Lipschitz continuous dynamic programming with discount

Jose M. Maroto[a] and Manuel Moran[b,*]

[a]Department of Estadística e Investigación Operativa II.

Universidad Complutense, 28223 (Madrid), Spain.

[b]Department of Fundamentos del Análisis Económico I.

Universidad Complutense, 28223 (Madrid), Spain.

*Corresponding author. PHONE: +34913942407, FAX: +34913942561.

E-mail addresses: mmoranca@ccee.ucm.es (M. Moran*), maroto@ccee.ucm.es.

April 13, 2005

Abstract

We show that if the return function, the technological constraints and the transition function of a standard problem of stochastic dynamic programming with discount satisfy Lipschitz regularity assumptions, then the value function is Lipschitz regular.

Keywords: Dynamic programming, optimization, optimal growth, renewable resources, non-smoothness.

1 Introduction

The results of this paper stand for a class of dynamic optimization problems with infinite horizon and discount, in a stochastic setting, as described by Stokey, Lucas and Prescott [29], chaps. 4, 9.

It is well known that, under topological assumptions (compactness and continuity) on the data of the problem (i.e., state space, return function and technological constraint correspondence), the existence, uniqueness and continuity of the value function is guaranteed.

The theory of dynamic programming with discount proceeds by completing the topological assumptions with a rather extensive block of assumptions, which we call standard assumptions, including concavity, smoothness and monotonicity of the data. Such assumptions guarantee the concavity, smoothness and numerical computability of the value function and optimal policy correspondence. In the non-random case they also guarantee the convergence of the optimal paths to an equilibrium state, and if noninterior optimal paths are ruled out, then a recursive computation of the optimal paths through Euler equations is possible.

The examples in Section 4 show that all these nice properties, with exception of the existence and continuity of the value function, fail to hold under small departures from the standard assumptions. A variety of phenomena emerge there related with nonconcavity of the objective function, as countably many points of discontinuity and non-uniqueness, following a systematic pattern, in the optimal policy correspondence; discontinuities in the form of jumps upwards in the marginal value function, synchronized with the discontinuities of the optimal policy correspondence (Example 18); asymptotic cyclic behavior of the optimal paths (Example 18). In these cases, not only do the properties derived from the standard assumptions are hopelessly lost, but also the numerical computation of the value function through Bellman operator iterates is not possible for states out of the grid used for the discretization of the phase space, since no rate of convergence can be derived from the standard theory for such states.

Is it possible to construct an alternative theoretical framework which does not require the extensive list of standard assumptions? Can such theory give useful information on some relevant problems that are intractable in the standard framework?

The aim of this paper is to give some partial answers to these questions. In particular we show that if the data of the problem satisfy Lipschitz continuous assumptions, then the value function is Lipschitz continuous (see Theorem 14). A first consequence of our result is that in that setting the value function and optimal policy correspondence are numerically computable (see Morán and Maroto [26]), thus giving a theoretical basis to our examples above, and to some numerical experiments that have recently raised interest in the literature (Dawid and Kopel [11, 12]).

The most direct antecedent of this paper is the result of Bertsekas [3]. There, in a setting of optimal stochastic control with a discrete state space for the random shocks and admissible controls, it is proved that under Lipschitz assumptions on the data of the problem, the value function can be computed. If we ignore the different settings of the problems, the main contribution of our results is that Bertsekas' method does not permit us to prove that the value function is Lipschitz regular, which is the key step to the obtention of the rate of convergence of the numerical algorithm for the computation of the value function (Morán and Maroto [26]).

The Lipschitz continuity of the value function has been analyzed by Yue [30, 31] in optimal control and optimal time control problems respectively. Montrucchio [25] proves that the policy function is Lipschitz continuous under assumptions of strong concavity.

Bardi and Capuzzo-Dolceta [2] proved that the value function is Lipschitz continuous in infinite horizon problems of optimal control with discount. This result does not allow dependence of the admissible controls on the state of the system, i.e., the existence of a technological constraint correspondence is ruled out. Such correspondence plays a central role in the problem.

Relevant research regarding applications of the results in this paper is focused on nonconcavity of the growth function of the resource in problems of optimal exploitation of renewable resources (Clark [8]). Notice that these problems are equivalent to optimal growth models with linear or strictly concave objective functions and convex-concave production functions (Majumdar and Mitra [22, 23], Dechert and Nishimura [13], Le Van and Dana [21]). In this sense, the results in this paper can also be applied to these economic problems.

A second field where the results of this paper find natural application is that of dynamic optimization problems with convex objective functions. This case has been shown to be relevant in the exploitation of renewable resources. In particular, there is empirical evidence in the literature about convex objective functions in fisheries management (Bj∅rndal [5], Bj∅rndal and Conrad [6]). See also Dasgupta and Mäler [10] for a recent overview on nonconvex ecosystems, and Dawid and Kopel [11, 12] for a numerical analysis of a renewable resource subject to a convex objective function. The case of a convex objective function is also relevant in capital accumulation models of the firm where the revenue is a convex function of the capital stock (Hartl and Kort [15]). See also Hartl and Kort [16] for capital accumulation models where the revenue is a convex-concave function of the capital stock.

Lastly, there are relevant economic problems that may be treated in the Lipschitz setting. See Arrow et al. [1], Brian [7], and Heal [17] for recent overviews on economic problems related to nonconcavities. See also references in Example (18).

2 Preliminaries

2.1 *The optimization problem*

We describe the stochastic dynamic optimization problem we deal with. Let $(X, \mathcal{X})$ be a measurable space with $X \subset \mathbb{R}^n$ and let $\mathcal{X}$ be the $\sigma-$algebra of Borel subsets of X. The space X is assumed to be the domain of an endogenous state variable x, and $Z \subset \mathbb{R}^m$ endowed with the $\sigma-$algebra $\mathcal{B}_m$ of Borel subsets of Z is the domain of a sequence $z_0, z_1, z_2, ...$ of exogenous random shocks.

The state of the system at time t is therefore described by a vector (x_t, z_t) ranging in the set $S := X \times Z$. As a topological space, S is endowed with the product topology, and as a measure space it is endowed with the product $\sigma-$algebra. The technological constraints of the problem are represented by a correspondence $\Gamma : S \to X$ which specifies the set $\Gamma(x_t, z_t)$ of feasible states x_{t+1}. We shall write Ω for the graph of Γ, and consider in Ω the topology and $\sigma-$algebra inherited from the product space $X \times X \times Z$.

Let the value z_0 of the first random shock be known, and for $t \geq 1$ assume that the sequence of random variables z_t are a Markov stochastic process with stationary transition function Q, which for $z \in Z$, A in the $\sigma-$algebra $\mathcal{B}_m$ of Borel subsets of $\mathbb{R}^m$ and $t \geq 1$, specifies the probability $Q(z, A)$ that $z_t \in A$ conditional to $z_{t-1} = z$. This defines in a standard way, for $z_0 \in Z$ and $t \geq 1$, the probability measure $\mu^t(z_0, \cdot)$ on the $t-$fold product space $Z^t = Z \times Z \times ... \times Z$ which specifies the (conditional to z_0) probabilities $\mu^t(z_0, A)$ that the sequence $z^t := (z_1, z_2, ..., z_t)$ of first t random shocks belongs to the sets A in the t-fold product $\sigma-$algebra $\mathcal{B}_m^t$.

Given $z_0 \in Z$ and $x_0 \in X$, a planner faces the problem of finding an optimal plan, that is, a constant π_0 and a sequence $\pi_1, \pi_2, \pi_3, ...$of measurable functions $\pi_t := Z^t \to X$ which solves the problem

$$\sup\{R(x_0, \pi_0, z_0) + \sum_{t=1}^{\infty} \beta^t \int_{Z^t} R(\pi_{t-1}(z^{t-1}), \pi_t(z^t), z_t)\mu^t(z_0, dz^t)\}, \quad (1)$$

where the supremum is taken over all plans satisfying the technological constraints; $\beta \in (0, 1)$ is a discount factor; and R is the return function, defined on the graph Ω of the correspondence Γ, so $R\ (x_t, x_{t+1}, z_t)$ is the return at time t if the state variable is set to be x_{t+1} at time $t+1$ and the current state of the system is (x_t, z_t).

2.2 *Notation and definitions*

We now discuss the required conditions on the data in the above problem, explore the scope of application of such conditions, and state several properties derived from such conditions, that are to be used later on.

2.2.1 *Lipschitz functions*

Given a metric space (Y, d) and a point $x \in Y$, we shall denote by $U(x, r)$ and $B(x, r)$ respectively, the open and the closed ball centered at x with radius r.

Recall that a mapping between two metric spaces $f : (Y, d) \to (Y', d')$ is said to be a Lipschitz mapping if it satisfies $d'(f(a), f(b)) \leq Kd(a, b)$ for all $a, b \in Y$ and some constant K. We shall write $f \in L(K)$ for such mapping (or $L_A(K)$ if we want to make explicit some subset $A \subset Y$ where the Lipschitz condition holds for f restricted to A). The function f is said to be locally Lipschitz on Y if for every $x \in Y$ there exists a constant K and an open ball $U(x, \varepsilon)$, $\varepsilon > 0$, such that $f \in L(K)$ on $U(x, \varepsilon)$, and we write $f \in L^{loc}(K)$ ($L_A^{loc}(K)$) if such condition holds on Y (for f restricted to $A \subset Y$). A function v will be said to belong to $BL_Y(\alpha, K)$ $(BL_Y^{loc}(\alpha, K))$ if $v \in L_Y(K)$ $(L_Y^{loc}(K))$ and it is bounded by α (in the supremum norm) on Y.

Definition 1 *: We call $L-$convex a subset $C \subset \mathbb{R}^p$ if it is a bilipschitz image* *of a convex set.*

The key property of $L-$convex sets is that functions which are locally Lipschitz on them, are also globally Lipschitz. The following lemma can be easily proved

Lemma 2 *: Let $C \subset S \subset \mathbb{R}^p$ be a $L-$convex set with $f(C)$ convex and f bilipschitz, and let $v{:}S \to \mathbb{R}$ with $v \in L_C^{loc}(K)$. Then $v \in L_C(\Lambda_C K)$ with*

$$\Lambda_C = K_{f^{-1}} K_f, \tag{2}$$

where K_f and $K_{f^{-1}}$ denote respectively the Lipschitz constants of f and f^{-1}.

Remark 3 *: If C is convex then we can set $f = id$, which shows $\Lambda_C = 1$. It is easy to see that in any case $\Lambda_C \geq 1$ (see note 1). The case $p = 1$ does not have interest, since a $L-$convex subset of R is an interval.*

2.2.2 *Correspondences and Hausdorff metric*

Given a metric space (Y, d) and a subset $A \subset Y$, we denote by $[A]_\delta$ the δ-parallel body of A, that is, $[A]_\delta = \cup_{x \in A} B(x, \delta)$. For compact non-empty subsets $A, B \subset Y$ we define $d_H(A, B) = \min\{\varepsilon : B \subset [A]_\varepsilon\}$, and $D_H(A, B) = \max\{d_H(A, B), d_H(B, A)\}$. With this definition D_H is a metric on the set $\mathcal{H}_Y$ of compact non-empty subsets of Y. D_H is known in the literature as the Hausdorff metric.

The results in this paper depend on the following qualification of continuity of a correspondence.

Definition 4 *: Let $\Delta : X \to Y$ be a correspondence between metric spaces (X, d) and (Y, d'). We say that Δ is topologically continuous at x if it is a compact non-empty valued correspondence, continuous at x and $D_H(\partial\Delta(x), \partial\Delta(y)) \to 0$ as $y \to x$, where $\partial\Delta$ denotes the boundary of Δ. Δ is said to be topologically continuous if it is topologically continuous at all $x \in X$.*

The topologically continuous correspondences most frequently used in the literature are compact, convex valued continuous correspondences. Indeed, it can be shown that compact and convex valued correspondences are topologically continuous.

The following lemma states the property of the type of correspondences used in the proof of the main result.

Lemma 5 *: Let (X, d) and (Y, d') be metric spaces and let $\Gamma : X \to Y$ be a correspondence topologically continuous at $x \in X$. Then*

i) For any $y \in int(\Gamma(x))$, there exists an open neighbourhood U of x and an open neighbourhood U' of y such that $U' \subset int(\Gamma(z))$ for all $z \in U$.

ii) (Mutual factibility condition) If $G : X \to Y$ is a compact valued u.h.c correspondence such that $G(x) \subset int(\Gamma(x))$, then there exists an open neighbourhood U of x such that $G(z) \subset int(\Gamma(x))$ and $G(x) \subset int(\Gamma(z))$ hold for any $z \in U$.

Proof. i) Let $\rho = d'(y, \partial\Gamma(x))$. We know that $\rho > 0$ because $y \notin \partial(\Gamma(x))$. By lower hemi-continuity of Γ, for a sufficiently small ε and for any z with $d(z, x) < \varepsilon$, there exists $z' \in \Gamma(z)$ with $d'(z', y) < \rho/2$. By topological continuity of Γ, we also may assume that $D_H(\partial\Gamma(x), \partial\Gamma(z)) < \rho/2$ holds if $d(z, x) < \varepsilon$. In this situation $B(y, \rho/2) \subset int(\Gamma(z))$ must hold. To check this, assume first that $B(y, \rho/2) \cap \Gamma(z)^c \neq \emptyset$. Then $B(y, \rho/2) \cap \partial\Gamma(z) \neq \emptyset$ since we know that $z' \in B(y, \rho/2) \cap \Gamma(z) \neq \emptyset$ *(see note 2)*. Let $z'' \in \partial\Gamma(z)$ with $d'(y, z'') \leq \rho/2$ and let $z''' \in \partial\Gamma(x)$ with $d'(z'', z''') < \rho/2$. We get the contradiction $d'(y, z''') < \rho$. This proves that $B(y, \rho/2) \subset \Gamma(z)$ hold. Since we have already proved that $B(y, \rho/2) \cap \partial\Gamma(z) \neq \emptyset$ cannot hold, this shows that if $d(z, x) < \varepsilon$, then $U(y, \rho/2) \subset B(y, \rho/2) \subset int(\Gamma(z))$.

ii) By upper hemicontinuity of G there exists an open neighbourhood U' of x such that, for $z \in U'$, $G(z) \subset [G(x)]_\varepsilon$, with $\varepsilon < d_H(G(x), \partial\Gamma(x))$. By the condition imposed on ε, $[G(x)]_\varepsilon \cap \partial\Gamma(x) = \emptyset$ holds, so $G(z) \subset int(\Gamma(x))$.

By part i), for any $y \in G(x)$ there exists an open neighbourhood U_y of x such that if $z \in U_y$ then $U(y, \varepsilon(y)) \subset int(\Gamma(z))$ holds for a sufficiently small positive $\varepsilon(y)$. By compactness we may find a cover $\{U(y', \varepsilon(y'))\}_{y' \in Y'}$ of $G(x)$, where $Y' \subset Y$ is a finite set. Let $U = U' \cap_{y' \in Y'} U_{y'}$, let $z \in U$ and let $y \in G(x)$. Since $z \in U'$, $G(z) \subset int(\Gamma(x))$ holds, and as $y \in U(y', \varepsilon(y'))$ for some $y' \in Y$ and $z \in U_{y'}$ we get $y \in U(y', \varepsilon(y')) \subset int(\Gamma(z))$, so $G(x) \subset int(\Gamma(z))$. ■

The Hausdorff metric gives also sense to the following definition, which allows us to introduce a second (metric) qualification of continuous correspondences

Definition 6 *: Let Δ be a compact non-empty valued correspondence from the metric space (Y, d) to the metric space (Y', d'). Then we write $\Delta \in L(K)$ if $\Delta_H \in L(K)$, so $D_H(\Delta(x), \Delta(y)) \leq K d(x, y), x, y \in Y$.*

Lipschitz correspondences arise in problems of optimal economic growth and exploitation of renewable resources, with technological constraints given by $\Gamma(x) = \{y : 0 \leq y \leq f(x)\}$ where $f(x)$ is a Lipschitz production function (law of growth of the resource, in problems of exploitation of renewable resources). It is easy to see that the Lipschitz constant of Γ coincides with that of f (see note 3).

2.2.3 *Lipschitz transition functions*

We shall use a metric on the set $\mathcal{M}_Z$ of Borel probability measures on the domain Z of the random shocks. Notice that the measures $Q(z, \cdot)$ defined by the transition function described above belong to $\mathcal{M}_Z$. We shall consider in $\mathcal{M}_Z$ the metric given by $d_\alpha(\mu, \nu) = \sup\{| \int f d\mu - \int f d\nu |: f \in BL_Z(\alpha, 1)\}$, where $BL_Z(\alpha, 1)$ is the set of functions $f : Z \to \mathbb{R}$ such that $f \in L_Z(1)$ and $\| f \| \leq \alpha$ (with $\| \cdot \|$ the supremum norm). The metrics d_α are equivalent metrics for any (positive) value of α. The metric d_1 is called Fortet-Mourier distance (see Dudley [14]). If Z is a compact subset of $\mathbb{R}^m$, it suffices to take the supremum, in the definition of the metric, over functions $f \in L_Z(1)$. We shall denote this metric by $d(\mu, \nu)$.

These distances allows us to introduce the following Lipschitz condition in the transition function Q of the Markov process which governs the random shocks

Definition 7 *: We say that the transition function satisfies the Lipschitz condition for the constant K_Q, and write $Q \in L_Z(K_Q)$ if the mapping $\mathcal{Q}(z) := Q(z, \cdot)$ from the metric space (Z, d), with d the Euclidean distance, to the metric space*

$(\mathcal{M}_Z, d_\alpha)$, *with the constant* α *specified later on (see (4)), satisfies* $\mathcal{Q} \in L_Z(K_Q)$.

The transition functions of many standard Markov processes satisfy the condition $Q \in L_Z(K_Q)$. As an example, consider the class of Markov processes defined by an initial random variable z_0 in $\mathbb{R}^m$ and the stochastic difference equation $z_{t+1} = g(z_t, \theta_t)$, with $g : Z \times \mathbb{R}^p \to Z \subset \mathbb{R}^m$, $\{\theta_t\}$ an i.i.d. process with $\theta_t \in \mathbb{R}^p$, all t, and g Borel-measurable (see Stokey, Lucas and Prescott [29], chap. 8). The transition function is here defined on the Borel subsets $A \subset Z$ by $Q(z, A) = P\{\theta \in \mathbb{R}^p : g(z, \theta) \in A\}$. Among the Markov processes in this class are AR and VAR models. The following proposition gives a simple sufficient condition for these processes to satisfy Assumption IV below, which in turn implies the convergence of the process to a unique invariant measure if the state space Z is compact.

Proposition 8 *: If the function* $g(\cdot, \theta) : Z \to Z$ *satisfies* $g(\cdot, \theta) \in L_Z(K)$ *then* $Q \in L_Z(K)$. *Furthermore, if* $K < 1$ *and* Z *is compact, given an arbitrary probability distribution (in* $\mathcal{M}_Z$*) for the random variable* z_0, *the probability distributions of the random variable* $\{z_t\}$ *converge (at a exponential rate, in the weak topology) to a unique invariant measure independent from* z_0.

Proof. Let $z \in Z$. For a Borel sets $A \subset Z$ we have $Q(z, A) = P\{\theta : g(z, \theta) \in A\}$,where P is the invariant probability distribution of the process $\{\theta_t\}$. This means that, if $g_z : Z \to Z$ is the $z-$section of g defined by $g_z(\theta) = g(z, \theta)$, then $Q(z, \cdot)$ is the image probability distribution of P under the mapping g_z. For $\alpha > 0$, $f \in BL_Z(\alpha, 1)$ and $z, z' \in Z$ the change of variable formula (see

Billingsley [4]) gives

$$\begin{aligned}
& \mid \int f(\theta)Q(z,d\theta) - \int f(\theta)Q(z',d\theta) \mid = \\
= & \mid \int f(g(z,\theta)) - f(g(z^{'},\theta))dP(\theta) \mid \leq \\
\leq & \int \mid f(g(z,\theta)) - f(g(z',\theta)) \mid dP(\theta) \leq K \parallel z - z' \parallel,
\end{aligned}$$

where it has been used that $f \in L_Z(1)$ and $g(\cdot,\theta) \in L_Z(K)$. This shows that $d_\alpha(Q(z,\cdot),Q(z',\cdot)) \leq K \parallel z - z' \parallel$, and the first assertion is proven.

Assume now that Z is a compact set. Recall that, in this case, $\mathcal{M}_Z$ endowed with the metrics $d(\mu,\nu) = \sup\{\mid \int f d\mu - \int f d\nu \mid : f \in L_Z(1)\}$ is a complete metric space, and the convergence defined by the metrics d is equivalent to the weak convergence. Let $T^* : \mathcal{M}_Z \to \mathcal{M}_Z$ be the adjoint operator associated to the transition function $Q(z,\cdot)$, defined by $T^*(\lambda)(A) = \int Q(z,A)d\lambda(\theta)$ for $\lambda \in \mathcal{M}_Z$ (see Stokey, Lucas and Prescott [29], chap. 8). We want to prove that, for any $\lambda \in \mathcal{M}_Z$, $T^{*^k}(\lambda) \to \mu \in \mathcal{M}_Z$, where the limiting measure μ is independent from λ. We only need to prove that T^* is a contracting mapping. Given $f \in L_Z(1)$ and $\lambda, \lambda' \in \mathcal{M}_Z$ we have

$$\begin{aligned}
& \mid \int f(\theta)dT^*\lambda(\theta) - \int f(\theta)dT^*\lambda'(\theta) \mid = \\
= & \mid \int\int f(\theta)Q(z,d\theta)d\lambda(z) - \int\int f(\theta)Q(z,d\theta)d\lambda'(z) \mid \leq \\
\leq & \int \mid \int f(g(z,\theta))d\lambda(z) - \int f(g(z,\theta))d\lambda'(z) \mid dP(\theta) = \\
= & K \int \mid \int K^{-1}f(g(z,\theta))d\lambda(z) - \int K^{-1}f(g(z,\theta))d\lambda'(z) \mid dP(\theta) \leq \\
\leq & K \int d(\lambda,\lambda')dP(\theta) \leq Kd(\lambda,\lambda'),
\end{aligned}$$

where it has been used that $K^{-1}f(g(\cdot,\theta)) \in L_Z(1)$ if $g(\cdot,\theta) \in L_Z(K)$.This gives

$d(T^*\lambda, T^*\lambda') \leq Kd(\lambda, \lambda')$, as it was to be shown. ■

The following lemma, whose proof is elementary, shows the role played by the requirement $Q \in L_Z(K_Q)$.

Lemma 9 *: If the transition function satisfies $Q \in L_Z(K_Q)$, then, for $z, z' \in Z$, $y \in X$, and $v \in BL_{y\times Z}(\alpha, K)$, the following inequality holds*

$$\int v(y,\theta)Q(z', d\theta) \leq \int v(y,\theta)Q(z, d\theta) + K_Q \max\{1, K\} \parallel z - z' \parallel .$$

2.3 Assumptions

We are now ready to formulate our assumptions on the data X, Z, R, Γ, and Q. Some straightforward consequences of these assumptions used later on are also analyzed in the remainder of this section.

Assumption I: *X is a closed subset of R^n and Z is a closed subset of R^m.*

Assumption II: *R is a real bounded function and $R \in L_\Omega(K_R)$.*

Assumption III: *Γ is a topologically continuous correspondence and $\Gamma \in L_S(K_\Gamma)$.*

Assumption IV: *$Q \in L_Z(K_Q)$ with $K_Q\beta < 1$.*

In this paper we shall consider only the case of interior optimal paths (see (10), in Theorem 14). We do not add this assumption to the list above because from that list and the results in Theorem 14, it can be obtained the Lipschitz continuity of the value function in the general case of eventually interior optimal plans (See Morán and Maroto [27]). Thus the list above may be considered the basic assumptions for the Lipschitz continuous dynamic programming setting.

Remark 10 : *Assumption IV embodies the case of i.i.d. random shocks if we take $Q(\cdot, A)$ to be constant for each $A \in B_m$. In this case $K_Q = 0$, and $\beta K_Q < 1$ always holds. The deterministic version of the optimization problem (1) is also embodied in our analysis if we think of Z as a singleton $\{z_0\}$ with $Q(z_0, \{z_0\}) = 1$ and $Q(z_0, \emptyset) = 0$. In this case we also have $K_Q = 0$. VAR models $z_t = Az_t + \vartheta_t$ with θ_t i.i.d. also satisfy Assumption IV if the eigenvalues of A are inside the unit ball, including VAR models with roots of the unity.*

Remark 11 : *In the most classical models of economic growth the marginal utility of capital tends to infinity when the capital tend to zero so the return function fails to be a Lipschitz function on a neighbourhood of the origin. These case can be treated in the above setting by considering compact subsets $Y = X - U$ which do not contain a small neighbourhood U of the origin. The high marginal value on U imposes a fast optimal growth. Then the space $Y \times Z$, where the assumptions above hold, may be taken as the phase space of the problem. See Examples 18 and 19 for cases in such situation.*

A consequence of Assumption *IV* is that $\mathcal{Q}(z) = Q(z, \cdot)$ is a continuous mapping, so the transition function Q enjoys the Feller property (See Stokey, Lucas and Prescott [29], chap.11).

Since X and Z are closed sets by Assumption I, so it is S, which, therefore, is also a complete metric space. Let $BC_S(\alpha)$ denote the set of real continuous functions on S bounded in supremum norm by the constant α. Bellman operator

T, defined by

$$T(v(x,z)) = \sup\{R(x,y,z) + \beta \int v(y,\theta)Q(z,d\theta) : y \in \Gamma(x,z)\}, \tag{3}$$

preserves the set BC_S of real continuous bounded functions on S, i.e. $T : BC_S \to BC_S$, and it is a contractive operator with respect to the supremum norm in BC_S (see Stokey, Lucas and Prescott [29], chap. 9). It is easy to check that if $v \in BC_S(\alpha)$ then $Tv \in BC_S(\| R \| + \beta\alpha)$. Thus, under Assumptions $I - IV$, setting

$$\alpha = \| R \| (1-\beta)^{-1} \tag{4}$$

we have $T : BC_S(\alpha) \to BC_S(\alpha)$.

We shall keep from now onwards the value of α given by (4). It follows from the completeness of $BC_S(\alpha)$ the existence of a unique $V \in BC_S(\alpha)$ such that $T(V) = V$. Moreover, if T^k denotes the $k-$th iterate of T, then $\| T^k(v) - V \| \to 0$ for any $v \in BC_S$ and V is the unique value function of the optimization problem (1).

It is well known that the optimal policy correspondence $G : S \to X$, given by

$$G(x,z) = \{y \in \Gamma(x,z) : V(x,z) = R(x,y,z) + \beta \textstyle\int V(y,\theta)Q(z,d\theta)\},$$

is a compact valued and u.h.c correspondence. For each $v \in BC_S$ a maximizing correspondence G_v may be defined by

$$G_v(x,z) = \{y \in \Gamma(x,z) : Tv(x,z) = R(x,y,z) + \beta \textstyle\int v(y,\theta)Q(z,d\theta)\}.$$

Observe that, as a consequence of Theorem of Maximum (see Stokey, Lucas and Prescott [29], Theorem 3.6), G_v is a u.h.c. and compact valued correspondence.

3 Lipschitz regularity of the value function

We first obtain an upper bound for the rate of growth of Lipschitz constants under Bellman operator T (see definition in expression (3)).

The following simple Lemma states a property of Lipschitz correspondences used in Lemma 13.

Lemma 12 *: Let $\Delta : (Y, d) \to (Y', d')$ be a Lipschitz correspondence with $\Delta \in L_Y(K)$. If $x_1, x_2 \in Y$ and $y \in \Delta(x_1)$ then there exists $y' \in \Delta(x_2)$ with $d'(y, y') \leq Kd(x_1, x_2)$.*

Proof. Assume on the contrary that for all $y' \in \Delta(x_2)$, $d'(y, y') > Kd(x_1, x_2)$ holds. Then $y \notin [\Delta(x_2)]_\varepsilon$, for $\varepsilon = Kd(x_1, x_2)$, giving the contradiction

$D_H(\Delta(x_1), \Delta(x_2)) > Kd(x_1, x_2)$. ■

Lemma 13 *: Let $v \in BC_S$ and $v \in L(M_0), M_0 \geq 0$. Then, under Assumptions I– IV, $Tv \in L(M_1)$ holds, with $M_1 = K_R(1+K_\Gamma)+\max\{1, M_0\}\beta K_Q+M_0\beta K_\Gamma$.*

Proof. *Let $(x, z), (x', z') \in S$. We may assume $Tv(x, z) \geq Tv(x', z')$. Let $y \in G_v(x, z)$, so that it holds*

$$Tv(x, z) = R(x, y, z) + \beta \int v(y, \theta)Q(z, d\theta). \tag{5}$$

Since $y \in \Gamma(x,z)$ and, by assumption III,

$$D_H(\Gamma(x,z),\Gamma(x',z')) \leq K_\Gamma \parallel (x,z)-(x',z') \parallel,$$

we may find (see Lemma 12) some $y' \in \Gamma(x',z')$ with

$$\parallel y-y' \parallel \leq K_\Gamma \parallel (x,z)-(x',z') \parallel . \tag{6}$$

Using Assumption II and (6) we get for the first summand in (5)

$$\begin{aligned} R(x,y,z) &\leq R(x',y',z')+K_R \parallel (x,y,z)-(x',y',z') \parallel \leq \\ &\leq R(x',y',z')+K_R(\parallel (x,z)-(x',z') \parallel + \parallel y-y' \parallel) \leq \\ &\leq R(x',y',z')+K_R(1+K_\Gamma) \parallel (x,z)-(x',z') \parallel . \end{aligned} \tag{7}$$

For the second summand in (5) using that: $v \in L_S(M_0)$; (6); and Lemma 9 we get

$$\begin{aligned} \beta \int v(y,\theta)Q(z,d\theta) &\leq \beta \int (v(y',\theta)+M_0 \parallel y-y' \parallel)Q(z,d\theta) \leq \\ &\leq \beta \int v(y',\theta)Q(z,d\theta)+\beta M_0 K_\Gamma \parallel (x,z)-(x',z') \parallel \leq \\ &\leq \beta \int v(y',\theta)Q(z',d\theta)+\max\{1,M_0\}K_Q\beta \parallel z-z' \parallel + \\ +\beta M_0 K_\Gamma &\parallel (x,z)-(x',z') \parallel \leq \end{aligned}$$

$$\leq \beta \int v(y',\theta)Q(z',d\theta)+\beta(\max\{1,M_0\}K_Q+M_0K_\Gamma) \parallel (x,z)-(x',z') \parallel . \tag{8}$$

Using that $Tv(x,z) \geq Tv(x',z')$, (7) and (8)

$$\begin{aligned} | \quad & Tv(x,z)-Tv(x',z') | = Tv(x,z)-Tv(x',z') \leq \\ \leq \quad & R(x',y',z')+\beta \int v(y',\theta)Q(z',d\theta)-Tv(x',z')+M_1 \parallel (x,z)-(x',z') \parallel . \end{aligned}$$

Lastly, Using that $y' \in \Gamma(x',z')$ we obtain

$$| Tv(x,z) - Tv(x',z') | \leq M_1 \parallel (x,z) - (x',z') \parallel . \ \blacksquare$$

This lemma shows that the Lipschitz constants M_k of the iterates $T^k v$ follow the difference equation

$$M_k = K_R(1 + K_\Gamma) + \max\{1, M_{k-1}\}\beta K_Q + M_{k-1}\beta K_\Gamma. \tag{9}$$

The result to be proved in this section is the following

Theorem 14 *: Let $C \subset S$ be a compact set and assume that*

$$G(x,z) \subset int(\Gamma(x,z)) \text{ for all } (x,z) \in C. \tag{10}$$

Let Assumptions $I - IV$ hold. Then $V \in L_C^{loc}(\alpha, K)$, with $K = \max\{1, K_R(1 - \beta K_Q)^{-1}\}$. Let $w \in BL_S(\alpha, M_0)$, for some given constant M_0, and let $\gamma > 0$. Then there exists a $j_0(\gamma)$ such that $T^j w \in BL_C^{loc}(\alpha, K + \gamma)$, all $j > j_0(\gamma)$. If C is an $L-$convex set, then $V \in BL_C(\alpha, \Lambda_C K)$ and $T^j w \in BL_C(\alpha, \Lambda_C(K + \gamma))$, all $j > j_0(\gamma)$, with Λ_C given by (2).

If we interpret the variable x as the stock level of some economic resource, the assumption (10) means that neither the exhaustion of the resource nor a null consumption will be optimal at any period. This is the relevant situation in problems of economic growth, if the extinction of the economy is ruled out, and in problems of exploitation of renewable resources, if the extinction of the resource and the paralysis of the exploitation for a period is too costly.

In order to prove Theorem 14, we first ensure that the correspondence G_v (see definition in Section 2.3) satisfies on C the condition required to G in the statement of the above theorem if v is a small perturbation of V.

Lemma 15 *: i) If BC_S is endowed with the supremum norm topology and $BC_S \times S$ is endowed with the product topology, then the correspondence $G^* : BC_S \times S \to X$ defined by $G^*(v, c) = G_v(c)$ is upper hemi-continuous.*

ii) There exists an open ball $U(V)$, centered at V, of the normed space BC_S, such that $G_v(x, z) \subset int(\Gamma(x, z))$ holds if $(x, z) \in C$ and $v \in U(V)$.

Proof. *i) By Theorem of Maximum, it is enough to check that $h(v, x, y, z) := R(x, y, z) + \beta \int v(y, \theta)Q(z, d\theta)$ is a continuous function which, in view of the continuity of R, reduces to check the continuity of the integral term. Let $(v_k, y_k, z_k) \in BC_S \times S$, all k, with $(v_k, y_k, z_k) \to (v, y, z) \in BC_S \times S$. We may write*

$$
\begin{aligned}
& \left| \int v(y,\theta)Q(z,d\theta) - \int v_k(y_k,\theta)Q(z_k,d\theta) \right| \leq \\
\leq & \left| \int v(y,\theta)Q(z,d\theta) - \int v(y_k,\theta)Q(z,d\theta) \right| + \\
+ & \left| \int v(y_k,\theta)Q(z,d\theta) - \int v(y_k,\theta)Q(z_k,d\theta) \right| + \\
+ & \left| \int v(y_k,\theta)Q(z_k,d\theta) - \int v_k(y_k,\theta)Q(z_k,d\theta) \right| .
\end{aligned}
$$

The third summand in the string tends to zero as k tends to infinity because $v_k \to v$ in the supremum norm; the second summand tends to zero because Q enjoys Feller's property; and the first summand tends to zero by Lebesgue theorem of dominated convergence, since the sequence of functions $v(y_k, \cdot)$ converges to $v(y, \cdot)$ and $\| v(y_k, \cdot) \| \leq \| v \| < \infty$.

ii) Assume, by contradiction, that for all $U(V)$, all $(x, z) \in C$ and all $v \in U(V)$, $\partial\Gamma(x, z) \cap G_v(x, z) \neq \varnothing$. Then, there exists a sequence $v_k \to V$ and sequences $\{(x_k, z_k)\}$ in C and $\{y_k\}$ in X with $y_k \in \partial\Gamma(x_k, z_k) \cap G_{v_k}(x_k, z_k) \neq \varnothing$. By compactness we may assume $(x_k, z_k) \to (x, z) \in C$. Using that Γ is topo-

logically continuous, so $D_H(\partial\Gamma(x_k, z_k), \partial\Gamma(x, z)) \to 0$, *and* $y_k \in \partial\Gamma((x_k, z_k))$, *we see that* $d(y_k, \partial\Gamma(x, z)) \to 0$. *This means that for all* k *there exists an* $y_k^* \in \partial\Gamma(x, z)$ *such* $d(y_k, y_k^*) \to 0$. *By compactness we may assume* $y_k^* \to y \in \partial\Gamma(x, z)$, *so we may assume* $y_k \to y \in \partial\Gamma(x, z)$. *But this is a contradiction, since from the upper hemi-continuity of* G^* *and from* $(v_k, x_k, z_k) \to (V, x, z)$ *it follows* $y \in G(x, z) \subset int(\Gamma(x, z))$. ■

Next we state an elementary property that provides the basic tool for the proof of Theorem 14.

Lemma 16 *: Under Assumptions* $I - IV$ *let* $v \in BL_S(\alpha, M)$, *and assume that* $Tv(x, z) \geq Tv(x', z')$ *holds, where* $(x', z'), (x, z) \in S$ *are such that* $G_v(x, z) \cap \Gamma(x', z') \neq \emptyset$. *Then* $\mid Tv(x, z) - Tv(x', z') \mid \leq (K_R + \max\{1, M\}\beta K_Q) \parallel (x, z) - (x', z') \parallel$.

Proof. *Let* $y \in G_v(x, z) \cap \Gamma(x', z')$. *We have*

$$\begin{aligned}
\mid & \quad Tv(x, z) - Tv(x', z') \mid = Tv(x, z) - Tv(x', z') = \\
= & \quad R(x, y, z) + \beta \int v(y, \theta) Q(z, d\theta) - Tv(x', z') \leq \\
\leq & \quad R(x', y, z') + K_R \parallel (x, z) - (x', z') \parallel + \beta \int v(y, \theta) Q(z', d\theta) + \\
+ \max\{1, M\}\beta K_Q \parallel & \quad z - z' \parallel - Tv(x', z') \leq \\
\leq & \quad (K_R + \max\{1, M\}\beta K_Q) \parallel (x, z) - (x', z') \parallel,
\end{aligned}$$

where the first inequality holds because $R \in L_\Omega(K_R)$ *and using Lemma 9, and the second inequality holds because* $y \in \Gamma(x', z')$. ■

Using this lemma we first analyze the local action of T. This shall give a recursive law for the Lipschitz constants of iterates under Bellman operator of functions in $U(V)$ which is as in (9) without the constant term $K_\Gamma K_R$ and the (potentially) growing term $M_{k-1}\beta K_\Gamma$. The mutual factibility condition (see Lemma 5 (ii)) guarantees that Lemma 16 can be applied to bounded Lipschitz functions in $U(V)$.

Lemma 17 *: Assume that $v \in U(V) \cap BL_S(\alpha, M)$ for some constant M, where $U(V)$ is as in part ii) of Lemma 15 and let $c \in C$. Then there exists an $\varepsilon > 0$ such that $Tv \in BL(\alpha, K)$ on $U(c, \varepsilon)$, with $K = K_R + \max\{1, M\}\beta K_Q$.*

Proof. *By Lemma 15 we know that $G_v(c) \subset int(\Gamma(c))$ holds for $c \in C$. By Lemma 5 applied to the correspondences Γ and G_v, we know that there exists an open ball $U(c, \varepsilon)$ such that if $c' \in U(c, \varepsilon)$, then $G_v(c') \subset \Gamma(c)$ and $G_v(c) \subset \Gamma(c')$ hold. If $Tv(c) \geq Tv(c')$, then $G_v(c) \subset \Gamma(c')$ and Lemma 16 give $\mid Tv(c) - Tv(c') \mid \leq K \parallel c - c' \parallel$, and we arrive at the same conclusion if $Tv(c) \leq Tv(c')$ using then that $G_v(c') \subset \Gamma(c)$.* ■

We have now all the ingredients needed in the proof of Theorem 14.

Proof. *Let $U(V)$ be as in Lemma 15 and let $w \in BL_S(\alpha, M_0)$ where M_0 is some constant. Then, as $T^k w \to V$, $T^{k_0-1}w \in U(V)$ holds if k_0 is large enough. Hence $T^{k_0}w \in U(V) \cap BL_S(\alpha, M_{k_0})$, with M_{k_0} as given by Lemma 13. Notice that, since T is a contractive operator, $U(V)$ is invariant under T. Reset M_0 equal to the constant M_{k_0}, let $c \in C$, and let $U(c, \varepsilon(c))$ be an open ball as that given by Lemma 17. We see that $T^{k_0+k}w \in BL(\alpha, M_k)$ on $U(c, \varepsilon(c))$, $k = 1, 2, ..,$with M_k following now the difference equation*

$$M_k = K_R + \max\{1, M_{k-1}\}\beta K_Q. \tag{11}$$

If $K := K_R(1-\beta K_Q)^{-1} > 1$ *and* $M_0 \leq 1$, *then* $M_1 = K_R + \beta K_Q$, *and since* $K_R > 1 - \beta K_Q$, *we get that* $M_1 > 1$. *Therefore any solution of equation (11) follows in turn, for* $k \geq 2$, *the difference equation*

$$M_k = K_R + M_{k-1}\beta K_Q. \tag{12}$$

which converges to its unique equilibrium point K *so, there is an integer* $k_1(\gamma)$ *such that*

$$T^{k_0+k}w \in U(V) \cap BL(\alpha, K+\gamma), k \geq k_1(\gamma) \tag{13}$$

on $U(c, \varepsilon(c))$. *If* $K = 1$ *(so* $K_R(1-\beta K_Q)^{-1} \leq 1$*) then, whenever a solution of (11) remains larger than* 1, *it follows the difference equation (12), which converges to its unique equilibrium point* $K_R(1-\beta K_Q)^{-1} \leq 1$. *If, for some* k, *we have* $M_k \leq 1$ *(which necessarily occurs in a finite number of steps if* $K_R(1-\beta K_Q)^{-1} < 1$, *see note 4), then* $M_{k+1} = K_R + \beta K_Q \leq 1$, *because now* $K_R \leq 1 - \beta K_Q$. *Thus (13) always holds.*

Set $j_0(\gamma) = k_0 + k_1(\gamma)$. *Then* $T^j w \in BL_{U(c,\varepsilon(c))}(\alpha, K+\gamma)$ *for* $j \geq j_0(\gamma)$, *which shows* $T^j w \in BL_C^{loc}(\alpha, K+\gamma)$, $j \geq j_0(\gamma)$, *and if* C *is a* $L-$*convex,* $T^j w \in BL_C(\alpha, \Lambda_C(K+\gamma))$, $j \geq j_0(\gamma)$ *by Lemma 2.*

Since the set of functions $BL_{U(c,\varepsilon(c))}(\alpha, K+\gamma)$ *is a closed set of functions and* $T^j w \in BL_{U(c,\varepsilon(c))}(\alpha, K+\gamma)$, $j \geq j_0(\gamma)$, *we see that* $V \in BL_{U(c,\varepsilon(c))}(\alpha, K+\gamma)$ *(see note 5). This shows that* $V \in L_C^{loc}(\alpha, K+\gamma)$, *and as* γ *was arbitrarily small, we get* $V \in L_C^{loc}(\alpha, K)$. *If* C *is* $L-$*convex, Lemma 2 gives* $V \in L_C(\alpha, \Lambda_C K)$. ∎

4 Examples

All data in the examples below were generated using a AXP-2100/AMP500 DIGITAL Computer, coded in standard FORTRAN 77.

4.1 *Deterministic example*

Example 18 ***An application of Theorem 14: Non-Concavity in Growth Models***

The optimization problem is

$$\max_{\{x_{t+1}\}_{t=0}^{\infty}} \left\{ \sum_{t=0}^{\infty} \beta^t U(f(x_t) - x_{t+1}) : \ 0 \leq x_{t+1} \leq f(x_t), \ t = 0, 1., 2... \right\}, \qquad (14)$$

where $\beta \in (0,1)$ is a discount factor; x_t is the capital stock at period t; $U(c_t)$ is the utility function of a private owner of a firm or household, where $c_t = f(x_t) - x_{t+1}$ is the consumption at period t; and $f(x_t)$ is a production function. Note that (14) is the classical problem of optimal growth in a one-sector model.

The Bellman equation associated to (14) is in this case

$$V(x) = \max_{0 \leq y \leq f(x)} \{U(f(x) - y) + \beta V(y)\}. \qquad (15)$$

We have analyzed the case of the utility function $U(x) = x^{0.7} + x^3 - x$ which is increasing, concave for low consumption levels, and it possesses an interval of convexity for high consumption levels. The numerical experiment shows that the optimal behaviour implied by such utility is rather common at a microeconomic level (see Rothenberg [28], and Arrow et al. [1], for theoretical justification of nonconcave utility functions in microeconomic problems). It is cyclical: It maintains low consumptions levels for some periods of time, and then makes a large consumption. The production function considered is $f(x) = 0.95x \exp(1 - x)$. It is concave and with strictly positive marginal productivity in the interval $[0, x_e]$, with $f(x_e) = x_e = 1 + \ln 0.95 < 1$, which is the relevant domain of f (see fig. 1-b).

In figs. 1-a and 1-b we can see the value function solution of (15) and the associated optimal policy correspondence respectively. There are countably many discontinuities in the optimal policy following a visible pattern synchronized with jumps upwards in the marginal value of capital stock. In fig. 1-b the production function is also plotted. There is not extinction of the economy, since all solutions are interior. In particular, the optimal capital stock satisfies $x_{t+1} \geq x' \simeq 0,03$ if $x_t \geq x'$. Since $x_{t+1} > x_t$ if $x_t \leq x'$, Theorem 14 applies to this example taking $Y = [x, 1]$, with $0 < x \leq x'$, as state space of the problem (see Remark 11). Fig. 1-b reveals that one strongly attractive period-six cycle, supports the long run behavior of the optimal paths. Research in progress is addressed to the analytical characterization of the threshold point, beyond which the main part of the savings are expended in a large consumption.

4.2 *An stochastic example*

Example 19 ***: Stochastic Optimal Growth***

We analyze here the Example (18) in a stochastic setting. Randomness enters in the problem through a multiplicative random shock which modifies the production function f, reflecting, for instance, the action of a exogenous shock which affects negatively to the production function. The intensity z_t of the multiplicative shock at period t is described by a stochastic i.i.d. process $\{z_n\}$ where $z_n = 0,5{+}0,5z'_n$ with z'_n distributed as a $\beta(0.5, 0.5)$. The output at period t corresponding to a resource level x is given by $z_t f(x)$. The Bellman equation is written now $V(x,z) = \max_{0 \leq y \leq zf(x)}\{U(zf(x) - y) + \beta \int_Z V(y,\theta)d\mu(\theta)\}$, where $Z = [0.5, 1]$ is the support of the probability distribution μ of z_n. Fig. 2 shows the value function of the problem. Notice that the value function is not a Lipschitz function on $X \times Z$ since the marginal return of capital stock tends to infinity when the stock tends to zero. The numerical analysis reveals that there exists a minimal capital stock $\overline{x} = 0.003$ such that $\overline{G}(x) \geq \overline{x}$ if $x \geq \overline{x}$, so we may use $Y \times Z$, with $Y := [\overline{x}, 1]$, as state space, as indicated in Remark 11. There is no extinction in this economy, as it can be checked that all optimal paths from a endogenous state $x_0 \geq \overline{x}$ are always interior, so Theorem 14 applies to the set $Y \times Z$.

If we compare the value function in this stochastic example with that in the deterministic case, Example 18, we can see that the uncertainty derived from the random shock produces an smoothing effect on the value function. Further numerical analysis reveals discontinuities in the $z-$sections of the optimal policy correspondence, which should cause non-smoothness of the value function.

5 Concluding remarks

There are three directions for the future extension of this research:

1) In Morán and Maroto [26], we analyze the convergence of the algorithm for the numerical computation of the value function. The Lipschitz continuity of the value function is the only assumption that we need in order to obtain a rate $O(\delta)$ of convergence of the numerical algorithm, where δ is the diameter of the discrete grid of points used for the computation.

2) The results in this paper for the case of interior optimal plans allow us to prove the Lipschitz continuity of the value function for the case of eventually interior optimal plans. This case is relevant to the area of optimal exploitation of renewable resources where the existence of nonconcavities has been largely admitted.

3) As a future perspective, the results in this paper allow us the use of powerful tools of non-smooth analysis (Clarke [9]). In the same spirit, an incipient development of a theory of dynamical systems with evolution law governed by correspondences (Lasota and Myjak [19, 20]), closely related to fractal geometry (Hutchinson [18]), open the perspective of the analysis of the long run behaviour of optimal policies and asymptotic stability of dynamical equilibria if nonuniqueness in the policy correspondence is allowed, so in the deterministic as in the stochastic settings. For recent research on chaotic behaviour of optimal paths, see also Majumdar et al. [24].

References

[1] K.J. Arrow, N. Yew-Kwang, X. Yang, Increasing returns and economic analysis, Macmillan Press, London, 1998.

[2] M. Bardi, I. Capuzzo-Dolcetta, Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations, Birkhäuser, Boston, 1997.

[3] D. P. Bertsekas, Convergence of Discretization Procedures in Dynamic Programming, IEEE Transactions on Automatic Control. AC-20 (1975) 415-419.

[4] P. Billingsley, Probability and Measure, John Wiley & Sons, New York, 1995.

[5] T. Bj$\emptyset$rndal, Production in a Schooling Fishery: The case of the North Sea Herring Fishery, Land Economics. 65 (1988) 49-56.

[6] T. Bj$\emptyset$rndal, J.M. Conrad, The dynamics of an open access fishery, Canadian journal of Economics. 20 (1987) 74-85.

[7] W. Brian, Increasing returns and path dependence in the economy, University of Michigan Press, 1994.

[8] C.W. Clark, Economically optimal policies for the utilization of biologically renewable resources, Mathematical Biosciences. 12 (1971) 245-260.

[9] F.H. Clarke, Optimization and Nonsmooth Analysis, Vol. 5 of Classics in Applied Mathematics, SIAM, Philadelphia, 1990.

[10] P. Dasgupta, K.-G. Mäler, Special issue: The economics of non-convex ecosystems, Environmental & Resource Economics. 26 (2003).

[11] H. Dawid, M. Kopel, On the economically optimal exploitation of a renewable resource: the case of a convex environment and a convex return function, Journal of Economic Theory. 76 (1997) 272-297.

[12] H. Dawid, M. Kopel, On optimal cycles in dynamic programming models with convex return function, Economic Theory. 13 (1999) 309-327.

[13] W.D. Dechert, K. Nishimura, A complete characterization of optimal growth paths in an aggregated model with a non-concave production function, Journal of Economic Theory. 31 (1983) 332-354.

[14] R.M. Dudley, Probabilities and metrics, Lecture Notes Series 45, Aarhus Universitet, 1976.

[15] R.F. Hartl, P.M. Kort, Optimal investments with increasing returns to scale: a further analysis, in: E.J. Dockner, R.F. Hartl, M. Luptacik, G. Sorger (Eds.), Optimization, dynamics, and economic analysis: essays in honor of Gustav Feichtinger, Physica-Verlag, New York, 2000, pp. 226-238.

[16] R.F. Hartl, P.M. Kort, Optimal investments with convex-concave revenue: a focus-node distinction, Optimal Control Applications & Methods. 25 (2004) 147-163.

[17] G. Heal, The economics of increasing returns, Edward Elgar, United Kindong, 1999.

[18] J.E. Hutchinson, Fractals and self-similarity, Indiana University Mathematical journal. 30 (1981) 713-747.

[19] A. Lasota, J. Myjak, Markov Operator and fractals, Bulletin of the polish Academy of Science. 45 (1997) 197-207.

[20] A. Lasota, J. Myjak, Attractors of Multifunctions, Bulletin of the polish Academy of Science. 48 (2000) 1-44.

[21] C. Le Van, R.-A. Dana, Dynamic Programming in Economics, Kluwer Academic Publishers, Netherlands, 2003.

[22] M. Majumdar, T. Mitra, On optimal exploitation of a renewable resource in a nonconvex environment and the minimum safe standard of conservation, Working paper No. 223, Cornell University, 1980.

[23] M. Majumdar, T. Mitra, Intertemporal allocation with a non-convex technology: the aggregative framework, Journal of Economic Theory. 27 (1982) 101-136.

[24] M. Majumdar, T. Mitra, K. Nishimura, Optimization and Chaos, Springer, Berlin, 2000.

[25] L. Montrucchio, Lipschitz Continuous Policy Functions for Strongly Concave Optimization Problems, Journal of Mathematical Economics. 16 (1987) 259-273.

[26] M. Morán, J.M. Maroto, Convergence in Dynamic Programming with Discount, Working paper, Universidad Complutense, 2003.

[27] M. Morán, J.M. Maroto, Lipschitz continuous dynamic programming with discount: The case of eventually interior optimal plans. Manuscript in preparation.

[28] J. Rothenberg, Non-convexity, agregation, and Pareto optimality, The Journal of Political Economy. 68 (1960) 435-468.

[29] N.L. Stokey, R. Lucas, E. Prescott, Recursive methods in economic dynamics, Harvard University Press, Cambridge, 1989.

[30] R. Yue, Lipschitz Continuity of Value Functions in a Class of Optimal Controls Problems Governed by Ordinary differential Equations, in: S. Chen, J. Yong (Eds.), Control Theory, Stochastic Analysis and Applications, World Scientific, Singapure, 1991, pp. 125-137.

[31] R. Yue, Properties of the Bellman Function in Time-Optimal Control Problems, Journal of Optimization Theory and Applications. 94 (1997) 155-168.

Notes

1.- $d(f(x), f(y)) \leq Kd(x, y)$, which shows that any Lipschitz constant for f^{-1} must be larger than K^{-1}. This shows that $K_f K_{f^{-1}} \geq KK^{-1} = 1$.

2.- The sets $B(y, \frac{\varrho}{2}) \cap adh(\Gamma(c))$ and $B(y, \frac{\varrho}{2}) \cap \Gamma(c)^c$ are non-empty closed sets whose union completes the connected set $B(y, \frac{\varrho}{2})$. We may write $B(y, \frac{\varrho}{2}) \cap adh(\Gamma(c)) = (B(y, \frac{\varrho}{2}) \cap \Gamma(c)) \cup (B(y, \frac{\varrho}{2}) \cap \partial(\Gamma(c)))$. If $B(y, \frac{\varrho}{2}) \cap \partial(\Gamma(c)) = \emptyset$ then $(B(y, \frac{\varrho}{2}) \cap adh(\Gamma(c)) \cap (B(y, \frac{\varrho}{2}) \cap \Gamma(c)^c) = \emptyset$, in contradiction with the connectedness of $B(y, \frac{\varrho}{2})$.

3.- Let $I = [0, a]$ and $J = [0, b]$. Then $D_H(I, J) = d(a, b)$. For the correspondences described in the text we have $\Gamma(x) = [0, f(x)]$. Therefore, if $f \in L(K)$ we have $D_H(\Gamma(x), \Gamma(y)) = d(f(x), f(y)) \leq Kd(x, y)$.

4.- If $K_R(1 - \beta K_Q)^{-1} = 1$ then M_k does not have to become smaller that 1 in a finite number of steps but, for any $\gamma > 0$, it becomes smaller than $1 + \gamma$ in a finite number of steps, as required.

5.- Let $v_n \to v$ with $v_n \in L_U(K + \gamma)$. Then, given $x, y \in U$ and $\varepsilon > 0$ there exists an n such that $d(v_n(x), v(x)) \leq \varepsilon$ and $d(v_n(y), v(y)) \leq \varepsilon$. Therefore, for $x, y \in U$,

$$d(v(x), v(y)) \leq 2\varepsilon + d(v_n(x), v_n(y)) \leq 2\varepsilon + (K + \gamma)d(x, y),$$

and, since ε was arbitrary, we get $v \in L_U(K + \gamma)$.

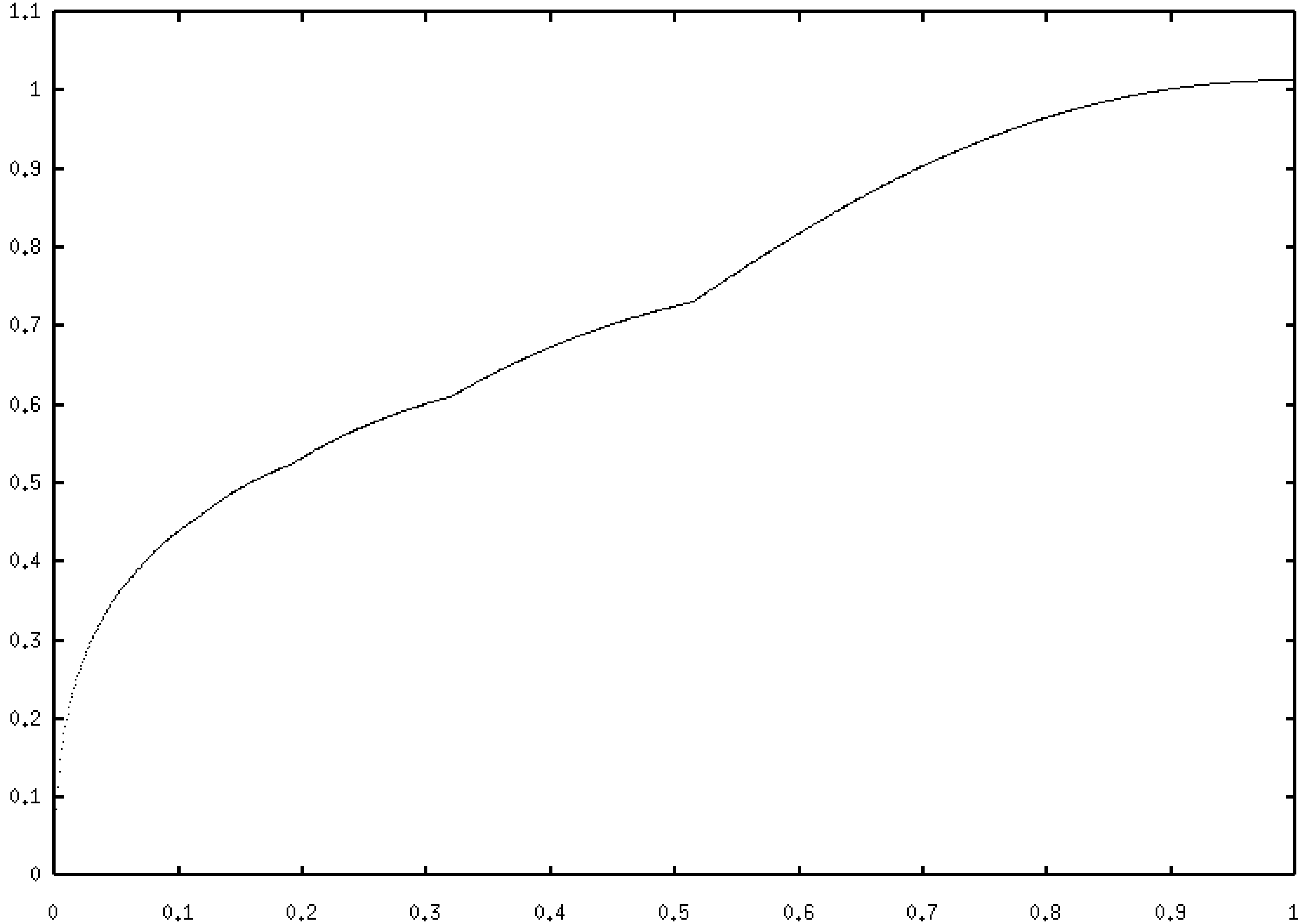

Figure 1-a. Value function.

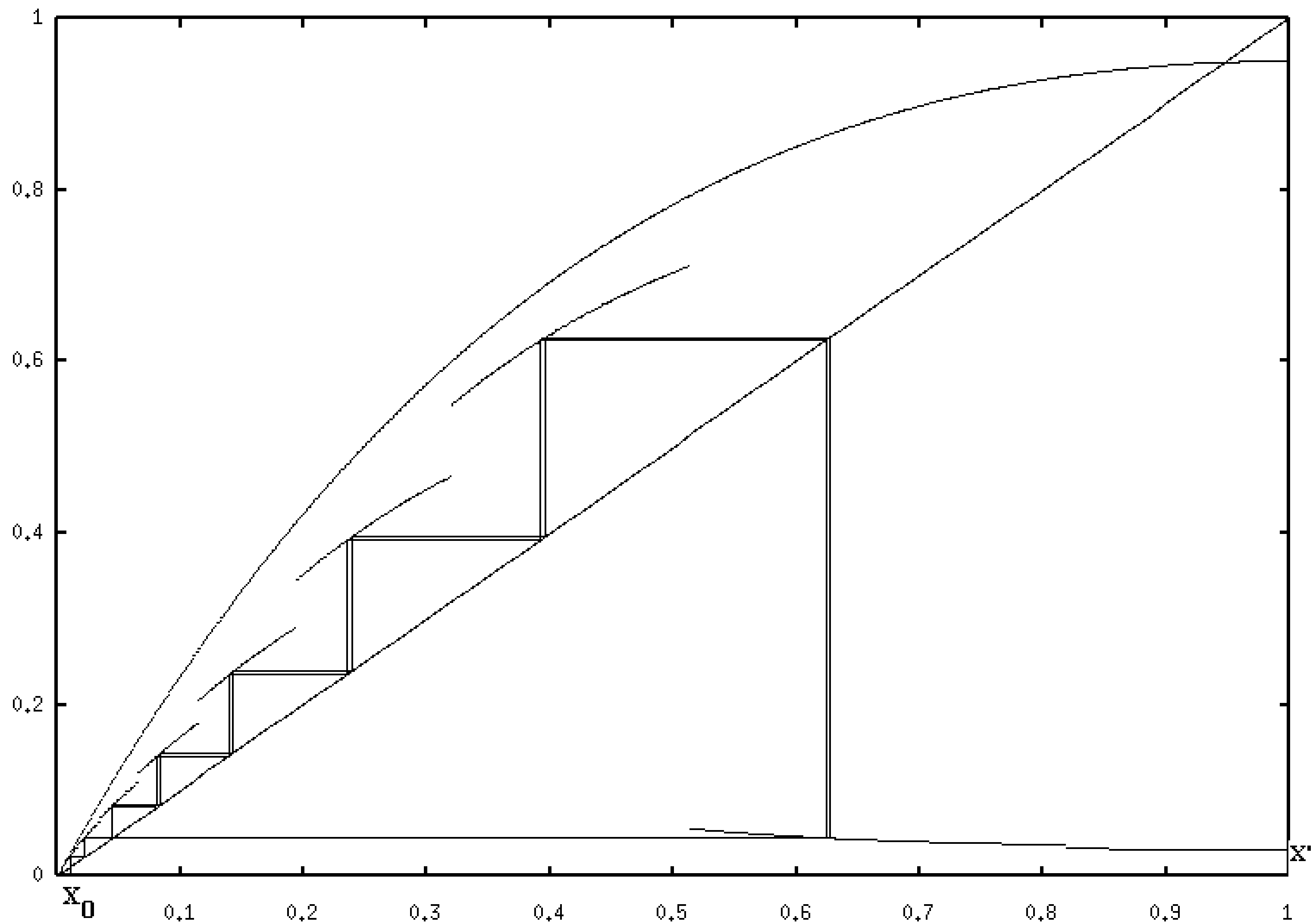

Figure 1-b. Production function, optimal policy correspondence and optimal dynamics.

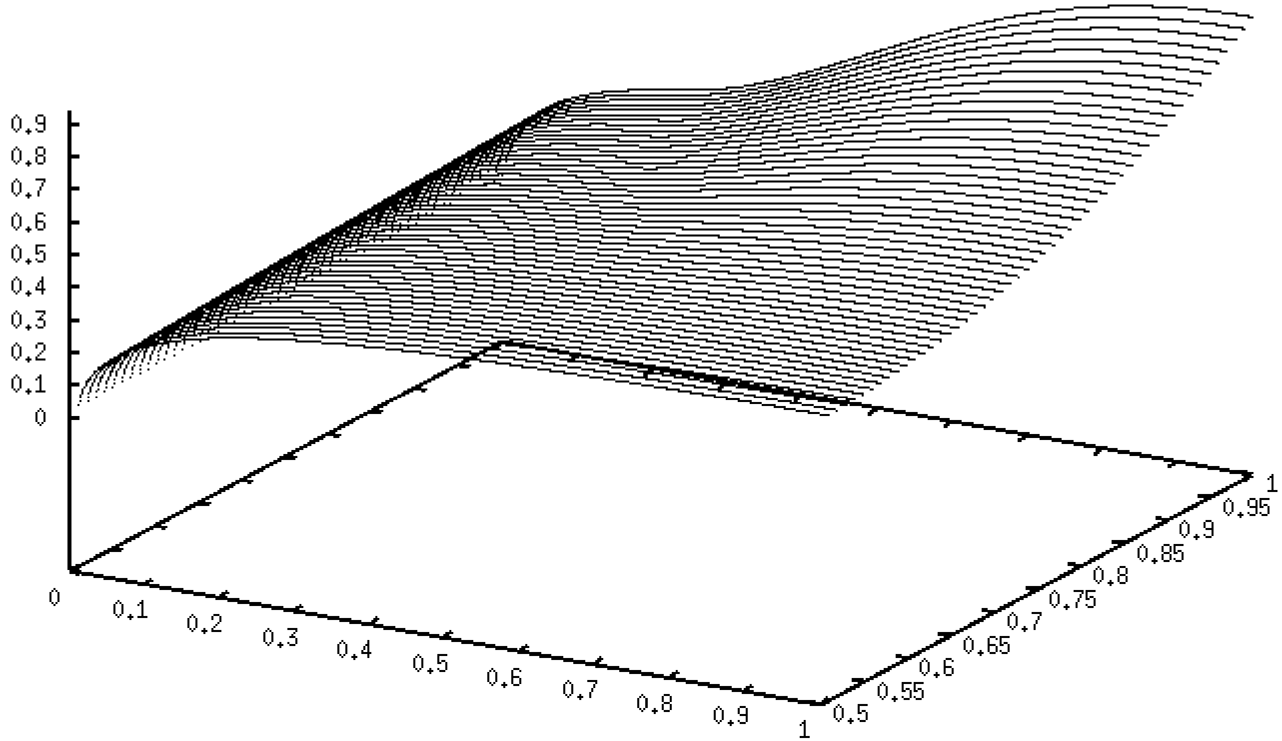

Figure 2. Value function.