Smooth Value and Policy Functions for Discounted Dynamic Programming

Andrew McLennan
University of Queensland*

June 13, 2015

Abstract

We consider a discounted dynamic program in which the spaces of states and actions are smooth (in a sense that is suitable for the problem at hand) manifolds. We give conditions that insure that the optimal policy and the value function are smooth functions of the state when the discount factor is small. In addition, these functions vary in a Lipschitz manner as the reward function-discount factor pair varies in a neighborhood of the pair consisting of a given reward function and zero.

Running Title: Smooth Value and Policy Functions

Keywords: Dynamic programming, stochastic optimal control, smooth value function, smooth policy function, Lipschitz functions.

1 Introduction

For optimization over a finite dimensional domain it is fundamental that, so long as the second order conditions hold strictly, the optimal action and the value of the problem vary in a smooth manner as one varies underlying parameters. With technical qualifications, this principle extends to optimal control, among other infinite dimensional settings. (The theory of minimal surfaces provides another example.) Furthermore, in optimal control the optimal policy and the value of the problem, as a function of the boundary conditions, are smooth functions, provided suitable technical conditions are satisfied.

*Email: a.mclennan@economics.uq.edu.au. This work was funded in part by Australian Research Council grant DP0773324. The hospitality of the Kyoto Institute for Economic Research is gratefully acknowledged.

This paper provides results of this sort for discounted dynamic programs. We are given spaces of states and actions that are smooth manifolds. The space of states is compact, and there is a correspondence mapping states into compact sets of feasible actions, that is continuous, so that its graph is compact. A feasible state-action pair results in an immediate reward, which is a smooth function of the pair, and a stochastic transition to the state in the next period, which will be assumed to be smooth in a suitable sense. Given an initial state and a discount factor, the problem is to maximize the expectation of the infinite sum of discounted rewards. There will always exist optimal policies that are stationary, in the sense that the choice of action depends only on the current state, independent of the date or prior history, and we will consider only such policies.

Suppose that for each state, the payoff function has a unique maximizing action, this state-action pair is in the interior of the graph of the feasibility correspondence, and the second order conditions for optimization hold strictly. Then the implicit function theorem implies that the payoff maximizing action, and the maximal payoff, are smooth functions of the state. These functions may be interpreted as the optimal policy and the value function of the myopic dynamic program corresponding to a discount factor of zero. This paper's main result gives conditions under which the optimal policy and the value function are smooth for reward function-discount factor pairs in some neighborhood of the pair consisting the given reward function and zero. Moreover, the optimal policy and the value function are continuous (in fact Lipschitz, relative to a metric for the relevant C^r topology) functions of these parameters.

One style of application of this result is to show that certain phenomena are possible with positive discount factors. As it happens, work on this paper began in the context of work on a learning problem described in the companion paper Loertscher and McLennan (2015). In the model studied in that paper, myopically optimal behavior can give rise to "learning traps" in which, with positive probability, beliefs converge to a point at which the myopically optimal action is uninformative, so learning ceases. Applied to that setting, the result of this paper contributes to an argument showing that learning traps are possible for small positive discount factors.

Smoothness of the policy and value functions for dynamic programs is a more difficult issue than naive intuition might suggest. Blume et al. (1982) study a somewhat different stochastic dynamic programming model, achieving C^r value functions and policy functions by means of the implicit function theorem for Banach spaces. For a model of capital accumulation the issue of smoothness of the policy and value functions was studied in Benveniste and Scheinkman (1979), Araujo (1991), and Santos (1991). On the one hand the model studied in those papers is less general, insofar as it is deterministic, and the set of actions available today is in effect a subset of tomorrow's state space that has nonempty interior. On the other hand our methods are restricted to discount factors near zero; Boldrin and Montrucchio (1986) study the capital accumulation model for discount factors in a neighborhood of zero, establishing the possibility of chaotic dynamics.

Our approach employs the methods associated in particular with Blackwell (1965) in which the value function is the fixed point of a contraction mapping that computes the value of the problem today for the given reward function and a value function tomorrow. The main source of technical difficulty is that we must arrange for this operator to be a Lipschitz function whose domain is complete. (Once it is Lipschitz, restricting to small discount factors makes it a contraction.) We do this by restricting to C^r functions whose r^{th} order partial derivatives are Lipschitz, and considering operators between spaces of such functions that are locally Lipschitz, and that map sets of functions satisfying a uniform Lipschitz condition to such sets. Another informal overview of our methods is given at the end of Section 3.

The remainder has the following organization. The next section defines fundamental concepts, of which the most novel is tameness: an operator taking functions to functions is tame if it is continuous, its restriction to any uniformly bounded and uniformly Lipschitz set of functions is Lipschitz, and the image of any such set is another set with these properties. Section 3 reviews fundamental facts concerning dynamic programming and states the paper's main result. As with other concepts of analysis, tameness has many simple and basic properties which are provided in Section 4. Section 5 provides the key technical results, which show that operators passing to implicit functions, and operators related to maximization, are tame. Section 6 completes the proof of our main result.

2 Basic Definitions

2.1 Locally Lipschitz Functions

Let X and Y be metric spaces. The space of continuous functions from X to Y is denoted by $C(X, Y)$. We write $C(X)$ in place of $C(X, \mathbb{R})$. Recall that a function $f : X \to Y$ is *Lipschitz* for the *Lipschitz constant* $\Lambda > 0$, or simply Λ-*Lipschitz*, if

$$d(f(x), f(x')) \leq \Lambda d(x, x')$$

for all $x, x' \in X$. (There will be many metric spaces, and when there is little potential for confusion we will denote a metric simply by d, with the space to be inferred from context.) Let $C_\Lambda(X,Y)$ be the set of such f. We say that f is *Lipschitz* if it is Lipschitz for some Lipschitz constant, and it is *locally Lipschitz* if each point has a neighborhood U such that $f|_U$ is Lipschitz.

We now assume that X is compact. There is a natural metric on $C(X,Y)$:

$$d(f,f') = \max_{x\in X} d(f(x), f'(x)).$$

Consider a set $S \subset C(X,Y)$. If $K \subset Y$ and $\Lambda \geq 0$, S is (K,Λ)-*Lipschitz bounded* if $S \subset C_\Lambda(X,Y)$ and the image of each $f \in S$ is contained in K. The set S is *compactly Lipschitz bounded* if it is (K,Λ)-Lipschitz bounded for some compact K and $\Lambda \geq 0$.

Let X' and Y' be two more metric spaces with X' compact.

Definition 1. *An operator $\gamma : S \to C(X',Y')$ is* tame *if it is continuous and for every compactly Lipschitz bounded $T \subset S$, $\gamma|_T$ is Lipschitz and $\gamma(T)$ is compactly Lipschitz bounded. If S is covered by (relatively) open sets U such that $\gamma|_U$ is tame, then γ is* locally tame.

2.2 Spaces of Smooth Functions

Fix a degree of differentiability r with $2 \leq r \leq \infty$. (A number of definitions and results are valid for $r = 0, 1$, but this will mostly be obvious, and unimportant in relation to our main goal.) Recall that for any $D \subset \mathbb{R}^m$, a function $f : D \to \mathbb{R}^n$ is C^r, by definition, if it has a C^r extension to an open superset of D. We say that $D \subset \mathbb{R}^m$ is a *differentiation domain* if, for any $f \in C^1(D)$ and any C^1 extensions $f', f'' : U \to \mathbb{R}$ of f to an open $U \supset D$, f' and f'' have the same partial derivatives at every point of D. For example, D is a differentiation domain if it is the closure of its interior. Of course the main point is that a C^1 function on D has well defined partial derivatives at each point of D.

If D is a differentiation domain and $U \subset \mathbb{R}^m$ is open, then $D \cap U$ is a differentiation domain. (If a C^1 function on $D \cap U$ had extensions to U with different partials at some point, the methods of Section 2.2 of Hirsch (1976) could be used to produce two C^1 functions that agreed on $D \cap U$, agreed with the extensions near the point in question, and vanished outside of some compact neighborhood $K \subset U$ of that point, so that they could be understood as extensions of a C^1 function on D to all of $\mathbb{R}^m$.) By continuity, the closure of $D \cap U$ in D is also a differentiation domain. Consequently any differentiation domain can be covered by a collection of relatively open or closed subsets of arbitrarily small diameter that are differentiation domains.

For the rest of this subsection we work with a fixed differentiation domain D. For each $f \in C^r(D, \mathbb{R}^n)$ there is a corresponding $\Psi^r(f) \in C(D, \mathbb{R}^N)$ that combines the components of f and all partial derivatives up to order r, where N is the number of such components and partials. For $\Lambda > 0$, we say that $f \in C^r(D, \mathbb{R}^n)$ is C^r *Λ-Lipschitz* if $\Psi^r(f)$ is Λ-Lipschitz. We say that $f \in C^r(D, \mathbb{R}^n)$ is C^{r+} if each $x \in D$ has a neighborhood U such that $f|_{D\cap U} \in C^r_\Lambda(D \cap U, \mathbb{R}^n)$ for some Λ. For $T \subset \mathbb{R}^n$ let $C^r_\Lambda(D, \mathbb{R}^n)$ be the set of C^r Λ-Lipschitz functions $f : D \to T$, and let $C^r(D, T)$ and $C^{r+}(D, T)$ be the sets of C^r and C^{r+} functions $f : D \to T$. We write $C^r(D)$ and $C^{r+}(D)$ in place of $C^r(D, \mathbb{R})$ and $C^{r+}(D, \mathbb{R})$.

Lemma 1. *If $U \subset \mathbb{R}^m$ and $V \subset \mathbb{R}^n$ are open and $f : U \to V$ and $g : V \to \mathbb{R}^p$ are C^{r+}, then $g \circ f$ is C^{r+}.*

Proof. Of course $g \circ f$ is C^r. By repeated applications of the chain rule, each r^{th} order partial of $g \circ f$ can be written as a polynomial function of terms of the form Q and $R \circ f$, where Q is a partial of f and R is a partial of g. The partials Q are locally Lipschitz by assumption, and the terms $R \circ f$ are locally Lipschitz because they are compositions of locally Lipschitz functions. Finally, sums and products of locally Lipschitz functions are locally Lipschitz. □

Now assume that D is compact. We endow $C^r(D, \mathbb{R}^n)$ with the metric d^r induced by Ψ^r: $d^r(f, f') = d(\Psi^r(f), \Psi^r(f'))$. The associated topology is called the *C^r topology.*

Consider a set of functions $S \subset C^{r+}(D, \mathbb{R}^n)$. We say that S is *compactly C^r Lipschitz bounded* if $\Psi^r(S)$ is compactly Lipschitz bounded.

Let $D' \subset \mathbb{R}^{m'}$ be another compact differentiation domain, and let $0 \leq s \leq \infty$ be another order of differentiability.

Definition 2. *An operator $\gamma : S \to C^s(D', \mathbb{R}^{m'})$ is* (r, s)-tame *if it is continuous and for every compactly C^r Lipschitz bounded $T \subset S$, $\gamma|_T$ is Lipschitz and $\gamma(T)$ is compactly C^s Lipschitz bounded. If S is covered by open sets U such that $\gamma|_U$ is tame, then γ is* locally tame.

That is, γ is (locally) (r, s)-tame if and only if $\Psi^s \circ \gamma \circ (\Psi^r)^{-1}|_{\Psi^r(S)}$ is (locally) tame.

2.3 Smooth Manifolds with Corners

We will work with C^{r+} manifolds with corners, which differ from C^r manifolds with boundary in two respects: 1) the transition functions between coordinate charts are required to be C^{r+}; 2) the images of coordinate charts are open subsets of the positive orthant $\mathbb{R}^m_{\geq}$. Manifolds with corners were first developed by Cerf (1961) and Douady (1961). They are infrequently encountered in the mathematical literature because the additional generality, beyond manifolds with boundary, is rarely necessary or useful. This is true here as well, insofar as the additional generality does not change the analysis in any way or lead to additional insights. We develop our results in this context only because manifolds with corners occur frequently in economic applications. In particular, in Bayesian models the state space is often the simplex of probability distributions on a finite set of possible parameter values.

Once we know that compositions of C^{r+} functions are C^{r+} (Lemma 1) the fundamental definitions and results for C^{r+} manifolds with corners are straightforward modifications of those for C^r manifolds with boundary, which can be found in many sources, and we will not repeat them. The reader is expected to be familiar with the notion of a C^{r+} atlas of coordinate charts, the tangent space at a point, and the tangent space of the manifold, the definition of the derivative of a C^{1+} function between C^{1+} manifolds, and basic results such as the chain rule.

Let M and N be m- and n-dimensional C^{r+} manifolds with corners. For any set $S \subset M$ a function $f : S \to N$ is (by convention) C^r or C^{r+} if it has a C^r or C^{r+} extension to some neighborhood of S. For $T \subset N$ let $C^r(S,T)$ ($C^{r+}(S,T)$) be the set of $f : S \to T$ that are C^r (C^{r+}).

A set $D \subset M$ is a *differentiation domain* if, for any C^1 coordinate chart $\varphi : U \to \mathbb{R}^m$ for M, $\varphi(D \cap U)$ is a differentiation domain. In view of the chain rule, D is a differentiation domain if it is covered by the domains of some system of coordinate charts with this property. If D is a differentiation domain, any C^1 function $f : D \to N$ has an unambiguously defined derivative at each point of D.

We will globalize the definitions of the preceeding subsection by covering D with a finite collection of coordinate charts. Let $\varphi : U_\varphi \to \mathbb{R}^m$ and $\psi : V_\psi \to \mathbb{R}^n$ be typical C^{r+} coordinate charts for M and N respectively. Consider a triple $\sigma = (K, \varphi, \psi)$ in which $K \subset D$ is a compact differentiation domain, φ and ψ are C^{r+} coordinate charts for M and N, and $K \subset U_\varphi$. Let

$$Z^r_\sigma = \{\, f \in C^r(D,N) : f(K) \subset V_\psi \,\},$$

and let $\Gamma^r_\sigma : Z^r_\sigma \to C^r(\varphi(K), \mathbb{R}^n)$ be the function that takes f to $\psi \circ f \circ \varphi^{-1}|_{\varphi(K)} \in C^r(\varphi(K), \mathbb{R}^n)$. There is a pseudometric[1] d^r_σ on Z^r_σ given by

$$d^r_\sigma(f,f') = d^r\big(\Gamma^r_\sigma(f'), \Gamma^r_\sigma(f'')\big) = d(\Psi^r(\Gamma^r_\sigma(f)), \Psi^r(\Gamma^r_\sigma(f'))).$$

A C^{r+} *metric configuration* for the tuple (M,N,D) is a finite collection of triples $B = \{\sigma_i = (K_i, \varphi_i, \psi_i)\}_{i=1}^k$ as above such that $\bigcup_i K_i = D$. When $N = \mathbb{R}$, we will always have

[1] A *pseudometric* is a function with all of the properties (values in $[0,\infty)$, symmetry, triangle inequality) of a metric except that the distance between distinct points may be zero.

each ψ_i be the identity function on $\mathbb{R}$ (which is omitted from the notation) so B has an associated C^{r+} metric configuration $\tilde{B} = \{\tilde{\sigma}_i = (K_i, \varphi_i)\}_{i=1}^k$ for $(M, \mathbb{R}, D)$.

For such a B let $Z_B^r = \bigcap Z_{\sigma_i}^r$, and for $f, f' \in Z_B^r$ let $d_B^r(f, f') = \max_i d_{\sigma_i}^r(f, f')$. This is evidently a pseudometric, and since the sets K_i cover D, it is actually a metric. We endow each Z_B^r with the topology induced by the metric d_B^r, and we endow $C^r(D, N)$ with the *C^r-topology*, which is, by definition, the coarsest topology for which each open subset of each Z_B^r is open. We endow $C^{r+}(D, N)$ with the induced relative topology.

The next result implies that the subspace topology of Z_B^r induced by its inclusion in $C^r(D, N)$ is not finer than the topology induced by d_B^r. Concretely, if $U \subset C^r(D, N)$ is open and $f \in U \cap Z_B^r$, then for some $\varepsilon > 0$ the ε-ball for d_B^r centered at f is contained in U. In addition, the next result implies that any two of the metrics d_B^r define the same class of locally Lipschitz functions on the intersection of their domains.

Proposition 1. *If $f \in Z_B^r \cap Z_{B'}^r$, where $B = \{\sigma_i = (K_i, \varphi_i, \psi_i)\}_{i=1}^k$ and $B' = \{\sigma'_j = (K'_j, \varphi'_j, \psi'_j)\}_{j=1}^{k'}$ are C^{r+} metric configurations for D, then there is an $\varepsilon > 0$ such that the ε-ball for d_B^r centered at f is contained in $Z_B^r \cap Z_{B'}^r$, and the identity function on this ball is Lipschitz when the domain has the metric d_B^r and the range has the metric $d_{B'}^r$.*

Proof. For each $i = 1, \ldots, k$ and $j = 1, \ldots, k'$ there is some $\varepsilon_{ij} > 0$ such that the ball of radius ε_{ij} around $\psi_i(f(K_i \cap K'_j))$ is contained in $\psi_i(V_\psi \cap V_{\psi'_i})$. If $f' \in Z_B^r$ and $d_B^0(f, f') < \min \varepsilon_{ij}$, then $f' \in Z_{B'}^r$. Of course

$$\psi'_j \circ f' \circ {\varphi'_j}^{-1}|_{\varphi'(K_i \cap K'_j)} = (\psi'_j \circ \psi_i^{-1}) \circ (\psi_i \circ f' \circ \varphi_i^{-1}) \circ (\varphi_i \circ {\varphi'_j}^{-1})|_{\varphi'(K_i \cap K'_j)},$$

so if $P(f')$ is a partial of $\psi'_j \circ f' \circ {\varphi'_j}^{-1}$ of order $\leq r$, then, using standard facts of differentiation, $P(f')$ can be expressed as a polynomial function of partials Q of $\varphi_i \circ {\varphi'_j}^{-1}$, functions of the form $R(f') \circ \varphi_i \circ {\varphi'_j}^{-1}$ where $R(f')$ is a partial of $\psi_i \circ f' \circ \varphi_i^{-1}$, and functions of the form $S \circ \psi_i \circ f' \circ {\varphi'_j}^{-1}$ where S is a partial of $\psi'_j \circ \psi_i^{-1}$. We may choose $\varepsilon \leq \min \varepsilon_{ij}$ such that the values of the functions $R(f') \circ \varphi_i \circ {\varphi'_i}^{-1}|_{\varphi'(K_i \cap K'_j)}$ and $S \circ \psi_i \circ f' \circ {\varphi'_j}^{-1}|_{\varphi'(K_i \cap K'_j)}$ are bounded uniformly for f' in the ε-ball for d_B^r centered at f. Let U be this ε-ball.

Consider a term $T(f') = X_1(f') \cdots X_\ell(f')$ in the polynomial expressing P, where each $X_h(f')$ is the restriction to $\varphi'(K_i \cap K'_j)$ of one of the functions of the forms Q, $R(f') \circ \varphi_i \circ {\varphi'_j}^{-1}$, or $S \circ \psi_i \circ f' \circ {\varphi'_j}^{-1}$. Then $T(f'') - T(f')$ is the sum of the expressions of the form

$$X_1(f'') \cdots X_{g-1}(f'')\big(X_g(f'') - X_g(f')\big)X_{g+1}(f') \cdots X_\ell(f'). \tag{$*$}$$

Each $X_h(f')$ and $X_h(f'')$ is bounded by virtue of our choice of U. A quantity of the form $(R(f'') - R(f')) \circ \varphi_i \circ {\varphi'_j}^{-1}$ is bounded by a multiple of $d_B^r(f', f'')$ by virtue of the definition of this metric. A quantity of the form $(S \circ \psi_i \circ f'' - S \circ \psi_i \circ f') \circ {\varphi'_j}^{-1}$ is bounded by a multiple of $d^r_B(f', f'')$ because ψ_i and S are both Lipschitz. Thus there is a constant such that for all $f', f'' \in U$, the absolute value of each expression of the form $(*)$ is bounded on $K_i \cap K'_j$ by that constant times $d^r_{K_i,\varphi,\psi}(f', f'')$. Aggregating, there is a constant $\Lambda > 0$ such $d^r_{B'}(f', f'') \leq \Lambda d^r_B(f', f'')$ for all $f', f'' \in U$. □

Following the pattern used to define manifolds, one could define a local Lipschitz atlas on a space to be a collection of metrics on the elements of an open cover of the space that are related to each other by local Lipschitz conditions on the overlaps of their domains, and a locally Lipschitz structure on the space would then be a maximal such atlas. In this way we can define a notion of a locally Lipschitz function that does not depend on the choice of particular metrics. Introducing such formalism would be cumbersome without easing the exposition, so we do not do so, but this is the correct conceptual orientation.

For $\Lambda > 0$ we say that $f \in Z^r_B$ is (C^r, B, Λ)-*Lipschitz bounded* if each $\Gamma^r_{\sigma_i}(f)$ is C^r Λ-Lipschitz bounded. Let $Z^r_{B,\Lambda}$ be the set of such f, and let $Z^{r+}_B = \bigcup_{\Lambda \geq 0} Z^r_{B,\Lambda}$ be the set of $f \in Z^r_B$ that are (C^r, B, Λ)-Lipschitz bounded for some $\Lambda > 0$.

Consider a set $S \subset C^r(D, N)$. We say that S is *metrically C^r Lipschitz bounded* if there is some B such that $S \subset Z^r_B$ and each $\Gamma^r_{\sigma_i}(S)$ is compactly C^r Lipschitz bounded. We say that S is *compactly C^r Lipschitz bounded* if it has a finite covering by metrically C^r Lipschitz bounded sets.

In addition to M, N, and D as above, let P and Q be C^{s+} manifolds with corners, and let $E \subset P$ be a compact differentiation domain. Let S be a subset of $C^r(D, N)$, and let $\gamma : S \to C^s(E, Q)$ be an operator. We say that γ is *locally Lipschitz* if, for any $f \in S$, any C^{r+} metric configuration $B = \{\sigma_i = (K_i, \varphi_i, \psi_i)\}_{i=1}^k$ with $f \in Z^r_B$, and any C^{s+} metric configation $C = \{\tau_j = (L_j, \alpha_j, \beta_j)\}_{j=1}^\ell$ with $\gamma(f) \in Z^s_C$, there is a neighborhood $U \subset Z^r_B$ of f such that $\gamma(S \cap U) \subset Z^s_C$ and $\gamma|_{S \cap U}$ is Lipschitz relative to the metrics d^r_B and d^s_C.

Definition 3. *The operator γ is (r, s)-*tame *if it is continuous and for every compactly C^r Lipschitz bounded $T \subset S$, $\gamma|_T$ is locally Lipschitz and $\gamma(T)$ is compactly C^s Lipschitz bounded. If S is covered by open sets U such that $\gamma|_U$ is tame, then γ is* locally tame.

3 Dynamic Programming

We begin with some general considerations related to dynamic programming. Let the space of *states* Ω and the space of *actions* A be metric spaces, with Ω compact. We endow the space $\Delta(\Omega)$ of probability measures on Ω with the weak* topology and the associated Borel σ-algebra.

In general, whenever S is a subset of a cartesian product $X \times Y$, $S(x) = \{\, y : (x, y) \in S\,\}$ will be the "slice" above $x \in X$. The correspondence mapping each state to the set of actions that are feasible at that state is expressed by a set $D \subset \Omega \times A$ such that $D(\omega) \neq \emptyset$ for all $\omega \in \Omega$. Let $P : D \to \Delta(\Omega)$ be a measurable *transition function*. For a bounded measurable *reward function* $u : D \to \mathbb{R}$ and a *discount factor* $\delta \in [0, 1)$ the dynamic program is to maximize the expectation of

$$\sum_{t=0}^{\infty} \delta^t u(\tilde{\omega}_t, \tilde{a}_t)$$

where $\tilde{\omega}_0 = \omega_0$ almost surely, $\tilde{\omega}_t$ is known at the time $\tilde{a}_t \in D(\tilde{\omega}_t)$ is chosen, and, conditional on $\tilde{\omega}_t$ and $\tilde{a}_t$, $\tilde{\omega}_{t+1}$ has the distribution $P(\tilde{\omega}_t, \tilde{a}_t)$.

A (stationary, deterministic) *policy* is a measurable function $\pi : \Omega \to A$ with $\pi(\omega) \in D(\omega)$ for all ω. For a given initial state ω_0, a policy π and the transition function P induce a probability measure on the space of infinite histories $(\omega_0, a_0), (\omega_1, a_1), \ldots$. Throughout we assume that u is bounded, so for any discount factor $\delta \in [0, 1)$ there is a well defined expected discounted sum of rewards associated with the induced measure on the space of histories. The supremum of this expectation, over all policies, is the *value* of ω_0 for discount factor δ, and the *value function* for δ is the function $V_{u,\delta} : \Omega \to \mathbb{R}$ taking each state in Ω to its value.

Our methods depend on the point of view developed by (among others) Blackwell (1965) according to which $V_{u,\delta}$ is the fixed point of a contraction mapping. As Blackwell's article describes, various pathologies are possible in a quite general setting, so we introduce assumptions that guarantee some regularity. In particular, we assume that D is compact, the correspondence $\omega \mapsto D(\omega)$ is continuous (that is, both upper and lower hemicontinuous) for each $\omega \in \Omega$ the set of $a \in D(\omega)$ such that (ω, a) is in the interior of D is nonempyt, and P is continuous, and we consider only continuous reward functions.

We now introduce the key operators:

(a) For $u \in C(D)$ let $J(u) : \Omega \to \mathbb{R}$ be the function

$$J(u)(\omega) = \max_{a \in D(\omega)} u(\omega, a).$$

Since $D(\cdot)$ is continuous, $J(u) \subset C(\Omega)$, and it is straightforward to show that $J : C(D) \to C(\Omega)$ is Lipschitz with Lipschitz constant one.

(b) For $V \in C(\Omega)$ let $K(V) : D \to \mathbb{R}$ be the function

$$K(V)(\omega, a) = \int_{\Omega} V(\omega')\, P(\omega, a; d\omega').$$

Since P is continuous, $K(V) \in C(D)$, and evidently the operator $K : C(\Omega) \to C(D)$ is Lipschitz with Lipschitz constant one.

(c) Let $L : C(D) \times \mathbb{R} \times C(\Omega) \to C(\Omega)$ be the operator

$$L(u, \delta, V) = J(u + \delta \cdot K(V)).$$

Evidently L is locally Lipschitz.

Lemma 2. *For each $u \in C(D)$ and $\delta \in [0, 1)$, $V_{u,\delta}$ is the unique fixed point of $L(u, \delta, \cdot)$.*

The proof is simple, and can be found in many sources. Obviously $V_{u,\delta}$ is a fixed point of $L(u, \delta, \cdot)$. Since $L(u, \delta, \cdot)$ is Lipschitz for the Lipschitz constant δ, and $C(\Omega)$ is complete, the contraction mapping implies that $L(u, \delta, \cdot)$ has a unique fixed point.

We now assume that Ω and A are C^{r+} manifolds. The transition function P is said to be r^{th} *order smoothing* if K is (r, r)-tame. This condition might hold for a variety of reasons, and must generally be verified in the context of each application, so we do not give conditions that imply it.

Next, we recall some basic calculus. If $V \subset \mathbb{R}^m$ is open and $u : V \to \mathbb{R}$ is C^2, we say that u *satisfies the second order conditions strictly* at $x \in V$ if $\frac{\partial u}{\partial x_i}(x) = 0$ for all i and the Hessian matrix H with entries $\frac{\partial^2 u}{\partial x_i \partial x_j}(x)$ is negative definite. If this is the case, then u attains a strict local maximum at x. If M is a C^r manifold, and $u : M \to \mathbb{R}$ is C^2, we say that u *satisfies the second order conditions strictly* at $x \in M$ if there is a C^2 coordinate chart $\varphi : U \to \mathbb{R}^m$ with $x \in U$ such that $u \circ \varphi^{-1}$ satisfies the second order conditions strictly at $\varphi(x)$. (If this is the case, then this condition holds for any other C^2 coordinate chart because if $V, W \subset \mathbb{R}^m$ are open, $u : V \to \mathbb{R}$ satisfies the second order conditions strictly at x, $h : W \to V$ is a C^2 diffeomorphism, and $h(y) = x$, then straightforward calculation shows that the Hessian matrix of $u \circ h$ at y is $D'HD$ where D is the matrix of partial derivatives of h.)

We say that $u \in C^r(D)$ satisfies the *standard conditions* if, for each $\omega \in \Omega$, there is a unique maximizer $\pi_u(\omega)$ of $u(\omega, \cdot)$, the point $(\omega, \pi_u(\omega))$ is in the interior of D, and $u(\omega, \cdot)$ satisfies the second order conditions strictly at $\pi_u(\omega)$. Let Σ^r be the set of $u \in C^r(D)$ that satisfy the standard conditions, and let $\Sigma^{r+} = \Sigma^r \cap C^{r+}(D)$. The following is our main result.

Theorem 1. *Suppose P is r^{th} order smoothing, $\Lambda > \Lambda' > 0$, and $u_0 \in \Sigma^{r+} \cap Z^r_{B,\Lambda'}$. Let $\tilde{B} = \{(K_i, \varphi_i)\}_{i=1}^k$ be a C^r metric configuration for $(M, \mathbb{R}, D)$, let $C = \{(L_j, \psi_j, \rho_j)\}_{j=1}^\ell$ be a C^{r+} metric configuration for (M, A, Ω) such that $\pi_{u_0} \in Z^r_C$, and let $\tilde{C} = \{(L_j, \psi_j)\}_{j=1}^\ell$ be the associated C^{r+} metric configuration for $(M, \mathbb{R}, \Omega)$. Then there is a neighborhood $N \subset Z^r_{\tilde{B},\Lambda} \times [0, 1)$ of $(u_0, 0)$ and numbers $\Lambda_\pi, \Lambda_V > 0$ such that for each $(u, \delta) \in N$:*

(a) $u \in \Sigma^{r+}$;

(b) there is a unique stationary optimal policy $\pi_{u,\delta}$ *for the discounted dynamic program payoff* u *and discount factor* δ;

(c) $\pi_{u,\delta} \in Z^r_{C,\Lambda_\pi}$;

(d) $V_{u,\delta} \in Z^r_{\tilde{C},\Lambda_V}$.

In addition, the restriction of the operator $(u, \delta) \mapsto \pi_{u,\delta}$ *to* N *is* $(r, r-1)$*-tame, and the restriction of the operator* $(u, \delta) \mapsto V_{u,\delta}$ *to* N *is* (r, r)*-tame.*

The rest of the paper provides the proof. The central thrust is to fulfill the conditions of Proposition 2, so we need a set of C^{r+} "candidate value functions" that is complete, such that the operator computing today's value function, given a value function for the problem tomorrow, is a contraction that maps this set to itself. Once we have domains for which J and K are Lipschitz operators, restricting to small discount factors reduces the Lipschitz constant of $L(u, \delta, \cdot)$ so that it becomes a contraction[2]. Since $V_{u,\delta}$ is the unique fixed point of $L(u, \delta, \cdot)$, it is C^{r+}.

We can now give some general explanation of the reason we restrict attention to functions whose r^{th} order partial derivatives are Lipschitz, which is that this restriction implies that some of the operators of interest are Lipschitz. Consider, for example, the operator that takes a pair of functions (f, g) to their composition $g \circ f$. This operator will not be Lipschitz (relative to the metric defining the C^r topology) if applied to functions g whose r^{th} order partial derivatives are not Lipschitz. Concretely, perturbing f perturbs the point $f(x)$ where the r^{th} order partials of g are evaluated, which changes the r^{th} order partial derivatives of $g \circ f$ at x in a manner that is not controlled by a Lipschitz constant. Seen in this light, the techniques we use to circumvent this difficulty are simple and natural. However, before the idea of the proof of Theorem 1 can be fulfilled, many elementary technical issues need to be addressed. This is the work of the next two sections.

4 Technical Foundations

This section contains numerous basic technical results underlying our analysis.

[2]This is the step that would fail if we tried to adapt our methods to prove the seemingly reasonable conjecture that if $u_0 \in Z^{r+}_{\tilde{B},\Lambda'}$, $0 < \delta_0 < 1$, $V_{u_0,\delta_0} \in C^{r+}(\Omega)$, and for each ω, $\pi_{u_0,\delta_0}(\omega)$ is the unique maximizer of $u_0(\omega, \cdot) + \delta_0 K(V_{u_0,\delta_0})(\omega, \cdot)$ and satisfies the second order conditions strictly, then these conditions hold also for (u, δ) in some neighborhood of (u_0, δ_0) in $Z^{r+}_{\tilde{B},\Lambda} \times (0, 1)$.

4.1 Lipschitz Facts

We will need the following parameterized version of the contraction mapping theorem. Let X and Y be metric spaces.

Proposition 2. *Suppose that Y is nonempty and complete and $f : X \times Y \to Y$ is a locally Lipschitz function such that for each $x \in X$ there is a neighborhood $V \subset X$ and a number $c \in [0, 1)$ such that $d(f(x', y), f(x', y')) \leq cd(y, y')$ for all $x' \in V$ and $y, y' \in Y$. Then for each $x \in X$ there is a unique fixed point $\varphi(x)$ of $f(x, \cdot) : Y \to Y$, and $\varphi : X \to Y$ is locally Lipschitz.*

Proof. The unique existence of $\varphi(x)$ is the assertion of the contraction mapping theorem, so we need only prove that φ is locally Lipschitz. Fix $x \in X$ and a neighborhood V and c as in the hypotheses. Also, fix a neighborhood $U \subset X \times Y$ of $(x, \varphi(x))$ such that $f|_U$ is Lipschitz for Lipschitz constant Λ. Consider $x' \in V$ such that $(x', \varphi(x)) \in U$. The sequence defined by setting $y_0 = \varphi(x)$ and $y_{i+1} = f(x', y_i)$ converges to $\varphi(x')$. We have

$$d(y_0, y_1) = d\big(f(x, \varphi(x)), f(x', \varphi(x))\big) \leq \Lambda d(x, x')$$

and

$$d(y_i, y_{i+1}) = d\big(f(x', y_{i-1}), f(x', y_i)\big) \leq cd(y_{i-1}, y_i)$$

for all $i \geq 1$, so $d(\varphi(x), \varphi(x')) \leq \Lambda d(x, x')/(1 - c)$. □

It is very easy to see that a composition of two (locally) Lipschitz functions is (locally) Lipschitz, and that the restriction of a (locally) Lipschitz function to a subset of the domain is (locally) Lipschitz. In fact one can say a bit more.

Lemma 3. *If $f : X \to Y$ is locally Lipschitz and $K \subset X$ is compact, then there is a neighborhood U of K such that $f|_U$ is Lipschitz.*

One can easily pass from the hypotheses of this result to the hypotheses of the following more precise formulation of the underlying technical situation, which will be needed later.

Lemma 4. *Suppose that $f : X \to Y$ is a function, $x_1, \ldots, x_k \in X$, $\varepsilon_1, \ldots, \varepsilon_k > 0$, $\Lambda_1, \ldots, \Lambda_k > 0$, and for each $i = 1, \ldots, k$ the restriction of f to the open ball of radius $2\varepsilon_i$ centered at x_i is Lipschitz with Lipschitz constant Λ_i. Let U be the union of the open balls of radius ε_i around the x_i. Suppose that $f(U)$ is contained in a subset of Y of radius M, let $\varepsilon = \min_i \varepsilon_i$, and let $\Lambda = M/\varepsilon$. Then $f|_U$ is Lipschitz with Lipschitz constant* $\max\{\Lambda_1, \ldots, \Lambda_k, \Lambda\}$.

Proof. For any $x, x' \in U$, if $d(x, x') < \varepsilon$, then $d(f(x), f(x')) \leq \Lambda_i d(x, x')$ for some i, and if $d(x, x') \geq \varepsilon$, then $d(f(x), f(x')) \leq M \leq \Lambda d(x, x')$. □

Throughout a cartesian product of metric spaces $(Y_1, d_1), \ldots, (Y_k, d_k)$ will automatically be endowed with the metric

$$\max_i d_i : ((y_1, \ldots, y_k), (y'_1, \ldots, y'_k)) \mapsto \max_i d_i(y_i, y'_i).$$

(In all of our work $\mathbb{R}$ will have the metric $(s, t) \mapsto |s - t|$, so that $\mathbb{R}^n$ is endowed with the metric $(x, y) \mapsto \max_i |x_i - y_i|$ derived from the norm $\|\cdot\|_\infty$.) Abusing notation slightly, for functions $f_1 : X \to Y_1, \ldots, f_k : X \to Y_k$ let $f_1 \times \cdots \times f_k$ be the function

$$f_1 \times \cdots \times f_k : x \mapsto (f_1(x), \ldots, f_k(x)).$$

With this convention the next two results are immediate.

Lemma 5. *If, for each i, $f_i : X \to Y_i$ is Lipschitz for Lipschitz constant Λ_i, then $f_1 \times \cdots \times f_k$ is Lipschitz for Lipschitz constant $\max_i \Lambda_i$.*

For the next two results we assume that X is compact. The following is obvious.

Lemma 6. *If, for each $i = 1, \ldots, k$, $S_i \subset C(X, Y_i)$ is (K_i, Λ_i)-Lipschitz bounded, then $\{ f_1 \times \cdots \times f_k : f_1 \in S_1, \ldots, f_k \in S_k \}$ is $(K_1 \times \cdots \times K_k, \max_i \Lambda_i)$-Lipschitz bounded.*

Recall that the pointwise limit of a uniformly convergent sequence of continuous functions is continuous. Obviously the pointwise limit of a uniformly convergent sequence of Λ-Lipschitz functions is Λ-Lipschitz. If Y is complete, then a Cauchy sequence $\{f_n\}$ in $C(X, Y)$ has a pointwise limit because each $\{f_n(x)\}$ is Cauchy. Thus:

Lemma 7. *For each $\Lambda \geq 0$, $C_\Lambda(X, Y)$ is a closed subset of $C(X, Y)$. If Y is complete, then $C(X, Y)$ is complete, and for any $\Lambda > 0$, $C_\Lambda(X, Y)$ is complete.*

4.2 Basic Properties of Tame Operators

Restriction to a subdomain, composition, and various forms of cartesian products are perhaps the most important elementary operations that are used to construct new functions from given functions, and our agenda in this subsection is largely a matter of establishing that these operations preserve the key properties. We work with metric spaces X, Y, X', Y', X'', and Y'', with X, X', and X'' compact. Fix a set $S \subset C(X, Y)$. The first result is trivial.

Lemma 8. *If $g \in C(X', Y')$ is Lipschitz and $\gamma : S \to C(X', Y')$ is the constant operator with value g, then γ is tame.*

Restriction to a subdomain is compatible with tameness in two senses: a) if $\gamma : S \to C(X', Y')$ is (locally) tame and $S' \subset S$, then $\gamma|_{S'}$ is (locally) tame; b) if $K \subset X$ is compact, then the operator $f \mapsto f|_K$ mapping $C(X, Y)$ into $C(K, Y)$ is a tame. (These claims follow immediately from the definitions.) There are also two senses in which composition preserves tameness.

Lemma 9. *If $S' \subset C(X', Y')$ and $\gamma : S \to S'$ and $\gamma' : S' \to C(X'', Y'')$ are (locally) tame, then $\gamma' \circ \gamma$ is (locally) tame.*

Proof. The claim for tame operators is an immediate consequence of the definition, so suppose that γ and γ' are locally tame. Fixing $f \in S$, let $U \subset S$ and $U' \subset S'$ be neighborhoods of f and $\gamma(f)$ such that $\gamma|_U$ and $\gamma'|_{U'}$ are tame. Then $\gamma|_U$ is continuous, so $U \cap \gamma^{-1}(U')$ is open, and $\gamma|_{U \cap \gamma^{-1}(U')}$ is tame so we may assume that $\gamma(U) \subset U'$. Now $(\gamma' \circ \gamma)|_U = \gamma'|_{U'} \circ \gamma|_U$ is tame. □

For $f \in C(X, Y)$ and $f' \in C(X', Y')$ there is the associated function $f \times f' : (x, x') \mapsto (f(x), f(x'))$ in $C(X \times X', Y \times Y')$. This identification provides the sense in which we understand the following result and similar assertions later.

Proposition 3. *If X, Y, and Z are metric spaces with X and Y compact, then $(f, g) \mapsto g \circ f$ is a tame operator from $C(X, Y) \times C(Y, Z)$ to $C(X, Z)$.*

Proof. We first show that the operator is continuous. Consider $f \in C(X, Y)$, $g \in C(Y, Z)$, and $\varepsilon > 0$. Since Y is compact, g is uniformly continuous, so there is $\delta > 0$ such that $d(g(y), g(y')) < \varepsilon/2$ whenever $d(y, y') < \delta$. If $d(f, f'), d(g, g') < \min\{\delta, \varepsilon/2\}$, then

$$d(g'(f'(x)), g(f(x))) \le d(g'(f'(x)), g(f'(x))) + d(g(f'(x)), g(f(x))) \le d(g, g') + \varepsilon/2 < \varepsilon$$

for all $x \in X$, so $d(g \circ f, g' \circ f') < \varepsilon$.

Let $T \subset C(X, Y) \times C(Y, Z)$ be compactly Lipschitz bounded. Then there is a compact $K \subset Y \times Z$ that contains the image of every element. Let $\hat{K}$ be the projection of K on Z. Then for every $(f, g) \in T$, then image of $g \circ f$ is contained in $\hat{K}$. There is some Λ such that $T \subset C_\Lambda(X, Y) \times C_\Lambda(Y, Z)$. It is easy to see that the image of $C_\Lambda(X, Y) \times C_\Lambda(Y, Z)$ is contained in $C_{\Lambda^2}(X, Z)$. Thus $\{ g \circ f : (f, g) \in T \}$ is $(\hat{K}, \Lambda^2)$-Lipschitz bounded.

For any $(f, g), (f', g') \in T$ and $x \in X$ we have

$$d(g'(f'(x)), g(f(x))) \le d(g'(f'(x)), g(f'(x))) + d(g(f'(x)), g(f(x)))$$

$$\le d(g, g') + \Lambda \cdot d(f'(x), f(x)) \le d(g, g') + \Lambda \cdot d(f, f') \le (\Lambda + 1) d((f, g), (f', g')).$$

Since this is true for any x, the restriction of the operator to T is $(\Lambda + 1)$-Lipschitz. □

Lemma 10. *If X is compact and $g : Y \to Z$ is locally Lipschitz, then $f \mapsto g \circ f$ is a tame operator from $C(X, Y)$ to $C(X, Z)$.*

Proof. If $K \subset Y$ is compact, Lemma 3 gives a neighborhood V of K such that $g|_V$ is Lipschitz, say with Lipschitz constant Λ'. If $f \in C(X, Y)$, then $f(X)$ is compact, so it is contained in such a V, $C(X, V)$ is a neighborhood of f, and the restriction of the operator to $C(X, V)$ is Λ'-Lipschitz. Since f was arbitrary we have shown that the operator is locally Lipschitz, hence continuous. If $T \subset C(X, Y)$ is (K, Λ)-Lipschitz bounded, then the restriction of the operator to T is Λ'-Lipschitz, and $\{ g \circ f : f \in T \}$ is $(g(K), \Lambda\Lambda')$-Lipschitz bounded. □

Lemma 11. *If $S \subset C(X, Y)$, then operators $\gamma_1 : S \to C(X', Y_1'), \ldots, \gamma_k : S \to C(X', Y_k')$ are (locally) tame if and only if the operator $\gamma : f \mapsto \gamma_1(f) \times \cdots \times \gamma_k(f)$ is a (locally) tame operator from S to $C(X', Y_1' \times \cdots \times Y_k')$.*

Proof. Of course γ is continuous if and only if $\gamma_1, \ldots, \gamma_k$ are continuous. Suppose $T \subset S$ is compactly Lipschitz bounded. Clearly $\gamma|_T$ is Lipschitz if and only each $\gamma_i|_T$ is Lipschitz. If $\gamma_1(T), \ldots, \gamma_k(T)$ are compactly Lipschitz bounded, then Lemma 6 implies that $\gamma(T)$ is compactly Lipschitz bounded. For each i, if $\pi_i : \prod_h Y_h' \to Y_i'$ is the projection, then Lemma 10 implies that the operator $g \mapsto \pi_i \circ g$ is tame, and γ_i is the composition of γ with this operator, so Proposition 3 implies that γ_i is tame.

If γ is locally tame, then from the above we see that each γ_i is locally tame. If each γ_i is tame, $f \in S$, $U_1, \ldots, U_k$ are open neighborhoods of f such that each $\gamma_i|_{U_i}$ is tame, and $U = \bigcap U_i$, then from the above $\gamma|_U$ is tame, so γ is locally tame. □

4.3 Basic Properties of Smooth Tameness

Let $D \subset \mathbb{R}^m$, $D' \subset \mathbb{R}^{m'}$, and $D'' \subset \mathbb{R}^{m''}$ be compact differentiation domains. Let $0 \le r, s, t \le \infty$ be orders of differentiability.

Lemma 12. *For any finite n and $\Lambda \ge 0$, $C^r(D, \mathbb{R}^n)$ is complete and $C^r_\Lambda(D, \mathbb{R}^n)$ is closed, hence complete.*

Proof. It is elementary, and easily proved, that if a sequence of univariate C^1 functions on an open subset of $\mathbb{R}$ and its sequence of derivatives converge uniformly, say to f and g respectively, then g is the derivative of f. In view of the definition of d^r, a Cauchy sequence $\{f_j\}$ in $C^r(D, \mathbb{R}^n)$ is a sequence which is mapped by Ψ^r to a Cauchy sequence in $C(D, \mathbb{R}^N)$ (where N is as before). Since $C(D, \mathbb{R}^N)$ is complete (Lemma 7) the sequence has a uniform limit f, which (by virtue of the fact just mentioned) is C^r and is mapped by

Ψ^r to the limit of the sequence $\{\Psi^r(f_j)\}$, so $f_j \to f$ in $C^r(D, \mathbb{R}^n)$. If $\{f_j\}$ is a sequence in $C^r_\Lambda(D, \mathbb{R}^n)$ that converges to f, then $\Psi^r(f_j) \to \Psi^r(f)$, so $\Psi^r(f) \in C_\Lambda(D, \mathbb{R}^N)$ by Lemma 7, and thus $f \in C^r_\Lambda(D, \mathbb{R}^n)$. □

There are several simple properties of smooth tameness that figure in our work. Fix $S \subset C^r(D, \mathbb{R}^n)$. The next two results follow from Lemmas 8 and 9 respectively.

Lemma 13. *If $g \in C^{s+}(D', \mathbb{R}^{n'})$ and $\gamma : S \to C^s(D', \mathbb{R}^{n'})$ is the constant operator with value g, then γ is (r, s)-tame.*

Lemma 14. *If $\gamma : S \to S'$ is (locally) (r, s)-tame and $\gamma' : S' \to C^t(D'', \mathbb{R}^{n''})$ is (locally) (s, t)-tame, then $\gamma' \circ \gamma : S \to C^t(D'', \mathbb{R}^{n''})$ is (locally) (r, t)-tame.*

The two next results follow easily from Definition 2 and Lemma 11.

Lemma 15. *If $s \le r$, then the inclusion $\iota^r_s : C^r(D, X) \hookrightarrow C^s(D, X)$ is an (r, s)-tame operator.*

Lemma 16. *Operators $\gamma_1 : S \to C^s(D', \mathbb{R}^{n_1}), \ldots, \gamma_k : S \to C^s(D', \mathbb{R}^{n_k})$ are (locally) (r, s)-tame if and only if the operator $\gamma : f \mapsto \gamma_1(f) \times \cdots \times \gamma_k(f)$ is a (locally) (r, s)-tame operator from S to $C^s(D', \mathbb{R}^{n_1+\cdots+n_k})$.*

If $f \in C^r(D, \mathbb{R}^n)$, let $Df \in C^{r-1}(D, \mathbb{R}^{m\times n})$ be the function that takes each x to the $m \times n$ matrix of partial derivatives of f. If $m_1 + m_2 = m$, so that elements of $\mathbb{R}^m$ are thought of as pairs $(x, y) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$, then we let $\partial_x f$ and $\partial_y f$ denote the respective "partial derivative" functions: $\partial_x f = Df(\cdot, y)$ and $\partial_y f = Df(x, \cdot)$.

Lemma 17. *The operators $f \mapsto \partial_x f$ and $f \mapsto \partial_y f$ are $(r, r-1)$-tame functions from $C^r(D, \mathbb{R}^n)$ to $C^{r-1}(D, \mathbb{R}^{m_1\times n})$ and $C^{r-1}(D, \mathbb{R}^{m_2\times n})$.*

Proof. We need to show that $\Psi^{r-1} \circ \partial_x \circ (\Psi^r)^{-1}|_{\Psi^r(C^{r+}(D,\mathbb{R}^n))}$ is tame, and this is a direct consequence of Definition 2 and Lemma 11. □

Proposition 4. *The operator $c : (f, g) \mapsto g \circ f$ is an (r, r)-tame function from $C^r(D, D') \times C^r(D', \mathbb{R}^n)$ to $C^r(D, \mathbb{R}^n)$.*

Proof. We need to show that $\Psi^r \circ c \circ (\Psi^r)^{-1}$ to $\Psi^r(C^r_\Lambda(D, D') \times C^r_\Lambda(D', \mathbb{R}^n))$ is tame. Proposition 3 implies that c is $(0, 0)$-tame. In view of Lemma 11 it suffices to show that if $1 \le s \le r$, $1 \le i_1, \ldots, i_s \le m$, and $1 \le h \le n$, then the partial derivative

$$P(f, g) = \frac{\partial^s(g_h \circ f)}{\partial x_{i_1} \cdots \partial x_{i_s}}$$

is a tame function of $(\Psi^r(f), \Psi^r(g))$. Repeated application of the basic facts of differentiation allows us to express $P(f, g)$ as a polynomial function of components $Q(f)$ of $\Psi^r(f)$ and functions of the form $R(g) \circ f$ where $R(g)$ is a component of $\Psi^r(g)$. Lemma 11 implies that the projections $(\Psi^r(f), \Psi^r(g)) \mapsto Q(f)$ and $(\Psi^r(f), \Psi^r(g)) \mapsto R(g)$ are tame. Proposition 3 implies that each $(f, R(g)) \mapsto R(g) \circ f$ is a tame operator, after which Lemma 9 implies that the functions $(\Psi^r(f), \Psi^r(g)) \mapsto R(g) \circ f$ are tame. Since a polynomial is differentiable, hence locally Lipschitz, the desired assertion follows from these facts and Lemmas 9, 10, and 11. □

Lemma 18. *If $U, U' \subset \mathbb{R}^m$ are open, $\rho : U' \to U$ is a C^{r+} diffeomorphism, $D \subset U$ is a compact differentiation domain, and $D' = \rho^{-1}(D)$, then the operator $\gamma : f \mapsto f \circ \rho|_{D'}$ is an (r, r)-tame operator mapping $C^r(D, \mathbb{R}^n)$ into $C^r(D', \mathbb{R}^n)$.*

Proof. We need to show that $\Psi^r \circ \gamma \circ (\Psi^r)^{-1}|_{\Psi^r(C^{r+}(D,V))}$ is tame, and by Lemma 11 it suffices to show that each component is tame. As in the last proof, the basic facts of differentiation allow the component to be expressed as a polynomial function of the partials of ρ and compositions of the components and partials of f with ρ. The partials of ρ are tame, because they are constant operators (Lemma 13). Lemma 11 implies that the partials of f are tame, as functions of $\Psi^r(f)$, and the constant operator with value ρ is tame. Now Proposition 3 implies that the process of forming the composition of ρ with a component or partial of f is a composition of tame operators, and thus itself a tame operator by virtue of Lemma 9. Lemma 11 allows us to gather up these elements into a single tame operator, to which the polynomial function is applied, and since polynomial functions are locally Lipschitz, Lemma 10 (and more applications of Lemma 9) imply that the component is indeed a tame operator of $\Psi^r(f)$. □

4.4 Smooth Tameness on Manifolds

Let M and N be m- and n-dimensional C^{r+} manifolds with corners, let $D \subset M$ be a compact differentiation domain, and let $B = \{\sigma_i = (K_i, \varphi_i, \psi_i)\}_{i=1}^q$ be a C^{r+} metric configuration for (M, N, D). Say that a topological space is *locally complete* if every point has a neighborhood whose relative topology is induced by a complete metric.

Lemma 19. *$C^r(D, N)$ is locally complete.*

Proof. Fix an $f \in C^r(D, N)$, which we may suppose is an element of Z^r_B. For a small $\varepsilon > 0$ consider a d^r_B-Cauchy sequence $\{f_j\}$ in the closed ε-ball around f. For each $i = 1, \dots, q$, the sequence $\{\Gamma^r_{\sigma_i}(f_j)\}$ is a Cauchy sequence in $C^r(\varphi_i(K_i), \mathbb{R}^n)$ and consequently has a limit, which is necessarily in $C^r(\varphi_i(K_i), V_i)$ if ε is sufficiently small. If this is the case for all i, then $\{f_j\}$ has a limit in the closed ε-ball for d^r_B centered at f. □

Lemma 20. *For any* $\Lambda \geq 0$, $Z^r_{B,\Lambda}$ *is closed in* Z^r_B.

Proof. If $\{f_j\}$ is a sequence in $Z^r_{B,\Lambda}$ that converges to $f \in Z^r_B$, then for each i, $\{\Gamma^r_{\sigma_i}(f_j)\}$ is a sequence in $C^r_\Lambda(\varphi_i(K_i), \mathbb{R}^n)$ that converges to $\Gamma^r_{\sigma_i}(f)$, so $\Gamma^r_{\sigma_i}(f) \in C^r_\Lambda(\varphi_i(K_i), \mathbb{R}^n)$ by Lemma 12. □

The following is an immediate consequence of the definitions.

Lemma 21. *For each* i, $\Gamma^r_{\sigma_i} : Z^r_B \to C^r(\varphi_i(K_i), \mathbb{R}^n)$ *is* (r,r)*-tame, and if* $q = 1$*, then its inverse is also* (r,r)*-tame.*

In addition to M, N, and D, let P and Q be C^{s+} manifolds with corners, and let $E \subset P$ be a compact differentiation domain. Let S be a subset of $C^{r+}(D,N)$, and let $\gamma : S \to C^{s+}(E,Q)$ be an operator. In addition, let R and S be C^{t+} manifolds with corners (the multiple use of 'S' is temporary, and should not result in confusion) and let $F \subset R$ be a compact differentiation domain. Since compositions of locally Lipschitz operators are locally Lipschitz, due to Proposition 1, with obvious modifications the proof of Lemma 9 establishes the following result.

Lemma 22. *If* $S \subset C^{r+}(D,N)$, $S' \subset C^{s+}(E,Q)$, $\gamma : S \to S'$ *is (locally)* (r,s)*-tame, and* $\gamma' : S' \to C^{t+}(F,S)$ *is (locally)* (s,t)*-tame, then* $\gamma' \circ \gamma$ *is (locally)* (r,t)*-tame.*

Proposition 5. *If* M, N*, and* P *are* C^{r+} *manifolds with corners with* N *compact, and* $D \subset M$ *is a compact differentiation domain, then the map* $(f,g) \mapsto g \circ f$ *is a* (r,r)*-tame function from* $C^r(D,N) \times C^r(N,P)$ *to* $C^r(D,P)$.

If B is a metric configuration for (M,N,D), the derived objects Z^r_B and d^r_B are unaffected if the indexing of the triples in B is changed, or if redundant copies of some triples are added. In addition, they are unaffected if σ_i is replaced by $(K_{i1}, \varphi_i, \psi_i), \ldots, (K_{i\ell}, \varphi_i, \psi_i)$ where $K_{i1}, \ldots, K_{i\ell} \subset D$ are compact differentiation domains that cover K_i. Below we will refer to these facts collectively as the *subdivision principle*.

Proof. We first show that if $S \subset C^r(D,N) \times C^r(N,P)$ is metrically C^r Lipschitz bounded, then the restriction of composition to this set is locally Lipschitz and $\{\, g \circ f : (f,g) \in S \,\}$ is compactly C^r Lipschitz bounded. Let $B = \{\sigma_i = (K_i, \varphi_i, \psi_i)\}_{i=1}^k$ and $C = \{\tau_j = (L_j, \alpha_j, \beta_j)\}_{j=1}^\ell$ be C^{r+} metric configurations for (M,N,D) and (N,P,N) respectively, such that $S \subset Z^r_B \times Z^r_C$. This inclusion will still hold if we expand the L_j slightly, so we may assume that the interiors of these sets cover N. Since the set $\{\, \Gamma^r_{\sigma_i}(f) : (f,g) \in S \,\}$ are compactly Lipschitz bounded, by applying the subdivision principle, we can make it the case that the K_i are small enough that for each $(f,g) \in S$ and each i there is some

j_i such that $f(K_i)$ is contained in the interior of L_{j_i}. It suffices to establish the claim with S replaced by an arbitrary element of some finite cover of S, so we may assume that for each i there is some j_i such that $f(K_i)$ is contained in the interior of L_{j_i} for *all* $(f, g) \in S$.

Now $A = \{\rho_i = (K_i, \varphi_i, \beta_{j_i})\}_{i=1}^k$ is a C^{r+} metric configuration for (M, P, D), and $\{\, g \circ f : (f, g) \in S \,\} \subset Z_A^r$. For $(f, g) \in S$ we have

$$\Gamma_{\rho_i}^r(g \circ f) = \Gamma_{\psi_{j_i}}^r(g) \circ \Gamma_{\sigma_i}^r(f),$$

so Proposition 4 implies that the restriction of composition to $\{\, (\Gamma_{\sigma_i}^r(f), \Gamma_{\psi_{j_i}}^r(g)) : (f, g) \in S \,\}$ is Lipschitz and $\{\, \Gamma_{\rho_i}^r(g \circ f) : (f, g) \in S \,\}$ is compactly C^r Lipschitz bounded. Since this is true for all i, the restriction of composition to S is locally Lipschitz and $\{\, g \circ f : (f, g) \in S \,\}$ is metrically C^r Lipschitz bounded.

To show that composition is continuous, consider $f \in Z_B^r$ and $g \in Z_C^r$ with $f(K_i)$ contained in the interior of L_{j_i} for all i. There is a neighborhood $U \subset Z_B^r \times Z_C^r$ of (f, g) such that for all $(f', g') \in U$, $f'(K_i)$ contained in the interior of L_{j_i} for all i. Proposition 4 implies that each map

$$(\Gamma_{\sigma_i}^r(f'), \Gamma_{\psi_{j_i}}^r(g')) \mapsto \Gamma_{\rho_i}^r(g' \circ f')$$

from $\{\, (\Gamma_{\sigma_i}^r(f'), \Gamma_{\psi_{j_i}}^r(g')) : (f', g') \in U \,\}$ to Z_A^r is continuous, which implies that $(f', g') \mapsto g' \circ f'$ is continuous. □

5 Implicit Functions and Maximization

The optimal policy for a dynamic program is the solution of a parameterized maximization problem given by the Bellman equation. We will characterize it as the unique solution of the first order conditions, and thus as the solution of a system of equations. In this section we study operators derived from the implicit function theorem and then operators derived from maximization of smooth functions.

5.1 Implicit Functions

Let $C \subset \mathbb{R}^m$ and $D \subset C \times \mathbb{R}^n$ be compact differentiation domains. Let Z be a compact subset of D such that for each $x \in C$, $Z(x)$ is convex and has a nonempty interior. Let Ξ^{r-1} be the set of $F \in C^{r-1}(D, \mathbb{R}^n)$ such that for each $x \in C$ there is a unique $f_F(x) \in Z(x)$ such that $F(x, f_F(x)) = 0$, $(x, f_F(x))$ is in the interior of Z, and $\partial_y F(x, f_F(x))$ is nonsingular. Let $\Xi^{(r-1)+} = \Xi^{r-1} \cap C^{(r-1)+}(D, \mathbb{R}^n)$.

Lemma 23. *The sets* Ξ^{r-1} *and* $\Xi^{(r-1)+}$ *are open in* $C^{r-1}(D, \mathbb{R}^n)$ *and* $C^{(r-1)+}(D, \mathbb{R}^n)$ *respectively.*

Proof. Since $\Xi^{(r-1)+}$ has the relative topology it inherits from Ξ^{r-1}, it suffices to show that Ξ^{r-1} is open. Consider $F \in \Xi^{r-1}$, $x \in C$, a compact neighborhood $A \subset C$ of x, and an open neighborhood V of $f_F(x)$. By taking A and V small we may insure that $f_F(x') \in V$ for all $x' \in A$, and that $\partial_y F(x', y')$ is arbitrary close to $\partial_y F(x, f_F(x))$ for all $(x', y') \in A \times V$. If A and V are sufficiently small and F' is sufficiently close to F, then for all $x' \in A$, $F(x', \cdot)$ will not have a zero outside of V, and it will not have more than one zero in V because for all $y' \in V$, $\partial_y F'(x', y')$ is very close to $\partial_y F(x, f_F(x))$, but it will have a zero $f_{F'}(x') \in V$ due to the topological theory of degree, and $\partial_y F'(x', f_{F'}(x'))$ will be nonsingular.

Since C is compact, we can find $A_1, V_1, \ldots, A_k, V_k$ as above such that $A_1, \ldots, A_k$ covers C. The set of F' that satisfies the conditions above for each $i = 1, \ldots, k$ is then a neighborhood of F that is contained in Ξ^{r-1}. ☐

Proposition 6. *For each* $F \in \Xi^{r-1}(Z)$*,* $f_F : \Omega \to A$ *is* C^{r-1}*, and if* $F \in \Xi^{(r-1)+}(Z)$*, then* $f_F \in C^{(r-1)+}(\Omega, A)$*. The function* $i(F) = f_F$ *is locally* $(r-1, r-1)$*-tame.*

Proof. Fix $F \in \Xi^{r-1}(Z)$. The implicit function theorem (applied to some C^{r-1} extension of F to a neighborhood of D in $\mathbb{R}^m \times \mathbb{R}^n$) implies that $f_F \in C^{r-1}(\Omega, A)$. We will work with continuous functions $\delta, \varepsilon, \zeta : \Omega \to (0, \infty)$ and an open neighborhood $U \subset \Xi^{(r-1)+}(Z)$ of F such that:

(a) $\delta(\omega) < \varepsilon(\omega)$ for all ω.

(b) (ω, a) is in the interior of D whenever $\|a - f_F(\omega)\| \leq \varepsilon(\omega)$.

(c) $\|f_{F'}(\omega) - f_F(\omega)\| < \delta(\omega)$ for all $F' \in U$ and $\omega \in \Omega$.

(d) There is some $b > 0$ such that $\|\partial_a F'(\omega, a) w\| \geq b\|w\|$ for all $F' \in U$, $(\omega, a) \in D$ such that $\|a - f_F(\omega)\| \leq \varepsilon(\omega)$, and $w \in \mathbb{R}^n$.

(e) There is some $c \in (0, b)$ such that $\|\partial_a F'(\omega, a) - \partial_a F'(\omega, a')\| < c$ for all $F' \in U$ and $(\omega, a) \in D$ such that $\|a - f_F(\omega)\| \leq \varepsilon(\omega)$.

(f) There is some $d > 0$ such that the determinant of $\partial_a F(\omega, a)^{-1}$ is greater than d whenever $\|a - f_F(\omega)\| < \delta(\omega)$.

Such objects exist: if ε is sufficiently small, then (b) holds, δ can be chosen to satisfy (a), and (c)-(f) hold if U is a sufficiently small neighborhood of F. We will show that $i|_U$ is $(r-1, r-1)$-tame.

We first show that $i|_U : U \to C^0(\Omega, A)$ is Lipschitz. For $F_0, F_1 \in U$ and $\omega \in \Omega$, if $c : [0,1] \to \mathbb{R}^n$ is the path $c(t) = (1-t)f_{F_0}(\omega) + t f_{F_1}(\omega)$, then

$$\begin{aligned} F_0(\omega, f_{F_1}(\omega)) &= \int_0^1 \partial_a F_0(\omega, c(t)) \cdot (f_{F_1}(\omega) - f_{F_0}(\omega))\, dt \\ &= \partial_a F_0(\omega, f_F(\omega)) \cdot (f_{F_1}(\omega) - f_{F_0}(\omega)) \\ &\quad + \int_0^1 \big(\partial_a F_0(\omega, c(t)) - \partial_a F_0(\omega, f_F(\omega))\big) \cdot (f_{F_1}(\omega) - f_{F_0}(\omega))\, dt. \end{aligned}$$

We can now apply (d) and (in view of (a) and (c)) (e) to obtain

$$\begin{aligned} \|F_1(\omega, f_{F_1}(\omega)) - F_0(\omega, f_{F_1}(\omega))\| &\geq \big\|\partial_a F_0(\omega, f_F(\omega)) \cdot (f_{F_1}(\omega) - f_{F_0}(\omega))\big\| \\ &\quad - \Big\| \int_0^1 \big(\partial_a F_0(\omega, c(t)) - \partial_a F_0(\omega, f_F(\omega))\big) \cdot (f_{F_1}(\omega) - f_{F_0}(\omega))\, dt \Big\| \\ &\geq (b-c)\|f_{F_1}(\omega)) - f_{F_0}(\omega)\|, \end{aligned}$$

so $d^0_\Omega(f_{F_0}, f_{F_1}) \leq \frac{1}{b-c} d^0_D(F_0, F_1)$.

Fixing $\omega \in \Omega$, there is a number $\zeta > 0$ small enough that $\delta(\omega_1) + \|f_F(\omega_1) - f_F(\omega_0)\| < \varepsilon(\omega_0)$ for all ω_0 and ω_1 in the open ζ-ball centered at ω. If $F' \in U$ and $c : [0,1] \to \mathbb{R}^n$ is the path $c(t) = (1-t)f_{F'}(\omega_0) + t f_{F'}(\omega_1)$, then

$$\begin{aligned} F'(\omega_0, f_{F'}(\omega_1)) &= \int_0^1 \partial_a F'(\omega_0, c(t)) \cdot (f_{F'}(\omega_1) - f_{F'}(\omega_0))\, dt \\ &= \partial_a F'(\omega_0, f_{F'}(\omega_0)) \cdot (f_{F'}(\omega_1) - f_{F'}(\omega_0)) \\ &\quad + \int_0^1 (\partial_a F'(\omega_0, c(t)) - \partial_a F'(\omega_0, f_{F'}(\omega_0))) \cdot (f_{F'}(\omega_1) - f_{F'}(\omega_0))\, dt. \end{aligned}$$

In view of (a) and (c) we can apply (d) to the first term. In view of (c) and our choice of ζ, $\|f_{F'}(\omega_1) - f_F(\omega_0)\| \leq \varepsilon(\omega_0)$, and (a) and (c) also give $\|f_{F'}(\omega_0) - f_F(\omega_0)\| \leq \varepsilon(\omega_0)$, so $\|c(t) - f_F(\omega_0)\| \leq \varepsilon(\omega_0)$ for all t, and we can apply (e) to the second term. Therefore

$$\|F'(\omega_0, f_{F'}(\omega_1)) - F'(\omega_1, f_{F'}(\omega_1))\| = \|F'(\omega_0, f_{F'}(\omega_1))\| \geq (b-c) \cdot \|f_{F'}(\omega_1) - f_{F'}(\omega_0)\|.$$

The point is that any Lipschitz bound on F' implies a proportional Lipschitz bound on the restriction of $f_{F'}$ to the ζ-ball centered at ω, and we can use Lemma 4 to infer that any Lipschitz bound on F' implies a proportional Lipschitz bound on $f_{F'}$. Since our set up constrains the images of the functions $f_{F'}$ to lie in a compact set, at this point we have shown that $i|_U$ is $(r, 0)$-tame.

A consequence of the implicit function theorem is the formula

$$Df_F(\omega) = -\partial_a F(\omega, f_F(\omega))^{-1} \cdot \partial_\omega F(\omega, f_F(\omega)),$$

which is derived by totally differentiating the formula $F(\omega, f_F(\omega)) = 0$ and solving for $Df_F(\omega)$. For each partial derivative of f_F of order s, where $1 \le s \le r-1$, repeated application of the rules of differentiation gives a formula for it as a rational function[3] of the various partials of F up to order r. Cramer's rule gives a formula for each entry of $\partial_a F(\omega, f_F(\omega))^{-1}$ as a rational function of the entries of $\partial_a F(\omega, f_F(\omega))$. The denominator in this formula is the determinant of $\partial_a F(\omega, f_F(\omega))$, and the denominators in the formulas for the higher order partials of f_F are powers of it. For $F' \in U$ the determinant of $\partial_a F'(\omega, f_{F'}(\omega))$ is bounded below by (c) and (f), so $i|_U : U \to C^{r-1}(\Omega, A)$ is continuous. By Lemma 17, each of the relevant partials of F' is a tame function of F', and Lemma 11 allows these to be bundled together into a tame function. Since a rational function is differentiable on its domain of definition, hence locally Lipschitz, Lemma 10 implies that each of the partials of $f_{F'}$ up to order $r-1$ are tame functions of the partials of F'. Another application of Lemma 11 bundles these together into a tame function, thereby showing that $i|_U$ is $(r-1, r-1)$-tame. □

5.2 Maximization on Euclidean Domains

Let C, D, and Z be as in the last subsection. If $u \in C^r(D)$, let $\partial_{yy} u \in C^{r-2}(U, \mathbb{R}^{n^2})$ be the function that assigns to each (x, y) the matrix of second partial derivatives of $u(x, \cdot)$ with respect to y. Let Θ^r be the set of $u \in C^r(D)$ such that for each $x \in C$ there is a unique maximizer $\pi_u(x)$ of $u(x, \cdot)$, $\pi_u(x)$ is in the interior of $Z(x)$, $\partial_{yy} u(x, \pi_u(x))$ is nonsingular, and there is no other $y \in Z(x)$ such that $\partial_y u(x, y) = 0$. Let $\Theta^{r+} = \Theta^r \cap C^{r+}(D)$.

Lemma 24. *The sets Θ^r and Θ^{r+} are open in $C^r(D)$ and $C^{r+}(D)$ respectively.*

Proof. Since $u \mapsto \partial_y u$ is continuous and Ξ^{r-1} is open (Lemmas 17 and 23) the set of u such that $\partial_y u \in \Xi^{r-1}$ is open. The set of u such that for each x the maximizers of $u(x, \cdot)$ lie in the interior of $Z(x)$ is also open. Since Θ^r is the intersection of these two sets, it is open. Since $C^{r+}(D)$ has the relative topology induced by its inclusion in $C^r(D)$, it follows that Θ^{r+} is open. □

For $u \in \Theta^{r+}$ let V_u be the function $V_u(x) = u(x, \pi_u(x))$.

Proposition 7. *The operator $u \mapsto \pi_u$ is $(r, r-1)$-tame. For $u \in \Theta^{r+}$, $V_u \in C^{r+}(C)$, and the operator $u \mapsto V_u$ is locally (r, r)-tame.*

Proof. By Proposition 6 and Lemmas 14 and 17 the composition $u \mapsto \partial_y u \mapsto i(\partial_y u) = \pi_u$ is locally $(r, r-1)$-tame. In view of Lemma 16, to show that the operator $u' \mapsto V_{u'}$ is

[3] A *rational function* is a quotient of polynomial functions.

locally (r, r)-tame it suffices to show that it is locally $(r, 0)$-tame and that $u' \mapsto DV_{u'}$ is locally $(r, r-1)$-tame. Since $u' \mapsto V_{u'}$ is $u' \mapsto u' \circ (\mathrm{Id}_C \times \pi_{u'})$, it is locally $(r, 0)$-tame by virtue of Lemmas 14, 15, 13, and 16, Proposition 4, and the fact that $u' \mapsto \pi_{u'}$ is locally $(r, r-1)$-tame. Because the first order condition holds we have $DV_{u'} = \partial_x u' \circ (\mathrm{Id}_C \times \pi_{u'})$. (Of course this is the envelope theorem.) Therefore $u' \mapsto DV_{u'}$ is locally $(r, r-1)$-tame by virtue of Lemmas 14, 13, 16, and 17, Proposition 4, and the fact that $u' \mapsto \pi_{u'}$ is locally $(r, r-1)$-tame. □

5.3 Maximization on Manifolds

We now return to the framework of Section 3: Ω and A are C^{r+} manifolds with corners, with Ω compact, and D is a compact subset of $\Omega \times A$ such that for each $\omega \in \Omega$, the set of $a \in D(\omega)$ such that (ω, a) is in the interior of D is nonempty.

Lemma 25. *Σ^r and Σ^{r+} are open subsets of $C^r(D)$ and $C^{r+}(D)$ respectively.*

Proof. Fix a $u \in \Sigma^r$ and a $\omega \in \Omega$. Let $\varphi : U \to \mathbb{R}^n$ and $\psi : V \to \mathbb{R}^n$ be C^{r+} coordinate charts for neighborhoods of ω and $\pi_u(\omega)$ respectively, and let $\tilde{u} = u \circ (\varphi \times \psi)^{-1}$. There is a compact neighborhood $C \subset U$ of ω such that $\pi_u(\omega') \in V$ for all $\omega' \in C$. Let $Z \subset C \times V$ be a compact neighborhood of the graph of $\pi_u|_C$, and let $\tilde{C} = \varphi(C)$ and $\tilde{Z} = (\varphi \times \psi)(Z)$. By taking C and V sufficiently small, we can insure that for all $(x, y) \in \tilde{Z}$, $\partial_{yy}\tilde{u}(x, y)$ is nonsingular, and for each $\omega' \in C$, $\psi(\pi_u(\omega'))$ is the only point in $\tilde{Z}(\varphi(\omega'))$ where $\partial_y \tilde{u}(\varphi(\omega'), \cdot)$ vanishes.

For $u' \in C^r(D)$ let $\tilde{u}' = u' \circ (\varphi \times \psi)^{-1}$. Let A be the set of u' such that for all $(x, y) \in \tilde{Z}$, $\partial_{yy}\tilde{u}'(x, y)$ is nonsingular, and for each $\omega' \in C$, $\psi(\pi_{u'}(\omega'))$ is the only point in $\tilde{Z}(\varphi(\omega'))$ where $\partial_y \tilde{u}'(\varphi(\omega'), \cdot)$ vanishes. Lemma 24 implies that A is open. Let B be the set of u' such that for each $\omega' \in C$, all of the maximizers of $u'(\omega, \cdot)$ are in the interior of $Z(\omega)$. It is easy to see that B is also open. Since Ω is covered by finitely many sets C as above implies that a neighborhood of u is contained in $\Sigma^r(D)$, Since u was an arbitrary element of $\Sigma^r(D)$, this set is open.

That Σ^{r+} is open follows from the fact $C^{r+}(D)$ has the relative topology inherited from $C^r(D)$. □

Proposition 8. *The restriction of $u \to \pi_u$ to Σ^{r+} is locally $(r, r-1)$-tame, and the restriction of J to Σ^{r+} is locally (r, r)-tame.*

Proof. Fix $u \in \Sigma^{r+}$. We choose C^{r+} coordinate charts $\varphi_i : U_i \to \mathbb{R}^n$ and $\psi_i : V_i \to \mathbb{R}^n$, $i = 1, \ldots, k$, for Ω and A respectively, and compact differentiation domains $E_i \subset U_i$ and $L_i \subset V_i$, such that:

(a) the interiors of $K_1 = E_1 \times L_1, \dots, K_k = E_k \times L_k$ cover D;

(b) for some $s \le k$:

  (i) the interiors of $E_1, \dots, E_s$ cover Ω;

  (ii) for each $i = 1, \dots, s$, $\pi_u(E_i)$ is contained in the interior of L_i, so that the interiors of $K_1, \dots, K_s$ cover the graph of π_u.

Evidently $B = \{\sigma_i = (E_i, \varphi_i, \psi_i)\}_{i=1}^s$ is a C^{r+} metric configuation for (Ω, A, D), and $\pi_u \in Z_B^{r+}$. For each $i = 1, \dots, k$ let $\kappa_i = \varphi_i \times \psi_i : U_i \times V_i \to \mathbb{R}^{m+n}$. Then $\tilde{C} = \{\tau_i = (K_i, \kappa_i)\}_{i=1}^k$ is a C^{r+} metric configuration for $(\Omega \times A, \mathbb{R}, D)$, and $u \in Z_{\tilde{C}}^{r+}$. For each $i = 1, \dots, k$ let $\tilde{E}_i = \varphi_i(E_I)$, $\tilde{L}_i = \psi_i(L_i)$, and $\tilde{K}_i = \kappa_i(K_i)$, and let $\tilde{u}_i = u \circ \kappa_i^{-1}|_{\tilde{K}_i} : \tilde{K}_i \to \mathbb{R}$.

Let Z be a compact neighborhood of the graph of π_u. For each i let $Z_i = Z \cap (E_i \times V_i)$, and let $\tilde{Z}_i = \kappa_i(Z_i)$. We claim that we can choose Z small enough that for each $i = 1, \dots, s$, Z_i is contained in the cartesian product of K_i and the interior of L_i, and for each $x \in \tilde{E}_i$, $\partial_y \tilde{u}_i(x, \cdot) : \tilde{Z}_i(x) \to \mathbb{R}^n$ is injective. (If we take a countable descending sequence of neighborhoods of the graph of π_u whose intersection is the graph, then some term of the sequence is satisfactory.)

For each $i = 1, \dots, s$ the operator $u' \mapsto \tilde{u}_i'$ is continuous, so Lemma 24 implies that there is a neighborhood N_i of u such that for each $u' \in N_i$, for all $x \in \tilde{E}_i$, there is a unique maximizer $\pi_{\tilde{u}_i'}(x)$ of $\tilde{u}_i'(x, \cdot)|_{\tilde{Z}_i(x)}$, $\pi_{\tilde{u}_i'}(x)$ lies in the interior of $\tilde{Z}_i(x)$, $\partial_{yy}\tilde{u}_i(x, \pi_{\tilde{u}_i'}(x))$ is nonsingular, and $\pi_{\tilde{u}_i'}(x)$ is the only point in $\tilde{Z}_i(x)$ where $\partial_y \tilde{u}_y'(x, \cdot)$ vanishes. By replacing N_i with a smaller neighborhood of u we can make it the case that for each $i = 1, \dots, s$ and each $x \in \tilde{K}_i$, $\psi_i^{-1}(\pi_{\tilde{u}_i}(x))$ is the unique maximizer of $u_i'(\varphi_i^{-1}(x), \cdot)$. For $u_i' \in N_i$ and $x \in \tilde{E}_i$ let $V_{\tilde{u}_i'}(x) = \tilde{u}_i'(x, \pi_{\tilde{u}_i'}(x))$. From the last subsection we know that the operator $\tilde{u}_i' \mapsto \pi_{\tilde{u}_i'}$ is $(r, r-1)$-tame and the operator $\tilde{u}_i' \mapsto V_{\tilde{u}_i'}$ is (r, r)-tame.

Let $N = \bigcap_{i=1}^s N_i$. Lemmas 13 and 14 now imply that for each $i = 1, \dots, s$ the operator taking $u' \in N$ to $\pi_{\tilde{u}_i'} \circ \varphi_i|_{C_i} = \pi_{u'}|_{C_i}$ is locally $(r, r-1)$-tame and the operator taking $u' \in N$ to $V_{\tilde{u}_i'} \circ \varphi_i|_{C_i} = V_{u'}|_{C_i}$ is locally (r, r)-tame. Piecing together the various definitions, it now follows that the operator $u' \mapsto \pi_{u'}$ is locally $(r, r-1)$-tame and the operator $u' \mapsto V_{u'}$ is locally (r, r)-tame. ◻

6 The Proof of Theorem 1

We now assume the context given by the statement of Theorem 1.

Proof of Theorem 1. For $(u, \delta, \tilde{u}) \in C(D) \times [0, 1) \times C(D)$ let $I(u, \delta, \tilde{u}) = u + \delta K(J(\tilde{u}))$. If $\tilde{u}$ is a fixed point of $I(u, \delta, \cdot)$, then $J(\tilde{u})$ is a fixed point of $L(u, \delta, \cdot)$. Conversely, if V

is a fixed point of $L(u,\delta,\cdot)$ and $\tilde{u} = u + \delta K(V)$, then

$$I(u,\delta,\tilde{u}) = u + \delta K(J(u + \delta K(V)) = u + \delta K(L(u,\delta,V)) = u + \delta K(V) = \tilde{u}.$$

Let $\tilde{u}_{u_0,0} = u_0 = I(u_0, 0, \tilde{u}_{u_0,0})$. Fix a $\Lambda'' > \Lambda$. Below we will show that there are neighborhoods $W \subset Z^r_{\tilde{B},\Lambda}$, $X \subset [0,1)$, and $Y \subset Z^r_{\tilde{B},\Lambda''} \cap \Sigma^r$ of u_0, 0, and $\tilde{u}_{u_0,0}$ respectively, such that:

(a) Y is complete;

(b) $I(W \times X \times Y) \subset Y$;

(c) $I|_{W\times X\times Y}$ is locally Lipschitz;

(d) there is some $c \in (0,1)$ such that $d^r_{\tilde{B}}(I(u,\delta,\tilde{u}), I(u,\delta,\tilde{u}')) \leq c d^r_{\tilde{B}}(\tilde{u},\tilde{u}')$ for all $u \in W$, $\delta \in X$, and $\tilde{u},\tilde{u}' \in Y$.

We first show that the desired assertions follow from this. Let $N = W \times X$. Since Σ^r is open (Lemma 25) this is a neighborhood of $(u_0, 0)$ in $Z^r_{\tilde{B},\Lambda} \times [0,1)$. Of course (a) of the assertion is satisfied by construction.

Proposition 2 implies that for each $(u,\delta) \in N$ there is a unique fixed point $\tilde{u}_{u,\delta}$ of $I(u,\delta,\cdot)$, and the map $(u,\delta) \mapsto \tilde{u}_{u,\delta}$ is locally Lipschitz. Since it is a subset of Y, the image of this map is uniformly C^r Lipschitz bounded, so this map is (r,r)-tame. For $(u,\delta) \in N$, $V_{u,\delta} = J(\tilde{u}_{u,\delta})$. Since the image of $(u,\delta) \mapsto \tilde{u}_{u,\delta}$ is contained in Σ^r, Proposition 8 and Lemma 22 imply that $(u,\delta) \mapsto V_{u,\delta}$ is (r,r)-tame.

It follows that the map $(u,\delta) \mapsto u + \delta V_{u,\delta}$ is (r,r)-tame, and its image is contained in Θ^r. We have $\pi_{u,\delta} = \pi_{u+\delta V_{u,\delta}}$, so Proposition 8 and Lemma 14 imply that the operator $(u,\delta) \mapsto \pi_{u,\delta}$ is $(r, r-1)$-tame. Thus (b)-(d) of the assertion hold, and the restrictions of the operators to N are tame as asserted.

It remains to construct satisfactory W, X, and Y. To begin with let these simply be neighborhoods of u_0, 0, and $\tilde{u}_{u_0,0}$ respectively, reserving the right to replace these with smaller neighborhoods as need be. For example, we may assume that $X = [0,\overline{\delta})$ for some $\overline{\delta}$. Since Σ^r is open, we may assume that Y is a subset. Clearly we may assume that Y is metrically C^r Lipschitz bounded. By Lemma 19 we may assume that Y is complete.

Since $J|_{\Sigma^{r+}}$ is locally (r,r)-tame (Proposition 8) it is continuous, so by taking Y small we can obtain $J(Y) \subset Z^r_{\tilde{C}}$. Of course local (r,r)-tameness implies that $J|_Y$ is locally Lipschitz and $J(Y)$ is compactly C^r Lipschitz bounded. We can force $J(Y)$ into a neighborhood of $V_{u_0,0}$ of the sort considered in Proposition 1, from which it follows that $J(Y) \subset Z^r_{\tilde{C},M_V}$ for some $M_V > 0$. Since K is (r,r)-tame, $K|_{J(Y)}$ is locally Lipschitz, and by replacing Y with a smaller neighborhood we can insure that $J|_Y$ and $K|_{J(Y)}$ are actually Lipschitz.

Since P is r^{th}-order smoothing, $K(J(Y))$ is compactly Lipschitz bounded, and by making Y small, $K(J(Y))$ can be forced into a neighborhood of $K(J(u_0))$ the sort considered in Proposition 1, in which case, as above, there is some $Q > 0$ such that $K(J(Y)) \subset Z^r_{\tilde{B},Q}$. Since we can make $\overline{\delta}$ as small as we like, we can assume that $\overline{\delta}K(J(Y)) \subset Z^r_{\tilde{B},\Lambda''-\Lambda}$. In view of the obvious additive property of Lipschitz constants for sums of Lipschitz functions, it follows that $I(W \times X \times Y) \subset Z^r_{\tilde{B},\Lambda''}$. If W and Y are closed balls around u_0 and $\tilde{u}_{u_0,0} = u_0$ respectively, and the radius is W is smaller, then making $\overline{\delta}$ small enough gives $I(W \times X \times Y) \subset Y$. Thus W, X, and Y have all required properties. □

References

Araujo, P. (1991). The once but not twice differentiability of the policy function. *Econometrica*, 59:1383–1393.

Benveniste, L. M. and Scheinkman, J. A. (1979). On the differentiability of the value function in dynamic models of economics. *Econometrica*, 47:727–732.

Blackwell, D. (1965). Discounted dynamic programming. *Annals of Mathematical Statistics*, 36:226–235.

Blume, L., Easley, D., and O'Hara, M. (1982). Characterization of optimal plans for stochastic dynamic programs. *Journal of Economic Theory*, 28:221–234.

Boldrin, M. and Montrucchio, L. (1986). On the indeterminacy of capital accumulation paths. *Journal of Economic Theory*, 40:26–39.

Cerf, J. (1961). Topologie de certains espaces de plongements. *Bull Soc. Math. France*, 89:227–380.

Douady, A. (1961). Variétés à bord anguleux et voisinages tubulaires. *Séminaire Henri Cartan*, 14: exp. 1, 1–11.

Hirsch, M. (1976). *Differential Topology.* Springer-Verlag, New York.

Loertscher, S. and McLennan, A. (2015). Some people never learn, rationally: Multidimensional learning traps. University of Queensland working paper.

Santos, M. S. (1991). Smoothness of the policy function in discrete time economic models. *Econometrica*, 59:1365–1382.