 0$ for $u > 0$ sufficiently small contradicts the hypotheses of the theorem, and part (c) of the theorem is proved.

(d) <u>$f(x)$ is concave.</u>[14]

---

13. This is an application of (1.7) to $\phi(u, v)$.

14. This is more general than the Kuhn-Tucker Theorem because the components of $g(x)$ are assumed to be quasi-concave rather than concave. See [4].

(1.3) and ([illegible].2) imply $f(x^0) \geq f(x^1)$ for all $x^1 \geq 0$, $g(x^1) \geq 0$. This completes the proof of Theorem 1.

Now we shall construct implicitly [illegible]ifferentiable (in fact, continuously differentiable) quasi-concave function which satisfies (KTL) at a point x^0 with $f^0_x \neq 0$, but which does not have a constrained maximum at that point. The example is designed to show that although the condition of twice differentiability, condition (c) of the theorem, can be weakened, it cannot be dispensed with altogether.

From the proof, it is clear that such an example must be [illegible] function, $f(x, y)$ with $f_x(0, 0) = 0$, $f(x, y)$ positive in a right-hand neighborhood, and $f_y(0, 0) = 0$.

The example will be chosen so that $f(0, y) = y$, $f_x(x, 0) = -1/\log x$ for $x \leq [illegible]$ and $1/\log x$ for $x \geq [illegible]$. Given the definition of $f(x, y)$ on the two axes, we complete the definition by requiring that all the level curves be straight lines, which insures the quasi-concavity of the example. Formally, for any fixed value of $f(x, y)$, say z, we define, $X(z)$ as the solution of the equation, $f(x, 0) = z$. Then the level curve $f(x, y) = z$ intersects the x-axis at $x = X(z)$ and the y-axis at $y = z$. If the level curve is to be a straight line, and (x, y) is any point on it, we have,

$$\frac{x}{X(z)} + \frac{y}{z} = 1. \tag{2.24}$$

For fixed x and y, then, $f(x, y)$ is the unique positive value of z for which (2.24) is satisfied (except that $f(x, y) = 0$ for $x = y = 0$).

Since $f_x(0, 0) = 0$, $f_y(0, 0) = 1$, (KTL) is satisfied at the origin for the constraint, $-y \leq 0$, with $\lambda^0 = 1$. But the origin is not a constrained maximum. It remains only to show that f_x and f_y are continuously differentiable. The construction makes clear, and it can be shown analytically, that

no difficulty could arise except possibly at the origin. The functions f_x and f_y can be evaluated from (...4) by implicit differentiation, and careful passage to the limit as x and y both approach zero shows that both are continuous, with f_x approaching zero and f_y approaching 1.

III. NECESSARY CONDITIONS FOR A CONSTRAINED MAXIMUM

Kuhn and Tucker [4] showed that (KTL) are necessary conditions for a constrained maximum provided the constraint functions $g(x)$ satisfy a condition termed by them the Constraint Qualification. To state the condition, we define a contained path in the direction $\xi = (\xi_1, \ldots, \xi_n)$ to be a vector function $\Psi(\theta)$, defined for the real variable $\theta \geq 0$ in an interval beginning at $\theta = 0$, whose values are points in the constraint set, and differentiable at $\theta = 0$ with $\Psi'(0) = \xi$. The Constraint Qualification then requires that for any $\underline{x}^o$ in the constraint set, there is a contained path with $\Psi(0) = \underline{x}^o$ in any direction ξ satisfying the conditions

(3.1) $\qquad$ if $g^j(x^o) = 0$ then $g_x^{jo}\xi \geq$,

(3.2) $\qquad$ if $x_i^o = 0$, then $\xi_i \geq 0$.[15]

To grasp the meaning of these conditions, consider any constraint, $g^j(x) \geq 0$, effective at x^o. The tangent hyperplane, $g_x^{jo}(x - x^o) = 0$, then divides the space into two half-spaces (provided $g_x^{jo} \neq 0$), one of which contains the constraint set. Then the direction satisfying (3.1) must point into or along the boundary of that half-space. A similar remark applies to the effective non-negativity constraints. Then the Constraint Qualification requires that for every direction from x^o which points into or along the boundaries of the appropriate half-spaces for each effective constraint, there is some path which begins at $\underline{x}^o$ in the direction ξ all of whose points in some neighborhood of $\underline{x}^o$ are in the constraint set. As Kuhn and Tucker point out, the Constraint Qualification is designed to rule out such singularities as outward pointing cusps at the boundary of the constraint set at which λ's satisfying (KTL) may not exist.

15. Kuhn and Tucker [4], p. 483, require the path to be differentiable but a careful reading of their proof (p. 484) shows that only differentiability at $\theta = 0$ is used.

In [1], some simpler conditions which, when satisfied, imply that the Constraint Qualification is satisfied were studied. One such condition is that $g(x)$ be linear. Another is that $g(x)$ be concave and that for some $\underline{x}^* \geq 0$, $g(x^*) > 0$ (that is, each coordinate is positive).[16] If the constraints $g(x)$ arise from a problem in activity analysis, then this condition means that it is possible to reduce all initial availabilities of primary commodities to some extent and still produce a positive amount of each intermediate and final good.

Since we are interested here in quasi-concave constraints, we shall generalize the latter condition.

<u>Theorem</u>: <u>Let $g(x)$ be an m-dimensional differentiable quasi-concave vector function. Let $g(x^*) > 0$ for some $x^* \geq 0$, and for each j let either</u>

(a) <u>$g^j(x)$ be concave, or</u>

(b) <u>for each x^0 in the constraint set, $g^j_{x^0} \neq 0$.</u>

<u>Then $g(x)$ satisfies the Constraint Qualification</u>. Therefore, if $\underline{x}^0$ maximizes any differentiable function $f(x)$ subject to $g(x) \geq 0$, (KTL) must be satisfied.

<u>Proof</u>: If $g(x^*) > 0$, then $g(x) > 0$ in some neighborhood of $\underline{x}^*$ in which we can find a point such that $\underline{x} > 0$. Hence, without loss of generality, we may suppose

16. This condition was used by M. Slater [6] in the case in which $f(x)$ is also assumed concave.

(3.3) $\quad$ [illegible]

Since $\underline{x}^*$ belongs to the constraint set, all variables are relevant. If $g^j(x^*) > g^j(x^o) = 0$, $g^j(x)$ is quasi-concave and $g_x^{j^o} \neq 0$, then by the argument used in the proof of parts (a) and (b) of Theorem 1, it follows that

$$g_x^{j^o}(x^* - x^o) > 0. \tag{3.4}$$

If $g^j(x)$ is concave, (3.4) follows from (1.3).

We now proceed along lines similar to the proof of Theorem 3 in [1]. For any ξ satisfying (3.1) and (3.2), define $\underline{x}(\theta) = \underline{x}^o + \theta\xi$. If $\underline{x}(\theta)$ belongs to the constraint set for some $\bar{\theta} > 0$, then from the linearity of $\underline{x}(\theta)$ and the convexity of the constraint set, it follows that $\underline{x}(\theta)$ belongs to the constraint set for all values of θ less than $\bar{\theta}$. Therefore, $\underline{x}(\theta)$ is a contained path in the direction ξ, and the requirement of the Constraint Qualification is satisfied. If $\underline{x}(\theta)$ does not belong to the constraint set for any $\theta > 0$, define $\mu(\theta)$ as the largest value of μ such that the point $(1 - \mu)\underline{x}^* + \mu\underline{x}(\theta)$ lies in the constraint set for $\theta > 0$. Let $\mu(0) = 1$. If $\underline{x}(\theta)$ does not belong to the constraint set for $\theta > 0$, clearly $\mu(\theta) < 1$ for $\theta > 0$. Now define

$$\psi(\theta) = [1 - \mu(\theta)]\underline{x}^* + \mu(\theta)\underline{x}(\theta), \qquad \theta \geq 0 \tag{3.5}$$

We shall show that $\psi(\theta)$ is a contained path in the direction ξ, whence again the requirement of the Constraint Qualification will be satisfied and our Theorem will be proved.

By construction $\psi(\theta)$ belongs to the constraint set for all θ, and $\psi(0) = \underline{x}^o$. Hence, we need only show that $\psi(\theta)$ is differentiable at $\theta = 0$ and $\psi'(0) = \xi$. Differentiating (3.5) and setting $\theta = 0$, we have

$$(3.6) \qquad \Psi'(0) = (x^0 - x^*)\mu'(0) + \mathcal{E}.$$

Thus, as was shown in section 4.4 of [1], showing that $\Psi'(0) = \mathcal{E}$ is equivalent to showing that $\mu(\theta)$ is differentiable at $\theta = 0$, and that $\mu'(0) = 0$.

We shall prove this by showing that the contrary would lead to a contradiction. Suppose then that $\mu(\theta)$ is not differentiable at $\theta = 0$ or that $\mu'(0) \neq 0$. Then, since $\mu(\theta) < 1 = \mu(0)$ for $\theta > 0$ there must exist a sequence of values of θ, $\{\theta_n\}$, approaching zero as $\underline{n}$ grows large such that

$$(3.7) \qquad \lim_{n \to \infty} \frac{\mu(\theta_n) - 1}{\theta_n} = -c < 0,$$

where $\underline{c}$ may possibly be infinite.

First, we shall show that $\mu(\theta)$ is continuous at $\theta = 0$, and therefore that $\mu(\theta_n)$ approaches 1 as θ_n approaches zero. Since $g(x^*) < 0$ and $\underline{x}^* \neq \underline{x}^0$, it is clear that for any given θ, in particular θ_1, $\mu_1 = \mu(\theta_1) > 0$. For any θ in the interval $0 \leq \theta \leq \theta_1$, form the convex combination (internal average) of the points $(1 - \mu_1)\underline{x}^* + \mu_1\underline{x}(\theta_1)$ and $\underline{x}^0$ with the weights

$$\theta/[\theta + \mu_1(\theta_1 - \theta)] \text{ and } \mu_1(\theta_1 - \theta)/[\theta + \mu_1(\theta_1 - \theta)]$$

respectively. The convex combination is

$$(3.8) \qquad \frac{\theta(1 - \mu_1)}{\theta + \mu_1(\theta_1 - \theta)}\,\underline{x}^* + \frac{\mu_1\theta_1}{\theta + \mu_1(\theta_1 - \theta)}\,\underline{x}(\theta).$$

Since, by the definition of $\mu(\theta)$, $(1 - \mu_1)\underline{x}^* + \mu_1\underline{x}(\theta_1)$ and $\underline{x}^0$ are both in the constraint set, the convexity of the constraint set implies that the convex combination (3.8) must also be in the constraint set. Again, by the definition of $\mu(\theta)$ and by (3.8),

$$(3.9) \qquad 1 - \mu(\theta) = \frac{\mu_1\theta_1}{\theta + \mu_1(\theta_1 - \theta)}.$$

As ϵ approaches zero, the right-hand side and therefore $\mu(\epsilon)$ approach 1. Therefore, $\mu(\epsilon)$ is continuous at $\epsilon = 0$. This implies that

(3.1[illegible])
$$\lim_{n \to \infty} \mu(\epsilon_n) = 1.$$

Now we proceed to show that (3.7) contradicts either (3.4) or (3.3). Choose a finite value of $\underline{b}$ such that $0 < b < \epsilon$, and define

(3.11)
$$x^n = [1 - \mu(\epsilon_n) - b\epsilon_n]x^* + [\mu(\epsilon_n) + b\epsilon_n]x(\epsilon_n).$$

Either there is an infinite sub-sequence on which

(3.12)
$$x^n \geq 0$$

or there is a sub-sequence on which, for each $\underline{n}$,

(3.13)
$$x_i^n < 0$$

for some $\underline{i}$.[17]

Suppose (3.12) holds. By restricting ourselves to the relevant sub-sequence, we may suppose without loss of generality that (3.12) holds for all $\underline{n}$. By the definition of $\mu(\epsilon_n)$, x^n cannot belong to the constraint set if $\epsilon_n > 0$. Therefore, for each $\underline{n}$ in the sub-sequence,

(3.14)
$$g^j(x^n) < 0$$

for some $\underline{j}$. Since the sub-sequence on which (3.12) holds is infinite, but there are only a finite number of constraints $g^j(x)$, there must be at least one constraint for which (3.14) holds for infinitely many $\underline{n}$. By restricting our attention to this sub-sequence, we may assume, again without loss of generality, that (3.14) holds for all $\underline{n}$ and some specific $\underline{j}$. Since $\underline{x}^n$ approaches $\underline{x}^0$ as $\underline{n}$ grows large, we must have $g^j(x^0) \leq 0$. Since $\underline{x}^0$ is in the constraint set,

---

17. Of course, both could hold also.

(3.15) $$g^j(x^o) = 0.$$

Now, select a direction ξ which satisfies the hypotheses of the Constraint Qualification. From (3.15) and (3.1)

(3.16) $$g_x^{j^o}\xi \geq 0.$$

From (3.14), (3.15) and quasi-concavity (1.7), it follows that

(3.17) $$g_x^j(x^n)(x^o - x^n) \geq 0,$$

or, if both sides are divided by $\theta_n > 0$,

(3.18) $$g_x^j(x^n)\ \frac{x^o - x^n}{\theta_n} \geq 0.$$

From the definition of $\underline{x}^n$, (3.11), we have

(3.19) $$\frac{x^o - x^n}{\theta_n} = \left\{\frac{1 - \mu(\theta_n)}{\theta_n} - b\right\}(x^o - x^*) - \left\{\frac{x(\theta_n) - x^o}{\theta_n}\right\}[\mu(\theta_n) + b\theta_n].$$

Recall from (3.7) that $\lim_{n\to\infty} \frac{\mu(\theta_n) - 1}{\theta_n} = -c$, while $\lim_{n\to\infty}[\mu(\theta_n) + b\theta_n] = 1$, and, from the definition of $x(\theta_n)$, $\frac{x(\theta_n) - x^o}{\theta_n} = \xi$. Thus, we have

(3.20) $$\lim_{n\to\infty} \frac{x^o - x^n}{\theta_n} = (c - b)(x^o - x^*) - \xi.$$

But since $\lim_{n\to\infty} \underline{x}^n = \underline{x}^o$, taking the limit of (3.19) as $\underline{n} \to \infty$ and applying (3.20) yields

(3.21) $$(c - b)g_x^{j^o}(x^o - x^*) - g_x^{j^o}\xi \geq 0.$$

Since $c > b$, and in view of (3.16), we must have $g_x^{j^o}(x^o - x^*) \geq 0$, which contradicts (3.4). Therefore, we have shown that when (3.12) holds, the supposition that $\mu(0) \neq 0$ leads to a contradiction.

If (3.12) does not hold, (3.13) must. Again by restricting ourselves to the relevant sub-sequence, we may suppose without loss of generality that

(3.13) holds for all $\underline{n}$, and for some $\underline{i}$. Since, by hypothesis, $x^o \geqq 0$, if (3.13) holds we must have $x_i^o = 0$. Thus, if ξ is to satisfy the hypotheses of the theorem,

$$\xi_i \geqq 0. \tag{3.22}$$

Also, from (3.13), for $\theta_n > 0$, recalling that $x_i^o = 0$,

$$\frac{x_i^o - x_i^n}{\theta_n} \geqq 0. \tag{3.23}$$

If we take the $\underline{i}$th component of (3.20), and apply (3.23), we have $(c - b)(x_i^o - x_i^*) - \xi_i \geqq 0$, or, from (3.22),

$$-(c - b)x_i^* \geqq 0, \tag{3.24}$$

or $x_i^* \leqq 0$, in contradiction to (3.3).

Thus, we have shown that the supposition that $\mu'(0) \neq 0$ leads to a contradiction. Therefore $\Psi(\theta)$ is a contained path in direction ξ, whence, given the hypotheses of the theorem a contained path exists and the theorem is proved.

If the hypotheses of Theorems 1 and 2 both hold, (KTL) are necessary and sufficient for a constrained maximum.

IV. EXTENSIONS OF THE THEOREMS

(1) Dropping the Non-negativity Constraints.

If $f(x)$ and $g(x)$ are defined for all $\underline{x}$, and not just for those values in the non-negative orthant, the conditions (a), (b), and (c) of Part II become merely $f_x^o \neq 0$, for in effect, all variables become relevant. That is, in the proof of condition (b), we can now construct $\underline{x}^2 = \underline{x}^1 + \alpha \underline{h}$, where $\underline{h}$ is the negative of the unit vector in the $\underline{i}_o^{th}$ direction, even if $\underline{x}^1_{i_o} = 0$, for the proof depends on the existence of $f(x^2)$, and not on $\underline{x}^2$ being in the constraint set. Thus we can say that (KTL) is sufficient for $\underline{x}^o$ to maximize $f(x)$ subject to $g(x) \geqq 0$, where $f(x)$ and $g(x)$ are differentiable quasi-concave functions provided that either (a) $f_x^o \neq 0$, or (b) $f(x)$ is concave.

In this case, the first two lines of (KTL) become simply $f_x^o + \lambda^o g_x^o = 0$.

The analogue of Theorem 2 also holds. If $g(x^*) > 0$ for some x^*, and for each j, $g^j(x)$ is concave or quasi-concave and $g_x^{j^o} \neq 0$ for all x^o in the constraint set, (KTL) are necessary for a constrained maximum.

(2) Equality Constraints.

The constraint $g(x) = 0$ can be expressed by the two inequality constraints $g(x) \geqq 0$ and $-g(x) \geqq 0$. Thus, if $g(x)$ and $-g(x)$ are both quasi-concave, as they will be, for example, if $g(x)$ is linear, Theorem 1 can be applied.

In this case, the last two lines of (KTL) become simply $g(x^o) = 0$.

There is no analogue of Theorem 2 here. However, we have already pointed out that if $g(x)$ is linear, (KTL) is necessary for a maximum.

(3) Unconstrained Maxima.

First, suppose that all variables must be non-negative, but that there are no other constraints. Since all variables are relevant, conditions (a), (b), and (c) of Part II become $f_x^o \neq 0$ as in (1) above. (KTL) becomes $f_x^o \leqq 0$, $x^o f_x^o = 0$. These statements together imply that x^o maximizes the quasi-concave function $f(x)$ for $x \geqq 0$ if either (a) $f_x^o \leqq 0$, $f_x^o \neq 0$, and $x^o f_x^o = 0$, or (b) $f_x^o = 0$ and $f(x)$ is concave. The first condition requires that the usual first-order conditions for a maximum be satisfied with at least one corner variable. In effect, the existence of the corner variable rules out such possibilities as that the apparent maximum was produced by a cubic transformation.

The Constraint Qualification is automatically satisfied in this case. Hence, for non-negative variables, $f_x^o \leqq 0$, $\underline{x}^o f_x^o = 0$ is necessary for an unconstrained maximum for any differentiable $f(x)$.

If the variables are not required to be non-negative, (KTL) becomes $f_x^o = 0$. As the examples at the end of Part I show, no conclusion can be drawn in general from (KTL) unless $f(x)$ is concave in which case the condition is clearly sufficient for a maximum.

V. ECONOMIC APPLICATIONS

(1) Consumer Demand.

The fundamental property of the utility function in the theory of consumer demand is that the indifference curves define convex sets or a diminishing marginal rate of substitution. Thus, the minimal property of all utility functions is quasi-concavity. The propositions of consumer demand theory such as the basic Weak Axiom of Revealed Preference follow directly from quasi-concavity without appeal to bordered determinants of partial derivatives, monotonic transformations and the like.

Let the utility function $u(x)$ be quasi-concave and assume non-satiation, that is $u^o_{x_i} > 0$ for some $\underline{x}_i$. Then the usual first order conditions are necessary and sufficient for a constrained maximum. Let $\underline{x}^o$ satisfy the conditions

$$
\begin{aligned}
& u^o_{x_i} - \lambda^o p_{x_i} \leqq 0 \qquad (i = 1, \ldots, n) \\
(5.1) \qquad & x^o_i (u^o_{x_i} - \lambda^o p_{x_i}) = 0 \\
& \lambda^o (B - \Sigma x^o_i p_{x_i}) = 0
\end{aligned}
$$

where $p_{x_i} > 0$ is the price of a unit of $\underline{x}_i$ and B is the consumer's budget. Then x^o maximizes $u(x)$ subject to the constraints $B - \Sigma x_i p_{x_i} \geqq 0$ and $\underline{x} \geqq 0$. Moreover, if $\lambda^o > 0$, and the assumption of non-satiation assures that it will be, $\underline{x}^o$ minimizes the cost of attaining $u(x^o)$, for it maximizes $-\Sigma x_i p_{x_i}$ subject to the constraints $u(x) - u(x^o) \geqq 0$ and $\underline{x} \geqq 0$.[18]

---

18. $u^o_{x_{i_o}} > 0$, $p_{x_{i_o}} > 0$, and $u^o_{x_{i_o}} - \lambda^o p_{x_{i_o}} \leqq 0$ imply $\lambda^o > 0$. The first two lines of (KTL) for the second maximum problem are $-p_{x_i} + \mu^o u^o_{x_i} \leqq 0$, and $x^o_i(-p_{x_i} + \mu^o u^o_{x_i}) = 0$, or (5.1) with $\mu^o = 1/\lambda^o$.

The sufficiency of (5.1) for consumers' demand theory is widely assumed; however, the only rigorous proof, under rather severe regularity conditions, is that of Wold [7], Theorem 6, p. 87.

(2) Production.

The theory of efficient production can now be extended to include production functions that are quasi-concave but not concave, that i to those cases in which there are increasing returns to scale but a diminishing marginal rate of substitution.

Suppose, for example, that an enterprise carries on production in a set of independent processes which transform purchased inputs into intermediate goods which are not traded on the market, and both into outputs. Let the scale or intensity of the $\underline{i}^{th}$ process be measured by the variable $\underline{x}_i$. Let the $\underline{j}^{th}$ output or input into the $\underline{i}^{th}$ process be a monotonic function $g_{ij}(x_i)$ that is positive if the commodity in question is an output of the process, negative if it is an input. Number the final outputs $j = 1, \ldots, m_1$, the purchased inputs, $j = m_1 + 1, \ldots, m_2$, the intermediate goods, $j = m_2 + 1, \ldots, m$, and let there be $\underline{n}$ processes. Then the net output or input of the $\underline{j}^{th}$ commodity will be

$$g_j(x) = \sum_{i=1}^{n} g_{ij}(x_i). \tag{5.2}$$

Now, consider the problem of deriving the minimum cost method of producing a fixed set of outputs at given input prices. Let the price of the $\underline{j}^{th}$ commodity be p_j. Then the problem is to maximize $\sum_{j=m_1+1}^{m_2} p_j g_j(x)$, subject to the output-level constraints $g_j(x) - g_j(x^o) \geq 0$, $j=1, \ldots, m_1$, and the constraints that the net outputs of the intermediate goods not be negative, or if we let ξ_j represent initial stocks, that the net consumption of intermediate goods not exceed the initial stocks, that is $g_j(x) + \xi_j \geq 0$, $j = m_2 + 1, \ldots, m$.

Under what conditions will this problem satisfy the hypotheses of Theorem 1? Since any monotonic function of one variable is quasi-concave, the functions $g_{ij}(x_i)$ are quasi-concave. But here we encounter a difference between concave and quasi-concave functions which is important from the point of view of applications to economic theory. While non-negative linear combinations of concave functions are also concave, non-negative linear combinations of quasi-concave functions are not necessarily quasi-concave. As a consequence, the hypothesis of quasi-concavity cannot replace the stronger hypothesis of concavity in many parts of economic theory.

Consider one of the output constraints, $g_j(x) - g_j(x^o) \geqq 0$, or

$$\sum_{i=1}^{n} g_{ij}(x_i) - g_j(x^o) \geqq 0. \tag{5.3}$$

For outputs, we have $g'_{ij}(x_i) \geqq 0$. If $g''_{ij}(x_i) \leqq 0$, $g_{ij}(x_i)$ is concave, and therefore, so is $g_j(x_i)$. If $g''_{ij}(x_i) > 0$ for one process, with $g''_{ij}(x_i) \leqq 0$ for all the others, $g_j(x)$ can be, though is not necessarily, quasi-concave. If $g''_{ij}(x_i) > 0$ for two or more activities, $g_j(x)$ cannot be quasi-concave. For, if $g_j(x)$ is quasi-concave, the marginal rate of substitution between any pair of inputs must be diminishing, all other inputs held constant. That is, from (1.10), holding, for example, $x_3, \ldots, x_n$ constant,

$$(g'_{2j})^2 g''_{1j} + (g'_{1j})^2 g''_{2j} \leqq 0.\text{[19]} \tag{5.4}$$

Thus, it is possible for either g''_{1j} or g''_{2j} to be positive without violating (1.10), but clearly both cannot be positive, and similarly for every other pair of processes. Therefore, $g''_{ij}(x_i) > 0$ for at most one process if $g_j(x)$ is to be quasi-concave. The same is true for the constraints on the use of

19. The inequality (5.4) is a necessary condition for (5.3) to be quasi-concave but it is not sufficient. For the corresponding sufficient condition, see Part VI below.

intermediate goods.

What about the maximand, $\sum_{j=m_1+1}^{m_2} p_j g_j(x)$? If the functions $g_j(x)$ are concave, their linear combination will be also. But if any are quasi-concave and not concave, the quasi-concavity of $\sum_{j=m_1+1}^{m_2} p_j g_j(x)$ cannot be guaranteed independently of the prices. Thus, the only way for Theorem 1 to be applicable for all sets of prices is for there to be diminishing or constant returns to scale in the use of the inputs. However, for any given set of prices we may have a limited amount of increasing returns in the use of inputs measured in money terms.

On the other hand, to apply Theorem 1 to profit maximization, we maximize $\sum_{j=1}^{m_2} p_j g_j(x)$ subject to the constraints $g_j(x) + \xi_j \geq 0$, $j = m_2+1, \ldots, m$. Now there can be a limited amount of increasing returns with regard to intermediate goods, but not, in general, with regard to outputs or to inputs purchased on the market (unless there is just one output and no purchased inputs). Again, for any given set of prices, a certain amount of increasing returns in outputs or purchased inputs measured in money can be tolerated.

Alternatively, let a firm's production function be

$$Y = K^{\alpha}L^{\beta} = [K^{\frac{\alpha}{\alpha+\beta}} L^{\frac{\beta}{\alpha+\beta}}]^{\alpha+\beta} \qquad (\alpha > 0,\ \beta > 0). \tag{5.5}$$

This function will be quasi-concave but not concave when $\alpha + \beta > 1$. Then, Theorem 1 will apply to the problem of determining the efficient combination of inputs, given any specified output, but it will not be applicable to the profit maximization problem. That is, the problem of minimizing $rK + wL$, or of maximizing $-(rK + wL)$, where r and w are the costs of a unit of K and L respectively, subject to the constraints $Y - Y^{o} \geq 0$, $L \geq 0$, $K \geq 0$, satisfies the hypotheses of Theorem 1. But the problem of maximizing

$\pi(K, L) = pK^{\alpha}L^{\beta} - rK - wL$, subject to $K \geqq 0$, $L \geqq 0$ does not satisfy the hypotheses of Theorem 1 because $\pi(K, L)$ is not quasi-concave.

(3) Welfare Economics.

Suppose that society's over-all production possibility function is quasi-concave. The problem of determining an efficient allocation of resources (a Pareto optimum) can then be formulated as the problem of maximizing the utility of one household (a quasi-concave function) subject to the constraints (also quasi-concave) that total output is within society's production possibilities and that the utilities of all other households are at least equal to specified levels.

VI. PROPERTIES OF QUASI-CONCAVE FUNCTIONS

In Part I, we gave several alternative definitions of quasi-concavity. Although the equivalence of these definitions, or their relationships when they are not strictly equivalent, seems to be rather generally understood, we have been unable to find in the literature either a proof of the equivalence of quasi-concavity and diminishing marginal rates of substitution (or increasing marginal rates of transformation), or a statement of the relationship between quasi-concavity and the signs of the bordered determinants of partial derivatives of quasi-concave functions.[20] Therefore, we provide both here.

Let $f(x)$ be a twice differentiable quasi-concave function, and let $\underline{x}^0$ and $\underline{x}^1$ be any two non-negative vectors, not zero and not proportional to each other. Let

$$g(u, v) = f(ux^0 + vx^1), \quad u, v \geqq 0. \tag{6.1}$$

Then $f(x)$ is quasi-concave if and only if $g(u, v)$ is quasi-concave for all $\underline{x}^0$ and $\underline{x}^1$. Clearly the quasi-concavity of $f(x)$ implies that of $g(u, v)$. On the other hand, if $g(u, v)$ is quasi-concave, then, in particular, for $0 \leqq \theta \leqq 1$, we have

$$f[\theta x^0 + (1 - \theta)x^1] = g(\theta, 1 - \theta) \geqq \min [g(0, 1), g(1, 0)] = \min [f(x^0), f(x^1)]. \tag{6.2}$$

If (6.2) holds for all $\underline{x}^0$ and x^1, we have the quasi-concavity of $f(x)$ by definition.

---

20. Wold [7], Theorem 5, pp. 85-86, states the relation between the signs of the bordered determinants and convexity of indifference surfaces to the origin, which is equivalent to quasi-concavity (see Theorem 4), under conditions more restrictive than those studied here.

Consider any locus of points on which $g(u, v)$ is constant. Along this locus, $\frac{du}{dv} = -\frac{g_v}{g_u}$. If $f_x > 0$ everywhere, as is normally the case in utility theory, $g_u > 0$ and $g_v > 0$, and $\frac{g_v}{g_u}$ is known as <u>the marginal rate of substitution</u> between the composite commodities $\underline{x}^0$ and $\underline{x}^1$. If $f(x)$ is quasi-concave, the marginal rate of substitution is diminishing. That is, $\frac{d}{dv}\left(\frac{g_v}{g_u}\right) \leq 0$. If $f_x^0 < 0$ everywhere, as is normally assumed in production theory ($-f_x$ being interpreted as marginal costs), $g_u < 0$ and $g_v < 0$, and $\frac{g_v}{g_u}$ is known as <u>the marginal rate of transformation</u> between the composite commodities $\underline{x}^0$ and $\underline{x}^1$, and if $f(x)$ is quasi-concave, the marginal rate of transformation is increasing. That is, $\frac{d}{dv}\left(\frac{g_v}{g_u}\right) \geq 0$.

In order to prove these statements, observe that

$$\frac{d}{dv}\left(\frac{g_v}{g_u}\right) = \frac{1}{g_u^3}\left[g_u^2 g_{vv} - 2g_u g_v g_{uv} + g_v^2 g_{uu}\right]. \tag{6.3}$$

Thus, if $g_u > 0$, we have a diminishing marginal rate of substitution, or if $g_u < 0$, we have an increasing marginal rate of transformation, if the expression in brackets in (6.3) is less than or equal to zero. Therefore, to prove our propositions, we shall prove the following theorem.

<u>**Theorem 3**</u>: <u>The twice differentiable function $g(u, v)$ with $g_u > 0$ and $g_v > 0$ everywhere, or $g_u < 0$ and $g_v < 0$ everywhere, is quasi-concave if and only if $g_u^2 g_{vv} - 2g_u g_v g_{uv} + g_v^2 g_{uu} \leq 0$.</u>

<u>Proof</u>: Since g_u and g_v are both positive or both negative, the implicit relation $g(u, v) = c$ defines u as a function of v. Let the function be

$$u = h(v). \tag{6.4}$$

Consider the case in which $g_u > 0$ and $g_v > 0$. By hypothesis, $\frac{d^2u}{dv^2} \geq 0$ so that $h(v)$ is a convex function.

Let (u^o, v^o) and (u^1, v^1) be any two points on the level curve $g(u, v) = c$. Then

$$u^o = h(v^o), \quad u^1 = h(v^1). \tag{6.5}$$

Let $(u^2, v^2) = (1 - \theta)(u^o, v^o) + \theta(u^1, v^1)$, for $0 \leq \theta \leq 1$. Then, from (6.5), and the convexity of $h(v)$,

$$h(v^2) \leq (1 - \theta)h(v^o) + \theta h(v^1) = (1 - \theta)u^o + \theta u^1 = u^2. \tag{6.6}$$

If $g_u > 0$, it follows from the definition of $h(v)$ that

$$c = g[h(v^2), v^2] \leq g(u^2, v^2), \tag{6.7}$$

so that $g(u^o, v^o) = g(u^1, v^1)$ implies

$$g[(1 - \theta)(u^o, v^o) + \theta(u^1, v^1)] \geq g(u^o, v^o). \qquad (0 \leq \theta \leq 1) \tag{6.8}$$

Quasi-concavity follows immediately. Suppose $g(u^1, v^1) > g(u^o, v^o)$. Let θ' be the largest value of θ for which

$$g[(1 - \theta)(u^o, v^o) + \theta(u^1, v^1)] = g(u^o, v^o).$$

Now, let $(u^2, v^2) = (1 - \theta')(u^o, v^o) + \theta'(u^1, v^1)$. If $0 \leq \theta \leq \theta'$, we can write $(1 - \theta)(u^o, v^o) + \theta(u^1, v^1) = (1 - t)(u^o, v^o) + t(u^2, v^2)$ where $t = \theta/\theta'$. Since $g(u^2, v^2) = g(u^o, v^o)$, we have shown that

$$g[(1 - \theta)(u^o, v^o) + \theta(u^1, v^1)] = g[(1 - t)(u^o, v^o) + t(u^2, v^2)] \geq g(u^o, v^o) \tag{6.9}$$

for $0 \leq \theta \leq \theta'$. On the other hand, by continuity and the definition of θ',

$$g[(1 - \theta)(u^o, v^o) + \theta(u^1, v^1)] > g(u^o, v^o) \tag{6.10}$$

for $\theta' < \theta \leq 1$. Thus g satisfies (1.5) and is therefore quasi-concave.

The theorem can be proved in a similar manner in the case in which $g_u < 0$ and $g_v < 0$ and there is an increasing marginal rate of transformation.

Finally, we shall prove that the quasi-concavity of $g(u, v)$ implies

$$g_u^2\, g_{vv} - 2g_u\, g_v\, g_{uv} + g_v^2\, g_{uu} \leq 0. \tag{6.11}$$

Consider any pair of points (u^o, v^o) and (u^1, v^1) such that $g(u^o, v^o) = g(u^1, v^1)$. From (1.7) we have

$$\begin{aligned} (u^1 - u^o)g_u^o + (v^1 - v^o)g_v^o &\leq 0 \\ (u^o - u^1)g_u^1 + (v^o - v^1)g_v^1 &\leq 0 \end{aligned} \tag{6.12}$$

which, when added, imply

$$(g_u^o - g_u^1)(u^1 - u^o) + (g_v^o - g_v^1)(v^1 - v^o) \leq 0. \tag{6.13}$$

Let $k = u^1 - u^o$. In the limit, for k small enough, $v^1 - v^o = -k\dfrac{g_u^o}{g_v^o}$. Substituting these relationships into (6.13) and dividing through by $-k^2$, we obtain

$$\frac{g_u\left(u^o+k,\ v^o-k\frac{g_u^o}{g_v^o}\right) - g_u(u^o,v^o)}{k} - \frac{g_u^o}{g_v^o}\,\frac{g_v\left(u^o+k,\ v^o-k\frac{g_u^o}{g_v^o}\right) - g_v(u^o,v^o)}{k} \leq 0. \tag{6.14}$$

Taking limits as k approaches zero, and multiplying both sides by g_v^2, we obtain (6.11).

Now consider the bordered determinant D_r defined by (1.11). The relationship between the property of quasi-concavity and the signs of D_r is given by the following theorem.

<u>Theorem 4</u>: <u>A sufficient condition for $f(x)$ to be quasi-concave for $x \geq 0$ is that the sign of D_r be the same as the sign of $(-1)^r$ for all x and all $r = 1, \ldots, n$. A necessary condition for $f(x)$ to be quasi-concave is that $(-1)^r D_r \geq 0$, for $r = 1, \ldots, n$, for all x.</u>

<u>Proof</u>: We shall begin by proving the sufficient condition. If $(-1)^r D_r > 0$ for all r for any point $\underline{x}^o$, then, by the usual second-order conditions for

a constrained maximum, $\underline{x}^o$ is a strict local maximum of $f(x)$ subject to the constraint $f_x^o x = f_x^o x^o$.[21] Let $\underline{x}^1 \geqq 0$ be any point $\underline{x}^1$ for which

$$f_x^o x^1 \geqq f_x^o x^o. \tag{6.15}$$

We shall prove that $f(x^1) \leqq f(x^o)$, that is, that $\underline{x}^o$ is a global constrained maximum subject to (6.15). Let

$$x(\theta) = (1 - \theta)x^o + \theta x^1, \tag{6.16}$$

and

$$F(\theta) = f[x(\theta)]. \tag{6.17}$$

Then let θ_o be the largest value of θ for which $F(\theta)$ takes its minimum in the interval $[0, 1]$. We shall show that $\theta_o < 1$ leads to a contradiction.

If $0 < \theta_o < 1$, then $F'(\theta_o) = 0$ because $F(\theta_o)$ is a minimum. If $\theta_o = 0$, then $F'(0) \geqq 0$, so that $f_x^o(x^1 - x^o) \geqq 0$. But from (6.15), $F'(0) \leqq 0$, so that $F'(0) = 0$. Hence, in either case, $F'(\theta_o) = 0$, or

$$f_x^{\theta_o}(x^1 - x^o) = 0 \qquad \text{if } 0 \leqq \theta_o < 1. \tag{6.18}$$

Since $x(\theta_o + h) - x(\theta_o) = h(x^1 - x^o)$, it follows from (6.18) that

$$f_x^{\theta_o}[x(\theta_o + h) - x(\theta_o)] = 0. \tag{6.19}$$

But, by assumption, (6.19) implies that $x(\theta_o)$ is a strict local maximum of $f(x)$ subject to $f_x^{\theta_o} x = f_x^{\theta_o} x(\theta_o)$, so that $f[x(\theta_o + h)] < f[x(\theta_o)]$, for $\underline{h}$ positive and sufficiently small. This contradicts the definition of θ_o as the minimum of $F(\theta)$. It follows that we cannot have $\theta_o < 1$, so that $\theta_o = 1$ and in particular, $F(1) \geqq F(0)$, or $f(x^1) \geqq f(x^o)$.

We have thus shown that any point $\underline{x}^o$ is a global constrained maximum of $f(x)$ subject to the constraints $x \geqq 0$ and

21. See, for example, [5], pp. 376-379.

(6.20) $$f_x^o x \leqq f_x^o x^o.$$

Now, let $\underline{x}^o$ and $\underline{x}^1$ be any two points, and let $\underline{x}^2$ be a convex combination (that is, an internal weighted average) of them. Since $f_x^2 x^2$ is an internal average of $f_x^2 x^o$ and $f_x^2 x^1$, it must be at least as great as the lesser. That is, we must have either $f_x^2 x^o \leqq f_x^2 x^2$ or $f_x^2 x^1 \leqq f_x^2 x^2$. Since x^2 maximizes $f(x)$ subject to $f_x^2 x \leqq f_x^2 x^2$, we must then have either $f(x^2) \geqq f(x^o)$ or $f(x^2) \geqq f(x^1)$, and, in either case

(6.21) $$f(x^2) \geqq \min [f(x^o), f(x^1)],$$

so that $f(x)$ is quasi-concave.

To prove the necessity condition, first consider any $x^o > 0$. If $f_x^o = 0$, $D_r = 0$ and the necessity condition is automatically satisfied. If $f_x^o \neq 0$, consider the maximization of $f(x)$ subject to the constraint (6.20). Since all variables are relevant and (KTL) is satisfied at $\underline{x}^o$ with $\lambda^o = 1$, it follows from Theorem 1 that $\underline{x}^o$ is the constrained maximum. Therefore, $\underline{x}^o$ certainly is a local constrained maximum of $f(x)$ subject to $f_x^o x = f_x^o x^o$, for which the conditions $(-1)^r D_r \geqq 0$ are necessary.[22] By continuity, this condition also holds when x^o has one or more components that are equal to zero.[23]

In Part V, we discussed a necessary condition for the quasi-concavity of a function of the form

(6.22) $$g(x) = \sum_{i=1}^{n} g_i(x_i).$$

---

22. See [5], ibid.

23. Let $x^1 > 0$ and $x(t) = (1 - t)x^o + tx^1$. Then $x(t) > 0$ for $t > 0$ whence $(-1)^r D_r(t) \geqq 0$ for $t > 0$, where $D_r(t)$ is D_r evaluated at the point $x(t)$. From this it follows that $(-1)^r D_r(0) \geqq 0$.

We now apply Theorem 4 to obtain a necessary and sufficient condition. In this case, $g_{x_i x_j} = 0$ for $i \neq j$. Let $g_{x_i} = g_i'$, and $g_{x_i x_i} = g_i''$, and let

$$P_r = \prod_{i=1}^{r} g_i''. \tag{6.23}$$

Then, by expansion of D_r, it is easy to see that

$$D_r = -(g_r')^2 P_{r-1} + g_r'' D_{r-1}. \tag{6.24}$$

If we assume, for simplicity, that $g_i'' \neq 0$ for all i, then

$D_r/P_r = -(g_r')^2/g_r'' + (D_{r-1}/P_{r-1})$.

Since $D_1/P_1 = -(g_1')^2/g_1''$, it easily follows by induction that,

$$(-1)^r D_r/(-1)^r P_r = -\sum_{i=1}^{r} (g_i')^2/g_i''. \tag{6.25}$$

If $g_i'' < 0$ for all $\underline{i}$, then $(-1)^r P_r > 0$ for all $\underline{r}$ and the right-hand side of (6.25) is positive, from which it follows that $(-1)^r D_r > 0$ for all $\underline{r}$, and $g(x)$ is quasi-concave, indeed concave.

Suppose $g_i'' > 0$ for two or more values of $\underline{i}$. By renumbering, we may suppose that $g_1'' > 0$, $g_2'' > 0$. Then $P_2 > 0$. From (6.25), with $r = 2$, we will have $D_2 < 0$, so that $g(x)$ is not quasi-concave.

In the remaining case, $g_i'' > 0$ for exactly one value of $\underline{i}$; we may suppose,

$$g_i'' < 0 \ (i < n), \ g_n'' > 0. \tag{6.26}$$

Then,

$$(-1)^r P_r > 0 \quad (r < n), \ (-1)^n P_n < 0. \tag{6.27}$$

The right-hand side of (6.25) is positive for $r = n$. With the aid of (6.27) we have that $(-1)^r D_r > 0$ for $r < n$. To insure quasi-concavity, it is

sufficient that, $(-1)^n D_n > 0$. In view of (6.27) and (6.25), this is equivalent to,

$$\sum_{i=1}^{n} (g_i')^2/g_i'' > 0. \tag{6.28}$$

Since the first n-1 terms are negative, this means that the last term must outweigh them all, or that g_n'' must be sufficiently small relative to $(g_n')^2$. This places an upper limit on the permissible rate of increasing returns in the nth process. The stronger the rate of diminishing returns in the other processes, the greater is the permissible rate of increasing returns in the nth process.

REFERENCES

1. Arrow, Kenneth J., Leonid Hurwicz, and Hirofumi Uzawa, "Constraint Qualifications in Maximization Problems," Naval Research Logistics Quarterly, forthcoming.

2. Courant, Richard, Differential and Integral Calculus, New York: Interscience Publishers, 1936.

3. Fenchel, Werner, Convex Cones, Sets and Functions, Logistics Research Project, Princeton University, September 1953 (mimeographed).

4. Kuhn, H. W., and A. W. Tucker, "Nonlinear Programming," in J. Neyman (ed), Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, 1951, pp. 481-492.

5. Samuelson, Paul A., Foundations of Economic Analysis, Cambridge, Harvard University Press, 1948.

6. Slater, Morton, "Lagrange Multipliers Revisited: A Contribution to Non-Linear Programming," Cowles Commission Discussion Paper, Math. 403, November, 1950.

7. Wold, Herman, Demand Analysis, New York: Wiley, 1953.