Discussion Paper No. 1067

# Monotonicity of Solution Sets for Parameterized Optimization Problems

by

**Leslie M. Marx***

First Version: December 1991
This Version: April 1993

ABSTRACT

This paper uses results of lattice theory, convex analysis, and nonsmooth analysis to establish conditions for the existence of pure- strategy Nash equilibria when payoff functions are continuous but not quasi-concave. A class of economic games for which this theory is important is given.

---

**Department of Economics, Northwestern University.*

The author acknowledges the support of an NSF Graduate Fellowship and is grateful to Stanley Reiter for numerous helpful discussions.

## 1. Introduction.

Many simple economic models fail to possess Nash equilibria, particularly Nash equilibria in pure strategies. The non-existence of a pure-strategy Nash equilibrium in an economic model may lead economists to doubt the validity or discount the usefulness of a model. Games that lack a pure-strategy Nash equilibrium must violate one or more of the hypotheses of the classical existence theorems (e.g., Debreu (1952), Fan (1952), and Glicksberg (1952)).[1] These hypotheses typically include continuity and a limited form of quasi-concavity of the payoff functions, in addition to the usual convexity and compactness assumptions on strategy sets. Dasgupta and Maskin (1986) show that one can still prove the existence of a pure-strategy Nash equilibrium when the continuity assumption is relaxed, as long as payoff functions still satisfy quasi-concavity. In this paper, I will show that the quasi-concavity assumption can be relaxed, as long as payoff functions are continuous. The relaxation of the quasi-concavity assumption facilitates the proof that pure-strategy Nash equilibria exist for a class of economic games and is based on results from lattice theory, convex analysis, and nonsmooth analysis. In order to relax the quasi-concavity assumption, I use a generalization of existing lattice-theoretic results in conjunction with techniques of convex and nonsmooth analysis. Thus, this paper generalizes the lattice approach of Milgrom and Roberts (1990) and Milgrom and Shannon (1991) and unifies the techniques of lattice theory (see Topkis (1978), Veinott (1989), Vives (1989)), convex analysis (see Rockafellar (1970, 1981)), and nonsmooth analysis (see Clarke (1983), Rockafellar (1981, 1990B)).

In order to study questions of the existence of pure-strategy Nash equilibria, I will begin by focusing on how one player's set of best responses varies with changes in her opponents' strategy choices. This basic problem involves studying how the set $m(t) \equiv \arg\min_{x \in X} f(t,x)$ varies with changes in the parameter t. When X is a lattice, we can use lattice theory to describe how the sets m(t) vary

---

[1]Also see Arrow and Debreu (1954).

with changes in t by defining a binary relation on the sets of optimizing values. One such binary relation is the strong set order, $\geq_S$, of Veinott (1989), which is defined as follows: for lattice P and $X,Y \subseteq P$, $X \geq_S Y$ if for all $x \in X$ and $y \in Y$, $x \wedge y \in Y$ and $x \vee y \in X$.[2] Thus, $X \geq_S Y$ if for every pair of elements, one in X and one in Y, the greatest lower bound is in the smaller set and the least upper bound is in the larger set.[3] The strong set order defines a notion of monotonicity of sets. The methods I propose in this paper are a generalization of the lattice approach in the sense that the notion of monotonicity I define is an extension of monotonicity in the strong set order. The more general definition of monotonicity of this paper is based on the concepts of convex and nonsmooth analysis. In this sense, the work of this paper brings together lattice theory, convex analysis, and nonsmooth analysis. By relying on concepts based in each of these three areas, we can analyze problems that do not fall wholly into any single area.

In the following section, I present some simple examples to illustrate how the lattice-theoretic conditions of decreasing differences and submodularity relate to the conditions required in simple comparative statics problems. In the third section, I define the concept of monotonicity and rate of change that I will use and present a theorem stating that this concept of monotonicity is indeed an extension of the strong set order. This section contains three simple examples. The fourth section shows how the concept of monotonicity introduced in this paper gives results that parallel those of the lattice approach as far as proving the existence of pure-strategy Nash equilibria. Since the definition of monotonicity used in this paper is a generalization of the lattice-theoretic notion, there exist classes of games which cannot be analyzed using the lattice-based approach, but which can be shown to have a pure-strategy Nash equilibrium using the methods of this paper. I describe such a

---

[2]The meet of x and y, denoted $x \wedge y$, is the greatest lower bound of x and y. The join of x and y, denoted $x \vee y$, is the least upper bound of x and y.

[3]When P is a subset of the real numbers and X and Y are convex, $X \geq_S Y$ if and only if $\min(X) \geq \min(Y)$ and $\max(X) \geq \max(Y)$.

class of games in Section 4. In Sections 5 and 6, I use results from convex and nonsmooth analysis to show how the monotonicity condition can be viewed as a condition on the subderivatives of a function. Some examples are given to illustrate these results.

**2. Results of the Lattice Approach.**

When the strategy sets of a game are sublattices, some simple lattice-theoretic conditions on the players' payoff functions guarantee the existence of a pure-strategy Nash equilibrium (see Milgrom and Roberts (1990)). These conditions imply that the set of maximizers of a player's payoff function is nondecreasing in the strong set order with respect to the opponents' strategies. Thus by Tarski's Fixed Point Theorem there exists a pure-strategy Nash equilibrium (see the Appendix for a statement and proof of Tarski's Theorem). I will introduce these game-theoretic conditions within the context of some simple examples. This will allow the reader to understand the role that each condition plays. But first I devote one paragraph to some definitions and observations.

A lattice is a partially ordered set with the property that for any two elements of the set, the greatest lower bound (meet, $\wedge$) and least upper bound (join, $\vee$) are also elements of the set. For example, the subset of $\mathbf{R}^2$ defined by $\{(x,y)|x\geq 0,y\geq 0\}$ is a lattice, while the set $\{(x,y)|x+y=1,x\geq 0,y\geq 0\}$ is not. To see why the latter set is not a lattice, note that (0,1) and (1,0) are elements of the set but their meet (0,0) and join (1,1) are not. A sublattice is a subset of a lattice that contains the meet and join of all pairs of its elements, calculated with regard to the inherited order. A complete lattice is a lattice that contains the infimum and supremum of all of its nonempty subsets. In particular, if X is a complete lattice, then it has an infimum and a supremum. Any compact interval in $\mathbf{R}^n$ is a complete lattice.

Now consider the following example. Suppose the function $f:\mathbf{R}x\mathbf{R}\rightarrow\mathbf{R}$ is of class $C^2$, and suppose we are interested in the behavior of the set arg $\min_{x\in\mathbf{R}}f(t,x)$. In particular, suppose that this set is a singleton, $\{x^*\}$, and that we are interested in conditions such that $dx^*/dt \geq 0$. The standard first and second order conditions tell us that $f_x(t,x^*)=0$ and $f_{xx}(t,x^*)\geq 0$. Differentiating the first order condition gives us

$$f_{tx}(t,x^*) + f_{xx}(t,x^*)\, dx^*/dt = 0.$$

Now note that if $f_{tx}(t,x^*)\leq 0$, then $dx^*/dt \geq 0$. The condition that $f_{tx}(t,x^*)\leq 0$ is one of the conditions used in the lattice approach and is labeled decreasing differences.

**Definition:** A function $f{:}TxX \rightarrow \mathbf{R}$, where T is a partially ordered set and X is a lattice, has decreasing differences in its two arguments x and t if for all $x' \geq x$, the difference $f(t,x')-f(t,x)$ is nonincreasing in t.

From Topkis (1978), we know that if f is a $C^2$ function $f{:}\mathbf{R}^m x \mathbf{R}^n \rightarrow \mathbf{R}$, then f has decreasing differences in x and t if and only if $f_{t_j x_i} \leq 0$ for all $1 \leq i \leq n$ and $1 \leq j \leq m$. So in our example, f has decreasing differences in x and t if and only if $dx^*/dt \geq 0$.

Now let us consider the function $f{:}\mathbf{R} x \mathbf{R}^2 \rightarrow \mathbf{R}$, also assumed to be $C^2$. Suppose that $f(t,x_1,x_2)$ has a unique minimizer x* for all values of t. As before we can write the first and second order conditions for the minimization problem:

$$\begin{pmatrix} f_{x_1} \\ f_{x_2} \end{pmatrix}(t,x*) = 0$$

$$\begin{pmatrix} f_{x_1x_1} & f_{x_1x_2} \\ f_{x_1x_2} & f_{x_2x_2} \end{pmatrix}(t,x^*) \text{ is positive semi-definite.}$$

Differentiating the first order condition, we get:

$$\begin{pmatrix} f_{x_1x_1} & f_{x_1x_2} \\ f_{x_1x_2} & f_{x_2x_2} \end{pmatrix}\begin{pmatrix} \frac{dx_1*}{dt} \\ \frac{dx_2*}{dt} \end{pmatrix} = \begin{pmatrix} -f_{tx_1} \\ -f_{tx_2} \end{pmatrix}.$$

Using Cramer's Rule, we can write:

$$\frac{dx_1*}{dt} = \frac{1}{\Delta}\begin{vmatrix} -f_{tx_1} & f_{x_1x_2} \\ -f_{tx_2} & f_{x_2x_2} \end{vmatrix}$$

where $\Delta$ is nonnegative by the second order condition. Rewriting the equation gives:

$$\frac{dx_1*}{dt} = \frac{1}{\Delta}\left[-f_{tx_1}f_{x_2x_2} + f_{tx_2}f_{x_1x_2}\right].$$

We can see from the above expression that if $f_{tx_1} \leq 0$ and $f_{tx_2} \leq 0$ and $f_{x_1x_2} \leq 0$, then $dx_1*/dt \geq 0$. The decreasing differences condition implies the first two inequalities, and the lattice approach employs an additional condition, submodularity, which implies the third.

**Definition:** A function $f:X\rightarrow\mathbf{R}$, where X is a lattice, is <u>submodular</u> if for all $x,y\in X$, $f(x)+f(y) \geq f(x\wedge y)+f(x\vee y)$.

From Topkis (1978), we know that if f is a $C^2$ function $f:\mathbf{R}^n\rightarrow\mathbf{R}$, then f is submodular if and only if $f_{x_i x_j}\leq 0$ for all $1\leq i<j\leq n$. There are no restrictions on $f_{x_i x_i}$.

So in our example, if f has decreasing differences in x and t and is submodular, then $dx_1^*/dt\geq 0$, and similarly for $dx_2^*$.

We also want to study cases where the set of minimizers is not a singleton. In order to say whether the set of minimizers is "increasing" with increases in the parameter, we must define an ordering on sets and the parameters must be drawn from a partially ordered set. In particular, using the strong set order defined in the Introduction, we have the following result:

**2.1 Theorem:** Let $f:TxX\rightarrow\mathbf{R}\cup\{\infty\}$, where T is a partially ordered set and X is a lattice, be order lower semicontinuous and submodular in x given t and satisfy decreasing differences in x and t. Let S be a sublattice of X. Then $m(t)\equiv\arg\min_{x\in S}f(t,x)$ is monotone nondecreasing in the strong set order in t.[4]

At this point it may not be clear why Theorem 2.1 requires that the constraint set for the minimization problem be a sublattice. To illustrate the need for this assumption, consider the following problem: $\min_{x\in R^2}f(t,x)$ subject to $B'x=0$, $B\neq 0$. The Lagrangian for this problem is $L=f(t,x)-\lambda B'x$. The first order conditions are $\nabla_x f(t,x^*)-\lambda B=0$ and $B'x^*=0$. The second order

[4]A more general version of this theorem states: Let $f:TxX\rightarrow\mathbf{R}\cup\{\infty\}$, where T is a partially ordered set and X is a lattice, be submodular in x given t and satisfy decreasing differences in x and t. Then $m(S,t)\equiv\arg\min_{x\in S}f(t,x)$ is monotone nondecreasing in the strong set order in (S,t).

condition requires that the following bordered Hessian have nonpositive determinant:

$$\begin{pmatrix} 0 & B_1 & B_2 \\ B_1 & \nabla_x^2 f(t,x*) & \\ B_2 & & \end{pmatrix}.$$

Differentiating the first order condition gives us:

$$\begin{pmatrix} 0 & B_1 & B_2 \\ B_1 & \nabla_x^2 f(t,x*) & \\ B_2 & & \end{pmatrix} \begin{pmatrix} \frac{d\lambda}{dt} \\ \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{pmatrix} = \begin{pmatrix} 0 \\ -f_{tx_1} \\ -f_{tx_2} \end{pmatrix}.$$

So by Cramer's rule $dx_1*/dt$ satisfies the following equation:

$$\frac{dx_1^*}{dt} = \frac{1}{\Delta} \begin{vmatrix} 0 & 0 & B_2 \\ B_1 & -f_{tx_1} & f_{x_1x_2} \\ B_2 & -f_{tx_2} & f_{x_2x_2} \end{vmatrix}.$$

Rewriting this gives us:

$$\frac{dx_1^*}{dt} = \frac{1}{\Delta} B_2 \left[ B_1 \left( -f_{tx_2} \right) + B_2 f_{tx_1} \right].$$

By the second order condition, $1/\Delta$ is nonpositive, and by the decreasing differences condition $f_{tx_1}$ and $f_{tx_2}$ are nonpositive. So the sign of $dx_1*/dt$ depends upon the signs of $B_1$ and $B_2$. The decreasing differences and submodularity assumptions are not sufficient to guarantee that $dx_1*/dt$ is nonnegative.

However, if $\{x|B'x=0\}$ is a sublattice, then either $B_2=0$ or $B_1/B_2 \leq 0$.[5] Note that this assumption in addition to the decreasing differences assumption is sufficient for the result that $dx_1^*/dt \geq 0$.

For convenience I will now define a "min-monotone" function. A min-monotone function is simply a parameterized function whose sets of minimizers are nondecreasing in the strong set order in the parameter. Theorem 2.2 states the now obvious result that functions satisfying submodularity and decreasing differences are min-monotone.

**Definition:** Let T be a partially ordered set, X a lattice, and S a subset of X. A function f(t,x), $f:TxX \rightarrow \mathbf{R} \cup \{\infty\}$, is min-monotone in t with respect to S if for all $t,t' \in T$ such that $t'>t$, $\arg\min_{x \in S} f(t',x) \geq_S \arg\min_{x \in S} f(t,x)$.

**2.2 Theorem:** Let T be a partially ordered set, X a lattice, and S a sublattice of X. Suppose $f:TxX \rightarrow \mathbf{R} \cup \{\infty\}$ is submodular in x given t and has decreasing differences in x and t. Then f is min-monotone in t with respect to S.

**3. Monotonicity Defined in terms of Nonsmooth Analysis.**

From the previous section, we know that the strong set order can be used to compare two sets. In particular, we can use the order to define monotonicity of a set-valued mapping. However, a more general definition of monotonicity will be sufficient to prove the existence of pure-strategy Nash equilibria. The concept of monotonicity that will be used is defined in this section. Theorem 3.2 states that this type of monotonicity is, in fact, a generalization of monotonicity in the strong set

---

[5]Note that $\{x|B'x=0\}$ is a line through the origin or all of $\mathbf{R}^2$. In order for the set to be a sublattice of $\mathbf{R}^2$, it must be a line with nonnegative slope or be equal to all of $\mathbf{R}^2$. Suppose $B_2 \neq 0$. Then the set is a line with slope $-B_1/B_2$. So the set is a sublattice if and only if $B_2=0$ or $B_1/B_2$ is nonpositive.

order. I conclude this section with some simple examples.

The following definitions will allow us to talk about the "derivative" of a set-valued mapping. I will give a simplified explanation before getting to the formal definitions. Consider a multifunction $m:\mathbf{R}^m \rightrightarrows P(\mathbf{R}^n)\cup\emptyset$ defined by $m(t)=\arg\min_{x\in S} f(t,x)$ where f is function mapping $TxX \subseteq \mathbf{R}^m x \mathbf{R}^n$ to $\mathbf{R}\cup\{\infty\}$ and S is a subset of $\mathbf{R}^n$.[6] For every t, m(t) gives the values for x that minimize f(t,x) on S. The graph of the multifunction m is a set in the space $\mathbf{R}^m x \mathbf{R}^n$. The characteristics of the graph of m provide the basis for the more general concept of the monotonicity of a set-valued mapping. For every point $(\bar{t},\bar{x})$ in the graph of m, we can define a set called the approximating cone -- roughly, the translation to the origin of the smallest cone with vertex at $(\bar{t},\bar{x})$ that encompasses (locally) the graph of m. This cone is a set in $\mathbf{R}^m x \mathbf{R}^n$. We can also define a multifunction whose graph is exactly this approximating cone. This multifunction will be called the proto-derivative of m. We can think of the proto-derivative as a type of directional derivative. The proto-derivative of m at $\bar{t}$ relative to $\bar{x}$ in direction t' is the set of changes x' in $\bar{x}$ such that the point (t',x') is an element of the approximating cone to gph m at $(\bar{t},\bar{x})$.

To proceed with the formal definition of the approximating cone and proto-derivative, I will need to define the contingent cone and the derivative cone. Let C be a closed set, and let $x\in C$. The contingent cone to C at x is defined as $\limsup_{t\downarrow 0} t^{-1}(C-x)$ and the derivative cone to C at x is defined as $\liminf_{t\downarrow 0} t^{-1}(C-x)$.[7] Both are always closed cones containing the origin. C is said to be approximable at x if the two cones coincide.[8]

---

[6]Unlike a correspondence, a multifunction may map a value in the domain onto the empty set.

[7]To formalize the notion of the lim sup and lim inf of a sequence of sets, we have: the set $\limsup_{t\downarrow 0} t^{-1}(C-x)$ is the set of cluster points to sequences of the form $(y^k)$, where $y^k \in (\tau^k)^{-1}(C-x)$, and the set $\liminf_{\tau\downarrow 0} \tau^{-1}(C-x)$ is the set of limits of such sequences.

[8]The contingent and derivative cones of a set C will always be equal when C is closed and convex. However, the following example shows that they need not always coincide. Let $C=\{0\}U\{1/k\,|\,k=1,2,...\}$. C is closed but not convex. Since there are no arcs entirely contained in

**Definition:** Let C be a closed subset of $\mathbf{R}^n$ and let $x \in C$. The <u>contingent cone</u> to C at x is the closed set $K_C(x) \equiv \lim\sup_{\tau \downarrow 0} \tau^{-1}(C - x)$.

**Definition:** Let C be a closed subset of $\mathbf{R}^n$ and let $x \in C$. The <u>derivative cone</u> to C at x is the closed set $V_C(x) \equiv \lim\inf_{\tau \downarrow 0} \tau^{-1}(C - x)$.

The derivative cone gets its name from its relationship with the derivatives of emanating arcs in C at x. For $C \subseteq \mathbf{R}^n$, an emanating arc in C at x is a function $y:[0,\tau) \to \mathbf{R}^n$ such that $y(0)=x$, $y(t) \in C$ for all $t \in [0,\tau)$, $y(t) \to x$ as $t \downarrow 0$, and $y_+'(0) \equiv \lim_{t \downarrow 0}[y(t)-y(0)]/t$ exists. Note that $y_+'(0)$ is the (right) derivative of y at x. It can be shown that the derivative cone consists of all vectors $\xi \in \mathbf{R}^n$ expressible as $\xi = y_+'(0)$ for the various emanating arcs y in C at x (see Rockafellar (1989B)).

As an illustration, if $f:\mathbf{R} \to \mathbf{R}$ is differentiable at a point x, then the contingent and derivative cones to gph f at $(x,f(x))$ are the same and are equal to the graph of the line through the origin with slope equal to $f'(x)$. If f has left and right derivatives at x, then the cones are again the same and are equal to the union of the ray emanating from the origin with slope $f_+'(x)$ and the ray emanating from the origin with slope $f_-'(x)$.

**Definition:** Let C be a closed subset of $\mathbf{R}^n$ and let $x \in C$. If $K_C(x)=V_C(x)$, then C is <u>approximable</u> at x. If $K_C(x)=V_C(x)$ for all $x \in C$, then C is <u>approximable</u>.

**Definition:** The multifunction $G:\mathbf{R}^d \rightrightarrows \mathbf{R}^n$ is <u>proto-differentiable</u> at t relative to $x \in G(t)$ if the set gph G is approximable at (t,x). The <u>proto-derivative</u> of G at t relative to x is the multifunction

---

C that include the point 0, $V_C(0)=\{0\}$. To see that $K_C(0)=\{x | x \geq 0\}$, note that $K_C(0)$ is a cone contained in $\mathbf{R}$ with no negative elements. It is sufficient to show that $1 \in K_C(0)$, i.e. that there exists $\tau^k \downarrow 0$ and $x^k \in C$ such that $x^k/\tau^k \to 1$. Defining $\tau^k \equiv 1/k$ and $x^k \equiv 1/k$ for $k=1,2,\ldots$, we are done.

$DG_{t|x}: R^d \rightrightarrows R^n$ defined by gph $DG_{t|x} = K_{gph\ G}(t,x) = V_{gph\ G}(t,x)$.[9]

The following definition introduces that notion of monotonicity of sets that I will use in this paper. Essentially, it requires that sets have positive proto-derivatives. In particular, whenever a set has a nonempty proto-derivative in a positive direction, the proto-derivative in that direction must contain a positive element. Note that the proto-derivative is defined for multifunctions defined on all of $\mathbf{R}^m$, but this is not restrictive since multifunctions may have the empty set as values.

**Definition:** Let $\bar{t} \in \mathbf{R}^m$. A multifunction $G: \mathbf{R}^m \rightrightarrows \mathbf{R}^n$ is monotone at $\bar{t}$ if $\forall \bar{x} \in \mathbf{R}^n$, $\forall t' \in \mathbf{R}^m$ with $t' \geq 0$ and $t' \neq 0$, $DG_{\bar{t}|\bar{x}}(t') \neq \emptyset \Rightarrow \exists x' \geq 0$ s.t. $x' \in DG_{\bar{t}|\bar{x}}(t')$. G is monotone if G is monotone at every t.

By definition, the set of minimizers of a parameterized function is monotone if for any increase in the parameter, all elements of the set of minimizers can be viewed (loosely) as changing by some nonnegative amount. If the elements of the set of minimizers always move in proportion to the change in the parameter, then we can define the rate of increase of the set. As made precise in Theorem 3.1, if the rate of increase of a multifunction is nonnegative, then the multifunction is monotone. The proof of Theorem 3.1 follows from directly the definitions of monotone and rate of increase.

**Definition:** Let $\bar{t} \in \mathbf{R}^m$, and let $G: \mathbf{R}^m \rightrightarrows \mathbf{R}^n$ be a multifunction. If there exists $r \in \mathbf{R}$ such that $\forall \bar{x} \in G(\bar{t})$ and $\forall t' \in \mathbf{R}^m$ with $t' \geq 0$ and $t' \neq 0$, $r \cdot t' \in DG(\bar{t}|\bar{x})(t')$, then G has rate of increase at $\bar{t}$ of r.

**3.1 Theorem:** Let $\bar{t} \in \mathbf{R}$, and let $G: \mathbf{R} \rightrightarrows \mathbf{R}$ be a multifunction. If G has nonnegative rate of increase at

[9]Note that for $x \notin G(t)$, $DG_{t|x}$ does not exist. Also, see Rockafellar (1989B).

$\bar{t}$, then G is monotone at $\bar{t}$.

The next theorem states that when f is min-monotone, the multifunction $m:\mathbf{R}^m \rightrightarrows \mathbf{R}^n$ defined by $m(t)=\arg\min_{x\in X} f(t,x)$ is monotone. Thus the assumptions of the lattice approach, submodularity and decreasing differences, are stronger than the assumption of monotonicity. In Example 3.3 I show that monotonicity does not imply submodularity and decreasing differences.

**3.2 Theorem:** Let $T\text{x}X\subseteq\mathbf{R}^m\text{x}\mathbf{R}^n$. If $f:T\text{x}X\rightarrow\mathbf{R}\cup\{\infty\}$ is min-monotone in t with respect to S, then the multifunction $m:\mathbf{R}^m \rightrightarrows \mathbf{R}^n$ defined by $m(t)=\arg\min_{x\in S} f(t,x)$ is monotone on T.[10]

Theorem 3.2 together with Theorem 2.2 proves that all functions satisfying the assumptions of the lattice approach will have associated multifunctions, $m(t)\equiv\arg\min_{x\in X} f(t,x)$, that are monotone:

**3.3 Theorem:** Let $T\text{x}X\subseteq\mathbf{R}^m\text{x}\mathbf{R}^n$. Let $f:T\text{x}X\rightarrow\mathbf{R}\cup\{\infty\}$, where X is a complete lattice. Suppose f(t,x) is order lower semicontinuous and submodular in x given t and has decreasing differences in x and t. Then the multifunction $m:\mathbf{R}^m \rightrightarrows \mathbf{R}^n$ defined by $m(t)=\arg\min_{x\in X} f(t,x)$ is monotone.

In order to get a better grasp of the concept of a monotone multifunction just introduced, I will now present three simple examples. For now I will not describe the actual calculation of the proto-derivatives in the different examples, but this will be presented in detail in Sections 5 and 6.

**Example 3.1.**

Consider the function $f:\mathbf{R}\text{x}\mathbf{R}\rightarrow\mathbf{R}$ defined by $f(t,x)=|x-t|$. It is easy to see that m(t)=arg

---

[10]Here, as in the remainder of this section, for $t\in\mathbf{R}^m\backslash T$, define $m(t)\equiv\emptyset$.

$\min_{x\in\mathbf{R}} f(t,x)=\{t\}$. Also, m is clearly nondecreasing in t with respect to the strong set order since for $t'\geq t$, $\{t'\}\geq_S\{t\}$. In Example 5.1, I calculate that $Dm_{\bar{t}|\bar{t}}(t')=\{t'\}$. For given $\bar{t}$, m has rate of increase at $\bar{t}$ of 1 since $\forall t'>0$, $1\cdot t'\in Dm_{\bar{t}|\bar{t}}(t')$. Thus, by Theorem 3.1, m is monotone.

**Example 3.2.**

This example is similar to Example 3.1, but it illustrates the case where the set of minimizers of a function is not a singleton set. Define the function f as follows:

$$f(t,x) = \begin{cases} t-x, & x<t \\ 0, & t\leq x\leq t+1 \\ x-(t+1), & t+1<x. \end{cases}$$

The function attains its minimum value when $x\in[t,t+1]$, so $m(t)=[t,t+1]$. As calculated in Example 5.2,

$$Dm_{\bar{t}|\bar{x}}(t') = \begin{cases} [t',\infty), & \bar{x}=\bar{t} \\ (-\infty,\infty), & \bar{t}<\bar{x}<\bar{t}+1 \\ (-\infty,t'], & \bar{x}=\bar{t}+1. \end{cases}$$

For given $\bar{t}$, m has rate of increase at $\bar{t}$ of 1 since $\forall\bar{x}\in m(\bar{t})$ and $\forall t'>0$, $1\cdot t'\in Dm_{\bar{t}|\bar{x}}(t')$. Thus, by Theorem 3.1, m is monotone.

**Example 3.3**

Consider the function $f(t,x)=(x-t)^4-2(x-t)^2$, which achieves its minimum value at $x=t-1$ and $x=t+1$, so $m(t)=\{t-1,t+1\}$. Note that f does not have decreasing differences in x and t.[11] Also,

[11]Here the decreasing differences condition is equivalent to the condition that $f_{tx}(t,x)\leq 0$ for all $(t,x)$. However, $f_{tx}(t,x) = -12(x-t)^2+4$, which is positive for $(x-t)^2 < 1/3$.

f is not min-monotone since, for instance, $m(1)=\{0,2\} \not\geq_S \{-1,1\}=m(0)$.[12] For given $\bar{t}$, Example 6.1 shows that for all $\bar{x}\in m(\bar{t})$, $Dm_{\bar{T}|\bar{x}}(t')=\{t'\}$. So m has rate of increase at $\bar{t}$ of 1. By Theorem 3.1, m is monotone.

## 4. Existence of Pure-Strategy Nash Equilibria

In this section I define a submodular game and give a key theorem regarding such games.[13] Then I present the definitions of two-player and multi-player monotone games and theorems on the existence of pure-strategy Nash equilibria in such games. Following this, I define a class of games which are not submodular but which are monotone, and thus can be shown to have a pure-strategy Nash equilibrium using the techniques of this paper.

Consider the normal form game $\Gamma=(S_n,f_n)_{n\in N}$. Here the function $f_n$ maps $S_{-n}xS_n$ onto $\mathbf{R}\cup\{\infty\}$, and player n maximizes utility by minimizing the value of $f_n$. The game $\Gamma$ is a submodular game if, for each $n\in N$:

(A1) $S_n$ is a compact interval in $\mathbf{R}^{k_n}$;

(A2) $f_n$ is continuous in $x_n$ (for fixed $x_{-n}$);

(A3) $f_n$ is submodular in $x_n$ (for fixed $x_{-n}$);

(A4) $f_n$ has decreasing differences in $x_n$ and $x_{-n}$.

The key theorem in the Milgrom and Roberts analysis (their Theorem 5) states that if $\Gamma$ is a supermodular game, then the game has a pure-strategy Nash equilibrium. It is rewritten below as

---

[12] $1\in m(0)$ and $0\in m(1)$, but $1\wedge 0=0\notin m(0)$, so $m(1) \not\geq_S m(0)$.

[13] The game as I define it is actually a hybrid of a submodular game and a smooth submodular game.

a theorem about submodular games.

**4.1 Theorem (Milgrom and Roberts (1990)):** A submodular game $\Gamma$ has a pure-strategy Nash equilibrium.

In summary, (A1)-(A4) guarantee that the "upper bound" of the set gph $m_n$ is a monotone nondecreasing function. If player n always chooses his or her maximal best response to the other $n-1$ players' strategies, then player n's strategies will increase as the other players' strategies increase. The existence of these monotone nondecreasing best-response functions for all the players in the game implies, by Tarski's Fixed Point Theorem, that there must be a pure-strategy Nash equilibrium of the game.

It is notable that it is exactly the characteristics of $m_n$ that make it monotone which allow the proof that the game has a pure-strategy Nash equilibrium. In the remainder of this section, we will see precisely in what sense this is true. I will start by defining a monotone game. For now I employ conditions that place restrictions directly on the multifunctions $m_n$. In Sections 5 and 6, I will discuss how the results of convex and nonsmooth analysis allow these conditions to be replaced by conditions on the payoff functions $f_n$.

First consider the very simple case of a two-player game with strategy sets that are compact intervals of the real line. The game $\Gamma=(S_n,f_n)_{n\in N}$ is a two-player monotone game if $N=\{1,2\}$ and for each $n\in N$:

(B1) $S_n$ is a compact interval of $\mathbf{R}$;

(B2) $f_n$ is continuous;

(B3) the multifunction $m_n: \mathbf{R} \rightrightarrows \mathbf{R}$ defined by $m_n(t) = \arg\min_{x \in S_n} f_n(t,x)$ is monotone;[14]

(B4) for all $t < \max(S_{-n})$, $D(m_n)_{t|\max m_n(t)}(1)$ exists and is nonempty.

**4.2 Existence Theorem for Two-Player Monotone Games:** A two-player monotone game $\Gamma$ has a pure-strategy Nash equilibrium.

Before continuing, I will briefly examine the conditions (B1)-(B4). (B1) and (B2) are the same conditions used in a submodular game. They imply that $m_n(t)$ is nonempty. (B3) replaces the decreasing differences assumption. The submodularity assumption is not needed in this case since $S_n$ is one-dimensional. (B4) is a technical assumption needed to eliminate problems caused by border solutions. More will be said on this later.

As Theorem 4.3 states, the conditions of a submodular game (A1)-(A4) imply that the players' best-response mappings are monotone, whereas in a monotone game, the monotonicity of the best-response mappings are assumed explicitly. Theorem 4.3 follows directly from the definition of a submodular game and Theorem 3.3.

**4.3 Theorem:** Let $\Gamma$ be a submodular game with $S_{-n} x S_n \subseteq \mathbf{R}^m x \mathbf{R}^n$. Then the multifunction $m_n: \mathbf{R}^m x \mathbf{R}^n \rightrightarrows \mathbf{R}$ defined by $m_n(t) = \arg\min_{x \in S_n} f_n(t,x)$ is monotone.

This begs the question, is every submodular game also a monotone game? The answer to this is no, as the following example illustrates. The example describes a submodular game which does not

[14]Here, as in the remainder of this section, $m_n(t)$ is defined to be the empty set for all t not contained in $S_{-n}$.

satisfy the border condition (B4).

Let $N=\{1,2\}$, $S_n=[0,2]$, and $f_n(t,x)=(x-t)^4-2(x-t)^2$. With $m_n(t)$ defined as $m_n(t)=\arg\min_{x\in S_n} f_n(t,x)$ for $t\in[0,2]$, we have

$$m_n(t)=\begin{cases}\{t+1\}, & 0\le t<1\\ \{0,2\}, & t=1\\ \{t-1\}, & 1<t\le 2.\end{cases}$$

This correspondence is monotone, but it has no fixed point. The problem in this example is that $D(m_n)_{1|2}(1)=\varnothing$, i.e. (B4) is not satisfied.

To understand how condition (B4) is related to corner solutions, consider Theorem 4.4. This theorem states that (B4) is satisfied if players' best responses are in the interior of their strategy sets and if a second derivative condition holds. The proof of Theorem 4.4 follows from Theorem 5.4 of Rockafellar (1989B).[15]

**4.4 Theorem:** Let $\Gamma$ be a submodular game with strategy sets that are subsets of the real numbers. Suppose that for all $\bar{t}$ and all $\bar{x}\in m_n(\bar{t})$, we have $\bar{x}\in\text{int } S_n$ and $\partial^2 f_n(\bar{t},\bar{x})/\partial x^2\neq 0$. Then (B4) is satisfied.

I have shown that (A1)-(A4) imply (B3) but do not imply (B4). So clearly any submodular game that satisfies (B4) is also a monotone game. The following example, similar to the previous

---

[15]Let $G:\mathbf{R}^d\rightrightarrows\mathbf{R}^n$ have the form $G(t)=\{x\in D\,|\,g(t,x)=\mathbf{0}\}$, where $g:\mathbf{R}^d\times\mathbf{R}^n\to\mathbf{R}^m$ is a continuously differentiable function and the set $D\subseteq\mathbf{R}^n$ is closed and convex. For t and $x\in G(t)$, G is proto-differentiable at t relative to x and the proto-derivative is given by $DG_{t|x}(t')=\{\xi\in K_D(x)\,|\,\nabla_t g(t,x)t'+\nabla_x g(t,x)\xi=\mathbf{0}\}$.

one, shows that not all monotone games are submodular games.

$N=\{1,2\}$, $S_1=[2,5]$, $S_2=[3,4]$, $f_1(t,x)=(x-t)^4-2(x-t)^2$, and $f_2(t,x)=x$. $m_1(t)=\{t-1,t+1\}$ and $m_2(t)=\{3\}$. $m_1$ and $m_2$ both satisfy (B3) and (B4), so $\Gamma$ is a two-player monotone game. However, $f_1$ does not satisfy the decreasing differences condition, so $\Gamma$ is not a submodular game.[16]

Now I will define a monotone game allowing more than two players. As before the definition of the game places restrictions directly on $m_n$, and I defer to Sections 5 and 6 the discussion of how these translate into restrictions on $f_n$.

The game $\Gamma=(S_n,f_n)_{n\in N}$ is a <u>monotone game</u> if, for each $n\in N$:

(C1) $S_n$ is a compact interval of $\mathbf{R}$;[17]

(C2) $f_n$ is continuous;

(C3) the multifunction $m_n:\mathbf{R}^{N-1}\rightrightarrows\mathbf{R}$, $m_n(t)=\arg\min_{x\in S_n} f_n(t,x)$, is proto-differentiable and monotone;

(C4) the function $M_n:S_{-n}\to S_n$ defined by $M_n(t)=\max m_n(t)$ is Lipschitz.

**4.5 Existence Theorem for Monotone Games:** A monotone game $\Gamma$ has a pure-strategy Nash equilibrium.

[16]To see that f does not satisfy decreasing differences, note that $f_1(3,4) = -1 < 0 = f_1(3,3)$, but $f_1(4,4) = 0 > -1 = f_1(4,3)$.

[17]Theorem 4.5 also holds when we allow $S_n$ to be a compact interval of $\mathbf{R}^{k_n}$. In this case (C4) should be read as: $\forall i\in\{1,2,\ldots,k_n\}$, the real-valued function mapping t onto the $i^{th}$ component of $M_n(t)$ is Lipschitz.

Note that in the definition of a monotone game, I have used a Lipschitz condition rather than a condition such as (B4). But Lemma A.3 in the Appendix shows that given (C1)-(C3), (C4) is just a higher dimensional version of (B4).

The following theorem describes the extent to which monotone games and submodular games are related. The proof follows from Theorem 4.3.

**4.6 Theorem:** Let $\Gamma$ be a submodular game with $S_n \subseteq \mathbf{R}$. Suppose $m_n(t)$ is proto-differentiable and that (C4) holds. Then $\Gamma$ is a monotone game.

Now I will describe a simple class of economic games that are monotone but not submodular. In this class of games, there are two competing players, 1 and 2. Each player chooses a parameter, call it quality, from his or her strategy set, which is a compact interval of the real line. The net profit player i receives on each item, $n_i(q_i)$, depends only on his own quality choice and decreases as his quality increases. Each player's total profit is equal to the number of items he or she sells times the net profit per item. The number of items player 1 sells depends upon the difference between his quality level $q_1$ and player 2's quality level $q_2$. Let the function $f_i:\mathbf{R}\rightarrow[0,\infty)$ be such that $f_i(q_i - q_{-i})$ specifies the number of customers who buy from player i when player i chooses $q_i$ and player i's opponent chooses $q_{-i}$. Assume (this is the key characteristic of this class of games) there are constants $k_i$ such that when $q_i$-$q_{-i}$>$k_i$, $f_i(q_i$-$q_{-i})$ is increasing and when $q_i$-$q_{-i}$<$k_i$, $f_i(q_i$-$q_{-i})$ is constant. In other words, the customers in this game have quality thresholds. As long as player i's quality level is sufficiently high relative to his opponent's quality level, then player i can attract more customers by increasing quality. However, if player i's quality level is low relative to his opponent's quality level, then small increases in quality do not enable him to attract more customers.

Games of this type are not submodular games (assume players minimize negative profits) because they do not satisfy the decreasing differences condition. In general, if player 1 chooses low quality, then player 2 will strictly prefer medium quality over low quality. But if player 1 chooses very high quality, then player 2 will weakly prefer low quality over medium quality -- in either case player 2 attracts the same number of customers, and his net profit per item is higher when he chooses low quality.

A more precise example within this framework is given in the Appendix.

**5. Convex Payoff Functions.**

In this section I will show how the conditions on the multifunctions $m_n$ used in Section 4 can be rewritten as conditions on payoff functions $f_n$ when those payoff functions are convex. I will deal with nonconvex payoff functions in Section 6.

The first step is simply to define $m_n(t)$ in terms of $f_n$. Before doing this, I will define the subgradient of a convex function and state a theorem regarding subgradients. Following Theorem 5.1, I give some examples to clarify these concepts.

**Definition:** Let $f:\mathbf{R}^n \to \mathbf{R} \cup \{\infty\}$ be a convex function, and let $x^*, x \in \mathbf{R}^n$. Then $x^*$ is a subgradient of f at x if $f(z) \geq f(x) + x^* \cdot (z-x)$, for all $z \in \mathbf{R}^n$. This inequality is referred to as the subgradient inequality. The set of all subgradients of f at x is the subdifferential of f at x and is denoted $\partial f(x)$.[18]

**5.1 Theorem (Rockafellar (1970)):** Let $f:\mathbf{R}^n \to \mathbf{R}$ be a convex differentiable function. Then $\partial f(x) = \nabla f(x)$.

To get a better understanding of the concept of a subgradient, consider the following

[18]Note that $\partial f$ is a multifunction mapping elements of $\mathbf{R}^n$ onto (possibly empty) subsets of $\mathbf{R}^n$.

examples. If a function $f:\mathbf{R}^n \to \mathbf{R}$ is differentiable at x, then f has only one subgradient at x, and it is equal to the gradient of f at x, $\partial f(x)=\nabla f(x)$. Now consider an example where f is not differentiable at a point. Suppose that f is the function from $\mathbf{R}$ to itself defined by $f(x)=|x|$. Then f is not differentiable at zero. In this case f has a continuum of subgradients at zero, $\partial f(0)=[-1,1]$. To see this, evaluate the subgradient inequality defined above for $x=0$. The definition says that $x^*$ is a subgradient if for all z, $f(z)\geq f(0)+x^*\cdot(z-0)$, or substituting the function definition, $|z|\geq x^*\cdot z$. Clearly this is true for all $x^*\in[-1,1]$. The subgradients include the left derivative and right derivative of the function, plus all points in between. The subgradients are all those values that can be used to define supporting hyperplanes for the epigraph of the function. In the above example, for each $x^*\in[-1,1]$, the vector $(x^*,-1)$ can be used as a normal vector to define a supporting hyperplane for the function f at the point (0,0). In this example those hyperplanes are defined by $h(z)=x^*\cdot z$, lines with slope $x^*$ that pass through the origin.

In general, when f is finite at x and $x^*\in\partial f(x)$, the graph of the affine function $h(z)=f(x) + x^*\cdot(z-x)$ is a non-vertical supporting hyperplane to the convex set epi f at the point (x,f(x)). If $x^*$ is a subgradient of f at x, then the vector $(x^*,-1)$ is normal to epi f at (x,f(x)).[19]

The following theorems state some properties of subgradients.

**5.2 Theorem:** Let $f:\mathbf{R}^n \to \mathbf{R}\cup\{\infty,-\infty\}$ be a convex function. $\partial f(x)$ is a closed convex set in $\mathbf{R}^n$.

**5.3 Theorem (Proposition 5A of Rockafellar (1981)):** If f is convex, the following conditions are equivalent to each other:

(a) f has its global minimum at x;

(b) f has a local minimum at x;

---

[19]If $x^*\in\partial f(x)$ and $(a,\beta)\in$epi f, then $((a,\beta)-(x,f(x)))\cdot(x^*,-1) = (a-x)\cdot x^* - \beta + f(x) \leq f(a) - \beta \leq 0$.

(c) $0 \in \partial f(x)$.

Now we have the tools to define the multifunction $m_n$ in terms of the payoff functions $f_n$. I will assume that the payoff functions are defined on all of $\mathbf{R}^m x \mathbf{R}^n$ since we can let $f_n$ have value $\infty$ for points outside the domain.

Applying Theorem 5.3, we have:

**5.4 Theorem:** Let $f:\mathbf{R}^m x \mathbf{R}^n \rightarrow \mathbf{R} \cup \{\infty\}$ be a convex function. Define multifunction $m:\mathbf{R}^m \rightrightarrows \mathbf{R}^n$ by $m(t) \equiv \arg\min_{x \in \mathbf{R}^n} f(t,x)$. Then $m(t) = \{x \mid (y,0) \in \partial f(t,x), \text{ some } y\}$.

Now we can prove the following theorem showing how the proto-derivative of $m_n$ can be written in terms of the function $f_n$.

**5.5 Theorem:** Let $f:\mathbf{R}^m x \mathbf{R}^n \rightarrow \mathbf{R} \cup \{\infty\}$ be a convex function. Define multifunction $m:\mathbf{R}^m \rightrightarrows \mathbf{R}^n$ by $m(t) = \arg\min_{x \in \mathbf{R}^n} f(t,x)$. Let $\bar{t}$ and $\bar{x}$ be given. Suppose $\bar{y} \in \partial f(\bar{t},\bar{x})$. Then the proto-derivative of m, if it exists, is given by

$$Dm_{\bar{t}|\bar{x}}(t') = \{x' \mid (y,0) \in D(\partial f)_{\bar{t},\bar{x}|\bar{y},0}(t',x'), \text{ some } y\}.$$

Using Theorem 5.5, we can formulate the conditions of Section 4, such as (B3), (B4), (C3), and (C4), in terms of $f_n$ rather than $m_n$ whenever the payoff functions are convex.[20] However, it would be preferable to have formulas for $Dm_n$ that are more easily calculated. Theorem 5.5 reduces our problem to finding formulas for the proto-derivative of the subgradient mapping of $f_n$. In the rest

[20]Replace references to $m_n$ by the expression given in Theorem 5.5.

of this section, we will develop such formulas. As a matter of notation, $\partial_t f(\bar{t},\bar{x})$ is the subderivative of the function of t, $f(t|x=\bar{x})$, at $t=\bar{t}$. Similarly for the gradient $\nabla_t f(\bar{t},\bar{x})$.

Theorem 5.6 uses the following definition to give one such formula and illustrates that the proto-derivative can be viewed as a type of second derivative.

**Definition:** Let $f:\mathbf{R}^n\to\mathbf{R}\cup\{\infty\}$ be a proper convex function.[21] Let $\bar{y}\in\mathbf{R}^n$. Define the <u>second-order epi-derivative</u> of f at $\bar{x}$ relative to $\bar{y}$ as follows:

$$D^2 f_{\bar{x}|\bar{y}}(x')=\lim_{\tau\downarrow 0}\frac{f(\bar{x}+\tau x')-f(\bar{x})-\tau(\bar{y}\cdot x')}{\frac{1}{2}\tau^2}.$$

**5.6 Theorem:** Let $f:\mathbf{R}^m x\mathbf{R}^n\to\mathbf{R}\cup\{\infty\}$ be a closed proper convex function and let $m:\mathbf{R}^m\rightrightarrows\mathbf{R}^n$ be defined by $m(t)=\arg\min_{x\in\mathbf{R}^n} f(t,x)$. Let $\bar{t}$ be given. Suppose $\bar{x}\in m(\bar{t})$ and $\bar{y}\in\partial_t f(\bar{t},\bar{x})$, and suppose f is twice epi-differentiable at $(\bar{t},\bar{x})$ relative to $(\bar{y},0)$.[22] Then

$$Dm_{\bar{t}|\bar{x}}(t')=\{x'|(y',0)\in\partial D^2 f_{\bar{t},\bar{x}|\bar{y},0}(t',x'),\text{ some } y'\}.$$

Applying Theorem 5.1 and Theorem 5.6, we have:

**5.7 Theorem:** Let $f:\mathbf{R}^m x\mathbf{R}^n\to\mathbf{R}\cup\{\infty\}$ be a closed proper convex function and let $m:\mathbf{R}^m\rightrightarrows\mathbf{R}^n$ be defined by $m(t)=\arg\min_{x\in\mathbf{R}^n} f(t,x)$. Let $\bar{t}$ be given. Suppose $\bar{x}\in m(\bar{t})$, and suppose f is differentiable with respect to t at $(\bar{t},\bar{x})$. Then

$$Dm_{\bar{t}|\bar{x}}(t')=\{x'|(y',0)\in\partial D^2 f_{\bar{t},\bar{x}|\nabla_t f(\bar{t},\bar{x}),0}(t',x'),\text{ some } y'\}.$$

---

[21]A function $f:S\to\mathbf{R}\cup\{\infty\}$, where $S\subseteq\mathbf{R}^n$, is <u>proper</u> if $f(x)<+\infty$ for at least one x.

[22]Section 3 of Rockafellar (1990A) gives examples of classes of functions that are epi-differentiable and that are twice epi-differentiable.

In order to get a better grasp of the concepts introduced in this section, I will now present two examples. These examples involve the same functions as examples 3.1 and 3.2; however, here I will show the actual calculation of the proto-derivatives.

**Example 5.1.**

Consider the function $f:\mathbf{R}x\mathbf{R}\to\mathbf{R}$ defined by $f(t,x)=|x-t|$. Since f is convex, by Theorem 5.4 we can write $m(t)=\{x\in\mathbf{R}\,|\,(y,0)\in\partial f(t,x),\text{ some } y\}$. Writing out the subdifferential of f, we have

$$\partial f(t,x) = \begin{cases} (1,-1), & t>x \\ [(1,-1),(-1,1)], & t=x \\ (-1,1), & t<x. \end{cases}$$

Clearly, $m(t)=\{t\}$ since the only value of x such that $(y,0)\in\partial f(t,x)$ is $x=t$. Also, note that if $(y,0)\in\partial f(t,t)$, then $y=0$. To calculate $Dm_{\bar{t}|\bar{x}}(t')$, we can use the formula of Theorem 5.6. Note that we need only consider cases with $\bar{t}=\bar{x}$ and $\bar{y}=0$. By definition, $D^2f_{\bar{t},\bar{t}|0,0}(t',x')$ equals

$$\lim_{\tau\downarrow 0}\frac{\tau\,|x'-t'|}{\frac{1}{2}\tau^2}.$$

This limit is zero if $t'=x'$ and $+\infty$ otherwise. So we have

$$D^2f_{\bar{t},\bar{t}|0,0}(t',x')=\begin{cases} 0, & t'=x' \\ \infty, & \text{otherwise} \end{cases}$$

and thus

$$\partial D^2f_{\bar{t},\bar{t}|0,0}(t',x')=\begin{cases} (0,0), & t'=x' \\ \text{DNE}, & \text{otherwise.} \end{cases}$$

Thus by Theorem 5.6, $Dm_{\bar{t}|\bar{t}}(t')=\{t'\}$ for all $\bar{t}\in\mathbf{R}$, and $Dm_{\bar{t}|\bar{x}}$ does not exist for $\bar{x}\neq\bar{t}$. As stated before, m is monotone.

**Example 5.2.**

Define the function f as follows:

$$f(t,x) = \begin{cases} t-x, & x<t \\ 0, & t\le x\le t+1 \\ x-(t+1), & t+1<x \end{cases}$$

This function attains its minimum value when $x\in[t,t+1]$. As in the first example, we first write out the subderivative of f.

$$\partial f(t,x) = \begin{cases} (1,-1), & x<t \\ [(1,-1),(0,0)], & x=t \\ (0,0), & t<x<t+1 \\ [(0,0),(-1,1)], & x=t+1 \\ (-1,1), & t+1<x. \end{cases}$$

By Theorem 5.4, $m(t)=\{x\,|\,(y,0)\in\partial f(t,x), \text{ some } y\} = [t,t+1]$. Note that if $(y,0)\in\partial f(t,x)$, then $y=0$. We need only consider $\bar{x}\in[\bar{t},\bar{t}+1]$ and $\bar{y}=0$ in the formula of Theorem 5.6. By definition, $D^2 f_{\bar{t},\bar{x}\,|\,0,0}(t',x')$ =

$$\lim_{\tau\downarrow 0}\frac{f(\bar{t}+\tau t',\bar{x}+\tau x')-f(\bar{t},\bar{x})}{\frac{1}{2}\tau^2}.$$

Note that if $\bar{x}=\bar{t}$, then this limit is zero if $x'\ge t'$ and $\infty$ otherwise. If $\bar{x}=\bar{t}+1$, then the limit is zero if $x'\le t'$ and $\infty$ otherwise. If $\bar{t}<\bar{x}<\bar{t}+1$, then the limit is zero for all values of x'. Using these observations and Theorem 5.6, we can write:

$$Dm_{\bar{t}\,|\,\bar{x}}(t') = \begin{cases} [t',\infty), & \bar{x}=\bar{t} \\ (-\infty,\infty), & \bar{t}<\bar{x}<\bar{t}+1 \\ (-\infty,t'], & \bar{x}=\bar{t}+1. \end{cases}$$

Thus, m is monotone.

**6. Nonconvex Payoff Functions.**

In this section I will reformulate the definition of a subgradient from Section 5 to hold for any lower semicontinuous, extended-real-valued function. As a special case, I will discuss functions that are twice continuously differentiable. In this section, I will introduce one more type of cone, the Clarke tangent cone. The new subgradient definition for general functions will be based on this cone and its polar, the Clarke normal cone. The generalized subgradient of a function f reduces to the subgradient defined in Section 5 when f happens to be convex.

**Definition:** Let C be a closed subset of $\mathbf{R}^n$ and let $x \in C$. The Clarke tangent cone to C at x is the closed convex cone,

$$T_C(x) = \liminf_{\substack{x' \to x \\ \tau \downarrow 0}} \tau^{-1}(C - x').$$

**Definition:** Let C be a closed subset of $\mathbf{R}^n$ and let $x \in C$. C is Clarke regular at x if $K_C(x)$, the contingent cone to C at x, is equal to $\lim \inf_{x'(\in C) \to x} K_C(x')$.

If C is Clarke regular at x, then the Clarke tangent cone to C at x is equal to the contingent cone of C at x. Clarke regularity holds at every point $x \in C$ when C is convex.

**Definition:** Let C be a closed subset of $\mathbf{R}^n$ and let $x \in C$. The Clarke normal cone to C at x is the closed convex cone $N_C(x)$ defined as the polar of the Clarke tangent cone to C at x.

If C is Clarke regular at x, then $N_C(x)$ is the polar of the contingent cone $K_C(x)$, and vice versa. If C is convex, Clarke regularity always holds and the contingent cone coincides with the Clarke tangent cone.

Now we get to our generalized definition of a subgradient.

**Definition:** Let $f:\mathbf{R}^n \to \mathbf{R}\cup\{\pm\infty\}$ be lower semicontinuous and let $\bar{x}\in\text{dom } f$. The set of generalized subgradients of f at $\bar{x}$ is defined by

$$\partial f(\bar{x}) = \{z\in\mathbf{R}^n \,|\, (z,-1)\in N_{\text{epi } f}(\bar{x}, f(\bar{x}))\}.$$

Formulas for calculating subgradients are provided in Clarke (1983).

Now I will define one more set, $\partial^\infty f(t,x)$. This set will allow us to define a condition analogous to a constraint qualification in regular optimization problems. The condition will eliminate situations where the value of min f(t,x) drops off at an infinite rate as t is perturbed slightly. If the basic constraint qualification holds for f at (t,x) and x is a local minimum of $f(t,\cdot)$, then $(y,0)\in\partial f(t,x)$ for some y.

**Definition:** Define $\partial^\infty f(t,x) = \{(t',x') \,|\, (t',x',0)\in N_{\text{epi } f}((t,x),f(t,x))\}$. The basic constraint qualification is said to hold for f at (t,x) if the only y with $(y,0)\in\partial^\infty f(t,x)$ is y=0.

Theorem 5.4 showed that when f is convex, we can write the multifunction $m(t)\equiv\arg\min_{x\in\mathbf{R}^n} f(t,x)$ in terms of f. Define the multifunction $\mu:\mathbf{R}^m \rightrightarrows \mathbf{R}$ by $\mu(t)=\{x\in\mathbf{R}^n \,|\, (y,0)\in\partial f(t,x)$, some y$\}$. If f is convex, then by Theorem 5.4, $m(t)=\mu(t)$. When dealing with nonconvex functions, this is no longer necessarily true. However, the inclusion $m(t)\subseteq\mu(t)$ holds as long as the basic constraint qualification holds. In fact, when the basic constraint qualification and an additional regularity condition hold, can we replace $m_n$ by $\mu_n$ in conditions (B3), (B4), (C3), and (C4) without altering the existence results of Theorems 4.2 and 4.5.

**6.1 Theorem:** Let $f:\mathbf{R}^m x \mathbf{R}^n \to \mathbf{R}$ be Clarke regular. Define the multifunction $\mu:\mathbf{R}^m \rightrightarrows \mathbf{R}$ by $\mu(t)=\{x\in\mathbf{R}^n|(y,0)\in\partial f(t,x), \text{ some } y\}$ and define the multifunction $m:\mathbf{R}^m \rightrightarrows \mathbf{R}$ by $m(t)=\arg\min_{x\in\mathbf{R}^n} f(t,x)$. Suppose the basic constraint qualification holds and (gph $\mu$)\(gph m) is closed in gph $\mu$. Then $\mu$ monotone implies m is monotone.[23]

Now I will state a theorem for nonconvex functions of class $C^2$. The theorem states that the proto-derivative of $\mu$ can be expressed in terms of the second derivatives of f.[24] The proof follows from Theorem 5.4 of Rockafellar (1989B).

**6.2 Theorem:** Let $f:\mathbf{R}^m x \mathbf{R}^n \to \mathbf{R}$ be a twice continuously differentiable function and let $\mu:\mathbf{R}^m \rightrightarrows \mathbf{R}$ be defined by $\mu(t)=\{x\in\mathbf{R}^n|(y,0)\in\partial f(t,x), \text{ some } y\}$. For $\bar{x}\in\mu(\bar{t})$, the proto-derivative of $\mu$ at $\bar{t}$ relative to $\bar{x}$ in direction t' is given by:

$$D\mu_{\bar{t}|\bar{x}}(t')=\{x'\,|\,\mathbf{0}=t'\cdot\nabla_{tx}^2 f(\bar{t},\bar{x})+x'\cdot\nabla_{xx}^2 f(\bar{t},\bar{x})\}.$$

**Example 6.1.**

Consider the function $f(t,x)=(x-t)^4-2(x-t)^2$. $f(t,\cdot)$ achieves its minimum value at $x=t-1$ and $x=t+1$, so $m(t)=\{t-1,t+1\}$. There is also a local maximum at $x=t$, and $\mu(t)=\{t-1,t,t+1\}$. f is not convex, but since it is twice continuously differentiable, we can apply Theorem 6.2 to get, $\forall\bar{x}\in\mu(\bar{t})$,

---

[23]In fact, the proof of Theorem 6.1 shows that $\forall\bar{x}\in m(\bar{t})$, $D\mu_{\bar{t}|\bar{x}}=Dm_{\bar{t}|\bar{x}}$.

[24]Notation: for $f:\mathbf{R}^m x \mathbf{R}^n \to \mathbf{R}$, $\nabla_{tx}^2 f(\bar{t},\bar{x})$ is the m×n matrix whose $(i,j)^{th}$ element is $f_{t_i x_j}(\bar{t},\bar{x})$.

$$D\mu_{\bar{t}|\bar{x}}(t')=\{x'|0=t'\cdot(-12(\bar{x}-\bar{t})^2+4)+x'\cdot(12(\bar{x}-\bar{t})^2-4)\}.$$

Note that for $\bar{x}\in\mu(\bar{t})$, $12(\bar{x}-\bar{t})^2-4\neq 0$, so $D\mu_{\bar{t}|\bar{x}}(t')=\{t'\}$. Thus $\mu$ has rate of increase at $\bar{t}$ of 1 and is monotone. By Theorem 6.1, m(t) is also monotone. Furthermore, $\forall\bar{x}\in m(\bar{t})$, $Dm_{\bar{t}|\bar{x}}(t')=D\mu_{\bar{t}|\bar{x}}(t')=\{t'\}$.

The following theorem gives a formula that will allow us to calculate the proto-derivative of the set $\mu(t)$. The proof is similar to that of Theorem 5.5.

**6.3 Theorem:** Let $f:\mathbf{R}^m \mathrm{x} \mathbf{R}^n \rightarrow \mathbf{R}\cup\{\pm\infty\}$ be Clarke regular. Suppose $\bar{y}\in\{y|(y,0)\in\partial f(t,x)\}$. The proto-derivative of the mapping $\mu(t)=\{x|(y,0)\in\partial f(t,x)$, some $y\}$ at $\bar{t}$ relative to $\bar{x}\in\mu(\bar{t})$ in direction t' is given by:

$$D\mu_{\bar{t}|\bar{x}}(t')=\{x'|(y',0)\in D(\partial f)_{\bar{t},\bar{x}|\bar{y},0}(t',x'),\ \text{some } y'\}.$$

Using Theorem 6.3, we can calculate $D\mu$ for any Clarke regular function and determine whether the multifunction $\mu$ is monotone. Under the conditions of Theorem 6.1, we can also determine whether the multifunction m is monotone. Thus, Theorems 6.1 and 6.3 allows us to write the conditions (C1)-(C4) in terms of the functions $f_n$ directly.

## Appendix

**Tarski's Fixed Point Theorem:** Let $(X,\geq)$ be a complete lattice and let $f:X\to X$ be monotone (say, nondecreasing). Then there exists $x^*\in X$ such that $f(x^*) = x^*$. Furthermore, the set of such fixed points is itself a complete lattice.

**Proof of Tarski's Fixed Point Theorem**

Let $T=\{z \mid f(x) \geq z\}$. Since $(X,\geq)$ is a complete lattice, X itself has a greatest lower bound $\underline{x}$ which is therefore also the minimum element of X. Hence, $f(\underline{x})\geq\underline{x}$, so T is non-empty and has a least upper bound $x^*$. Since f is monotone nondecreasing, $f(x^*)\geq f(x)\geq x$ for all $x\in T$. Hence, $f(x^*)$ is an upper bound for T, so $f(x^*)\geq x^*$. Then, since f is monotone nondecreasing, $f(f(x^*))\geq f(x^*)$. Hence, $f(x^*)\in T$. It then follows that $x^*\geq f(x^*)$. So, by the antisymmetry of the order, $x^*=f(x^*)$. It remains to be shown that the set of fixed points is a complete lattice. Suppose x and y are fixed points of f. Let $R=\{z \mid f(z)\geq z \text{ and } x\leq x\wedge y\}$ and $S=\{z \mid f(z)\leq z \text{ and } z\geq x\vee y\}$. Note that since $x\wedge y\leq x$, $f(x\wedge y)\leq f(x)=x$, and since $x\wedge y\leq y$, $f(x\wedge y)\leq f(y)=y$. Thus $f(x\wedge y)\leq x\wedge y$. As before we can define the largest fixed point of f, $x^*\in S$, and the smallest fixed point of f, $x_*\in R$. So R and S are nonempty. $x\wedge y$ is an upper bound of R, and so $f(x\wedge y)$ is also an upper bound of R. Let $z^*=\text{lub } R$. For all $z\in R$, $f(z^*)\geq f(z)\geq z$, so $f(z^*)\geq z^*$. $x\wedge y\geq f(x\wedge y)\geq z^*$ and hence $x\wedge y\geq f(z^*)$, and $f(f(z^*))\geq f(z^*)$, so $f(z^*)\in R$. Thus $z^*\geq f(z^*)$, so $z^*=f(z^*)$. So $z^*$ is the greatest fixed point less than $x\wedge y$. Hence for any two fixed points, there exists a least upper bound and by analogous argument a greatest lower bound. ‖

**Proof of Theorem 2.1** (proof of version in footnote 4)

Let $t,t'\in T$ be such that $t'\geq t$ and let $S,S'\subseteq X$ be such that $S'\geq_S S$. We must show that $m(S',t')\geq_S m(S,t)$. If $m(S,t)$ or $m(S',t')$ is empty, this holds trivially, so suppose $x\in m(S,t)$ and $x'\in m(S',t')$. By definition of the strong set order, $x\wedge x'\in S$ and $x\vee x'\in S'$. We must show that

$x \wedge x' \in m(S,t)$ and $x \vee x' \in m(S',t')$. Since f is submodular in x, $f(x,t)+f(x',t) \geq f(x \wedge x',t)+f(x \vee x',t)$. Rearranging this inequality, we have $f(x,t)-f(x \wedge x',t) \geq f(x \vee x',t)-f(x',t)$. Since $x \vee x' \geq x'$, the decreasing differences condition says that $f(x \vee x',t)-f(x',t) \geq f(x \vee x',t')-f(x',t')$. Combining the last two inequalities we have,

(1) $$f(x,t)-f(x \wedge x',t) \geq f(x \vee x',t')-f(x',t').$$

Since x minimizes $f(\cdot,t)$ on S and $x \wedge x' \in S$, $0 \geq f(x,t)-f(x \wedge x',t)$; and since x' minimizes $f(\cdot,t')$ on S' and $x \vee x' \in S'$, $f(x \vee x',t')-f(x',t') \geq 0$. These inequalities together with (1) give us the result that $0 \geq f(x,t)-f(x \wedge x',t) \geq f(x \vee x',t')-f(x',t') \geq 0$. Hence $f(x,t)=f(x \wedge x',t)$, implying $x \wedge x' \in m(S,t)$. The proof that $x \vee x' \in m(S',t')$ is similar. ‖

**Proof of Theorem 3.2**

Let TxX be a subset of $\mathbf{R}^m x \mathbf{R}^n$. Let $\bar{t} \in \mathbf{R}^m$ and $t' \in \mathbf{R}^m$ with $t'>0$. Suppose $Dm_{\bar{t}|\bar{x}}(t')$ exists and is nonempty. By the definition, this implies that $\bar{x} \in m(\bar{t})$. Let $x' \in Dm_{\bar{t}|\bar{x}}(t')$ be such that $x' \ngeq 0$.[25] Then there exist sequences $(\tau^k)$, $(t^k)$, and $(x^k)$ such that $\tau^k \downarrow 0$, $\forall k$ $x^k \in (\tau^k)^{-1}[m(\bar{t}+\tau^k t^k)-\bar{x}]$, and $(t',x')$ is a cluster point of the sequence $(t^k,x^k)$. Let $\delta_1=\min\{t_1',...,t_m'\}$. Since $t'>0$, $\delta_1>0$. Let $\delta_2=\min_{i \text{ s.t. } x_i' \neq 0} |x_i'|$. Since $x' \neq 0$, $\delta_2$ is well defined and positive. Let $\epsilon$ be such that $\min\{\delta_1,\delta_2\}>\epsilon>0$.

Since $(t',x')$ is a cluster point of $(t^k,x^k)$, there exist subsequences of $(\tau^k)$, $(t^k)$, and $(x^k)$, denoted by $(\tau^j)$, $(t^j)$, and $(x^j)$, such that for all j, $\max_{i \in \{1,...,m\}} |t_i^j-t_i'|<\epsilon$ and $\max_{i \in \{1,...,n\}} |x_i^j-x_i'|<\epsilon$. Note that this means $t^j>0$ for all j.

Since $t^j>0$ and $\tau^j>0$, $m(\bar{t}+\tau^j t^j) \geq_S m(\bar{t})$ and thus $(\tau^j)^{-1}[m(\bar{t}+\tau^j t^j)-\bar{x}] \geq_S (\tau^j)^{-1}[m(\bar{t})-\bar{x}]$. Note that $0 \in (\tau^j)^{-1}[m(\bar{t})-\bar{x}]$. By definition of the strong set order, $x^j \vee 0 \in (\tau^j)^{-1}[m(\bar{t}+\tau^j t^j)-\bar{x}]$.

Now we show that $\max_{i \in \{1,...,n\}} |(x_i^j \vee 0)-(x_i' \vee 0)|<\epsilon$, meaning that $x' \vee 0$ is a cluster point of the

[25]If no such x' exists, then we are done.

sequence ($x^k \vee 0$), by considering the following three cases.

Let $p \in \{1,...,n\}$. $x_p' \vee 0 = \max\{x_p', 0\}$.

(a) $x_p' < 0$. Then $x_p' \vee 0 = 0$ and $\epsilon < -x_p'$. We know $|x_p^j - x_p'| < \epsilon$ for all j. Suppose $x_p^j \geq 0$. Then $x_p^j > \epsilon + x_p'$ and we deduce that $x_p^j - x_p' > \epsilon$, a contradiction. So it must be that $x_p^j < 0$. But then $x_p^j \vee 0 = 0$ and $|(x_p^j \vee 0) - (x_p' \vee 0)| = |0-0| = 0$.

(b) $x_p' > 0$. Then $x_p' \vee 0 = x_p'$ and $\epsilon < x_p'$. We know $|x_p^j - x_p'| < \epsilon$ for all j. Suppose $x_p^j \leq 0$. Then $x_p^j < x_p' - \epsilon$ and we deduce that $x_p^j - x_p' < -\epsilon$, a contradiction. So it must be that $x_p^j > 0$. But then $x_p^j \vee 0 = x_p^j$ and $|(x_p^j \vee 0) - (x_p' \vee 0)| = |x_p^j - x_p'| < \epsilon$.

(c) $x_p' = 0$. Then $x_p' \vee 0 = x_p' = 0$. We know $|x_p^j - x_p'| < \epsilon$ for all j. If $x_p^j \leq 0$, then $|(x_p^j \vee 0) - (x_p' \vee 0)| = |0-0| = 0$. Otherwise, $x_p^j > 0$, and then $|(x_p^j \vee 0) - (x_p' \vee 0)| = |x_p^j - x_p'| < \epsilon$.

We also know that for all j, $x^j \vee 0 \in (\tau^j)^{-1}[m(\bar{t} + \tau^j t^j) - \bar{x}]$. Since $\tau^j \downarrow 0$, and $\forall j$, $x^j \vee 0 \in (\tau^j)^{-1}[m(\bar{t} + \tau^j t^j) - \bar{x}]$, and $(t', x' \vee 0)$ is a cluster point of $(t^j, x^j \vee 0)$, then $(t', x' \vee 0) \in \text{gph} \lim \sup_{\tau \downarrow 0} \tau^{-1}[m(\bar{t} + \tau t') - \bar{x}]$, or in other words, $x' \vee 0 \in Dm_{\bar{t}|\bar{x}}(t')$. Since $x' \vee 0 \geq 0$, we are done. ‖

**Proof of Theorem 4.2**

**A.1 Lemma (Rockafellar (1989A)):** Suppose $\mu: \mathbf{R}^m \rightrightarrows \mathbf{R}^n$ is proto-differentiable. Then $x' \in D\mu_{\bar{t}|\bar{x}}(t')$ if and only if for all s in some interval $(0, \delta)$ there exist arcs x(s) and t(s), $x(s) \in \mu(t(s))$ with $t(0) = \bar{t}$ and $x(0) = \bar{x}$, such that $t_+'(0) = t'$ and $x_+'(0) = x'$ (right derivatives).[26]

**A.2 Lemma (See Theorem 2 of Milgrom and Roberts (1990)):** Let $f: \mathbf{R} \to \mathbf{R}$ be lower semi-continuous and let S be a compact interval of **R**. Then the set $\arg\min_{x \in S} f(x)$ is a complete lattice.

By Lemma A.2 and (B1)-(B2), for all $t \in S_{-n}$, the set $m_n(t)$ is a nonempty complete lattice.

[26] The proof is a specialization of the property of proto-differentiability given in Proposition 2.3 of Rockafellar (1989B).

By Lemma A.2, $m_n(t)$ has a maximal element, and the function $M_n$:$S_{-n}$→$S_n$ defined by $M_n(t)=\max_{x\in S_n} m_n(t)$ is well defined.

**Claim:** Suppose (B1)-(B2) hold and suppose that for every $t<\max S_{-n}$, there exists $x\in D(m_n)_{t|M_n(t)}(1)$ such that $x>0$. Then $t''\geq t'$ implies $M_n(t'')\geq M_n(t')$.

**Proof of Claim:** Suppose $t',t''\in S_{-n}$ with $t''\geq t'$ and $M_n(t'')<M_n(t')$. Since is $M_n$ is finite and upper semicontinuous on [t',t''], it attains its maximum on the nonempty compact set [t',t''].[27] Let $y\equiv\arg\max_{t\in[t',t'']}M_n(t)$. Note that $y<t''\leq\max S_{-n}$. By assumption, there exists $x\in D(m_n)_{y|M_n(y)}(1)$ such that $x>0$. By Lemma A.1, for all s in some interval $(0,\delta)$ there exist arcs x(s) and t(s), $x(s)\in m_n(t(s))$ with $t(0)=y$ and $x(0)=M_n(y)$, such that $t_+'(0)=1$ and $x_+'(0)=x>0$. Since $t(0)=y<t''$ and $t_+'(0)=1$, for all s in some interval $(0,\delta')$, $t(s)\in[y,t'']$. Thus for all s in $(0,\delta')$, $M_n(t(s))\leq M_n(y)$. Since $x(s)\in m_n(t(s))$, then $x(s)\leq M_n(t(s))$. So for all s in $(0,\delta')$, $x(s)\leq M_n(y)$, implying that $x_+'(0)\leq 0$, a contradiction.

Assume (B1)-(B4) hold. Let $\epsilon>0$ be given. Define $m_n^*$ by $m_n^*(t)=m_n(t)+\epsilon\cdot t$. Then $M_n^*(t)=\max_{x\in S_n} m_n^*(t)=M_n(t)+\epsilon\cdot t$. Let $t<\max S_{-n}$. By (B4) there exists $x\geq 0$ such that $x\in D(m_n)_{t|M_n(t)}(1)$. By Lemma A.1, for all s in some interval $(0,\delta)$ there exist arcs x(s) and t(s),

---

[27]By definition, $M_n$ is upper semicontinuous at $t^*\in[t',t'']$ if, given any $\epsilon>0$, there exists a neighborhood of $t^*$ in which $M_n(t)<M_n(t^*)+\epsilon$. Suppose $M_n$ is not upper semicontinuous. Then there exists $t^*\in[t',t'']$ and $\epsilon>0$ such that for all neighborhoods N of $t^*$ there exists $t\in N$ with $M_n(t)\geq M_n(t^*)+\epsilon$. Choose $t^k$ such that $t^k\in B(t^*,1/k)$ (open ball centered at $t^*$ with radius 1/k) and $M_n(t^k)\geq M_n(t^*)+\epsilon$. Then $t^k\to t^*$. $(M_n(t^k))$ is a sequence contained in the compact interval $S_n$, so it has a convergent subsequence $(M_n(t^p))$, $M_n(t^p)\to y$. Note that $(t^p,M_n(t^p))$ is a convergent sequence contained in gph $m_n$. Since f is continuous and S is compact, $m_n$ is an upper semicontinuous correspondence and thus has closed graph. This implies that $(t^*,y)\in$gph $m_n$, i.e. $y\in m_n(t^*)$. Thus $y\leq M_n(t^*)$. There exists P such that $\forall p'>P$, $M_n(t^{p'})\in B(y,\epsilon)$, implying that $M_n(t^{p'})<y+\epsilon\leq M_n(t^*)+\epsilon$, a contradiction.

$x(s)\in m_n(t(s))$ with $t(0)=t$ and $x(0)=M_n(t)$, such that $t_+'(0)=1$ and $x_+'(0)=x\geq 0$. For all s in $(0,\delta)$ define $x^*(s)=x(s)+\epsilon\cdot t(s)$. Then $x^*(0)=x(0)+\epsilon\cdot t(0)=M_n(t)+\epsilon\cdot t=M_n^*(t)$ and for all s in $(0,\delta)$, $x^*(s)=x(s)+\epsilon\cdot t(s)\in m_n(t(s))+\epsilon\cdot t(s)\in m_n^*(t(s))$ and $x_+^{*\prime}(0)=x_+'(0)+\epsilon\cdot t_+'(0)=x_+'(0)+\epsilon\geq\epsilon$. Therefore $x_+^{*\prime}(0)\in D(m_n^*)_{t|M_n^*(t)}(1)$. So for all $t<\max S_{-n}$, $\exists x\geq\epsilon>0$ such that $x\in D(m_n^*)_{t|M_n^*(t)}(1)$. By the Claim, $t''\geq t'$ implies $M_n^*(t'')\geq M_n^*(t')$. But this says $M_n(t'')+\epsilon\cdot t''\geq M_n(t')+\epsilon\cdot t'$. Letting $\epsilon$ approach zero, we have $M_n(t'')\geq M_n(t')$. Therefore, $M_n$ is a nondecreasing function. Apply Tarski's Fixed Point Theorem to obtain a fixed point of the game $\Gamma$. ∎

**Proof of Theorem 4.5**

**A.3 Lemma:** Let $f:\mathbf{R}^m\to\mathbf{R}$. If f is Lipschitz, then for all $t\in\mathbf{R}^m$ and all $t'\in\mathbf{R}^m$ such that $t'>0$ there exists $\zeta$ such that $(t',\zeta)\in K_{\text{gph } f}(t,f(t))$.

**Proof of Lemma A.3:** Let $t,t'\in\mathbf{R}^m$ with $t'>0$. Since f is Lipschitz, there exists $\epsilon>0$ and $K>0$ such that whenever $x',x''\in B(x,\epsilon)$, $|f(x'')-f(x')|\leq K|x''-x'|$. Let $(\tau^k)$ be a sequence with $\tau^k\downarrow 0$ and $\tau^k<\epsilon/\max\{t_1',...,t_m'\}$. Note that since $t'>0$, $\max\{t_1',...,t_m'\}>0$. Also note that $t+t'\cdot\tau^k\in B(t,\epsilon)$. Define $y^k\equiv(\tau^k)^{-1}[(t+t'\cdot\tau^k,f(t+t'\cdot\tau^k))-(t,f(t))]$. Then $(y_1^k,...,y_m^k)=(\tau^k)^{-1}(t'\cdot\tau^k)=t'$ and $y_{m+1}^k=(\tau^k)^{-1}[f(t+t'\cdot\tau^k)-f(t)]$. Hence $|y_{m+1}^k|=(\tau^k)^{-1}|f(t+t'\cdot\tau^k)-f(t)|$. Using the Lipschitz condition, $|y_{m+1}^k|\leq(\tau^k)^{-1}K|t+t'\cdot\tau^k-t|=(\tau^k)^{-1}K(t'\cdot\tau^k)=K\cdot t'$. This implies that $(y_{m+1}^k)$ has a convergent subsequence with limit $\zeta$. $(t',\zeta)$ is a cluster point of $(y^k)$, so $(t',\zeta)\in K_{\text{gph } f}(t,f(t))$.

By Lemma A.2, for all $t\in S_{-n}$, the set $m_n(t)$ is a nonempty complete lattice. Hence $m_n(t)$ has a maximal element, and the function $M_n:S_{-n}\to S_n$ of (C4) is well defined.

**Claim:** Suppose (C1), (C2), and (C4) hold. Suppose that for all $t<\max S_{-n}$ and all $t'>0$, there exists $x>0$ such that $x\in D(m_n)_{t|M_n(t)}(t')$. Then $t''\geq t'$ implies $M_n(t'')\geq M_n(t')$.

**Proof of Claim:** Suppose $t''\geq t'$ and $M_n(t'')\ngeq M_n(t')$. Since $M_n$ is Lipschitz, it is continuous and thus attains its maximum on the nonempty compact set $[t',t'']$. So let $y=\arg\max_{t\in[t',t'']}M_n(t)$. Note that $y<t''\leq \max S_{-n}$.

*case (i):* $y<<t''$. Define $\Delta\equiv t''-y$. By supposition, there exists $x>0$ such that $x\in \mathbf{D}(\mathbf{m}_n)_{y|M_n(y)}(\Delta)$. By Lemma A.1, for all s in some interval $(0,\delta)$ there exist arcs $x(s)$ and $t(s)$, $x(s)\in m_n(t(s))$ with $t(0)=y$ and $x(0)=M_n(y)$, such that $t_+'(0)=\Delta$ and $x_+'(0)=x>0$. Then for all s in some interval $(0,\delta')$, $t(s)\in[t',t'']$. Thus for all s in $(0,\delta')$, $M_n(t(s))\leq M_n(y)$. Since $x(s)\in m_n(t(s))$, $x(s)\leq M_n(t(s))$. So for all s in $(0,\delta')$, $x(s)\leq M_n(y)$, implying that $x_+'(0)\leq 0$, a contradiction.

*case (ii):* there exists j such that $y_j=t''_j$. Define $\Delta\equiv t''-y$. By supposition there exists $x>0$ such that $x\in \mathbf{D}(\mathbf{m}_n)_{y|M_n(y)}(\Delta)$. As before for all s in some interval $(0,\delta)$ there exist arcs $x(s)$ and $t(s)$, $x(s)\in m_n(t(s))$ with $t(0)=y$ and $x(0)=M_n(y)$, such that $t_+'(0)=t''-y$ and $x_+'(0)=x>0$. Since $x_+'(0)>0$, there exists a sequence $(s^k)$ such that for all k, $x(s^k)>x(0)$. Thus for all k, $M_n(t(s^k))>M_n(y)$. If $t(s^k)$ were contained in $[t',t'']$, then this would be impossible. So for all k, $t(s^k)\notin[t',t'']$. Define $\tau^k=(y_j,t_{-j}(s^k))$. Since $M_n$ is Lipschitz, there exists $K>0$ and $\epsilon>0$ such that for all r',r'' in an open ball with center r and radius $\epsilon$, $|M_n(r')-M_n(r'')|\leq K|r''-r'|$.[28] Pick $s^k$ such that $|t(s^k)-y|<\epsilon$. Then by construction it is also true that $|\tau^k-y|<\epsilon$. By the Lipschitz condition $|M_n(t(s^k))-M_n(\tau^k)|\leq K|t(s^k)-\tau^k|$. We know that $M_n(t(s^k))>M_n(y)\geq M_n(\tau^k)$. We also know that $t_{-j}(s^k)=\tau^k_{-j}$ and $t_j(s^k)>t''_j=y_j=\tau^k_j$. So we can write the Lipschitz condition as $M_n(t(s^k))-M_n(\tau^k)\leq K(t_j(s^k)-t^k_j)$. Now note that

[28] Here $|r|=\max_{i\in\{1,\dots,N-1\}}|r_i|$.

$$x_+'(0) = \lim_{\lambda\downarrow 0}\frac{x(\lambda)-x(0)}{\lambda} = \lim_{\lambda\downarrow 0}\frac{M_n(t(\lambda))-M_n(y)}{\lambda}.$$

But using the Lipschitz condition, we have

$$\lim_{\lambda\downarrow 0}\frac{M_n(t(\lambda))-M_n(y)}{\lambda} \le \lim_{\lambda\downarrow 0}\frac{M_n(\tau^\lambda)-M_n(y)+K(t_j(\lambda)-\tau^\lambda_j)}{\lambda}.$$

The last expression can be rewritten as

$$\lim_{\lambda\downarrow 0}\frac{M_n(\tau^\lambda)-M_n(y)}{\lambda} + \lim_{\lambda\downarrow 0} K\left[\frac{t_j(\lambda)-y_j}{\lambda}+\frac{y_j-\tau^\lambda_j}{\lambda}\right].$$

But $\lim_{\lambda\downarrow 0}\frac{t_j(\lambda)-y_j}{\lambda} = t'_{+j}(0)=0$ and $y_j=\tau^\lambda_j$ so $\lim_{\lambda\downarrow 0}\frac{y_j-\tau^\lambda_j}{\lambda}=0.$ We can conclude that

$$x_+'(0) \le \lim_{\lambda\downarrow 0}\frac{M_n(\tau^\lambda)-M_n(y)}{\lambda} \le 0,$$ which is a contradiction.

Assume (C1)-(C4) hold. Let $\epsilon>0$ be given. Define $\xi\in\mathbf{R}^{N-1}$ by $\xi\equiv(\epsilon,\ldots,\epsilon)$. Define $m_n^*$ by $m_n^*(t)=m_n(t)+\xi\cdot t$. Then $M_n^*(t)=\max_{x\in S_n} m_n^*(t)=M_n(t)+\xi\cdot t$. Let $t'>0$. Define j such that $t_j'=\min_{i\in\{1,\ldots,N-1\}}\{t_i'\,|\,t_i'\neq 0\}$ and define $\Delta\in\mathbf{R}^{N-1}$ by $\Delta\equiv t'/t_j'$. Then $\Delta>0$, and for all i either $\Delta_i=0$ or $\Delta_i\ge 1$. By (C4) and Lemma A.3 for all $t^*>0$, there exists $\zeta$ such that $(t^*,\zeta)\in K_{\mathrm{gph}\,M_n}(t,M_n(t))$. Since $K_{\mathrm{gph}\,M_n}(t,M_n(t))\subseteq K_{\mathrm{gph}\,m_n}(t,M_n(t))$, this implies that $K_{\mathrm{gph}\,m_n}(t,M_n(t))$ is nonempty, and thus using the proto-differentiability of $m_n$, for all $t^*>0$, $\mathbf{D}(\mathbf{m_n})_{t|M_n(t)}(\mathbf{t}^*)\neq\emptyset$. Using (C3), there exists $x\ge 0$ such that $x\in\mathbf{D}(\mathbf{m_n})_{t|M_n(t)}(\Delta)$. By Lemma A.1, for all s in some interval $(0,\delta)$ there exist arcs x(s) and t(s), $x(s)\in m_n(t(s))$ with $t(0)=t$ and $x(0)=M_n(t)$, such that $t_+'(0)=\Delta$ and $x_+'(0)=x\ge 0$. For all s in $(0,\delta)$

define $x^*(s)=x(s)+\xi\cdot t(s)$. Then $x^*(0)=x(0)+\xi\cdot t(0)=M_n(t)+\xi\cdot t=M_n^*(t)$ and for all s in $(0,\delta)$, $x^*(s)=x(s)+\xi\cdot t(s)\in m_n(t(s))+\xi\cdot t(s)\in m_n^*(t(s))$ and $x_+^{*\prime}(0)=x_+'(0)+\xi\cdot t_+'(0)=x_+'(0)+\xi\cdot\Delta\geq\epsilon$. Therefore $x_+^{*\prime}(0)\in D(m_n^*)_{t|M_n^*(t)}(\Delta)$. Since $D(m_n^*)_{t|M_n^*(t)}$ is a cone,

$$t_j'\cdot x_+^{*\prime}(0)\in D(m_n^*)_{t|M_n^*(t)}(t_j'\cdot\Delta)=D(m_n^*)_{t|M_n^*(t)}(t').$$

So for all t and all t'>0, $\exists x>0$ such that $x\in D(m_n^*)_{t|M_n^*(t)}(t')$. By the Claim, $t''\geq t'$ implies $M_n^*(t'')\geq M_n^*(t')$, which implies $M_n(t'')+\xi\cdot t''\geq M_n(t')+\xi\cdot t'$. Letting $\epsilon$ approach zero, we have $M_n(t'')\geq M_n(t')$. Therefore, $M_n$ is a nondecreasing function. Apply Tarski's Fixed Point Theorem to obtain a fixed point of the game $\Gamma$. ||

**Example of a Monotone Game**

Two firms compete for customers by choosing a level of quality q. Quality level q has cost c(q) per customer (or per item sold). The price per customer is assumed to be fixed at p. Assume firm i has $N_i$ loyal customers. Let f(x) be the fraction of a firm's loyal customers that desert it when its quality level is at least x below the quality of the competing firm. For example, $f(q_2-q_1)$ is the fraction of firm 1's customers that switch to firm 2 if $q_2>q_1$. Then firm 1's profits are given by:

$$\Pi_1(q_1,q_2)=\begin{cases}[N_1+N_2(1-f(q_1-q_2))](p-c(q_1)), & q_1\geq q_2\\ N_1 f(q_2-q_1)(p-c(q_1)), & q_1<q_2.\end{cases}$$

Assume:

$$N_i=50,\ q_i\in[0,5],\ c(q)=q/10,\ p=1,\ \text{and}\ f(x)=\begin{cases}\frac{(x-3)^4}{3^4}, & x\in[0,3]\\ 0, & x\in(3,5].\end{cases}$$

Then $\Pi_1$ is twice continuously differentiable, but $-\Pi_1$ does not have decreasing differences in $q_1$ and $q_2$. Hence this is not a submodular game. However, this game is a monotone game and thus has a pure-strategy Nash equilibrium.[29]

**Proof of Theorem 5.2**

Suppose $\{x^k\}$ is a sequence of points of $\partial f(x)$ converging to $x^*$. Then for all k and for all z, $f(z) \geq f(x) + x^k \cdot (z-x)$. Rewriting this, for all k and for all z, $x^k \cdot (z-x) \leq f(z) - f(x)$. Since $x^k \rightarrow x^*$, $x^o \cdot (z-x) \leq f(z) - f(x)$ for all z, implying $x^* \in \partial f(x)$. Therefore, $\partial f(x)$ is closed. Suppose $w, y \in \partial f(x)$ and $\lambda \in [0,1]$. Then for all z, $f(z) \geq f(x) + w \cdot (z-x)$ and $f(z) \geq f(x) + y \cdot (z-x)$. But this implies that $f(z) \geq f(x) + (\lambda w + (1-\lambda) y) \cdot (z-x)$, so $\lambda w + (1-\lambda) y \in \partial f(x)$. Therefore, $\partial f(x)$ is convex. ‖

**Proof of Theorem 5.5**

**Definition:** For multifunction $S: \mathbf{R}^m \rightrightarrows \mathbf{R}^n$ with closed graph, define the <u>quotient mapping</u> $\Delta_\tau S_{\bar{t}|\bar{x}}: \mathbf{R}^m \rightrightarrows \mathbf{R}^n$ by

$$\Delta_\tau S_{\bar{t}|\bar{x}}(t') = \tau^{-1}[S(\bar{t}+\tau t') - \bar{x}] = \{x' \mid \bar{x} + \tau x' \in S(\bar{t}+\tau t')\}.$$

It is clear from the definition of proto-derivative and of quotient mapping that the limit as $\tau \downarrow 0$ of $\Delta_\tau S_{\bar{t}|\bar{x}}(t')$ is the proto-derivative of S at $\bar{t}$ with respect to $\bar{x}$ in direction t'. Thus we have:

**A.4 Lemma:** For multifunction $S: \mathbf{R}^m \rightrightarrows \mathbf{R}^n$ with closed graph, the proto-derivative of S at $\bar{t}$ with respect to $\bar{x}$ in direction t' is given by

$$DS_{\bar{t}|\bar{x}}(t') = \lim_{\tau \downarrow 0} \Delta_\tau S_{\bar{t}|\bar{x}}(t').$$

---

[29]This is easily verified using Theorem 5.4 of Rockafellar (1989B) since either both $\partial^2\Pi_1/\partial q_1 \partial q_2$ and $\partial^2\Pi_1/\partial q_1 \partial q_1$ are zero or $\left[\dfrac{\partial^2 \Pi_1}{\partial q_1 \partial q_2}\right]\left[\dfrac{\partial^2 \Pi_1}{\partial q_1 \partial q_1}\right]^{-1} \leq 0$.

Using the definition of a quotient mapping and Theorem 5.4, we have the following Lemma.

**A.5 Lemma:** Let $f:\mathbf{R}^m x\mathbf{R}^n \to \mathbf{R}\cup\{\infty\}$ be a convex function. Define multifunction $m:\mathbf{R}^m \rightrightarrows \mathbf{R}^n$ by $m(t)=\arg\min_{x\in\mathbf{R}^n} f(t,x)$. Then

$$\Delta_\tau m_{\bar{t}|\bar{x}}(t')=\{x'|(y,0)\in\partial f(\bar{t}+\tau t',\bar{x}+\tau x'),\text{ some } y\}.$$

**A.6 Lemma:** Let $f:\mathbf{R}^m x\mathbf{R}^n \to \mathbf{R}\cup\{\infty\}$. Suppose $(\bar{y},0)\in\partial f(\bar{t},\bar{x})$. $\exists y$ such that $(y,0)\in\partial f(\bar{t}+\tau t',\bar{x}+\tau x')$ if and only if there exists y' such that $(y',0)\in\Delta_\tau\partial f(\bar{t},\bar{x}|\bar{y},0)(t',x')$.

**Proof of Lemma A.6:** $\exists y$ s.t. $(y,0)\in\partial f(\bar{t}+\tau t',\bar{x}+\tau x') \leftrightarrow \exists y'$ s.t. $(\bar{y}+\tau y',0)\in\partial f(\bar{t}+\tau t',\bar{x}+\tau x')$

$\leftrightarrow \exists y'$ s.t. $(\bar{y},0)+\tau(y',0)\in\partial f(\bar{t}+\tau t',\bar{x}+\tau x')$

$\leftrightarrow \exists y'$ s.t. $(y',0)\in\Delta_\tau\partial f(\bar{t},\bar{x}|\bar{y},0)(t',x')$.

By Lemmas A.4 and A.5, $Dm_{\bar{t}|\bar{x}}(t')=\lim_{\tau\downarrow 0}\{x'|(y,0)\in\partial f(\bar{t}+\tau t',\bar{x}+\tau x'),\text{ some } y\}$. So by Lemma A.6, $Dm_{\bar{t}|\bar{x}}(t')=\lim_{\tau\downarrow 0}\{x'|(y',0)\in\Delta_\tau(\partial f)_{\bar{t},\bar{x}|\bar{y},0}(t',x'),\text{ some } y'\}$. Since we assume $Dm_{\bar{t}|\bar{x}}(t')$ exists, $\Delta_\tau(\partial f)_{\bar{t},\bar{x}|\bar{y},0}$ must converge as $\tau\downarrow 0$. Thus, $\lim_{\tau\downarrow 0}\Delta_\tau(\partial f)_{\bar{t},\bar{x}|\bar{y},0}=D(\partial f)_{\bar{t},\bar{x}|\bar{y},0}$.[30] Taking x' as given, there is a set of y values, $Y(\tau,x')$, defined by $Y(\tau,x')\equiv\{y'|(y',0)\in\Delta_\tau(\partial f)_{\bar{t},\bar{x}|\bar{y},0}(t',x')\}$. Define $Y^*(x')\equiv\{y'|(y',0)\in D(\partial f)_{\bar{t},\bar{x}|\bar{y},0}(t',x')\}$. Note that $Y(\tau,x')$ must converge to $Y^*(x')$ as $\tau\downarrow 0$. The set of x values such that $(y,0)\in\Delta_\tau(\partial f)_{\bar{t},\bar{x}|\bar{y},0}(t',x')$ for some y is just the set $\{x'|Y(\tau,x')\neq\emptyset\}$. Since $Y(\tau,x')$ converges to $Y^*(x')$ as $\tau\downarrow 0$, this set converges to $\{x'|Y(x')\neq\emptyset\}$ as $\tau\downarrow 0$. In other words, $\{x'|(y,0)\in\Delta_\tau(\partial f)_{\bar{t},\bar{x}|\bar{y},0}(t',x'),\text{ some } y\}$ converges to $\{x'|(y,0)\in D(\partial f)_{\bar{t},\bar{x}|\bar{y},0}(t',x'),\text{ some } y\}$, completing the proof. ‖

**Proof of Theorem 5.6**

[30]Note that this is convergence in the sense that the graphs of the mappings $\Delta_\tau\partial f$ converge.

**A.7 Lemma (Theorem 2.2 of Rockafellar (1990A)):** Let $f:\mathbf{R}^n \to \mathbf{R} \cup \{\infty\}$ be a closed proper convex function. f is twice epi-differentiable at x relative to a vector v if and only if $v \in \partial f(x)$ and $\partial f$ is proto-differentiable at x relative to v. Moreover the subdifferential of $\frac{1}{2}D^2 f(x|v)$ is the proto-derivative of $\partial f$ at x relative to v: $\partial(\frac{1}{2}D^2 f)(x|v)(\xi) = D(\partial f)(x|v)(\xi)$ for all $\xi$.

First note that f being twice epi-differentiable implies $\partial f$ is proto-differentiable by Lemma A.7. By Theorem 5.5, $Dm(\bar{t}|\bar{x})(t') = \{x' | (y,0) \in D(\partial f)(\bar{t},\bar{x}|\bar{y},0)(t',x')\}$. We must show that $(y,0) \in D(\partial f)(\bar{t},\bar{x}|\bar{y},0)(t',x')$ if and only if $(y',0) \in \partial[D^2 f(\bar{t},\bar{x}|\bar{y},0)](t',x')$. By Lemma A.7, $(y,0) \in D(\partial f)(\bar{t},\bar{x}|\bar{y},0)(t',x')$ if and only if $(y,0) \in \partial(\frac{1}{2}D^2 f)(\bar{t},\bar{x}|\bar{y},0)(t',x')$. But this holds if and only if $(2y,0) \in \partial(D^2 f)(\bar{t},\bar{x}|\bar{y},0)(t',x')$, and we are done. ‖

**Proof of Theorem 5.10**

**A.8 Lemma (Rockafellar (1990A)):** Let $f:\mathbf{R}^n \to \mathbf{R} \cup \{\infty\}$ be a closed proper convex function. f is twice epi-differentiable at x relative to a vector v if and only if $v \in \partial f(x)$ and $\partial f$ is proto-differentiable at x relative to v. Moreover the subdifferential of $\frac{1}{2}D^2 f_{x|v}$ is the proto-derivative of $\partial f$ at x relative to v: $\partial(\frac{1}{2}D^2 f_{x|v})(\xi) = D(\partial f)_{x|v}(\xi)$ for all $\xi$.

By Lemma A.8, f being twice epi-differentiable implies $\partial f$ is proto-differentiable, and $Dm_{\bar{t}|\bar{x}}(t') = \{x' | (y,0) \in D(\partial f)_{\bar{t},\bar{x}|\bar{y},0}(t',x')\}$. We must show that $(y,0) \in D(\partial f)_{\bar{t},\bar{x}|\bar{y},0}(t',x')$ if and only if $(y',0) \in \partial(D^2 f_{\bar{t},\bar{x}|\bar{y},0})(t',x')$. By Lemma A.8, $(y,0) \in D(\partial f)_{\bar{t},\bar{x}|\bar{y},0}(t',x')$ if and only if $(y,0) \in \partial\frac{1}{2}D^2 f_{\bar{t},\bar{x}|\bar{y},0}(t',x')$. But this holds if and only if $(2y,0) \in \partial D^2 f_{\bar{t},\bar{x}|\bar{y},0}(t',x')$, and we are done. ‖

**Proof of Theorem 6.1**

Since the basic constraint qualification holds, $(\text{gph } m) \subseteq (\text{gph } \mu)$, so $K_{\text{gph } m}(\bar{t},\bar{x}) \subseteq K_{\text{gph } \mu}(\bar{t},\bar{x})$. I claim that $K_{gph\,\mu}=K_{gph\,m}$.

**Proof of claim:** Let $\bar{x}\in m(\bar{t})$. Suppose $y\in K_{gph\,\mu}(\bar{t},\bar{x})\backslash K_{gph\,m}(\bar{t},\bar{x})$. Then there exist sequences$\bar{x}$ $(y^k)$ and $(\tau^k)$ such that $y^k\to y$ and $\forall k$, $y^k\in(\tau^k)^{-1}[\text{gph}\ \mu - (\bar{t},\bar{x})]$. If the sequences are such that $y^k\in(\tau^k)^{-1}[\text{gph}\ m - (\bar{t},\bar{x})]$ infinitely often, then $y\in K_{gph\,m}(\bar{t},\bar{x})$, a contradiction. So there exist subsequences of $(y^k)$ and $(\tau^k)$, denoted $(y^j)$ and $(\tau^j)$, such that for all j

$$y^j\in(\tau^j)^{-1}[(\text{gph}\ \mu)\backslash(\text{gph}\ m) - (\bar{t},\bar{x})].$$

So for all j, $(\bar{t},\bar{x})+\tau^j y^j\in(\text{gph}\ \mu)\backslash(\text{gph}\ m)$. Since $((\bar{t},\bar{x})+\tau^j y^j)$ is a sequence in $(\text{gph}\ \mu)\backslash(\text{gph}\ m)$ that converges to $(\bar{t},\bar{x})$, then $(\bar{t},\bar{x})\in(\text{gph}\ \mu)\backslash(\text{gph}\ m)$. But this contradicts $\bar{x}\in m(\bar{t})$.

Since monotonicity depends only upon the contingent cone of a proto-differentiable multifunction, and since the contingent cones of $\mu$ and m are the same when $\mu$ is monotone, the proof is complete. ‖

## References


Arrow, Kenneth J., and Gerard Debreu (1954), "Existence of an Equilibrium for a Competitive Economy," *Econometrica*, **22**, 265-290.

Attouch, H., *Convergence de fonctions convexes, des sous-différentiels, et semi-groupes associés*, C.R. Acad. Sci. Paris, **284**, 539-542.

Clarke, F. H. (1983), *Nonsmooth Analysis and Optimization*, Wiley.

Dasgupta, Partha and Eric Maskin (1986), "The Existence of Equilibrium in Discontinuous Economic Games, I: Theory," *Review of Economic Studies*, **53**, 1-26.

Debreu, Gerard (1952), "A Social Equilibrium Existence Theorem," *Proceedings of the National Academy of Sciences*, **38**, 886-893.

Fan, Ky (1952), "Fixed-Point and Minimax Theorems in Locally Convex Topological Linear Spaces," *Proceedings of the National Academy of Sciences*, **38**, 121-126.

Glicksberg, I.L. (1952), "A Further Generalization of the Kakutani Fixed Point Theorem, with Applications to Nash Equilibrium Points," *Proceedings of the American Mathematical Society*, **3**, 170-174.

Milgrom, Paul and John Roberts (1990), "The economics of modern manufacturing: Technology, strategy, and organization," *American Economic Review*, **80**, 3, 511-528.

Milgrom, Paul and Christina Shannon (1991), "Monotone Comparative Statics," SITE Technical Report No. 11.

Rockafellar, R. Tyrrell (1990A), "Generalized Second Derivatives of Convex Functions and Saddle Functions, "*Transactions of the American Mathematical Society*, **322**, 1, 51-77.

Rockafellar, R. Tyrrell (1990B), "Nonsmooth Analysis and Parametric Optimization," in *Methods on Nonconvex Analysis* (A. Cellina, ed.), Springer-Verlag Lecture Notes in Mathematics, no.

1446, 137-151.

Rockafellar, R. Tyrrell (1989A), "Perturbation of Generalized Kuhn-Tucker Points in Finite-Dimensional Optimization," *Nonsmooth Analysis and Related Topics*, F.H. Clarke et al. (editors), Plenum Press, 393-402.

Rockafellar, R. Tyrrell (1989B), "Proto-Differentiability of Set-Valued Mappings and its Applications in Optimization," *Analyse Non Linéaire*, H. Attouch, J.-P. Aubin, F. H. Clarke, I. Ekeland (eds.), Gauthier-Villars: Paris, 449-482.

Rockafellar, R. Tyrrell (1981), *The Theory of Subgradients and its Applications to Problems of Optimization: Convex and Nonconvex Functions*, Heldermann Verlag: Berlin.

Rockafellar, R. Tyrrell (1970), *Convex Analysis*, Princeton University Press: Princeton, New Jersey.

Topkis, Donald M. (1978), "Minimizing a submodular function on a lattice," *Operations Research*, March-April 1978, **26**, 305-21.

Veinott, Arthur F. (1989), "Lattice Programming," unpublished notes from lectures delivered at Johns Hopkins University.

Vives, Xavier (1989), "Nash equilibrium with strategic complementarities," mimeo, Universitat Autonoma de Barcelona. To appear in the *Journal of Mathematical Economics*.