View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

ADVANCES IN APPLIED MATHEMATICS 5, 398–415 (1984)

Equivalence of Saddle-Points and Optima for Non-concave Programmes

GEOFFREY HEAL

Columbia Business School, New York, New York 10027

A basic result of optimisation theory is that a saddle-point of the Lagrangian is an optimum of the associated programming problem, independently of any concavity assumptions. It is also well known that under concavity assumptions the two are equivalent; i.e., an optimum is always a saddle-point. It is demonstrated that this basic equivalence of saddle-points and optima in fact holds for a much larger class of problems, which are not necessarily concave, but are equivalent to concave programmes up to a diffeomorphism. This class generalises the class of geometric programmes. © 1984 Academic Press, Inc.

1. INTRODUCTION

The analysis of the relationship between constrained optima and saddle-points has played a crucial role both in the economics and in the optimisation literature. In economics, one uses this relationship both to characterise optima and to provide a basis for adjustment processes designed to reach them. (See, for example, Arrow and Hurwicz [2], Arrow, Hurwicz, and Uzawa [3], and Heal [18, Chap. 4].) In optimisation, the association between optima and saddle-points is also used both to provide a characterisation of optima and to provide the basis of gradient methods for locating them.

It is well known that one part of this association, namely, the fact that a saddle-point is an optimum, does not depend on concavity (see, e.g., Uzawa [30]). However, the converse, i.e., the equivalence of an optimum to a saddle-point, is usually proven via a proof in which the separating hyperplane theorem, and hence concavity, plays a vital role (see again Uzawa [30], or Intriligator [19]).

There are a number of earlier works which in some way generalise this. For example, Arrow and Hurwicz [1, 2] show that under certain circumstances the Lagrangian of a non-concave programme may be transformed in such a way that a local optimum is a local saddle-point, and Rockafellar

0196-8858/84 $7.50
Copyright © 1984 by Academic Press, Inc.
All rights of reproduction in any form reserved.

[27] gives more general results of this type. The analysis of this paper applies to a more restricted class of cases than these, but for that class gives much stronger results. In particular, I show below that in fact the equivalence between saddle-points and optima is not restricted to concave programming problems, but continues to hold for a much larger class which may be strongly non-concave. These are programmes which, under a suitable diffeomorphism of the choice space, become concave programmes. They can therefore be described as topologically equivalent (in the sense of a diffeomorphism) to concave programming problems, and I use the term *concave transformable* to describe them. For any such programme, a point is an optimum if and only if it is a saddle-point of the normal (untransformed, unaugmented) Lagrangian. Hence all of the duality results associated with concave programmes are available, a point of considerable economic significance. I also show that, as in concave programming problems, all critical points of these problems are global maxima. In addition, the maximum value of the objective function, as a function of the constraint, is shown to be a concave function for these problems. In economic terms, they thus display a form of diminishing returns, in spite of the fact that the various defining functions need not be concave.

A particular class of programmes with this property has been identified in the literature as *geometric* or *extended geometric* programmes (see, for example, Duffin, Peterson, and Zener [15] or Peterson [25]). These are problems in which, typically, a polynomial is minimised subject to constraints involving products of the variables, and they have a natural change of variable which renders them concave programmes. Standard Kuhn–Tucker theory can then be shown to apply to these, so that the results of the present paper can be obtained. However, it should be noted that concave transformability is a much weaker condition than that of being a geometric programme, so that although geometric programmes provide a useful illustration of the class of concave-transformable programmes, they by no means exhaust this class. A particular example of a concave-transformable programme is given in Section 4. The idea of concave transformability is also related to that of generalised convexity, and this connection is discussed in Section 5.

At an intuitive level, it seems plausible that the property that any critical point is a global maximum is invariant under a diffeomorphism of the space, and indeed this part of the proof is relatively standard. What is more surprising is that the property of being a saddle-point of the Lagrangian is invariant in this way, as a saddle-point is a very geometric structure and so seems less robust. In terms of applications, this is clearly an important part of the result, because, as already mentioned, it makes a range of powerful duality results involving shadow prices available for a significant class of non-concave problems.

This idea of concave transformability, or of topological equivalence to a concave problem, is closely related to a concept discussed by Brown and Heal in [7]. That paper addresses the issue of whether, in a general equilibrium model with economies of scale in production and hence non-convex production sets, there exists a Pareto-efficient marginal cost pricing general equilibrium. It is shown that this may not be the case. Several sufficient conditions are given for the existence of an efficient equilibrium, one of which is that the economy be topologically equivalent to a convex economy. This concept of topological equivalence to a convex economy formalises the idea that, although there are increasing returns in production, in some basic sense this does not matter, as returns to scale increase less rapidly than marginal utility diminishes. It is therefore the topological rather than the geometric structure of the economy that is at stake when deriving conditions for Pareto efficiency.

In both Brown and Heal [7] and the present paper, the mathematical point which underlies the basic results is that the structure of the critical points of a real-valued function defined on a manifold depends on the topological characteristics of that manifold, and so may be an invariant preserved by topological equivalence. It is of course this kind of consideration which forms the basis of Morse theory (see, for example, Milnor [22] and Morse [23]).

There are a number of earlier works which are related to the present paper, either in that they are concerned with the saddle-point-optimum equivalence for non-concave problems or in that they bring topological techniques to bear on non-concave optimisation problems. In the first category are the works of Arrow and Hurwicz and Rockafellar, already cited, which show that under certain circumstances the Lagrangian of a non-concave programming problem may be transformed in such a way that a local optimum is a local saddle-point. Note that in comparison, the results in this paper are global, and *do not require* that the Lagrangian be transformed.

Other papers which rely on topological techniques rather than on convexity in the study of optimisation problems are those of Chichilnisky and Kalman [12], Brown, Heal, and Westhoff [9], Fujiwara [16], Spingarn and Rockafellar [28], and Yun [31]. These papers apply techniques from differential topology to the characterisation of solutions to non-concave optimisation problems, and to the study of their comparative static properties.

The remainder of the paper is organised as follows. The next section presents the basic notation and definitions to be used. In Section 3 the main theorems are established. Section 4 contains an example of a non-concave but concave-transformable programme. Section 5 relates the idea of concave transformability to that of generalised convexity, Section 6 contains a

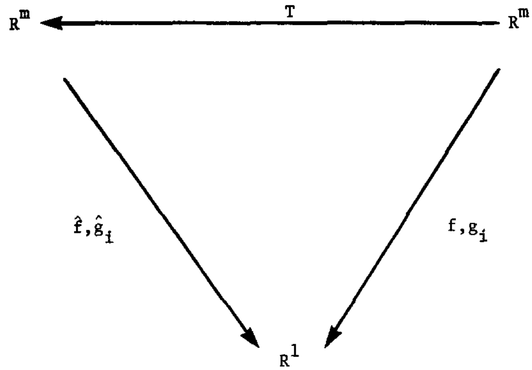

FIG. 1. Definition of the transformed problem.

discussion, and Section 7 concludes with some conjectures about possible extensions of the present result.

2. NOTATION AND DEFINITIONS

We shall assume x, the choice variable, to be a vector in Euclidean m-space R^m. Our concern is with the problem

$$\text{maximise } f(x) \text{ subject to } g_i(x) \geq 0, \qquad i = 1, 2, \ldots, n, \tag{1}$$

where f: $R^m \to R^1$, f is C^1 and g_i: $R^m \to R^1$, g_i is C^1, $i = 1, \ldots, n$. It is assumed that problem (1) satisfies the constraint qualification

$$\text{There exists } \bar{x} \text{ such that } g_i(\bar{x}) > 0, i = 1, \ldots, n. \tag{2}$$

Now let T be a regular C^1 diffeomorphism[1] of R^m to itself. Define the following functions, which will serve as objective and constraints in the transformed problem

$$\hat{f}(x') = f \circ T^{-1}(x') = f(T^{-1}(x'))$$

$$\hat{g}_i(x') = g_i \circ T^{-1}(x') = g_i(T^{-1}(x')), \qquad i = 1, 2, \ldots, n.$$

The commutative diagram shown in Fig. 1 illustrates these definitions.

The next step is to define the transformed programming problem:

$$\text{Maximise } \hat{f}(x') \text{ subject to } \hat{g}_i(x') \geq 0, \qquad i = 1, \ldots, n. \tag{3}$$

[1]A C^1 diffeomorphism is a continuous map which is injective, surjective, and has a continuous inverse, and is once continuously differentiable. A function is regular if any of its values is a regular value, i.e., if for all x in $f^{-1}(y)$, y in the range of f, $Df(x)$ is onto.

The functions defining this problem are the composition of those defining (1) with a diffeomorphism of the choice space. We can now define the basic concept of concave transformability:

DEFINITION. The programming problem (1) is said to be *concave transformable* if there exists a regular C^1 diffeomorphism T: $R^m \to R^m$ such that the transformed problem (3) is a concave programming problem; i.e., $\hat{f}$ and $\hat{g}_i$, $i = 1, 2, \ldots, n$, are all concave functions.

The geometric intuition underlying this analysis is given in Fig. 2 for the case where $n = 1$ and $m = 2$. The left-hand side shows the untransformed problem (1): although the feasible set is non-convex, there is a unique point of tangency between a contour of f and its boundary, and this is the global maximum. The transformation T sends the feasible set and the contours of f into the standard convex configuration shown on the right: here the image under T of for example $f^{-1}(\alpha)$ becomes the inverse image of α under a new function $\hat{f}$, which clearly is $f \circ T^{-1}$. The basic idea, then, is that if we have a "well-behaved" non-convex problem as on the left, it can be deformed into a convex problem. For this, the standard results are available, and can be pulled back to the non-convex problem. Figure 3 shows a non-convex problem which is not convex transformable.

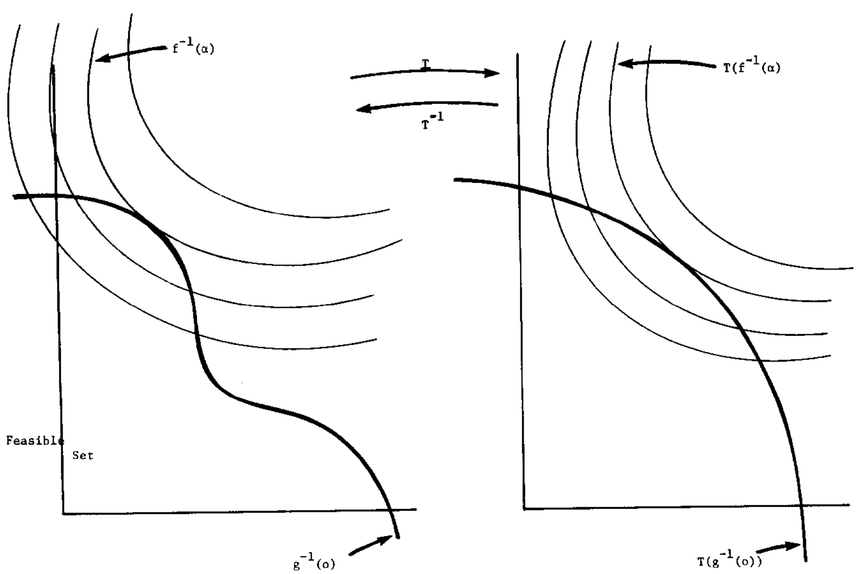

FIG. 2. Concave transformability in a simple case.

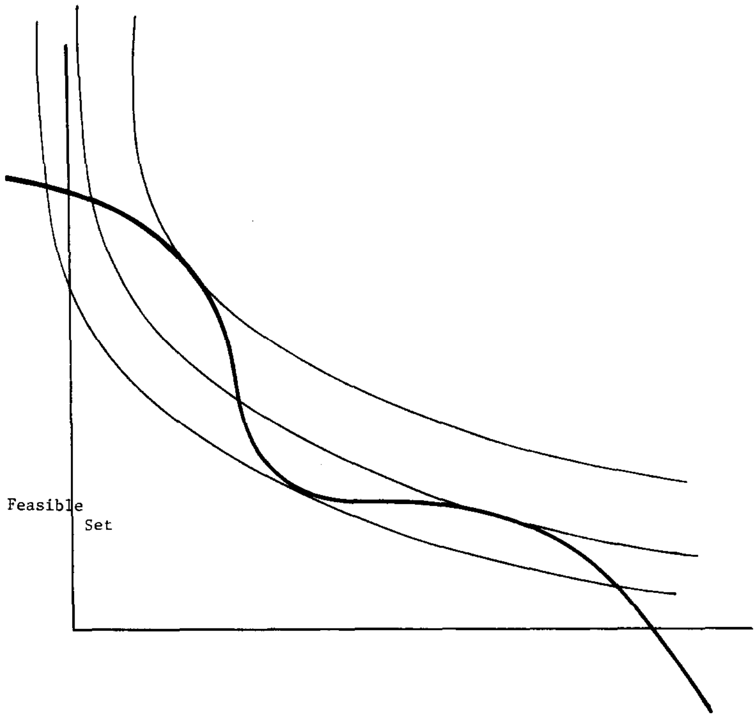

FIG. 3. A problem which is not concave transformable.

3. RESULTS

If (3) is a concave programming problem, then known results can be applied. The idea of the proofs below is that these results, applicable to the problem on R^m defined by $\hat{f}$ and $\hat{g}_i$ $(i = 1, 2, \ldots, n)$, can then be "pulled back" via the inverse of T to the problem defined by f and g_i $(i = 1, 2, \ldots, n)$. The maximum value function of a programming problem is defined as follows. Consider the problem

$$\text{Maximise } f(x) \text{ subject to } g_i(x) \geq b_i, \qquad i = 1, \ldots, n.$$

Let $b = (b_i)$, $i = 1, \ldots, n$, and

$$V(b) = \operatorname*{Max}_x f(x), x \in \{x/g_i(x) \geq b_i, i = 1, \ldots, n\}.$$

This gives the maximum value of the objective function attainable for various values of the constraints b_i.

We can now state and prove the main result.

THEOREM 1. *Let problem* (1) *be concave transformable and satisfy the constraint qualification* (2). *Then*

(i) *any critical point is a global maximum*;

(ii) x^* *is a solution to* (1) *if and only if there exists* $\lambda^* \in R^n$, $\lambda^* \geq 0$, *such that* (x^*, λ^*) *form a saddle-point of the Lagrangian*

$$L = f(x) + \sum_i \lambda_i g_i(x);$$

(iii) *the maximum value function* $V(b)$ *of* (1) *is a concave function.*

Proof. First note that any x' in the domain of $(\hat{f}, \hat{g}_i)$ is the unique image under T of a single x in the domain of (f, g_i). Write

$$x' = T(x). \tag{4}$$

Hence

$$\begin{aligned} \hat{f}(x') &= f(T^{-1}(x')) = f(T^{-1}(T(x)) = f(x) \\ \hat{g}_i(x) &= g_i(T^{-1}(x')) = g_i(T^{-1}(T(x)) = g_i(x). \end{aligned} \tag{5}$$

Thus the $\hat{f}$ and $\hat{g}_i$ assume at x' the same value as f and g at x, if $x' = T(x)$. This follows immediately from the commutativity of Fig. 1.

Next note that, as problem (1) satisfies the constraint qualification (2), by setting

$$\bar{x}' = T(\bar{x})$$

it is clear that

$$\hat{g}_i(\bar{x}') > 0, \qquad i = 1, 2, \dots, n, \tag{6}$$

and problem (3) also satisfies the constraint qualification.

Now we show that if x'^* solves the transformed problem (3), then $T^{-1}(x'^*) = x^*$ solves the original problem (1). By assumption,

$$\hat{f}(x'^*) \geq \hat{f}(x') \qquad \forall x' \in \{x'/\hat{g}_i(x') \geq 0,\ i = 1, \dots, n\}.$$

By (5),

$$x' \in \{x'/\hat{g}_i(x') \geq 0, \text{ all } i\}$$

if and only if

$$T^{-1}(x') = x \in \{x/g_i(x) \geq 0, \text{ all } i\}.$$

We now assert that $x^* = T^{-1}(x'^*)$ satisfies

$$f(x^*) \geq f(x) \qquad \forall x \in \{x/g_i(x) \geq 0, \text{all } i\}.$$

Suppose the contrary. Then there exists $\tilde{x}$ satisfying

$$f(\tilde{x}) > f(x^*), \qquad \tilde{x} \in \{x/g_i(x) \geq 0, \text{all } i\}.$$

But by (5),

$$\hat{f}(T(\tilde{x})) = \hat{f}(\tilde{x}') = f(\tilde{x})$$
$$\hat{f}(x'^*) = f(x^*)$$

so that $\hat{f}(\tilde{x}') > \tilde{f}(x^*)$ and $\tilde{x}'$ is feasible for the transformed problem. This is a contradiction: hence $x^* = T^{-1}(x'^*)$ solves (1).

We have proven that

$$(x^{*\prime} \text{ solves } (3)) \Rightarrow (T^{-1}(x^{*\prime}) = x^* \text{ solves } (1)).$$

Now we establish the converse. Let x^* solve (1). Then by (5)

$$f(x^*) \geq f(x) \qquad \forall x \text{ s.t. } T(x) = x' \in \{x' : \hat{g}_i(x') \geq 0,\ i = 1, \ldots, n\}.$$

Again by (5) it follows that

$$\hat{f}(x^{*\prime}) \geq \hat{f}(x') \qquad \forall x' \in \{x' : \hat{g}_i(x') \geq 0,\ i = 1, \ldots, n\}.$$

Hence we have that

$$(x^{*\prime} \text{ solves } (3)) \Leftrightarrow (x^* \text{ solves } (1)). \tag{7}$$

We next investigate the relationship between critical points of (3) and those of (1). At a critical point of (3), the derivative of the Lagrangian

$$\hat{\mathscr{L}} = \hat{f}(x') + \sum_i \lambda_i \hat{g}_i(x')$$

vanishes. Hence

$$D\hat{f}(x') + \sum_i \lambda_i D\hat{g}_i(x') = 0, \tag{8}$$

where D is the derivative operator. Now $x' = T(x)$ for some x, so that $\hat{f}(x') = f(T^{-1}(x'))$ so that (8) is equivalent to

$$Df(T^{-1}(x')) + \sum_i \lambda_i Dg_i(T^{-1}(x')) = 0.$$

By the chain rule this is equivalent to

$$Df(x) \cdot DT^{-1}(x') + \sum_i{}' \lambda_i Dg_i(x) \cdot DT^{-1}(x') = 0. \tag{9}$$

But as T is regular $DT^{-1}(x')$ is a square matrix of full rank. Post-multiplying (9) by the inverse of $DT^{-1}(x')$ gives

$$Df(x) + \sum_i \lambda_i Dg_i(x) = 0, \tag{10}$$

which is precisely the condition which characterises a critical point of problem (1). Hence we have, for $x' = T(x)$,

$$(x' \text{ is a critical point of } (3)) \Leftrightarrow (x \text{ is a critical point of } (1)). \tag{11}$$

Because (3) is a concave programming problem, we know also that

$$(x' \text{ is a critical point of } (3)) \Leftrightarrow (x' \text{ is a global maximum of } (3)). \tag{12}$$

Conditions (7), (11), and (12) imply (i) of Theorem 1, namely, that any critical point of (1) is a global maximum.

To prove part (ii), we note that because (3) is a concave programming problem satisfying the constraint qualification, there exists a $\lambda^* \in R^n, \lambda^* \geq 0$, such that x'^* is a solution to (3) if and only if (x'^*, λ^*) form a saddle-point of the Lagrangian, i.e., if and only if

$$\begin{aligned} \hat{f}(x') + \Sigma\lambda_i^* \hat{g}_i(x') &\leq \hat{f}(x'^*) + \Sigma\lambda_i^* g_i(x'^*) \\ &\leq \hat{f}(x'^*) + \Sigma\lambda_i g_i(x'^*) \end{aligned}$$

for any x' and any $\lambda \in R^n, \lambda \geq 0$. By (5) we can rewrite this as

$$f(x) + \Sigma\lambda_i^* g_i(x) \leq f(x^*) + \Sigma\lambda_i^* g_i(x^*) \leq f(x^*) + \Sigma\lambda_i g_i(x^*)$$

for any x and $\lambda \in R^n, \lambda \geq 0$. Hence (x^*, λ^*) form a saddle-point of the Lagrangian of the untransformed problem (1), so we have established that for this problem, an optimum is a saddle-point. The converse is well known (see Uzawa [30]). This proves (ii) of Theorem 1.

To prove (iii), we consider next the set A' in R^{n+1}

$$A' = \left\{ \begin{matrix} y \in R^1 \\ \\ z \in R^n \end{matrix} \;\middle|\; \text{for some } x', \quad \begin{matrix} y \leq \hat{f}(x') \\ \\ z_i \leq \hat{g}_i(x'), \text{ all } i \end{matrix} \right\}.$$

As f and g_i are concave functions, this is a convex set. (For $n = 1$, see Fig. 4.)

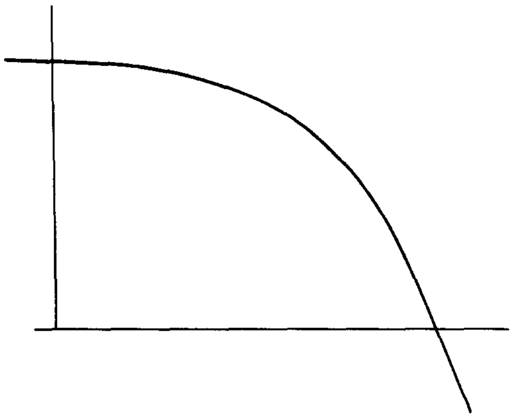

FIG. 4. The set A', which is convex for both concave and concave-transformable (non-concave) problems.

Now define the following real-valued function on R^n:

$$\text{For } z \in R^n, v'(z) = \text{maximum } y : (y, z) \in A'.$$

$v'(z)$ is obviously identical to the maximum value function for the transformed problem, and by standard arguments is a concave function. Next consider the set equivalent to A' for the untransformed problem (1):

$$A = \left\{ \begin{matrix} y \in R^1 \\ z \in R^n \end{matrix} \quad \text{for some } x, \quad \begin{matrix} y \leq f(x) \\ z_i \leq g_i(x) \text{ all } i \end{matrix} \right\}.$$

Likewise define:

$$\text{For } z \in R^n, v(z) = \text{maximum } y : (y, z) \in A.$$

By (5), $A = A'$. Hence $v(z)$, the maximum value function for the untransformed problem, is also concave.

This completes the proof of Theorem 1. □

It is worth noting that the fact that an optimum is a saddle-point could also have been proved by the usual approach of separating the set A from the set $B \subset R^{n+1}$ defined as

$$B = \left\{ \begin{matrix} y \in R^1 \\ z \in R^n \end{matrix} \middle| \begin{matrix} y > f(x^*) \\ z > 0 \end{matrix} \right\}$$

as both A and B can be shown to be convex sets for a concave-transformable programme.

4. An Example

In this section we give an extremely simple example, merely to provide some intuitive grasp for the kind of problem that falls within the scope of the above theorem. The example is

$$\text{maximise } \left(\alpha x_1^3(\beta x_2^3 - \gamma)\right)^{1/2}$$
$$\text{subject to } A - \delta x_1^3 - \varepsilon x_2^3 \geq 0.$$

Now let

$$T_1(x_1) = \alpha x_1^3 = x'_1$$
$$T_2(x_2) = \beta x_2^3 - \gamma = x'_2$$

so that the transformed problem is

$$\text{maximise } (x'_1 x'_2)^{1/2}$$
$$\text{subject to } A - x'_1\frac{\delta}{\alpha} - x'_2\frac{\varepsilon}{\beta} - \frac{\gamma\varepsilon}{\beta} \geq 0.$$

It is readily verified that the original problem is non-concave, whereas the transformed problem is concave. The solution to the transformed problem is

$$x'^*_1 = A\frac{\alpha\varepsilon}{\delta\beta}\Big/\left(1 + \frac{\alpha\varepsilon}{\delta\beta}\right), \qquad x'^*_2 = A\Big/\left(1 + \frac{\alpha\varepsilon}{\delta\beta}\right)$$
$$\text{Lagrange multiplier } \lambda^* = \frac{1}{2}\left(\frac{\alpha\beta}{\delta\varepsilon}\right)^{1/2}.$$

It can be readily verified that the implied values (x_1^*, x_2^*) of x_1 and x_2 solve the untransformed problem, and that $(x_1^*, x_2^*, \lambda^*)$ forms a saddle-point for the untransformed problem.

5. Relationship with Generalised Convexity

Various concepts of generalised convexity have been introduced in the programming literature, for example by Avriel [4], Zang [32], and Ben-Tal [5]. Here we review briefly the connection between concave transformability

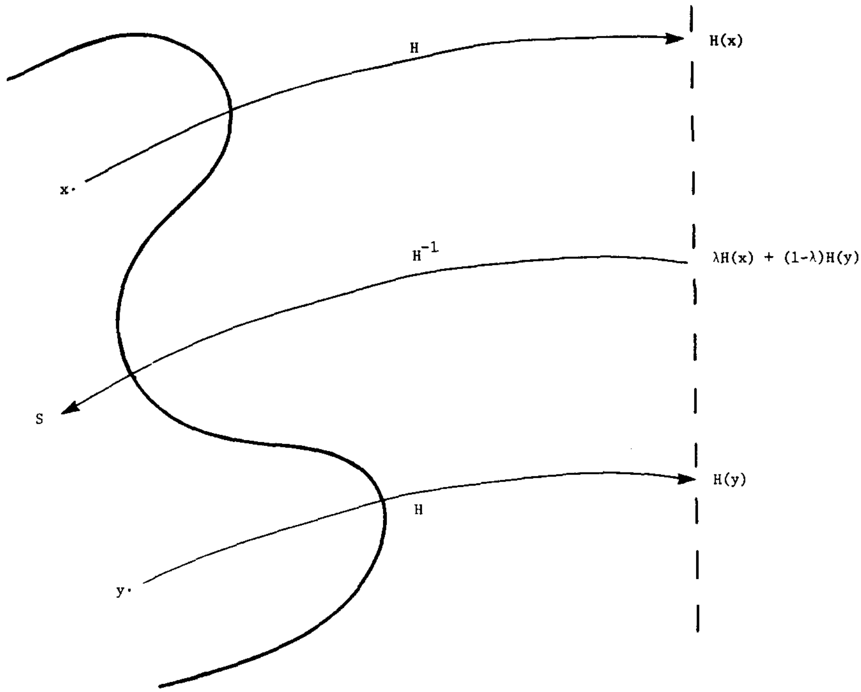

FIG. 5. The concept of H-convexity.

and generalised convexity. Following Ben-Tal [5], consider a set $S \subset R^m$, and a function $H: R^m \to R^m$, H 1-to-1, onto, and possessing an inverse H^{-1}. (Note that continuity is not required of H and H^{-1}.) Then for any $x, y \in S$ and any $\lambda \in [0,1]$, define $M_H((x, y), \lambda)$, the H-weighted mean of x and y, as

$$M_H((x, y), \lambda) = H^{-1}(\lambda H(x) + (1 - \lambda)H(y)).$$

We then say that S is *H-convex* if for any $x, y \in S$ and $\lambda \in [0,1]$, $M_H((x, y), \lambda) \in S$. Figure 5 illustrates this idea. In relating this to the idea of concave transformability, the following remark is useful.

Remark. S is H-convex if and only if $H(S)$ is convex.

Proof. Suppose S is H-convex. Then if $x, y \in S, H^{-1}(\lambda H(x) + (1 - \lambda)H(g)) \in S$. This implies that $H(x)$, $H(y)$ and $\lambda H(x) + (1 - \lambda)H(y)$ are all in $H(S)$, which is therefore convex. Now suppose $H(S)$ is convex. Then $H(x), H(y) \in H(S)$ implies $\lambda H(x) + (1 - \lambda)H(y) \in H(S)$, which in turn implies $H^{-1}(\lambda H(x) + (1 - \lambda)H(g)) \in S$, which is H-convex. This completes the proof. □

The concept of H-convexity therefore amounts to being transformable into a convex set: if the requirements of continuity are imposed on H and H^{-1}, then H-convexity amounts to being homeomorphic to a convex set, which in turn implies contractibility[2] (see Kuhn [21] and Chichilnisky [10] for a discussion of results obtainable using contractibility instead of convexity).

The property of concave transformability can now be interpreted in this framework. Let

$$S = \{x : g_i(x) \geq 0,\ i = 1, \ldots, n\}.$$

S is thus the feasible set of problem (1).

We can now establish:

THEOREM 2. *Problem* (1) *is concave transformable if and only if S is T-convex for a regular C^1 diffeomorphism T such that $f \circ T^{-1}$ is concave.*

Proof. Suppose first that (1) is concave transformable. Then there exists T as specified such that $f \circ T^{-1}$ is concave, and $g_i \circ T^{-1}$ are also concave. Hence $\{x' : g_i \circ T^{-1}(x') \geq 0,\ i = 1, \ldots, n\}$ is convex. But

$$\{x' : g_i \circ T^{-1}(x') \geq 0,\ i = 1, \ldots, n\} = T(S).$$

Hence by the remark above S is T-convex for a regular C^1 diffeomorphism such that $f \circ T^{-1}$ is concave.

Suppose next that S is T-convex, with T as specified and $f \circ T^{-1}$ concave. Then

$$\{x' : g_i \circ T^{-1}(x') \geq 0,\ i = 1, \ldots, n\} = T(S)$$

is convex, so that the functions $g_i \circ T^{-1}$ are concave. As $f \circ T^{-1}$ is concave by assumption, (1) is concave transformable, proving the result.

6. ECONOMIC APPLICATIONS

We have demonstrated that the equivalence of an optimum and a saddle-point of the untransformed Lagrangian holds, not only for concave programmes, but for concave-transformable programmes, defined as pro-

[2]A space S is contractible if there exists a continuous function $F: S \times [0,1] \to S$ and a point $s_0 \in S$ such that $F(x,0) = s\ \forall s \in S$, $F(s,1) = s_0\ \forall s \in S$.

grammes which become concave under a suitable diffeomorphism of the choice space. Section 4 gives an extremely simple example where the appropriate diffeomorphism is readily apparent and indeed is just a change of variables. It should be emphasised, however, that a diffeomorphism does not in general correspond to a change of variables or coordinates, except in a local sense, though coordinate changes are of course diffeomorphisms.

It is worth noting that non-convex problems which become convex with a change of variables have been studied in economics. Both Chipman [13] and Quinzii [26] study economies whose feasible sets are non-convex, but become convex in logarithms.

As mentioned in the Introduction, there exist other results connecting optima and saddle-points for non-concave problems. The earliest are those of Arrow and Hurwicz [1]; subsequently there is the work of Rockafellar [27] and Fujiwara [16]. But in all of these cases, the equivalence of optima is with a saddle-point of a transformed or augmented Lagrangian: here it has been shown to hold for the standard Lagrangian of concave programming, without augmentation or transformation. In particular, it has been shown that if it is possible to transform the choice space diffeomorphically so that the problem becomes concave, *then* for that problem (without the use of the transformation) an optimum is equivalent to the saddle-point of the (ordinary, untransformed) Lagrangian.

As the original saddle-point–optimum equivalence had substantial economic implications, particularly in the field of decentralisation (see, e.g., Arrow and Hurwicz [27] and Heal [18]), it is clearly of interest to investigate the analogous implications of the present work. Suppose the problem

$$\text{maximise } f(x) \text{ subject to } g_i(x) \geq 0, \qquad i = 1, 2, \ldots, n, \tag{1}$$

to be concave transformable and to satisfy a constraint qualification. Then we know that if x^* is the solution, and λ^* is the associated vector of Lagrange multipliers, x^* maximises with respect to x:

$$f(x) + \sum \lambda_i^* g_i(x).$$

Hence the optimal solution is the solution of an unconstrained maximisation problem, given an appropriate vector of shadow prices to value the constraints. We therefore have a category of non-concave optimisation problem in which ordinary or "linear" prices can be used to value the constraints and convert to an unconstrained problem. Transformation or augmentation of the Lagrangian corresponds of course to using "non-linear" prices, or to using dual functions rather than dual variables, and this is the usual approach outside of a concave environment; see, for example Brown and Heal [6, 8] and Tind and Wolsey [29].

Although conventional linear shadow prices can be used to value the constraints of a concave-transformable programme, it does not follow that the usual economic decentralisation results (see, e.g., Koopmans [20]) hold in a concave-transformable environment. These results say that the solution to the overall optimisation problem may be achieved as the sum of solutions to a number of independent optimisation problems, linked only by the use of the same prices. Such a result rests on the fact that the overall optimum can be characterised as the maximum of a linear function over a sum of sets. As set addition and maximisation of a linear function commute, this in turn equals the sum of its maxima over the individual sets, which gives rise to the possibility of decentralisation. This possibility does not in general occur in a concave-transformable problem, because the solution of this problem cannot be described as the maximum of a linear function over the feasible set. This would be the case only if the optimum lay on the boundary of the convex hull of the feasible set. Figure 6 illustrates these different cases.

Although the results here seem to have no clear precedent in the optimisation literature, they are clearly related to a result of Chichilnisky and Heal [11]. This paper analyses the minimum dimensions of the message space needed to achieve efficiency in various resource allocation problems. For the standard convex environment, the minimum dimension of the message space is that of the commodity space. (An earlier derivation of this result is given by Mount and Reiter [24].) However, Chichilnisky and Heal show that there is a class of non-convex environments in which the minimum dimension of the message space is also that of the commodity space. This seems to

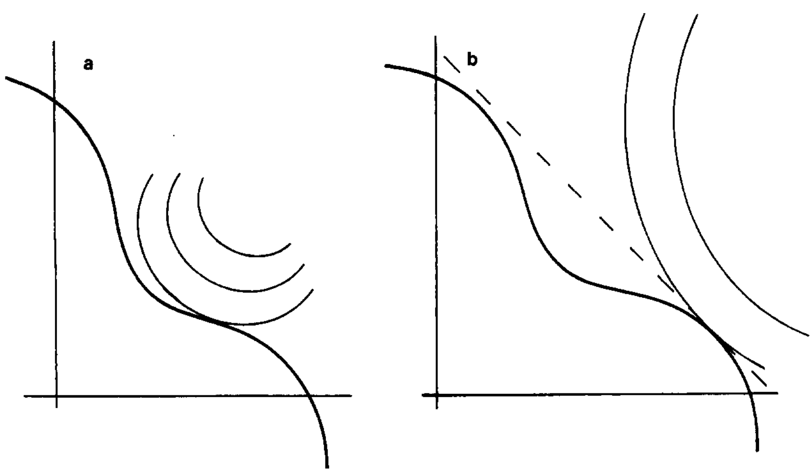

FIG. 6. (a) and (b) illustrate concave-transformable problems. —, The boundary of feasible set; —, a contour of the maximand. In case (b) the solution is the maximum of a linear function over the feasible set: in case (a) it is not.

correspond to the class of environment analysed by Brown and Heal [7], and to the class of non-concave programme discussed here.

7. Possible Extensions

Finally, I turn to the possibility of extending the results given here. I have given above a condition, weaker than concavity, which is sufficient for the equivalence of saddle-points and optima. It seems that this condition, though significantly weaker than those already known, is not necessary and it would naturally be interesting to have a necessary and sufficient condition for the saddle-point–optima equivalence. Another aspect of Theorem 1 is that it gives sufficient conditions for any critical point of a constrained maximisation problem to be a global maximum. Again, a necessary and sufficient condition for this property would be of general interest. It seems in fact likely that such a condition can be established by arguments similar to those used in the development of Morse theory [22].

For problem (1), define the set

$$P_\alpha = \{x : f(x) \geq \alpha, g_i(x) \geq 0, i = 1, \ldots, n\}.$$

Then one might conjecture that under certain regularity conditions on the functions f and g_i, and certain compactness conditions, a necessary and sufficient condition for any critical point of (1) to be a global maximum is that for each real number α, the set P_α is contractible. The problems illustrated in Fig. 2 satisfy this property, whereas that in Fig. 3 does not.

An alternative way of approaching this problem would be in terms of the properties of the set-valued mapping H,

$$H : R^1 \to 2^{R^n},$$
$$H : \alpha \to P_\alpha.$$

Zang *et al.* [33] show that functions all of whose critical-points are global minima can be characterised in terms of the continuity properties of this mapping. More recently, Dolecki [14] has applied the same approach to constrained maximisation problems. However, a topological approach via the topological characteristics of the sets P_α seems preferable, as it would lead to a connection with a substantial body of powerful results.

Acknowledgments

I am grateful to Don Brown, Graciela Chichilnisky, Okitsu Fujiwara, Lawrence Wolsey, and Israel Zang for valuable comments.

REFERENCES

1. K. J. ARROW AND L. HURWICZ, The reduction of constrained maxima to saddle-points, *in* "Proceedings, Third Berkeley Symposium on Probability and Statistics" (J. Neyman, Ed.), Univ. of California Press, Berkeley, 1956.
2. K. J. ARROW AND L. HURWICZ, Decentralisation and computation in resource allocation, *in* "Essays in Economics and Econometrics in Honour of Harold Hotelling" (R. Pfouts, Ed.), Univ. of North Carolina Press, Chapel Hill, 1961.
3. K. J. ARROW, L. HURWICZ, AND H. UZAWA, "Studies in Linear and Nonlinear Programming," Stanford Univ. Press, Stanford, Calif., 1958.
4. M. AVRIEL, "Nonlinear Programming," Prentice–Hall, Englewood Cliffs, N. J., 1976.
5. A. BEN-TAL, On generalised means and generalised convex functions, *J. Optim. Theory Anal.* **21**, No. 1 (1977).
6. D. J. BROWN AND G. M. HEAL, "The Existence of Equilibrium in an Economy with Increasing Returns," Cowles Foundation Discussion Paper, Yale University, 1976.
7. D. J. BROWN AND G. M. HEAL, Equity, efficiency and increasing returns, *Rev. Econom. Stud.* (Oct. 1979).
8. D. J. BROWN AND G. M. HEAL, Marginal cost pricing and two part tariffs in a general equilibrium model with increasing returns, *J. Public Econom.* (Feb. 1980).
9. D. J. BROWN, G. M. HEAL, AND F. WESTHOFF, "Regular Non-Linear Programmes," Cowles Foundation Discussion Paper, Yale University, 1979.
10. G. CHICHILNISKY, "Intersecting Families of Sets, a Topological Characterisation," Essex University Economics Papers, 1981, *Topology*, in press.
11. G. CHICHILNISKY AND G. M. HEAL, "Aggregation, Information and Message Spaces," Essex University Economics Discussion Paper, 1980.
12. G. CHICHILNISKY AND P. J. KALMAN, The comparative statics of less neoclassical economic agents, *Internat. Econom. Rev.* (1978).
13. J. CHIPMAN, External economies of scale and competitive equilibrium, *Quart. J. Econom.* (1978).
14. S. DOLECKI, The role of lower semicontinuity in optimality theory, *in* "Proceedings, Game Theory and Economics" (O. Moeschler, Ed.), Springer-Verlag Lecture Notes, Springer-Verlag, New York/Berlin, 1981.
15. R. J. DUFFIN, E. L. PETERSON, AND C. ZENER, "Geometric Programming: Theory and Applications," Wiley, New York, 1967.
16. O. FUJIWARA, "Duality in Non-convex Programs," Ph.D. thesis, Yale University, 1980.
17. O. FUJIWARA, "Morse Program," Cowles Foundation Discussion Paper, Yale University, 1979.
18. G. M. HEAL, "The Theory of Economic Planning," North-Holland, Amsterdam, 1973.
19. M. D. INTRILIGATOR, "Mathematical Optimisation and Economic Theory," Prentice–Hall, Englewood Cliffs, N. J., 1971.
20. T. C. KOOPMANS, The price system and resource allocation, *in* "Three Essays on the State of Economic Science," by T. C. Koopmans, McGraw–Hill, New York, 1957.
21. H. KUHN, Contractibility and convexity, *Proc. Amer. Math. Soc.* **5** (1954).
22. J. MILNOR, "Morse Theory," Princeton Univ. Press, Princeton, N. J., 1963.
23. M. MORSE, "The Calculus of Variations in the Large," Amer. Math. Soc., New York, 1934.
24. K. MOUNT AND S. REITER, The informational size of message spaces, *J. Econom. Theory* **8** (1974), 161–192.
25. E. L. PETERSON, Geometric programming and some of its extensions, *in* "Optimisation and Design" (M. Avriel, M. J. Rickaert, and D. J. Wilde, Eds.), Prentice–Hall, Englewood Cliffs, N. J., 1973.
26. M. QUINZII, "An Existence Theorem for the Core of an Economy with Increasing Returns

in Production," Discussion Paper, Ecole Polytechnique, Paris, 1980.

27. R. T. ROCKAFELLAR, Lagrange multipliers in optimisation, *SIAM–AMS Proc.* **9** (1976).
28. J. E. SPINGARN AND W. T. ROCKAFELLAR, The generic nature of optimality conditions in non-linear programming, *Math. Oper. Res.* (1979).
29. J. TIND AND L. WOLSEY, "A Unifying Framework for Duality Theory in Mathematical Programming," Discussion Paper No. 7834, C.O.R.E., Université Catholique de Louvain, 1978.
30. H. UZAWA, The Kuhn–Tucker theorem for concave programming, *in* "Studies in Linear and Non-Linear Programming," by K. J. Arrow, L. Hurwicz, and H. Uzawa, Stanford Univ. Press, Stanford, Calif., 1958.
31. K. K. YUN, "Uniqueness conditions for Kuhn–Tucker points on a disk," Discussion Paper, State University of New York, Albany, Department of Economics. *J. Optim. Theory Appl.*, in press.
32. I. ZANG, "Generalised Convex Programming," Ph.D. thesis, Department of Chemical Engineering, Technion, Haifa, 1974.
33. I. ZANG, E. U. CHOO, AND M. AVRIEL, On functions whose stationary points are global minima, *J. Optim. Theory Appl.* **22**, No. 2 (1977).