Revue française d'automatique informatique recherche opérationnelle. Mathématique

Bernard Bereanu

Quasi-convexity, strictly quasi-convexity and pseudo-convexity of composite objective functions

Revue française d'automatique informatique recherche opérationnelle. Mathématique, tome 6, nº R1 (1972), p. 15-26

<http://www.numdam.org/item?id=M2AN_1972__6_1_15_0>

R A.I.R.O.
(6e annee, R-1, 1972, p. 15-26)

QUASI-CONVEXITY, STRICTLY QUASI-CONVEXITY AND PSEUDO-CONVEXITY OF COMPOSITE OBJECTIVE FUNCTIONS (1)

by Bernard BEREANU (2)

Résumé. — *Dans un article précédent on a donné une condition nécessaire et suffisante qui assure qu'un certain type de composition de fonctions fait aboutir à des fonctions convexes. Dans cet article on étend ce résultat aux fonctions convexes généralisées (quasi-convexes, strictement quasi-convexes ou pseudo-convexes) fournissant ainsi une méthode pratique de reconnaissance des fonctions ayant ces propriétés. On discute des applications a l'économie mathématique et à la programmation mathématique.*

1. In practice the following situation occurs. One has some freedom to choose a real function f in m variables which will appear in the mathematical model through its composition $f \circ u$ with a vector valued function $u = (u_1, ..., u_m)$ of n variables. The economic interpretation of the components of u may justify the assumption that some are concave (for instance utilities), while other are convex (describing situations of increasing returns per unit increase of input), the type being specified for each component, but the functions themselves having the possibility to differ from one application to another. In this case it may be important to choose f so that $f \circ u$ be convex (concave) for any feasible u. This happens for instance in optimization problems if the analyst has to choose an overall criterion function f to incorporate several conflicting objectives represented by $u_i (i = 1, ..., m)$.

In Bereanu [4] it was proved that the composite function $F = f \circ u$ is convex for every vector function u with the range in the domain of f and having its components, separately convex or concave *if and only if* f is convex and partially monotone, increasing in the convex components and decreasing in the concave components (3). An analogous result is valid when convex is replaced by concave.

(1) This paper is a revised version of Preprint 7101, Centre of Mathematical Statistics, Bucharest, April 1971.
(2) Centre of Mathematical Statistics of the Academy of the Romanian Socialist Republic, Bucharest.
(3) In [4] this theorem was proved assuming only mid-point convexity (convexity in the sense of Jensen).

Recently Mangasarian [17] proved that the *sufficiency* part of this theorem remains valid when *convexity*, as an attribute of F and f, is replaced by *quasi-convexity* or *pseudo-convexity*.

The purpose of this note is to prove the mentioned theorem of [4] when *quasi-convexity* or *strictly quasi-convexity* or *pseudo-convexity* of F and f is substituted to *convexity*. A characteristic property of quasi-convex functions is given in Theorem 1.

Applications to mathematical economics and mathematical programming are briefly discussed.

2. Generalized convexity. There are various extensions of classical convexity which maintain certain basic properties of convex functions important for mathematical programming. We shall recall some of the definitions to be used and we shall give here a characterization of quasi-convex functions. Although some results remain valid under more general conditions, the real valued functions considered will be defined, if not otherwise stated, on subsets of Euclidean spaces.

The function $f : D \to R$, where D is a convex set in R^m and R denotes the real line, is *convex* on D provided that for every x and y in D and for every $\lambda \in [0, 1]$

$$f(\lambda x + (1 - \lambda)y) \leqq \lambda f(x) + (1 - \lambda)f(y). \tag{1}$$

(To simplify the notations we shall use the symbol $+$ for addition of both vectors and scalars.)

The function f is said *quasi-convex* on D [9, pp. 117-118] if

$$f(\lambda x + (1 - \lambda)y) \leqq \max(f(x), f(y)) \, (x, y \in D), \lambda \in [0, 1], \tag{2}$$

or, equivalently, if for arbitrary real α the set $\{ x \mid f(x) \leqq \alpha \}$ is either empty or convex.

If

$$f(y) < f(x) \text{ implies } f(\lambda x + (1 - \lambda)y) < f(x) \text{ for every } \lambda \in (0, 1), \tag{3}$$

$x, y \in D$, then the function f is *strictly quasi-convex* (1) [16].

The above definitions apply also when D is a convex subset of a linear space.

Let now D be an open set, not necessarily convex, in R^m, and $f : D \to R$ be differentiable and denote by ∇f its gradient. Using the terminology of

---

(1) M. A. Hanson in [11] uses functionals with this property but does not employ the term strictly quasi-convex, while Stoer and Witzgall [21, p. 169] call pseudo-convex a function having property (3). In the present paper we shall use the terminology of Mangasarian [16].

Mangasarian [16] the function f is said pseudo-convex if for every x and y in D,

$$(y-x)' \nabla f(x) \geqq 0 \text{ implies } f(y) \geqq f(x). \tag{4}$$

It is easily seen that (1) implies (2) and (3). Hence a convex function is strictly quasi-convex and quasi-convex.

However a strictly quasi-convex function need not be quasi-convex (1). This verbal anomaly disappears when in addition, lower semicontinuity is assumed [14]. A differentiable convex function is obviously pseudo-convex. For differentiable functions the hierarchy, in order of increasing generality is the following : convex, pseudo-convex, strictly quasi-convex and quasi-convex [16].

A function g is respectively concave, pseudo-concave, strictly quasi-concave or quasi-concave if $-g$ is convex, pseudo-convex, strictly quasi-convex or quasi-convex.

A characteristic property of quasi-convex functions is given by the following theorem.

Theorem 1. *Let X be a convex set in an arbitrary linear space and let $f : X \to R$ be a real valued function. The following two statements are equivalent.*

(i) *The function $f : X \to R$ is quasi-convex.*

(ii) If Y is an arbitrary subset of X and $co\,Y$ is its convex hull, then (2)

$$\sup_{co\,Y} f(x) = \sup_{Y} f(x) \tag{5}$$

Proof. (i) $\to$ (ii). It is enough to prove that for an arbitrary set $Y \subset X$ we have

$$\sup_{co\,Y} f(x) \leqq \sup_{Y} f(x) \tag{6}$$

If $\sup_Y f(x) = +\infty$, is nothing to prove. So we suppose that $\sup_Y f(x) = a$, where a is a real number, and let $X_a = \{ x \mid f(x) \leqq a \}$.

Obviously $Y \subset X_a$. But X_a is a convex set because (i). Furthermore $X_a \supset co\,Y$ because the minimal property of the convex hull.

---

(1) See [21, p. 170] for examples of functions which are strictly quasi-convex, but not quasi-convex.

(2) After this paper was circulated as a preprint it was discovered that property (5) is contained in a paper by L. BRAGARD, « Programmation quasi-concave », *Bull. Soc. Royale des Sciences de Liège*, 9-10, 1970, pp. 478-485, where a different proof is given. However in that paper it is not shown that this property is characteristic for quasi-convex functions.

Hence

$$\sup_{\mathrm{co}\,Y} f(x) \leqq \sup_{X_a} f(x) \leqq a = \sup_{Y} f(x) \tag{7}$$

(ii) $\rightarrow$ (i). We must show that for an arbitrary real a for which X_a is not empty, X_a is convex.

Let x^1, $x^2 \in X_a$ and $\lambda \in [0,1]$. Because (ii) we have

(8) $f(\lambda x^1 + (1 - \lambda)x^2) \leq \max[f(x^1), f(x^2)] \leqq a$. Hence $\lambda x^1 + (1, - \lambda)x^2 \in X_a$ and the proof is completed.

Corollary 1. *The following two statements are equivalent :*

(i) *The function $f : X \rightarrow R$ is quasi-concave.*

(ii) *If Y is an arbitrary subset of X and $\mathrm{co}\,Y$ is its convex hull then*

$$\inf_{\mathrm{co}\,Y} f(x) = \inf_{Y} f(x) \tag{9}$$

Corollary 2. (Martos [18]). *A real-valued function defined on a convex set $L \subset R^m$ is quasi-concave if and only if it attains its global minimum on each polytope contained in L, in one of the vertices.*

As seen from above, quasi-convex functions share with convex functions property (ii) and also the convexity of level sets. It can be easily verified that every local minimum of a strictly quasi-convex function is a global minimum. The same is true for pseudo-convex functions [16, p. 284].

Extensions of the Kuhn-Tucker theorem [13] to quasi-convex and pseudo-convex programming are given in [1] and [16].

3. Invariance of generalized convexity under composition with affine functions

Let D and D_1 be bounded convex sets in R^m, respectively R^n. We consider the real valued function $f : D \rightarrow R$ and the family $\mathfrak{L}$ of all vector valued affine functions $l : D_1 \rightarrow R^m$. The set $D_l = \{ x \mid x \in D_1, l(x) \in D \}$ is convex. We further denote

$$F(x) = f \circ l = f(l_1(x), ..., l_m(x)), \ x \in D_l. \tag{10}$$

We shall say that a function is G-convex (generalized convex) if it is quasi-convex or strictly quasi-convex or pseudo-convex. An analogous definition will be used for G-concave functions.

With these notations we have the following theorem which generalizes lemma 2 of [4].

Theorem 2. *The function $F(x) = f \circ l$ $(x \in D_l)$ is G-convex for every $l \in \mathfrak{L}$ if and only if the function f is G-convex in D.*

Proof. When G-convex means quasi-convex, or strictly quasi-convex the theorem follows from the fact that these types of G-convexity are equivalent to G-convexity on every line.

Suppose now that $F(x)$ is pseudo-convex in D_l for every $l \in \mathfrak{L}$, while $f(y)$ which is differentiable in D is not pseudo-convex there. Hence there are points y^1, $y^2 \in D$ which satisfy

$$(y^2 - y^1) \nabla f(y^1) \geqq 0 \tag{11}$$

and

$$f(y^2) < f(y^1). \tag{12}$$

There are $l \in \mathfrak{L}$ and x^1, $x^2 \in D_l$ which satisfy

$$l(x^1) = y^1 \text{ and } l(x^2) = y^2.$$

If $\nabla l(x)$ is the matrix $\dfrac{\partial l_i(x)}{\partial x_j}$ $(i = 1, ..., m; j = 1, ..., n)$ and $'$ indicates transposition, we have from (11) :

$$(x^2 - x^1)' \nabla F(x^1) = (x^2 - x^1)' \nabla f(y^1) \nabla l(x^1) = (y^2 - y^1)' \nabla f(y^1) \geqq 0. \tag{13}$$

Because $F(x)$ is supposed pseudo-convex, (13) imply

$$F(x^2) \geqq F(x^1), \tag{14}$$

which contradicts (12).

The direct implication follows immediately. Indeed if

$$(x^2 - x^1)' \nabla F(x^1) \geqq 0,$$

from (13) follows $f(y^2) \geqq f(y^1)$ where $y^i = l(x^i)$, $i = 1, 2$, i.e. $F(x^2) \geqq F(x^1)$.

The if-part of the theorem for the quasi-convex and pseudo-convex cases is contained in Mangasarian [17].

We have the following obvious corollary.

Corollary 3. *Theorem 2 remains valid if G-convex stands for lower semi-continuous strictly quasi-convex.*

Corollary 4. *The function $F(x) = f \circ l$ $(x \in D_l)$ is G-concave for every $l \in \mathfrak{L}$ if and only if the function f is G-concave on the convex set D.*

4. *Generalized convex, composite functions.* A function $f : D \to R(D \subset R^m)$ is said *partially monotone* in D if the partial functions

$$f_{(i)} : x_i \to f(x_1, ..., x_i, ..., x_m)$$

are, separately for each i, monotone increasing (non-decreasing) or monotone decreasing (non-increasing). Let I_1 and I_2 be two sets of indices such that $I_1 \cap I_2 = \emptyset$ and $I_1 \cup I_2 = \{1, ..., m\}$. Denote by $U(D ; I_1, I_2)$ the family of vector valued functions $u = (u_1, ..., u_m)$ defined on some non-empty convex set D_1 in R^n, with their range in D, the components u_i being convex if $i \in I_1$, and concave if $i \in I_2$. In particular $U(D ; I_1, I_2)$ contains the set $\mathcal{L}$ of functions, affine component-wise on D_1.

Theorem 3. *The function $F = f \circ u$ is G-convex for all $u \in U(D ; I_1, I_2)$ if and only if the function f is G-convex in D, and partially monotone, i-increasing if $i \in I_1$, i-decreasing if $i \in I_2$. Here G-convex stands for quasi-convex, or strictly quasi-convex and lower-semi-continuous, or, pseudo-convex and f and u restricted to be differentiable.*

Proof. Sufficiency when G-convex means quasi-convex or pseudo-convex follows from [17]. It remains to be proved for f, strictly quasi-convex.

Suppose we have for some x^1, x^2 in some convex set D_1 on which $F = f \circ u$ is defined :

$F(x^2) < F(x^1)$, i.e. $f(u(x^2)) < f(u(x^1))$. Because f is strictly quasi-convex the last inequality implies

$$(15) \qquad f(\lambda u(x^1) + (1 - \lambda)u(x^2)) < f(u(x^1)) \text{ for every } \lambda \in (0, 1).$$

But because u_i is convex when $i \in I_1$ and concave when $i \in I_2$, and the monotonicity assumptions concerning f, we have

$$(16) \qquad f(u(\lambda x^1 + (1 - \lambda)x^2) \leqq f(\lambda u(x^1) + (1 - \lambda)u(x^2)).$$

Hence from (15) and (16) follows

$F(\lambda x^1 + (1 - \lambda)x^2) < F(x^1)$, i.e. $F(x)$ is strictly quasi-convex.

Necessity. From theorem 2 follows that if $F(x)$ is G-convex for all $u \in U(D ; I_1, I_2)$, then f must be G-convex because we can take l defined there as the function u (the affine components of l may be considered, as required, convex or concave).

Partial monotonicity of f. It is enough to prove it for one of the variables, say the first one. Let assume that $1 \in I_1$. Without reducing the generality we may take $m = 1$ and let D be an interval on the real line. We must prove that f is monotone increasing in this interval if F and f are G-convex for all

$u \in U(D\,; I_1, I_2)$. Suppose that this is not true. Hence there are y^* and $\Delta^* > 0$ with $y^* \in D$, $y^* + \Delta^* \in D$, such that

$$f(y^*) - f(y^* + \Delta^*) = \Delta_1 > 0. \tag{17}$$

We can take as function $u_1 : D_1 \to R$, the convex function $u^*(x)$ defined by :

$$u^*(x) = y^* + \frac{2\Delta^*}{b-a}\left| x - \frac{1}{2}(a+b)\right| \qquad (x \in D_1), \tag{18}$$

where D_1 is an interval on the real line, $a, b, \in D_1$, $a < b$ and $|.|$ represents absolute value.

We have

$$u^*(a) = u^*(b) = y^* + \Delta^* \tag{19}$$

and

$$u^*(1/2(a+b)) = y^*. \tag{20}$$

a) G-convex : quasi-convex. We should have from (2)

$$F(1/2(a+b))) \leqq \max\,[F(a), F(b)] \qquad \text{i.e.} \tag{21}$$

$$f(u^*(1/2(a+b)) \leqq f(u^*(a)).$$

But following relations (19) and (20), this contradicts (17).

b) G-convex : lower semicontinuous strictly quasi-convex. Since a lower semicontinuous strictly quasi-convex function is quasi-convex [14], partial monotonicity of f follows from *a*). However the direct proof, using the same function (18), throws some light on the necessity of the lower semicontinuity not yet used in the proof.

From (17), (19) and (20) follows

$$F(a) = F(b) < F(1/2(a+b)). \tag{22}$$

But there are points $x \in (a, 1/2(a+b))$ which satisfy $F(x) > F(a)$, because otherwise the function F would not be lower semicontinuous in $1/2(a+b)$. Let x^* be such a point. Thus $F(b) < F(x^*)$. But this implies $F(y) < F(x^*)$ for every $y \in (x^*, b)$. However we cannot have $F(1/2(a+b)) < F(x^*)$ because of (22) and the strict quasi-convexity of F. Hence the contradiction is proved.

c) G-convex : pseudo-convex. The proof follows from the case *a*) since a pseudo-convex function is quasi-convex and in proving *a*) we could have used a differentiable function $u(x)$, say a parabola satisfying (19) and (20).

We supposed that the component considered, $u^*(x)$, is convex. A similar proof is valid in the concave case. We only have to use instead of the function defined in (18) a concave v-shaped function, or a corresponding parabola. Thus the proof of theorem 3 is completed.

REMARK : The sufficiency part of the theorem is valid without assuming that the strictly quasi-convex function be lower semicontinuous.

REMARK (1) : Theorem 3 remains valid if G-convex stands for *explicit quasi-convex* in the sense of Martos [18].

Corollary 5. *When $u_i(x)$ is linear, the if-part of theorem 3 is valid without i-monotonicity.*

Proof. This follows from the proofs of theorems 2 and 3.

Corollary 6. *The function $F = f \circ u$ is G-concave for all $u \in U(D\,;I_1, I_2)$ if and only if the function f is G-concave in D and partially monotone, i-decreasing if $i \in I_1$, and i-increasing if $i \in I_2$. Here G-concave stands for quasi-concave, strictly quasi-concave and lower semicontinuous or, pseudo-concave and f and u restricted to be differentiable.*

APPLICATIONS

5. Social welfare functions. The definition of optimum welfare in terms of Pareto optimality maintains a certain indeterminacy which may be removed by explicitly introducing a social welfare function [12, 20]. This is a function of the utility levels of all individuals or of groups of individuals and measures the social welfare. There are other considerations which lead to the introduction of a social welfare function to be maximized subject to various restrictions [12, 20]. Let m be the number of individuals or more realistically, of determined groups which compose the society and n be the number of commodities and productive services. We consider the vector function $u : D_1 \to R^m$, where D_1 is a bounded subset of R^{mn}. The components of u are the utility functions of the m individuals or groups (2). We suppose that u belongs to the family of all concave functions defined on D_1, denoted by $U(D_1)$. Let $f : R^m \to R$ be a real valued function called an *utility aggregation function.* For a given utility vector function u and a given aggregation function f, the social welfare function w is defined by

$$w = f \circ u$$

While the components of u are chosen by the individuals or groups, the aggregation function f is established by some planning agency and depends on the institutional framework in which such decision is taken. Such a function will

(1) I am indebted to B. Martos for this remark.
(2) The case of dependent utilities, i.e. the utility of an individual (group) depends also on the consumption and production of others, is thus included.

have to be choosen without full information about the function u which may also change during the time period in which f remains effective. It is therefore a noteworthy problem to find conditions to impose on f, guaranteeing that the resulting social welfare function will have certain utility-like properties, independent of the choice of u within the family $U(D_1)$.

As remarked by Arrow and Enthoven [1, p. 792] the minimal property of all utility functions is quasi-concavity. We have the following immediate consequence of Corollary 6.

Theorem 4. *The social welfare function $w = f \circ u$ is quasi-concave for every vector of utilities $u \in U(D_1)$ if and only if the aggregation function f is a quasi-concave function partially increasing in each component.*

Analogous results are valid if we impose on w the condition to be strictly quasi-concave or pseudo-concave and consequently in this case each local maximum is a global maximum.

6. Non-linear, non-convex programming

In non-linear, non-convex programming it is often important to establish that a local minimum is also a global minimum. Mangasarian shows in [17] that some non-convex programming problems recently considered by various authors have pseudo-convex objective function and thus every local minimum is a global minimum. But this argument cannot be used in the case of non-differentiable objective function.

However from the *if* — parts of theorem 3 and corollary 6 and from the strict quasi-convexity of the function $f(y, z) = y/z$ on either of the convex sets $\{ (y, z) \mid (y, z) \in E^2, z > 0 \}$ or $\{ (y, z) \mid (y, z) \in E^2, z < 0 \}$ and its strict quasi-concavity on $\{ (y, z) \mid (y, z) \in E^2, y > 0 \}$ or $\{ (y, z) \mid (y, z) \in E^2, y < 0 \}$, follows a result analogous to (A) of [17]. Let $\theta(x) = \rho(x)/\sigma(x)$, where $\rho(x)$, $\sigma(x)$ are functions defined on a convex set $D_1 \in E^n$.

Suppose that one of the following assumptions hold in D_1

	ρ	σ
1)	convex $\geqq 0$	concave > 0
2)	concave $\leqq 0$	convex < 0
3)	convex $\leqq 0$	convex > 0
4)	concave $\geqq 0$	concave < 0
5)	linear $\leqq 0$	convex $\neq 0$
6)	linear $\geqq 0$	concave $\neq 0$
7)	convex	linear > 0
8)	concave	linear < 0

Then $\theta(x)$ is strictly quasi-convex on D_1 and every local minimum is a global minimum. The result remains valid if we replace convex by concave and conversely, and minimum by maximum.

If $\theta(x) = \rho(x)\sigma(x)$, then the following implications hold on D_1 (analogous to C of [17]).

	ρ	σ	θ
9)	convex $\leqq 0$	concave $> 0 \Rightarrow$	strictly quasi-convex
10)	convex < 0	concave $\geqq 0 \Rightarrow$	strictly quasi-convex
11)	convex < 0	convex $\leqq 0 \Rightarrow$	strictly quasi-concave
12)	concave $\geqq 0$	concave $> 0 \Rightarrow$	strictly quasi-concave.

As an example of applications of these results to non-convex non-differentiable programming we can give the programming problems investigated by Bector [5]. It follows from 12) that the problems I, III and IV considered there, have strictly quasi-concave objective and linear restrictions and consequently a local maximum is a global maximum.

7. Stochastic programming

a) Consider the non-convex programming problem investigated by various authors [3, 6, 7, 8, 10, 15] which appears in relation to the minimum risk solution of stochastic programming (or P-model)

$$\max_x \left\{ \frac{c'x - \alpha}{\sqrt{x'Vx}} \,\middle|\, Ax \leqq b \quad , \quad x \geqq 0 \right\},$$

where A is an $m \times n$ matrix, V is an $n \times n$ positive definite matrix, c and x are n-dimensional vectors, b is an m-dimensional vector and α is a scalar.

It follows from 5) of § 6, that if on the set $X = \{ x | Ax \leqq b,\ x \geqq 0 \}$ we have $c'x - \alpha \leqslant 0$, then the maximand is strictly quasi-convex. It is obviously lower semicontinuous and hence it is also quasi-convex. From theorem 1 it follows that if the set X is bounded, then the optimum is achieved in one of its vertices.

If on X we have $c'x - \alpha \geqslant 0$, then by replacing in 6) concave by convex, we obtain that the maximand is strictly quasi-concave and a local maximum is a global maximum.

b) Suppose that in the context of a given family of programming problems our information about the state of nature is represented by the probability distribution function $F(z)$ of a certain random vector $b = (b_1, ..., b_m)$. The family of programming problems we have in view is the following :

$$\min_x f(x) \tag{23}$$

subject to

$$P \left\{ \begin{array}{c} g_1(x) \geqq b_1 \\ \cdots\cdots\cdots \\ g_m(x) \geqq b_m \end{array} \right\} \geqq p \quad , \qquad 0 < p < 1 \tag{24}$$

where $f(x)$ is convex and $g_i(x)$, $i = 1, \ldots, m$ are concave functions of n variables and P stands for probability.

In practice $F(z)$ is obtained by fitting some theoretical probability distribution to statistical data and there is some freedom to choose the type of distribution to be fitted.

Therefore it is interesting to obtain conditions which, if satisfied by $F(z)$, assure that the non-linear programming problem (23), (24) has required convexity properties (i.e. it is a convex programming problem or at least (24) defines a convex set).

We have the following

Theorem 5. *A sufficient condition that the function*

$$\varphi(x) = P\{ g_1(x) \geqq b_1, \ldots, g_m(x) \geqq b_m \}$$

be concave (quasi-concave) is that $F(z)$ be concave (quasi-concave). The condition is also necessary if $\varphi(x)$ is to be concave (quasi-concave) for every

$$g(x) = (g_1(x), \ldots, g_m(x))$$

with components defined and concave on a given bounded convex set in E^n.

Proof. $\varphi(x) = F(g_1(x), \ldots, g_m(x))$. Hence the statement of the theorem follows from [4] and theorem 3.

Corollary 7. *The function $\varphi(x)$ is quasi-concave if $F(z)$ is logarithmic concave (i.e.* log *$F(z)$ is a concave function).*

Proof. This follows from the fact that

$$F(\lambda z_1 + (1 - \lambda)z_2) \geqq [F(z_1)]^{\lambda} [F(z_2)]^{1-\lambda} \text{ implies}$$
$$F(\lambda z_1 + (1 - \lambda)z_2) \geqq \min [F(z_1), F(z_2)].$$

REMARK. When the random vector b has a multinormal distribution, $F(z)$ is logarithmic concave ([2, 19]).

Aknowledgements. I am indebted to Werner Œttli, Cristian Bergthaller and the referee for useful remarks on an earlier draft of this paper.