On the Generic Structure and Stability of Stackelberg Equilibria

Alberto Bressan and Yilun Jiang

Department of Mathematics, Penn State University.
University Park, PA 16802, USA.

e-mails: axb62@psu.edu, yxj141@psu.edu

July 19, 2019

Abstract

We consider a non-cooperative Stackelberg game where the two players choose their strategies within domains $X \subseteq \mathbb{R}^m$ and $Y \subseteq \mathbb{R}^n$. Assuming that the cost functions F, G for the two players are sufficiently smooth, we study the structure of the best reply map for the follower and the optimal strategy for the leader. Two main cases are considered: either $X = Y = [0,1]$, or $X = \mathbb{R}$, $Y = \mathbb{R}^n$ with $n \geq 1$. Using techniques from differential geometry, including a multi-jet version of Thom's transversality theorem, we prove that for an open dense set of cost functions $F \in \mathcal{C}^2$ and $G \in \mathcal{C}^3$ the Stackelberg equilibrium is unique and is stable w.r.t. small perturbations of the two cost functions.

Key words: Non-cooperative game, Stackelberg equilibrium, stability, generic property.

1 Introduction

In the theory of non-cooperative games, the concept of Stackelberg equilibrium [18] has been widely investigated, due to its several applications to economic models [2]. In a basic setting, a game for two players can be formulated as follows.

- Player 1 (the leader) chooses $x \in X$ and seeks to minimize his cost $F(x, y)$.
- Player 2 (the follower) chooses $y \in Y$ and seeks to minimize his cost $G(x, y)$.

Here X, Y are topological spaces, while $F, G : X \times Y \mapsto \mathbb{R}$ are continuous functions.

For a given $x \in X$, the set of **best replies** for the follower is defined as

$$R(x) \doteq \Big\{ y^* \in Y\,; \quad G(x, y^*) \leq G(x, y) \ \text{ for all } y \in Y \Big\}. \tag{1.1}$$

We say that a couple $(x^*, y^*) \in X \times Y$ is a **Stackelberg equilibrium** if $y^* \in R(x)$ and

$$F(x^*, y^*) \ \le\ F(x,y) \qquad \text{for all } x \in X \text{ and } y \in R(x). \tag{1.2}$$

This models a situation where the leading player announces his strategy $x \in X$ in advance, and the follower chooses a reply $y \in Y$ which minimizes his own cost $G(x,y)$.

In the literature, the existence of a Stackelberg equilibrium is known under fairly general assumptions [2, 4, 10, 13, 17]. A major related issue is the uniqueness and stability of this equilibrium. Namely, if the cost functions F, G are slightly perturbed, does the new game still have a unique solution, close to the original one? This problem has been investigated in [10, 12, 13], within the general class of continuous cost functions. As pointed out in [10], it is not possible to obtain, under sufficiently general conditions, existence and stability results for the exact Stackelberg solutions. For this reason, in the above papers, a weaker concept of ϵ-solution was used.

Aim of the present paper is to study stability for the best reply map and for exact Stackelberg solutions, within a class of smooth functions $F, G : \mathbb{R}^m \times \mathbb{R}^n \mapsto \mathbb{R}$. In this setting, examples of games with multiple equilibria are easy to construct. However, our main results show that, for "most" functions F, G (in a topological sense), the Stackelberg equilibrium is unique and is stable under small perturbations. While the results in [10, 12, 13] are based general topological principles, our stability results rely on completely different techniques, stemming from differential geometry; namely: Sard's theorem and a multi-jet version of Thom's transversality theorem [3, 7].

The analysis of Stackelberg equilibria can be accomplished in two steps.

(i) Study the graph of the best reply map $R(\cdot)$, namely

$$\text{Graph}(R) \ =\ \{(x,y)\,;\ y \in R(x)\} \ \subset\ \mathbb{R}^m \times \mathbb{R}^n. \tag{1.3}$$

Show that, for a generic function $G \in \mathcal{C}^3(R^{m+n})$, this graph can be expressed in terms of finitely many equalities or inequalities, in generic position.

(ii) Study the constrained minimization problem for the function F, restricted to Graph(R). Show that, for a generic function $F \in \mathcal{C}^2(R^{m+n})$, a unique global minimum exists, which is stable under small perturbations.

We recall that a property is said to be **generic** if it holds on the intersection of countably many open dense sets. The main goal of the present paper is to study the stability of the equilibrium (x^*, y^*) under perturbations of the cost functions F, G, in the following sense.

Definition. *Given the cost functions F, G, we say that the Stackelberg equilibrium (x^*, y^*) is* **strongly stable** *if, for any $\varepsilon > 0$ there exists $\delta > 0$ such that the condition*

$$\|\widetilde{F} - F\|_{\mathcal{C}^2} \ \le\ \delta, \qquad\qquad \|\widetilde{G} - G\|_{\mathcal{C}^3} \ \le\ \delta, \tag{1.4}$$

implies that the perturbed game, with F, G replaced by $\widetilde{F}, \widetilde{G}$, has a unique Stackelberg equilibrium $(\tilde{x}, \tilde{y}) \in X \times Y$. Moreover,

$$|\tilde{x} - x^*| \ \le\ \varepsilon, \qquad\qquad |\tilde{y} - y^*| \ \le\ \varepsilon. \tag{1.5}$$

We first consider the basic case where $X = Y = [0,1]$. Using Thom's transversality theorem we prove that the set of couples $(F,G) \in \mathcal{C}^2 \times \mathcal{C}^3$ that yield a unique, strongly stable Stackelberg equilibrium is open and dense. In a later section we prove similar results in the case where the strategy of the follower takes values in a multi-dimensional set. Namely: $X = \mathbb{R}$ while $Y = \mathbb{R}^n$.

The remainder of the paper is organized as follows. Section 2 collects some basic stability results on the stability of the minimizer of a $\mathcal{C}^2$ function F restricted to a set $\mathcal{A} \subset \mathbb{R}^N$ which is defined in terms of finitely many equalities or inequalities. In Section 3 we begin the analysis of the best reply map. When $X = \mathbb{R}^m$, $Y = \mathbb{R}^n$, the necessary conditions for optimality imply that

$$\mathrm{Graph}(R) \;\subseteq\; \mathcal{M} \;\doteq\; \{(x,y) \in \mathbb{R}^{m+n}\,;\;\; \nabla_y G(x,y) = 0\}. \tag{1.6}$$

As a first step, we observe that, by Sard's theorem, for a generic function $G \in \mathcal{C}^3$ the set $\mathcal{M}$ in (1.6) is a $\mathcal{C}^2$ manifold. Theorem 4.1 in Section 4 describes in detail the generic structure of the best reply map, in the one-dimensional case where $X = Y = [0,1]$. In turn, this yields the generic stability of the Stackelberg equilibrium, proved in Section 5.

The analysis in Section 6 shows that similar results on the structure of the best reply map and on the stability of the Stackelberg equilibrium still hold, in the case where the follower chooses his strategy within a multi-dimensional space. Namely, $X = \mathbb{R}$ while $Y = \mathbb{R}^n$. Section 7 contains some concluding remarks. In particular, we discuss the possible extension of our results to the case where the strategy of the leader also lies in a multi-dimensional set $X \subseteq \mathbb{R}^m$. Finally, an Appendix collects some basic results from differential geometry. A multi-jet version of Thom's transversality theorem is proved, which provides a key ingredient for our analysis.

In connection with Nash equilibria for special classes of non-cooperative games, generic properties of solutions have been studied in [9, 14].

2 Minima of generic functions on generic sets

In a Stackelberg game, the leader seeks to minimize his own cost $F(x,y)$, restricted to the graph of the best reply map $R(\cdot)$. As it will be shown in a later section, for a generic function G in (1.1), this graph can be expressed in terms of a finite set of equalities and inequalities.

We thus consider here a constrained minimization problem of the general form

$$\min_{x \in \mathcal{A}} f(x), \tag{2.1}$$

under the following assumptions.

(B1) *$f : \mathbb{R}^N \mapsto \mathbb{R}$ is twice continuously differentiable. Moreover,* $\lim_{|x| \to \infty} f(x) = +\infty$.

(B2) *$\mathcal{A} \subset \mathbb{R}^N$ is a nonempty closed set, described in terms of finitely many equalities and inequalities:*

$$\mathcal{A} \;=\; \Big\{x \in \mathbb{R}^N\,;\;\; \phi_i(x) = 0, \;\; \psi_j(x) \geq 0, \qquad i \in \mathcal{I}, \;\; j \in \mathcal{J}\Big\}. \tag{2.2}$$

(B3) *All functions ϕ_i, ψ_j are in $\mathcal{C}^2(\mathbb{R}^N)$. Moreover, for any subset $\mathcal{J}' \subseteq \mathcal{J}$ and any point x such that*

$$\phi_i(x) = \psi_j(x) = 0 \qquad \textit{for all} \quad i \in \mathcal{I}, \quad j \in \mathcal{J}', \tag{2.3}$$

the gradients

$$\nabla\phi_i(x), \quad \nabla\psi_j(x), \qquad\qquad i \in \mathcal{I}, \quad j \in \mathcal{J}', \tag{2.4}$$

are linearly independent.

Roughly speaking, we would like to prove that, for all functions f, ϕ_i, ψ_j in an open dense set of $\mathcal{C}^2$, the problem (2.1) admits a unique minimizer, which is stable under perturbations. With this goal in mind, we denote by $\mathcal{F}^\infty$ the family of all functions f satisfying **(B1)**. Since $\mathcal{F}^\infty$ is not a vector space, a suitable topology must first be defined.

We recall that $\mathcal{C}^k(\mathbb{R}^N)$ is a Banach space with the norm

$$\|f\|_{\mathcal{C}^k} \ \doteq\ \sup_{|\alpha|\leq k}\ \sup_{x\in\mathbb{R}^N} \Big|D^\alpha f(x)\Big|. \tag{2.5}$$

Using a standard notation (see for example [8]), here

$$D^\alpha f(x) \ \doteq\ \frac{\partial^{|\alpha|}}{\partial_{x_1}^{\alpha_1}\cdots\partial_{x_N}^{\alpha_N}} f(x)$$

denotes a partial derivative of order $|\alpha| = \alpha_1 + \cdots + \alpha_N$. On the set $\mathcal{F}^\infty$ we now consider the distance

$$d_2(f,g) \ \doteq\ \min\Big\{1,\ \|f-g\|_{\mathcal{C}^2}\Big\}. \tag{2.6}$$

With this distance, $(\mathcal{F}^\infty, d_2)$ is a complete metric space. Notice that the convergence $f_n \to f$ holds iff $\|f_n - f\|_{\mathcal{C}^2} \to 0$.

Next, we introduce

Definition 2.1 *Under the assumptions* **(B2)-(B3)***, we say that the global minimum (2.1) is attained at a point $\bar{x}$* **in generic position** *if the following conditions hold.*

(i) $\bar{x} \in \mathcal{A}$ is the unique point where the global minimum is attained.

(ii) Setting $\mathcal{J}' \doteq \{j \in \mathcal{J}\,;\ \ \psi_j(\bar{x}) = 0\}$, there exists constants α_i, β_j, $i \in \mathcal{I}$, $j \in \mathcal{J}'$ such that

$$\nabla f(\bar{x}) \ =\ \sum_{i\in\mathcal{I}} \alpha_i\,\nabla\phi_i(\bar{x}) + \sum_{j\in\mathcal{J}'} \beta_j\,\nabla\psi_j(\bar{x}), \tag{2.7}$$

with $\beta_j > 0$ for all $j \in \mathcal{J}'$.

(iii) There exists $\rho, \varepsilon > 0$ such that

$$f(x) - f(\bar{x}) \ \geq\ \varepsilon\,|x-\bar{x}|^2 \qquad \textit{for all } x \in \mathcal{A} \textit{ with } |x-\bar{x}| \leq \rho\,. \tag{2.8}$$

We remark that, by the assumption **(B3)**, the equations (2.3) define a manifold $\mathcal{M}_{\mathcal{I}\cup\mathcal{J}'} \subset \mathbb{R}^N$ of dimension $N - |\mathcal{I}| - |\mathcal{J}'|$. By assumption, the restriction of f to this manifold has a local minimum at $\bar{x}$. According to (2.8), the Hessian matrix $D^2 f(\bar{x})$ has full rank (hence it is strictly positive definite) at $\bar{x}$. We remark that this rank condition is invariant under coordinate changes on the submanifold $\mathcal{M}_{\mathcal{I}\cup\mathcal{J}'}$.

As shown by the following theorem, minima attained in generic position (according to Definition 2.1) are stable w.r.t. small $\mathcal{C}^2$ perturbations in the cost function f or in the constraints φ_i, ψ_j.

Theorem 2.1 *Let $f, \phi_i, \psi_j \in \mathcal{C}^2(\mathbb{R}^N)$, $i \in \mathcal{I}$, $j \in \mathcal{J}$, be such that the assumptions* **(B1)-(B3)** *hold. Moreover, assume that the global minimum (2.1) is attained at a point $\bar{x}$ in generic position. Then there exists $\delta > 0$ such that, if*

$$\|\tilde{f} - f\|_{\mathcal{C}^2} \le \delta, \qquad \|\tilde{\phi}_i - \phi_i\|_{\mathcal{C}^2} \le \delta, \qquad \|\tilde{\psi}_j - \psi_j\|_{\mathcal{C}^2} \le \delta \tag{2.9}$$

for all $i \in \mathcal{I}$, $j \in \mathcal{J}$, then the corresponding optimization problem

$$\min_{x\in\widetilde{\mathcal{A}}} \tilde{f}(x) \tag{2.10}$$

has a unique minimizer $\tilde{x}$, also in generic position. Here $\widetilde{\mathcal{A}}$ is the set defined as in (2.2), with ϕ_i, ψ_j replaced by $\tilde{\phi}_i, \tilde{\psi}_j$, respectively. Moreover, for some constant C one has

$$|\tilde{x} - \bar{x}| \le C \cdot \max\Big\{\|\tilde{f} - f\|_{\mathcal{C}^2}, \ \|\tilde{\phi}_i - \phi_i\|_{\mathcal{C}^2}, \ \|\tilde{\psi}_j - \psi_j\|_{\mathcal{C}^2}; \quad i \in \mathcal{I}, \ j \in \mathcal{J}\Big\}. \tag{2.11}$$

Proof. 1. As a first step we claim that, for every $r > 0$, there exists $\delta > 0$ such that the inequalities in (2.9) imply that (2.10) has a minimizer $\tilde{x}$ with

$$|\tilde{x} - \bar{x}| \le r. \tag{2.12}$$

Indeed, by **(B1)-(B3)** and since $\bar{x}$ is the unique global minimizer of f, by choosing $\delta_1 > 0$ sufficiently small we have the implication

$$\|\tilde{f} - f\|_{\mathcal{C}^2} \le \delta_1 \qquad \Longrightarrow \qquad \sup_{|x-\bar{x}|\le\delta_1} \tilde{f}(x) < \inf_{|x-\bar{x}|\ge r} \tilde{f}(x). \tag{2.13}$$

Next, by choosing $\delta_2 > 0$ small enough, we have the implication

$$\begin{array}{l}\|\tilde{\phi}_i - \phi_i\|_{\mathcal{C}^2} \le \delta_2, \qquad \|\tilde{\psi}_j - \psi_j\|_{\mathcal{C}^2} \le \delta_2 \qquad \text{for all } i \in \mathcal{I}, \quad j \in \mathcal{J} \\ \qquad \Longrightarrow \qquad \widetilde{\mathcal{A}} \cap B(\bar{x}, \delta_1) \neq \emptyset \qquad \text{and} \qquad \tilde{\psi}_j(x) > 0 \quad \text{for all } x \in B(\bar{x}, \delta_1), \quad j \notin \mathcal{J}'.\end{array} \tag{2.14}$$

Moreover, $B(\bar{x}, \delta_1)$ denotes the open ball centered at $\bar{x}$ with radius δ_1.

Now choose any $x_1 \in B(\bar{x}, \delta_1) \cap \widetilde{\mathcal{A}}$. For any x such that $|x - \bar{x}| \ge r$ by (2.13) it follows $\tilde{f}(x) > \tilde{f}(x_1)$. Hence the global minimum of $\tilde{f}$ cannot be attained outside $B(\bar{x}, r)$.

2. We now prove (2.11). Let $\tilde{x} \in B(\bar{x}, \delta_1)$ be a global minimizer of $\tilde{f}$, restricted to $\widetilde{\mathcal{A}}$. The first order necessary condition for optimality imply

$$\nabla \tilde{f}(\tilde{x}) \;=\; \sum_{i\in\mathcal{I}} \tilde{a}_i \nabla \widetilde{\phi}_i(\tilde{x}) + \sum_{j\in\mathcal{J}^\sharp} \tilde{b}_j \nabla \widetilde{\psi}_j(\tilde{x}), \tag{2.15}$$

for some coefficients $\tilde{a}_i, \tilde{b}_j$. Here $\mathcal{J}^\sharp \subseteq \mathcal{J}'$ is the set of indices $j \in \mathcal{J}$ such that $\widetilde{\psi}_j(\tilde{x}) = 0$.

By continuity, choosing $0 < \delta < \min\{\delta_1, \delta_2\}$ small enough, we obtain

$$\mathcal{J}^\sharp \;=\; \mathcal{J}', \qquad\qquad \tilde{b}_j \;>\; 0 \quad \text{for all } j \in \mathcal{J}'.$$

3. We now apply the implicit function theorem to the map

$$\Lambda : \;(\phi, \psi, x) \;\mapsto\; (\phi_i(x),\, \psi_j(x))_{i\in\mathcal{I}, j\in\mathcal{J}'}\,, \tag{2.16}$$

from a space $\mathcal{C}^2 \times \mathcal{C}^2 \times \mathbb{R}^N$ into $\mathbb{R}^{|\mathcal{I}|+|\mathcal{J}'|}$. We choose coordinates

$$x \;=\; (x', x'') \;=\; (x_1, \ldots, x_\nu\,,\; x_{\nu+1}, \ldots, x_N), \tag{2.17}$$

with $\nu = N - |\mathcal{I}| - |\mathcal{J}'|$ such that the equation $\Lambda(\phi, \psi, x', x'') = 0$ defines a $\mathcal{C}^2$ function

$$x'' \;=\; \varphi(x'),$$

for x' in a neighborhood of $\bar{x}'$.

By the implicit function theorem, one has

$$D_{x'}\varphi \;=\; -\left[\frac{\partial \Lambda(\phi,\psi)}{\partial x''}\right]^{-1} \frac{\partial \Lambda(\phi,\psi)}{\partial x'}.$$

By continuity, for all $\delta > 0$ small enough the matrix of partial derivatives $\frac{\partial \Lambda(\tilde{\phi},\tilde{\psi})}{\partial x''}$ still has full rank for all $x = (x', x'')$ in a neighborhood of $\bar{x} = (\bar{x}', \bar{x}'')$. Hence, the vector equation

$$\Lambda(\widetilde{\phi}, \widetilde{\psi}, x', x'') \;=\; 0,$$

defined as in (2.16), determines a $\mathcal{C}^2$ function

$$x'' \;=\; \tilde{\varphi}(x'), \qquad \text{with} \qquad D_{x'}\tilde{\varphi} = -\left[\frac{\partial \Lambda(\tilde{\phi},\tilde{\psi})}{\partial x''}\right]^{-1} \frac{\partial \Lambda(\tilde{\phi},\tilde{\psi})}{\partial x'},$$

for x' in a neighborhood of $\bar{x}'$.

Computing the second order derivatives of the implicit functions φ and $\tilde{\varphi}$, by (2.9) we obtain an estimate of the form

$$\|\tilde{\varphi} - \varphi\|_{\mathcal{C}^2(B(\bar{x}',\rho))} \;\le\; C_0\delta, \tag{2.18}$$

for some constant C_0.

4. Now consider the map $x' \mapsto F(x') \doteq f(x', \varphi(x'))$. By the assumption **(B1)**, this has a strict minimum at $x' = \bar{x}'$. Moreover, by (2.8) the Hessian matrix $D^2_{x'}F(\bar{x}')$ is strictly positive definite.

By continuity, for $\delta > 0$ small enough, if (2.9) holds then the corresponding function $\widetilde{F}(x') = \tilde{f}(x', \tilde{\varphi}(x'))$ has strictly positive definite Hessian matrix $D^2_{x'}\widetilde{F}(x')$ at every point x' in a neighborhood of $\bar{x}'$. By possibly shrinking the value of δ, we conclude that $D^2_{x'}\widetilde{F}(\tilde{x}')$ is strictly positive definite. Observing that (2.15) holds with $\tilde{b}_j > 0$ for all $j \in \mathcal{J}'$, we obtain

$$\tilde{f}(x) - \tilde{f}(\tilde{x}) \;\geq\; \tilde{\varepsilon}\,|x - \tilde{x}|^2 \qquad \text{for all } x \in \tilde{\mathcal{A}} \text{ with } |x - \tilde{x}| \leq \tilde{\rho}\,, \tag{2.19}$$

for some $\tilde{\varepsilon}, \tilde{\rho} > 0$. Hence the global minimum of $\tilde{f}$ on $\widetilde{\mathcal{A}}$ is attained at a point $\tilde{x}$ in generic position.

5. It remains to prove the inequality (2.11), showing that the minimizer depends in a Lipschitz continuous way on the functions f, ϕ_i, ψ_j, w.r.t. the $\mathcal{C}^2$ norm. This follows from the first order necessary conditions for optimality, together with the implicit function theorem. Indeed, the point $\bar{x} = (\bar{x}', \bar{x}'')$ where the constrained minimum is attained is uniquely determined by the N equations

$$\phi_i(\bar{x}) = 0, \qquad \psi_j(\bar{x}) \;=\; 0, \qquad \nabla_{\bar{x}'} F(\bar{x}', \varphi(\bar{x}')) \;=\; 0. \tag{2.20}$$

In other words, $\bar{x}$ can be represented as a zero of the map

$$\Gamma : (f, \phi, \psi, x) \;\mapsto\; \Big(\phi_i(x),\, \psi_j(x),\, \partial_{x_k} F(x', \varphi(x'))\Big) \;\in\; \mathbb{R}^N. \tag{2.21}$$

Here $i \in \mathcal{I}$, $j \in \mathcal{J}'$, $k \in \{1, \ldots, \nu\}$ with $\nu = N - |\mathcal{I}| - |\mathcal{J}'|$, as in (2.17). We regard (2.21) as a map from $\mathcal{C}^2 \times \mathcal{C}^2 \times \mathcal{C}^2 \times \mathbb{R}^N$ into $\mathbb{R}^N$. In order to apply the implicit function theorem on a Banach space [6] and achieve the estimate (2.11), it suffices to check that the $N \times N$ matrix of partial derivatives $\frac{\partial \Gamma}{\partial x}$ has full rank in a neighborhood of $(f, \phi, \psi, \bar{x})$. Indeed, by the condition (2.8) one has

$$F(x', \varphi(x')) \;\geq\; F(\bar{x}', \varphi(\bar{x}')) + \varepsilon |x' - \bar{x}'|^2$$

for all x' in a neighborhood of $\bar{x}'$. This implies that the $\nu \times \nu$ Hessian matrix $D^2_{x'} F(x', \varphi(x'))$ has full rank at $\bar{x}'$. Furthermore, since the $\nu \times (N - \nu)$ matrix of partial derivatives

$$\left(\frac{\partial [\phi_i, \psi_j]}{\partial x''}\right)_{i \in \mathcal{I}, j \in \mathcal{J}'}$$

also has full rank at $\bar{x}''$, the $N \times N$ matrix of the partial derivatives $\frac{\partial \Gamma}{\partial x}$ has full rank at $(f, \phi, \psi, \bar{x})$. By continuity, $\frac{\partial \Gamma}{\partial x}$ has full rank in a neighborhood of $(f, \phi, \psi, \bar{x})$. Observing that Γ is Lipschitz continuous w.r.t. f, ϕ_i, ψ_j (in the $\mathcal{C}^2$ distance), this achieves the proof of (2.11). □

We now show that the conditions (i)–(iii) in Definition 2.1 are "generic". Indeed, they hold for an open, dense set of $\mathcal{C}^2$ functions f, φ, ψ.

Theorem 2.2 *Let $\phi_i, \psi_j \in \mathcal{C}^2(\mathbb{R}^N)$, $i \in \mathcal{I}$, $j \in \mathcal{J}$ be such that the assumptions* **(B2)-(B3)** *hold. Then, for an open dense set of functions $f \in \mathcal{F}^\infty$, the global minimum in (2.1) is attained in generic position.*

Proof. 1. Consider any $f \in \mathcal{F}^\infty$. Let $\bar{x} \in \mathcal{A}$ be a point where f attains its global minimum. Consider the set of indices

$$\mathcal{J}' \;\doteq\; \{j \in \mathcal{J}\,;\;\; \psi_j(\bar{x}) = 0\}.$$

Introduce two smooth functions $\rho, \eta : \mathbb{R}_+ \mapsto [0,1]$, satisfying

$$\rho(s) \;=\; \begin{cases} 1 & \text{if } \; s \in [0, r_0], \\ 0 & \text{if } \; s \geq 2r_0 \,, \end{cases} \qquad\qquad \rho'(s) \leq 0 \quad \text{for all } s > 0, \tag{2.22}$$

$$\eta(s) \;=\; \begin{cases} s^2 & \text{if } \; s \in [0, 1/2], \\ 1 & \text{if } \; s \geq 1 \,, \end{cases} \qquad\qquad \eta'(s) \geq 0 \quad \text{for all } s > 0, \tag{2.23}$$

We then define a family of perturbed functions

$$f_\varepsilon(x) \;=\; f(x) + \varepsilon \rho(|x - \bar{x}|) \cdot \sum_{j \in \mathcal{J}'} \psi_j(x) + \varepsilon \eta(|x - \bar{x}|) \tag{2.24}$$

Choosing $r_0 > 0$ suitably small, it follows that $f_\varepsilon(x) > f(x)$ for all $x \neq \bar{x}$ and $\varepsilon > 0$.

Thanks to the properties of the cut-off functions ρ, η we have $\|f_\varepsilon - f\|_{\mathcal{C}^2} \to 0$ as $\varepsilon \to 0$.

2. It remains to check that, for every $\varepsilon > 0$, the global minimum of f_ε, which is attained at the single point $\bar{x}$, is in generic position. If $\bar{x}$ is a global minimizer of $f(\cdot)$ on $\mathcal{A}$, the first order necessary condition (2.7) implies

$$\nabla f_\varepsilon(\bar{x}) \;=\; \nabla f(\bar{x}) + \varepsilon \sum_{j \in \mathcal{J}'} \nabla \psi_j(\bar{x}) \;=\; \sum_{i \in \mathcal{I}} \alpha_i \nabla \phi_i(\bar{x}) + \sum_{j \in \mathcal{J}'} (\beta_j + \varepsilon) \nabla \psi_j(\bar{x}).$$

Therefore, for every $\varepsilon > 0$, the condition (2.7) is satisfied.

On the other hand, the inequality (2.8) follows from

$$f_\varepsilon(x) \;\geq\; f(x) + \varepsilon |x - \bar{x}|^2 \;\geq\; f_\varepsilon(\bar{x}) + \varepsilon |x - \bar{x}|^2 \qquad \text{for all } x \in \mathcal{A} \text{ with } |x - \bar{x}| \leq 1/2.$$

3. The two previous steps show that the set of functions f, for which the minimun is attained in generic position, is dense on $\mathcal{F}^\infty$. The fact that it is open follows from Theorem 2.1. □

Corollary 2.1 *Consider the set $\mathcal{F}^\sharp$ of functions $(f, \phi_i, \psi_j)_{i \in \mathcal{I}, j \in \mathcal{J}}$ such that either (i) the domain $\mathcal{A}$ in (2.2) is empty, or else (ii) $\mathcal{A} \neq \emptyset$ and the minimization problem (2.1) has a unique minimizer in generic position. Then $\mathcal{F}^\sharp$ is open and dense in $\mathcal{F}^\infty \times \mathcal{C}^2(\mathbb{R}^N) \times \ldots \times \mathcal{C}^2(\mathbb{R}^N)$.*

Proof. 1. The openness of the set of functions (f, ϕ_i, ψ_j) that satisfy the conditions in Definition 2.1 is again a consequence of Theorem 2.1.

2. To prove that this set is dense in the $\mathcal{C}^2$ topology, in view of Theorem 2.2 it suffices to show that the set of $\mathcal{C}^2$ functions $(\phi_i, \psi_j)_{i \in \mathcal{I}, j \in \mathcal{J}}$ that satisfy the condition **(B3)** is dense. Toward this goal, let functions $\phi_i, \psi_j \in \mathcal{C}^2(\mathbb{R}^N)$ be given. For each subset $\mathcal{J}' \subset \mathcal{J}$ consider the map

$$x \;\mapsto\; (\phi_i(x),\, \psi_j(x))_{i \in \mathcal{I}, j \in \mathcal{J}'}$$

from $\mathbb{R}^N$ into $\mathbb{R}^{|\mathcal{I}| + |\mathcal{J}'|}$. By Sard's theorem [3, 7], a.e. $(y, z) = (y_i, z_j) \in \mathbb{R}^{|\mathcal{I}|} \times \mathbb{R}^{|\mathcal{J}'|}$ is a regular value of this map. Hence, taking the perturbations

$$\tilde{\phi}_i(x) \;=\; \phi_i(x) + y_i \,, \qquad\qquad \tilde{\psi}_j(x) \;=\; \psi_j(x) + z_j \,,$$

one has the alternative:

- either $(\tilde\phi_i(x), \tilde\psi_j(x))_{i\in\mathcal{I}, j\in\mathcal{J}'} \neq (0,\ldots,0) \in \mathbb{R}^{|\mathcal{I}|+|\mathcal{J}'|}$,
- or else $(\tilde\phi_i(x), \tilde\psi_j(x))_{i\in\mathcal{I}, j\in\mathcal{J}'} = (0,\ldots,0)$ and the matrix of partial derivatives

$$\left(\frac{\partial[\phi_i, \psi_j]}{\partial x_k}\right)_{i\in\mathcal{I}, j\in\mathcal{J}', 1\le k\le N}$$

  has rank $|\mathcal{I}| + |\mathcal{J}'|$.

3. We now observe that the second alternative cannot hold if $|\mathcal{I}| + |\mathcal{J}'| > N$. Since the vector (y_i, z_j) can be taken arbitrarily small, we conclude that there exists an open dense set of functions $(\phi_i, \psi_j)_{i\in\mathcal{I}, j\in\mathcal{J}}$ such that, for any $\mathcal{J}' \subseteq \mathcal{J}$, the following conditions hold.

- If $|\mathcal{I}| + |\mathcal{J}'| > N$, then $(\phi_i(x), \psi_j(x))_{i\in\mathcal{I}, j\in\mathcal{J}'} \neq (0,\ldots,0) \in \mathbb{R}^{|\mathcal{I}|+|\mathcal{J}'|}$ for every $x \in \mathbb{R}^N$,
- If $(\phi_i(x), \psi_j(x))_{i\in\mathcal{I}, j\in\mathcal{J}'} = (0,\ldots,0) \in \mathbb{R}^{|\mathcal{I}|+|\mathcal{J}'|}$, then the $|\mathcal{I}| + |\mathcal{J}'|$ gradients $\nabla\phi_i(x)$, $\nabla\psi_j(x)$ are linearly independent.

This achieves the proof. □

Remark 2.1 Although the family of sets $\mathcal{A}$ that can be represented as in **(B2)-(B3)** is quite general, there are simple examples where the graph of best reply map does not fit within this framework. For example, let $X = Y = [0, 1]$ and take $G(x, y) = y^2 - 4xy$. Then the graph of the best reply map is

$$\text{graph}(R) = \{x \in [0, 1/2], \quad y = 2x\} \cup \{x \in [1/2, 1], \quad y = 1\}. \tag{2.25}$$

For this reason, we need to extend the previous results to this slightly more general setting. In this direction, we observe that the set in (2.25) can be equivalently written as

$$\text{graph}(R) = \{1 - y \ge 0, \quad 2x - y = 0\} \cup \{1 - y = 0, \quad 2x - y \ge 0\}. \tag{2.26}$$

In the following, we shall thus consider more general domains $\mathcal{A}$ which admit the following characterization.

(C) *There exists a finite open covering $\mathbb{R}^N = V_1 \cup \cdots \cup V_m$ such that, for each $k \in \{1, \ldots, m\}$, the intersection $\mathcal{A} \cap V_k$ admits a representation in terms of finitely many functions $\phi_i \in \mathcal{C}^2(\mathbb{R}^N)$, $i \in I$. Namely, the following properties hold:*

(i) For any $x \in V_k$ and any subset $I' \subseteq I$, if $\phi_i(x) = 0$ for all $i \in I'$ then the gradients $\nabla\phi_i(x)$, $i \in I'$, are linearly independent.

(ii) There exists finitely many subsets of indices $I_1, \ldots, I_\nu \subseteq I$, such that

$$\mathcal{A} \cap V_k \doteq \mathcal{A}_1 \cup \cdots \cup \mathcal{A}_\nu \tag{2.27}$$

where, for each $\ell = 1, \ldots, \nu$,

$$\mathcal{A}_\ell = \{x \in V_k\,; \quad \phi_i(x) = 0, \quad \phi_j(x) \ge 0 \quad \textit{for all} \quad i \in I_\ell, \quad j \notin I_\ell\}. \tag{2.28}$$

Notice that the example (2.26) fits in this framework, with $\phi_1 = 1 - y$, $\phi_2 = 2x - y$.

We now extend Definition 2.1 to this more general setting.

Definition 2.2 *Consider a domain $\mathcal{A}$ which admits the characterization in* **(C)**. *We say that the global minimum (2.1) is attained at a point $\bar{x}$* **in generic position** *if the following conditions hold.*

(i) *$\bar{x} \in \mathcal{A}$ is the unique point where the global minimum is attained.*

(ii) *Let $\bar{x} \in V_k$, so that (2.27)-(2.28) holds. Then, defining $J \doteq \{j \in I\,;\quad \phi_j(\bar{x}) = 0\}$, for any ℓ such that $\bar{x} \in \mathcal{A}_\ell$, there exists constants α_i, β_j, $i \in I_\ell$, $j \in J \setminus I_\ell$ such that*

$$\nabla f(\bar{x}) \;=\; \sum_{i \in I_l} \alpha_i \,\nabla \phi_i(\bar{x}) + \sum_{j \in J \setminus I_l} \beta_j \,\nabla \phi_j(\bar{x}), \tag{2.29}$$

with $\beta_j > 0$ for all $j \in J \setminus I_\ell$.

(iii) *There exists $\rho, \varepsilon > 0$ such that*

$$f(x) - f(\bar{x}) \;\geq\; \varepsilon\,|x - \bar{x}|^2 \qquad \text{for all } x \in \mathcal{A} \text{ with } |x - \bar{x}| \leq \rho\,. \tag{2.30}$$

The results proved in Theorem 2.1 and in Theorem 2.2 can be easily extended to this more general setting.

Corollary 2.2 *Let $f \in \mathcal{C}^2(\mathbb{R}^N)$ and let $\mathcal{A} \subset \mathbb{R}^N$ be a compact domain which admits the characterization* **(C)**. *Moreover, assume that the global minimum (2.1) is attained at a point $\bar{x}$ in generic position. Then the conclusions in Theorem 2.1 remain valid.*

Proof. Assume that the global minimum is attained at a point $\bar{x} \in \mathcal{A} \cap V_k$, so that (2.27)-(2.28) holds. If $\bar{x} \in \mathcal{A}_\ell$, then the proof of Theorem 2.1 shows that, for any sufficiently small perturbations $\tilde{f}, \tilde{\phi}_i$, the corresponding minimum is still achieved at a point in $\mathcal{A}_\ell$. The conclusion is thus obtained by applying Theorem 2.1 to each domain $\mathcal{A}_\ell$ which contains $\bar{x}$. □

Corollary 2.3 *Let $\mathcal{A} \subset \mathbb{R}^N$ be a compact domain which admits the characterization* **(C)**. *Then, for an open dense set of functions $f \in \mathcal{F}^\infty$, the global minimum in (2.1) is attained in generic position.*

Proof. Given the open covering $\mathbb{R}^N = V_1 \cup \cdots \cup V_m$, we can find a new covering $\mathbb{R}^N = V_1' \cup \cdots \cup V_m'$, where each $V_k' \subset V_k$ is a closed set. For any given $k \in \{1, \ldots, m\}$, consider the representation (2.27)-(2.28). Using Theorem 2.2 we obtain an open dense set of functions $\mathcal{F}_{\ell,k}$ such that, if $f \in \mathcal{F}_{\ell,k}$ and the global minimum of f on $\mathcal{A}$ is achieved at a point $\bar{x} \in \mathcal{A}_\ell \cap V_k'$, then it is attained in generic position. Taking the intersection of these finitely many open dense sets of functions, the result is proved. □

3 Generic properties of the best reply map

Starting with this section, we consider a cost function for the follower $G : \mathbb{R}^m \times \mathbb{R}^n \mapsto \mathbb{R}$, and study the structure of the **best reply map:** $x \mapsto R(x) \subset \mathbb{R}^n$. We seek a description of the graph

$$\mathcal{A} \doteq \text{Graph}(R) = \{(x,y)\,;\; y \in R(x)\} \tag{3.1}$$

for a generic function $G \in \mathcal{C}^3$. The eventual goal is to show that, for a generic function G, the above graph can be expressed in terms of finitely many equalities or inequalities, as in (2.2)–(2.4). In the following, our basic assumptions will be

(A1) *The function $G = G(x,y)$ lies in $\mathcal{C}^3(\mathbb{R}^m \times \mathbb{R}^n)$.*

(A2) *There exists $\rho > 0$ such that*

$$G(x,y) > G(x,0) \qquad \textit{for all } x \in \mathbb{R}^m,\; |y| \geq \rho, \tag{3.2}$$

We shall denote by $\mathcal{G}$ the family of all functions $G = G(x,y)$ which satisfy **(A1)-(A2)**. Notice that, if $G \in \mathcal{G}$, then by (3.2), for each $x \in \mathbb{R}^m$ the set of best replies $R(x)$ is a nonempty compact set contained in the open ball centered at the origin with radius ρ, namely

$$R(x) \subseteq B_\rho \subset \mathbb{R}^n. \tag{3.3}$$

As a consequence, the admissible set $\mathcal{A}$ in (3.1) is closed. We seek to understand the structure of this set $\mathcal{A}$, for a generic function $G \in \mathcal{C}^3(\mathbb{R}^{m+n})$. As a preliminary, we observe that the necessary conditions for optimality imply

$$\mathcal{A} \subseteq \mathcal{M} \doteq \Big\{(x,y) \in \mathbb{R}^{m+n}\,;\; \nabla_y G(x,y) = 0\Big\}. \tag{3.4}$$

Here and in the sequel we write $\nabla_y G = (G_{y_1}, \ldots, G_{y_n})$. The next lemma establishes the generic regularity of $\mathcal{M}$.

Lemma 3.1 *Let $\kappa > 0$ be given. There exists an open, dense subset $\mathcal{G}^\sharp \subset \mathcal{G}$ such that, for every $G \in \mathcal{G}^\sharp$, the set*

$$\mathcal{M}_\kappa \doteq \Big\{(x,y)\,;\; \nabla_y G(x,y) = 0, \quad |x| < \kappa\Big\} \tag{3.5}$$

is an m-dimensional $\mathcal{C}^2$ manifold, embedded in $\mathbb{R}^{m+n}$.

Proof. 1. Given $G \in \mathcal{G}$ and $\varepsilon > 0$, by a mollification procedure we can construct $g \in \mathcal{G} \cap \mathcal{C}^\infty$ with $\|g - G\|_{\mathcal{C}^3} < \varepsilon$.

2. Consider the map $(x,y) \mapsto \nabla_y g(x,y)$ from $\mathbb{R}^{m+n}$ into $\mathbb{R}^n$. By Sard's theorem [3, 7], the set of critical values of this map has measure zero. As a consequence, we can find $\theta = (\theta_1, \ldots, \theta_n)$ with $|\theta| < \varepsilon$ such that, at every point where $\nabla_y g(x,y) = \theta$, the $n \times (m+n)$ Jacobian matrix $D(\nabla_y g)$ has full rank, namely

$$\text{rank} \begin{pmatrix} \partial_{x_1}\partial_{y_1} g & \cdots & \partial_{x_m}\partial_{y_1} g & \partial_{y_1}\partial_{y_1} g & \ldots & \partial_{y_n}\partial_{y_1} g \\ \vdots & & \cdots & & & \vdots \\ \partial_{x_1}\partial_{y_n} g & \cdots & \partial_{x_m}\partial_{y_n} g & \partial_{y_1}\partial_{y_n} g & \ldots & \partial_{y_n}\partial_{y_n} g \end{pmatrix} = n. \tag{3.6}$$

3. We can now consider a smooth function $\tilde{g} : \mathbb{R}^{m+n} \mapsto \mathbb{R}$ such that

$$\tilde{g}(x,y) \;=\; g(x,y) - \sum_{i=1}^{n} \theta_i y_i \qquad\qquad |x| \leq \kappa, \quad |y| \leq \rho.$$

If $\varepsilon > 0$ was chosen sufficiently small, this can be extended to the entire space, still remaining in $\mathcal{G}$.

The above construction yields a function $\tilde{g}$, arbitrarily close to G in the $\mathcal{C}^3$ norm, for which the following implication holds:

$$|x| \leq \kappa, \quad |y| \leq \rho, \quad \nabla_y \tilde{g}(x,y) \;=\; 0 \qquad \Longrightarrow \qquad \mathrm{rank}\big(D(\nabla_y \tilde{g})(x,y)\big) \;=\; n. \tag{3.7}$$

By the implicit function theorem, this implies that the set

$$\widetilde{\mathcal{M}}_\kappa \;\doteq\; \Big\{(x,y)\,;\; \nabla \tilde{g}_y(x,y) = 0, \quad |x| < \kappa\Big\}$$

is still a smooth manifold.

4. Let now $\mathcal{G}^\sharp$ be the set of all functions $G \in \mathcal{G}$ for which the implication (3.7) holds. By the previous steps, this set is dense in $\mathcal{G}$. It remains to show that $\mathcal{G}^\sharp$ is open.

Assume, on the contrary, that there exists a sequence of functions $G_k \notin \mathcal{G}^\sharp$, with $G_k \to G$ in $\mathcal{C}^3$ and $G \in \mathcal{G}^\sharp$. This implies that, for each $k \geq 1$, there exists (x_k, y_k) with

$$|x_k| \leq \kappa, \quad |y_k| \leq \rho, \quad \nabla G_k(x_k, y_k) \;=\; 0, \qquad \mathrm{rank}\big(D(\nabla_y G_k)(x_k, y_k)\big) \;<\; n.$$

Taking a subsequence we can assume the convergence $x_k \to x$, $y_k \to y$. By continuity, it follows

$$|x| \leq \kappa, \quad |y| \leq \rho, \quad \nabla G(x,y) \;=\; 0, \qquad \mathrm{rank}\big(D(\nabla_y G)(x,y)\big) \;<\; n.$$

against the assumptions. This contradiction completes the proof. □

4 The best reply map with one-dimensional strategies

In this section we study the generic structure of the best reply map, and the stability of the Stackelberg equilibrium, starting with a simple one-dimensional framework. Namely, we assume that the strategies x, y for both the leader and the follower range over a closed interval, say

$$x \in X = [0,1], \qquad\qquad y \in Y = [0,1]. \tag{4.1}$$

Given a function $G \in \mathcal{C}^3(\mathbb{R}^2)$, consider the best reply map

$$R(x) \;\doteq\; \Big\{y^* \in [0,1]\,;\;\; G(x,y^*) \leq G(x,y) \quad \text{for all} \;\; y \in [0,1]\Big\}. \tag{4.2}$$

By Lemma 3.1 there exists an open dense subset $\mathcal{G}^\sharp \subset \mathcal{C}^3(\mathbb{R}^2)$ such that, for every $G \in \mathcal{G}^\sharp$, the set

$$\mathcal{M} \;\doteq\; \Big\{(x,y)\,;\; G_y(x,y) = 0, \quad |x| < 2, \;\; |y| < 2\Big\} \tag{4.3}$$

is a $\mathcal{C}^2$ manifold. Indeed, in analogy with (3.7), for a generic function $G \in \mathcal{C}^3$ we have

$$|x| \leq 2, \quad |y| \leq 2, \quad G_y(x,y) = 0 \qquad \Longrightarrow \qquad \nabla G_y(x,y) \neq 0. \tag{4.4}$$

We now prove a structure theorem for the best reply map, valid for a generic cost function G of the follower.

Theorem 4.1 *There exists an open dense subset $\mathcal{G} \subset \mathcal{C}^3(\mathbb{R}^2)$ of cost functions such that, if $G \in \mathcal{G}$ then the best reply map (4.2) has the following structure.*

There exists finitely many points $0 = x_0 < x_1 < \cdots < x_\nu = 1$, and functions $\varphi_k \in \mathcal{C}^2(\mathbb{R})$ such that

$$\Big\{(x,y)\,; \quad y \in R(x)\,, \quad x \in [0,1]\Big\} \;=\; \bigcup_{k=1}^{\nu} \Big\{(x, \varphi_k(x))\,; \quad x \in [x_{k-1}, x_k]\Big\}. \tag{4.5}$$

For each $k = 1, \ldots, \nu - 1$, one has

$$G\big(x_k, \varphi_k(x_k)\big) \;=\; G\big(x_k, \varphi_{k+1}(x_k)\big), \qquad \frac{d}{dx} G\big(x, \varphi_k(x)\big)\bigg|_{x=x_k} > \frac{d}{dx} G\big(x, \varphi_{k+1}(x)\big)\bigg|_{x=x_k}. \tag{4.6}$$

Either $\varphi_k(x_k) \neq \varphi_{k+1}(x_k)$, or else

$$\varphi_k(x_k) \;=\; \varphi_{k+1}(x_k) \;\in\; \{0,1\}, \qquad \varphi_k'(x_k) \;\neq\; \varphi_{k+1}'(x_k). \tag{4.7}$$

Moreover, for each $k = 1, \ldots, \nu$ one of the following three cases must occur.

(i) $\varphi_k(x) \equiv 0$ and $G_y(x,0) > 0$ for all $x \in]x_{k-1}, x_k[$,

(ii) $\varphi_k(x) \equiv 1$ and $G_y(x,1) < 0$ for all $x \in]x_{k-1}, x_k[$,

(iii) $0 < \varphi_k(x) < 1$ for all $x \in]x_{k-1}, x_k[$. In this case one has $G_y(x, \varphi_k(x)) = 0$ and $G_{yy}(x, \varphi_k(x)) > 0$ for all $x \in [x_{k-1}, x_k]$.

Proof. 1. We introduce a set of conditions such that, if none of them is satisfied, (for any choice of x, y, y_1, y_2, y_3 in $[0,1]$), then the representation (4.5) holds (Fig. 1). By showing that each of these conditions is NOT satisfied by all functions G in an open dense subset of $\mathcal{C}^3$, the theorem will be proved.

(i) $G_y(x,y) = 0$, $G_{yy}(x,y) = 0$, and $x \in \{0,1\}$ or $y \in \{0,1\}$.

(ii) $G_y(x,y) = 0$, $G_{yx}(x,y) = 0$, and $x \in \{0,1\}$ or $y \in \{0,1\}$.

(iii) $G_y(x,y) = 0$, $x \in \{0,1\}$ and $y \in \{0,1\}$.

(iv) $G_y(x,y) = G_{yy}(x,y) = G_{yyy}(x,y) = 0$.

(v) $G_y(x,y) = G_{yy}(x,y) = G_{xy}(x,y) = 0$.

(vi) $G_y(x,y_1) = G_y(x,y_2) = 0$, $G(x,y_1) = G(x,y_2)$, $G_{yy}(x,y_1) = 0$, for some $y_1 \neq y_2$.

(vii) $G_y(x,y_1) = G_y(x,y_2) = 0$, $G(x,y_1) = G(x,y_2)$, $G_x(x,y_1) = G_x(x,y_2)$, for some $y_1 \neq y_2$.

(viii) $G_y(x,y_1) = 0$, $G_{yy}(x,y_1) = 0$, $G(x,y_1) = G(x,y_2)$, $y_2 \in \{0,1\}$, for some $y_1 \neq y_2$.

(ix) $G_y(x,y_1) = 0$, $G(x,y_1) = G(x,y_2)$, $G_x(x,y_1) = G_x(x,y_2)$, $y_2 \in \{0,1\}$, for some $y_1 \neq y_2$.

(x) $G_y(x,y_1) = G_y(x,y_2) = 0$, $G(x,y_1) = G(x,y_2)$, $y_2 \in \{0,1\}$, for some $y_1 \neq y_2$.

(xi) $G_y(x,y_1) = 0$, $G(x,y_1) = G(x,y_2)$, $y_1, y_2 \in \{0,1\}$, for some $y_1 \neq y_2$.

(xii) $G(x,y_1) = G(x,y_2)$, $G_x(x,y_1) = G_x(x,y_2)$, $y_1, y_2 \in \{0,1\}$, for some $y_1 \neq y_2$.

(xiii) There are three distinct points (x,y_1), (x,y_2), (x,y_3) such that $G(x,y_1) = G(x,y_2) = G(x,y_3)$ and for each $i = 1,2,3$ one has either $G_y(x,y_i) = 0$ or $y_i \in \{0,1\}$.

As the reader will easily check, each of these conditions involves a number of identities which is strictly larger than the corresponding number of variables. Hence, for "most" functions G, this set of equations will have no solution. As shown in the following steps, a rigorous proof of this fact can be achieved thanks to a multi-jet version of Thom's transversality theorem.

2. The conditions (i)–(v) are all handled in a similar way. Given a function $G \in \mathcal{C}^\infty(\mathbb{R}^2)$, its third order jet prolongation is the vector function whose components are all its derivatives up to order three:

$$j^3G(x,y) \;=\; (G,\; G_x, G_y,\; G_{xx}, G_{xy}, G_{yy},\; G_{xxx}, G_{xxy}, G_{xyy}, G_{yyy})(x,y). \tag{4.8}$$

The map j^3G is thus a section of the vector bundle $J^3(\mathbb{R}^2,\mathbb{R})$ of all third order jets of maps from $\mathbb{R}^2$ into $\mathbb{R}$.

For each of the conditions in (i)–(v) we shall consider a smooth submanifold $W \subset J^3(\mathbb{R}^2,\mathbb{R})$. This will be defined in terms of three independent equalities, hence it will have codimension 3. By Thom's transversality theorem, there is a dense set of $\mathcal{C}^\infty$ functions G whose prolongation j^3G is transversal to W. Since j^3G is a section of $J^3(\mathbb{R}^2,\mathbb{R})$, it is a two-dimensional manifold. In this case, transversality implies that the intersection is empty. In other words, for a dense set of $\mathcal{C}^\infty$ functions G, the three identities that define W are never simultaneously satisfied.

For example, for condition (i) we consider four distinct sub-manifolds. Each of them is defined by the two identities

$$G_y = 0, \qquad G_{yy} = 0,$$

together with one of the four equalities $x = 0$, $x = 1$, $y = 0$, or $y = 1$. Condition (ii) is entirely similar.

For condition (iii) we need again to consider four distinct sub-manifolds. Each of them is defined by the identity $G_y = 0$, plus a choice of $x \in \{0,1\}$ and $y \in \{0,1\}$.

To handle condition (iv), it suffices to consider the linear sub-manifold $W \subset J^3(\mathbb{R}^2;\mathbb{R})$ consisting of all jets such that $G_y = 0$, $G_{yy} = 0$, $G_{yyy} = 0$. Finally, condition (v) is handled by defining $W \subset J^2(\mathbb{R}^2;\mathbb{R})$ to be the linear manifold of all jets such that $G_y = 0$, $G_{yy} = 0$, $G_{xy} = 0$.

3. Conditions (vi)–(xi) refer to the values of G and its first two derivatives at two different points. For this reason, we shall need a multi-jet transversality theorem, proved in the Appendix. We start by introducing the manifold

$$Z^{(2)} \doteq \big\{(x, y_1, y_2)\,;\quad y_1 \neq y_2\big\}.$$

On $Z^{(2)}$ we consider the multi-jet bundle $\widehat{J}^2_2(\mathbb{R}^2, \mathbb{R})$, consisting of couples of 2-jets of maps from $\mathbb{R}^2$ to $\mathbb{R}$ with sources (x, y_1), (x, y_2), $y_1 \neq y_2$. Notice that every function $G \in \mathcal{C}^\infty(\mathbb{R}^2, \mathbb{R})$ determines a map

$$\widehat{j}^2_2 G : (x, y_1, y_2) \;\mapsto\; \big(j^2 G(x, y_1)\,,\; j^2 G(x, y_2)\big). \tag{4.9}$$

Each of the conditions (vi) and (vii) yields a manifold $W \subset \widehat{J}^2_2(\mathbb{R}^2, \mathbb{R})$, consisting of multijets which satisfy the four given identities. By Theorem 8.1, there is a dense set of functions $G \in \mathcal{C}^\infty(\mathbb{R}^2, \mathbb{R})$ whose second order jet prolongation $\widehat{j}^2_2 G$ is transversal to W.

Since W is defined in terms of four identities, it has codimension 4. On the other hand, the graph of $\widehat{j}^2_2 G$ has dimension 3. In this case, transversality implies that the intersection is empty. In other words, for a dense set of functions G, the four conditions in (v) or in (vi) are not simultaneously satisfied at any couple of distinct points (x, y_1), (x, y_2).

For each of the conditions (viii), (ix), and (x) we obtain two distinct sub-manifolds $W_0, W_1 \subset \widehat{J}^2_2(\mathbb{R}^2, \mathbb{R})$, imposing the equality $y_2 = 0$ or $y_2 = 1$, respectively. Again, by Theorem 8.1, there is a dense set of functions $G \in \mathcal{C}^\infty(\mathbb{R}^2, \mathbb{R})$ whose second order jet prolongation $\widehat{j}^2_2 G$ is transversal to W_0 or W_1, respectively. By dimensionality this implies that, for such functions G, for every couple of points $(x, y_1) \neq (x, y_2)$ at least one of the four conditions in (viii) fails. Similarly, at least one of the four conditions in (ix) and at least one in (x) must fail.

Condition (xi) leads to four affine sub-manifolds W, each of codimension 4, depending on the choices of $y_1, y_2 \in \{0, 1\}$. The analysis is entirely similar to the previous cases. Condition (xii) is entirely straightforward.

4. Condition (xiii) refers to the values of G and its first derivatives at three different points. For this reason, we introduce the manifold

$$Z^{(3)} \doteq \{(x, y_1, y_2, y_3)\,;\quad y_i \neq y_j \;\text{ for }\; i < j\}.$$

On $Z^{(3)}$ we consider the multi-jet bundle $\widehat{J}^1_3(\mathbb{R}^2, \mathbb{R})$, consisting of couples of 1-jets of maps from $\mathbb{R}^2$ to $\mathbb{R}$ with three distinct sources (x, y_1), (x, y_2), and (x, y_3). Notice that every function $G \in \mathcal{C}^\infty(\mathbb{R}^2, \mathbb{R})$ determines a map

$$\widehat{j}^1_3 G : (x, y_1, y_2, y_3) \;\mapsto\; \Big(j^1 G(x, y_1)\,,\; j^1 G(x, y_2)\,,\; j^1 G(x, y_3)\Big).$$

On $\widehat{J}^1_3(\mathbb{R}^2, \mathbb{R})$ we consider a finite number of sub-manifolds W, each defined by 5 identities. The first two identities are

$$G(x, y_1) \,=\, G(x, y_2), \qquad\qquad G(x, y_1) \,=\, G(x, y_3).$$

The remaining three identities are obtained by choosing, for each $i = 1, 2, 3$, either $G_y(x, y_i) = 0$, or $y_i = 0$, or else $y_i = 1$.

By Theorem 8.1, there is a dense set of functions $G \in \mathcal{C}^\infty(\mathbb{R}^2, \mathbb{R})$ whose first order jet prolongation $\widehat{j}^1_3 G$ is transversal to any of the above manifolds W.

We now observe that each W is defined in terms of five identities, and thus has codimension 5. On the other hand, the graph of $\widehat{j}^1_3 G$ has dimension 4. Once again, transversality implies that the intersection is empty. In other words, for a dense set of functions G, the five conditions in (xiii) are not simultaneously satisfied at any triple of distinct points (x, y_1), (x, y_2), (x, y_3).

5. In this step we prove that the set of functions $G \in \mathcal{C}^3(\mathbb{R}^2, \mathbb{R})$ for which none of the conditions (i)–(xiii) is satisfied on the domain $Q = \{(x, y) \in [0,1] \times [0,1]\}$ is open in $\mathcal{C}^3$.

Let $(G^{(n)})_{n\geq 1}$ be a sequence of functions converging to G in $\mathcal{C}^3$. Assume that, for every n, at least one of the conditions (i)–(xiii) is satisfied, within the domain Q. We need to show that the same holds for G.

We start with the easy case where, for some sequence $(x^{(n)}, y^{(n)}) \in Q$, each $G^{(n)}$ satisfies one of the conditions (i)–(v). By taking a subsequence, we can assume $(x^{(n)}, y^{(n)}) \to (\bar{x}, \bar{y}) \in Q$. By continuity, the limit function G satisfies the same condition at $(\bar{x}, \bar{y})$, proving our claim.

Concerning the remaining conditions (vi)–(xiii), a more careful analysis is needed, because these conditions involve two or three distinct points.

Assume that, for each $n \geq 1$, the function $G^{(n)}$ satisfies one of the conditions (vi)–(xiii), for distinct points $(x^{(n)}, y_i^{(n)})$, $i = 1, 2, 3$. By taking a subsequence, we can assume the convergence

$$(x^{(n)}, y_i^{(n)}) \;\to\; (\bar{x}, \bar{y}_i), \qquad\qquad i = 1, 2, 3.$$

If the limit points $(\bar{x}, \bar{y}_i)$ are distinct, then by continuity G still satisfies the same condition, and we are done. Notice that this is certainly the case for (xi) and (xii), because here we require $y_1^{(n)}, y_2^{(n)} \in \{0, 1\}$ with $y_1^{(n)} \neq y_2^{(n)}$.

To complete the proof, we need to consider the cases where two of the points $(\bar{x}, \bar{y}_i)$ coincide, say,

$$\lim_{n\to\infty} y_1^{(n)} \;=\; \lim_{n\to\infty} y_2^{(n)} \;=\; \bar{y} \;\in [0, 1]. \tag{4.10}$$

- If $G^{(n)}$ satisfies all identities in (vi), for every $n \geq 1$, then by taking limits we conclude

  $$G_y(\bar{x}, \bar{y}) \;=\; G_{yy}(\bar{x}, \bar{y}) \;=\; G_{yyy}(\bar{x}, \bar{y}) \;=\; 0. \tag{4.11}$$

  Hence the limit function G satisfies (iv).

- If $G^{(n)}$ satisfies all identities in (vii), for every $n \geq 1$, then

  $$G_y(\bar{x}, \bar{y}) \;=\; G_{yy}(\bar{x}, \bar{y}) \;=\; G_{xy}(\bar{x}, \bar{y}) \;=\; 0.$$

  Hence the limit function G satisfies (v).

- Next, assume that $G^{(n)}$ satisfies all identities in(viii), or in (ix), or in (x), for every $n \geq 1$. Taking the limit, in all cases we conclude

  $$G_y(\bar{x}, \bar{y}) \;=\; G_{yy}(\bar{x}, \bar{y}) \;=\; 0, \qquad \bar{y} \in \{0, 1\}.$$

  Hence (i) holds.

- Finally, assuming that all functions $G^{(n)}$ satisfy (xiii), different cases need to be considered.

  If all three sequences of points $y_1^{(n)}$, $y_2^{(n)}$, $y_3^{(n)}$, converge to the same limit $\bar{y}$, then the convergence $G^{(n)} \to G$ in $\mathcal{C}^3$ implies (4.11). Hence the limit function G satisfies (iv).

  The remaining possibility is that

$$y_1^{(n)}, y_3^{(n)} \to \bar{y}_1 , \qquad\qquad y_2^{(n)} \to \bar{y}_2 \neq \bar{y}_1 . \tag{4.12}$$

  Four sub-cases must be considered.

  - If $\bar{y}_1 \in \{0,1\}$ and $\bar{y}_2 \in \{0,1\}$, then

$$G_y(\bar{x}, \bar{y}_1) \;=\; 0, \qquad G(\bar{x}, \bar{y}_1) \;=\; G(\bar{x}, \bar{y}_2), \qquad \bar{y}_1, \bar{y}_2 \in \{0,1\}.$$

  Hence all identities in (xi) hold.

  - If $\bar{y}_1 \notin \{0,1\}$ and $\bar{y}_2 \in \{0,1\}$, then

$$G_y(\bar{x}, \bar{y}_1) = 0, \qquad G_{yy}(\bar{x}, \bar{y}_1) = 0, \qquad G(\bar{x}, \bar{y}_1) = G(\bar{x}, \bar{y}_2),$$

  for some $y_1 \neq y_2 \in \{0,1\}$. Hence (viii) holds.

  - If $\bar{y}_1 \in \{0,1\}$ and $\bar{y}_2 \notin \{0,1\}$, then

$$G_y(\bar{x}, \bar{y}_1) = G_y(\bar{x}, \bar{y}_2) = 0, \qquad\qquad G(\bar{x}, \bar{y}_1) = G(\bar{x}, \bar{y}_2),$$

  for some $y_1 \neq y_2 \in \{0,1\}$. Hence (x) holds.

  - If $\bar{y}_1 \notin \{0,1\}$ and $\bar{y}_2 \notin \{0,1\}$, then

$$G_y(\bar{x}, \bar{y}_1) = G_{yy}(\bar{x}, \bar{y}_1) = G_y(\bar{x}, \bar{y}_2) = 0, \qquad\qquad G(\bar{x}, \bar{y}_1) = G(\bar{x}, \bar{y}_2),$$

  Hence (vi) holds.

6. To conclude the proof of the theorem, consider a function $G \in \mathcal{C}^3(\mathbb{R}^2, \mathbb{R})$ in the open, dense set where none of the conditions (i)–(xiii) holds. We claim that, in this case, the best reply map satisfies the conclusions of the theorem.

By the necessary conditions for optimality, the graph of the best reply map is a closed set, contained in the union of the three sets

$$\{(x,y)\,;\; G_y(x,y) = 0, \quad x, y \in [0,1]\} \;\cup\; \{(x,0)\,;\; x \in [0,1]\} \;\cup\; \{(x,1)\,;\; x \in [0,1]\}.$$

Consider any point $\bar{x} \in [0,1]$. By (xii), the minimum of the function $y \mapsto G(\bar{x}, y)$ over $[0,1]$ can be attained at most at two distinct points. Various cases will be considered in the remaining steps.

7. We first assume that the global minimum is attained at a single point $\bar{y}$. Two main cases can occur.

CASE 1: $\bar{y} \in]0,1[$. We claim that in this case $G_{yy}(\bar{x}, \bar{y}) \neq 0$, hence for x in a neighborhood of $\bar{x}$ the best reply map is single valued:

$$R(x) \;=\; \{\phi(x)\},$$

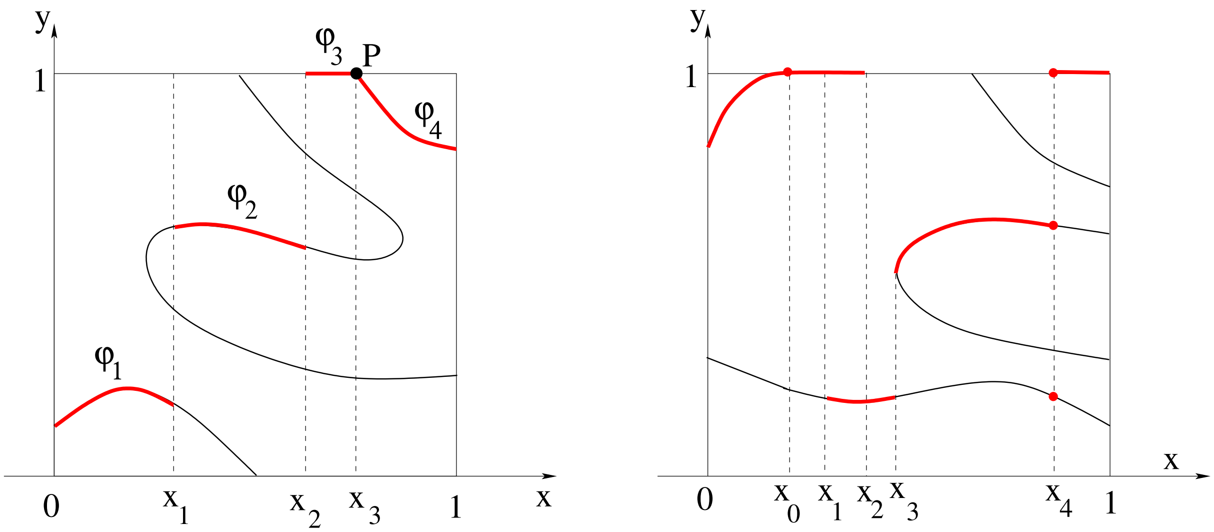

Figure 1: Left: the graph of the "best reply map" (in red), for a generic cost function G. Right: a sample of non-generic cases. At x_0 the curve where $G_y = 0$ touches the line $y = 1$ tangentially, so that (ii) holds. On the whole interval $[x_1, x_2]$ the function $G(x, \cdot)$ has two equal minimizers, and (ix) holds. At the point x_3 two global minima are attained, where one of these is along a curve where $G_y = 0$, with vertical tangent. This happens when $G_{yy} = 0$, so that (iv) holds. At x_4 the function $G(x_4, \cdot)$ achieves the minimum at three distinct points, hence (xiii) holds.

where $y = \phi(x)$ is the function implicitly defined by

$$G_y(x,y) \;=\; 0. \tag{4.13}$$

Indeed, assume on the contrary that $G_{yy}(\bar{x}, \bar{y}) = 0$. Since we also have $G_y(\bar{x}, \bar{y}) = 0$, by (i) it follows that $\bar{x} \notin \{0, 1\}$. Moreover, from (iv) and (v) it follows that

$$G_{xy}(\bar{x},\bar{y}) \;\neq\; 0, \qquad G_{yyy}(\bar{x},\bar{y}) \;\neq\; 0. \tag{4.14}$$

Hence, by (4.14), the equation (4.13) can be solved in a neighborhood of $(\bar{x}, \bar{y})$ in terms of a function $x = \psi(y)$, with

$$\psi(\bar{y}) \;=\; \bar{x}, \qquad \psi'(\bar{y}) \;=\; -\frac{G_{yy}(\bar{x},\bar{y})}{G_{xy}(\bar{x},\bar{y})} \;=\; 0, \qquad \psi''(\bar{y}) \;=\; -\frac{G_{yyy}(\bar{x},\bar{y})}{G_{xy}(\bar{x},\bar{y})} \;\neq\; 0.$$

To fix the ideas, assume $\psi''(\bar{y}) > 0$, the other case being entirely similar. Since $\bar{x} > 0$, we can find a strictly increasing sequence $x_n \uparrow \bar{x}$, with $x_n > 0$ for every n. Call $y_n \in [0, 1]$ a point where $G(x_n, \cdot)$ attains its global maximum. This implies that either $y_n \in \{0, 1\}$ or else $G_y(x_n, y_n) = 0$. Choosing a subsequence, we achieve the convergence $(x_n, y_n) \to (\bar{x}, y^*)$ for some $y^* \in [0, 1]$. By continuity, $G(x^*, \cdot)$ attains its global minimum at y^*. Hence, by uniqueness, $y^* = \bar{y}$. Since $\bar{y} \notin \{0, 1\}$, we conclude that $G_y(x_n, y_n) = 0$ for all sufficiently large n. This yields a contradiction, because the equation (4.13) does not admit any solution with $x < \bar{x}$, in a suitably small neighborhood of $(\bar{x}, \bar{y})$.

CASE 2: $\bar{y} = 0$. (The case where where $\bar{y} = 1$ is entirely similar.)

In this case, a necessary condition is $G_y(\bar{x}, 0) \geq 0$. If $G_y(\bar{x}, 0) > 0$, then we immediately conclude that $R(x) = \{0\}$ on an entire neighborhood of $\bar{x}$.

The remaining case is where $G_y(\bar{x},0) = 0$. By (i) and (ii) we then have $G_{yy}(\bar{x},0) \neq 0$ and $G_{xy}(\bar{x},0) \neq 0$. Hence, in a neighborhood of $(\bar{x},0)$, the equation (4.13) uniquely defines a function $y = \phi(x)$, with

$$\phi(\bar{x}) \;=\; 0, \qquad\qquad \phi'(\bar{x}) \;=\; -\,\frac{G_{xy}(\bar{x},0)}{G_{yy}(\bar{x},0)} \neq 0.$$

To fix the ideas, let $\phi'(\bar{x}) > 0$. Then, in a neighborhood of $\bar{x}$, the best reply map is single-valued and has the form

$$R(x) \;=\; \begin{cases} \{0\} & \text{if} \quad x \leq \bar{x}, \\ \{\phi(x)\} & \text{if} \quad x > \bar{x}. \end{cases} \tag{4.15}$$

8. Finally, we assume that the global minimum of $G(\bar{x},\cdot)$ is attained at the two distinct points $\bar{y}_1 \neq \bar{y}_2$.

CASE 1: $\bar{y}_1 = 0$, $\bar{y}_2 = 1$.

Using (xi) and (xii), together with the necessary conditions for optimality, we obtain

$$G_y(\bar{x},0) \;>\; 0, \qquad G_y(\bar{x},1) \;<\; 0, \qquad G_x(\bar{x},0) \neq G_x(\bar{x},1).$$

To fix the ideas, assume $G_x(\bar{x},0) < G_x(\bar{x},1)$, the other case being similar. Then, for all x in a neighborhood of $\bar{x}$, the best reply map has the form

$$R(x) \;=\; \begin{cases} \{1\} & \text{if} \quad x < \bar{x}, \\ \{0,1\} & \text{if} \quad x = \bar{x}, \\ \{0\} & \text{if} \quad x > \bar{x}. \end{cases} \tag{4.16}$$

CASE 2: $0 < \bar{y}_1 < \bar{y}_2 = 1$. (The case $0 = \bar{y}_1 < \bar{y}_2 < 1$ is entirely similar.)

Since $G_y(\bar{x},\bar{y}_1) = 0$ and (viii) fails, we must have $G_{yy}(\bar{x},\bar{y}_1) \neq 0$. Hence, in a neighborhood of $(\bar{x},\bar{y}_1)$ the equation (4.13) is solved by a function $y = \phi(x)$.

At the point $x = \bar{x}$ we now compute

$$\frac{d}{dx}G(x,\phi(x)) \;=\; G_x + G_y\phi'(x) \;=\; G_x(\bar{x},\bar{y}) \;\neq\; G_x(\bar{x},1)$$

Note that the last inequality stems from the fact that (ix) fails. To fix the ideas, assume $G_x(\bar{x},\bar{y}_1) < G_x(\bar{x},1)$, the other case being similar. Then, for all x in a neighborhood of $\bar{x}$, the best reply map has the form

$$R(x) \;=\; \begin{cases} \{1\} & \text{if} \quad x < \bar{x}, \\ \{\bar{y}_1,1\} & \text{if} \quad x = \bar{x}, \\ \{\phi(x)\} & \text{if} \quad x > \bar{x}. \end{cases}$$

CASE 3: $0 < \bar{y}_1 < \bar{y}_2 < 1$.

By the necessary conditions for a minimum, this implies $G_y(\bar{x}, \bar{y}_1) = G_y(\bar{x}, \bar{y}_2) = 0$. Since both (vi) and (vii) fail, this implies

$$G_{yy}(\bar{x}, \bar{y}_1) \neq 0, \qquad G_{yy}(\bar{x}, \bar{y}_2) \neq 0, \tag{4.17}$$

$$G_x(\bar{x}, \bar{y}_1) \neq G_x(\bar{x}, \bar{y}_2). \tag{4.18}$$

By (4.17), near the points $(\bar{x}, \bar{y}_1)$ and $(\bar{x}, \bar{y}_1)$ the equation (4.13) implicitly defines two functions $y = \phi_1(x)$ and $y = \phi_2(x)$. By (4.18) at $x = \bar{x}$ we have

$$\frac{d}{dx}G(x, \phi_1(x))\bigg|_{x=\bar{x}} = G_x(\bar{x}, \bar{y}_1) + G_y(\bar{x}, \bar{y}_1)\phi_1'(\bar{x}) = G_x(\bar{x}, \bar{y}_1) \neq G_x(\bar{x}, \bar{y}_2) = \frac{d}{dx}G(x, \phi_2(x))\bigg|_{x=\bar{x}}.$$

To fix the ideas, assume $G_x(\bar{x}, \bar{y}_1) < G_x(\bar{x}, \bar{y}_2)$, the other case being similar. Then, for all x in a neighborhood of $\bar{x}$, the best reply map has the form

$$R(x) = \begin{cases} \{\phi_2(x)\} & \text{if} \quad x < \bar{x}, \\ \{\bar{y}_1, \bar{y}_2\} & \text{if} \quad x = \bar{x}, \\ \{\phi_1(x)\} & \text{if} \quad x > \bar{x}. \end{cases} \tag{4.19}$$

9. By the previous analysis, if $G \in \mathcal{C}^3(\mathbb{R}^2, \mathbb{R})$ is a function that does not satisfy any of the conditions (i)–(xiii), then for x in a neighborhood of any point $\bar{x} \in [0, 1]$ the best reply map has the structure described in (4.5)-(4.7). Covering the compact domain $[0, 1]$ with a finite number of open interval, the theorem is proved. □

5 Generic stability of one-dimensional Stackelberg equilibria

We again consider a noncooperative game, where the players choose strategies in $X = Y = [0, 1]$. In order to apply the results in Section 2, we shall need

Lemma 5.1 *Assume that the graph of the best reply map has the structure described at (4.5)-(4.7) in Theorem 4.1. Then this graph can be also written in the form (2.27)-(2.28), where the functions φ_i satisfy property (i) in the characterization* **(C)**.

Proof. Let (4.5)-(4.7) hold. By suitably covering the square $[0, 1] \times [0, 1]$ with finitely many open sets $V_1, \ldots, V_m$, for each $q \in \{1, \ldots, m\}$ the intersection $V_q \cap \text{graph}(R)$ takes one of the forms described below.

CASE 1 (see Fig. 2, left). Assume that

$$\mathcal{A} = \{y = \varphi_k(x), \quad a < x_k \leq x_k\}.$$

To treat this case, it suffices to consider the function

$$\phi_1(x, y) \doteq y - \varphi_k(x)$$

and an additional function ϕ_2, with the following properties:

$$\begin{cases} \phi_2(x,\varphi_k(x)) > 0 & \text{for} \quad x \in \,]a, x_k[\,, \\ \phi_2(x,\varphi_k(x_k)) = 0\,, \end{cases} \tag{5.1}$$

$$\frac{d}{dx}\phi_2(x,\varphi_k(x))\bigg|_{x=x_k} \;<\; 0. \tag{5.2}$$

We then have the representation

$$\mathcal{A} \;=\; \{(x,y)\,;\quad \phi_1(x,y)=0, \qquad \phi_2(x,y)\geq 0\}. \tag{5.3}$$

CASE 2 (see Fig. 2, center). Assume that

$$\mathcal{A} \;=\; \{y=\varphi_k(x), \quad a < x_k \leq x_k\} \;\cup\; \{y=1, \quad x_k \leq x < b\},$$

with $\varphi_k(x_k)=1$, $\varphi_k'(x_k)>1$. Without loss of generality, we can assume $\varphi_k(x)>1$ for $x>x_k$. To handle this case we consider the two functions

$$\phi_1(x,y) \;\doteq\; \varphi_k(x)-y, \qquad \phi_2(x,y) \;=\; 1-y.$$

We then have the representation

$$\mathcal{A} \;=\; \mathcal{A}_1 \cup \mathcal{A}_2 \;=\; \Big\{(x,y)\,;\;\; \phi_1(x,y)=0,\;\; \phi_2(x,y)\geq 0\Big\} \cup \Big\{(x,y)\,;\;\; \phi_1(x,y)\geq 0,\;\; \phi_2(x,y)=0\Big\}.$$

All other cases (see for example Fig. 2, right) are entirely similar. □

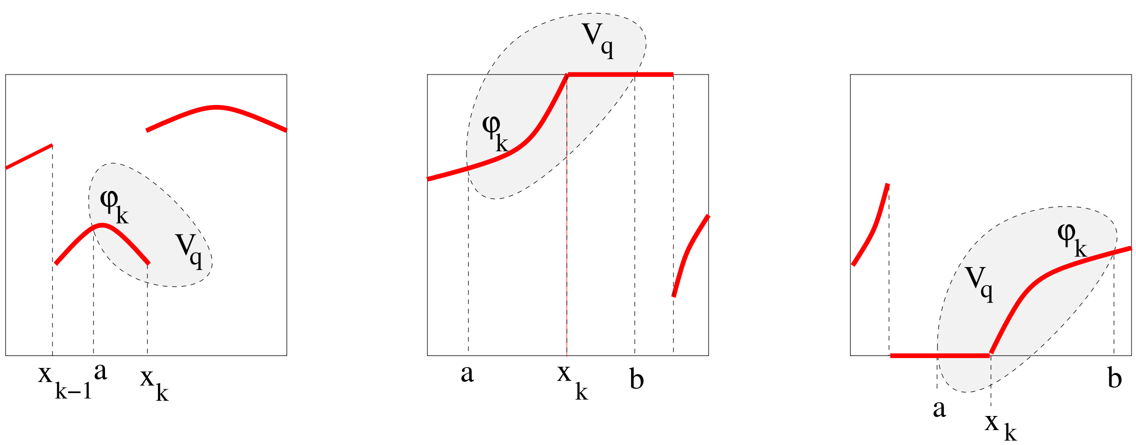

Figure 2: Left and center: the two main cases considered in the proof of Lemma 5.1. Right: another configuration, entirely similar to the one considered in CASE 2.

As an immediate consequence of Corollaries 2.2 and 2.3 we now obtain the generic stability of the Stackelberg equilibrium, in this particular setting (see Fig. 3).

Theorem 5.1 *Let $G \in \mathcal{C}^3(\mathbb{R}^2)$ be a function in the open dense set considered in Theorem 4.1, for which the best reply map R has the structure described at (4.5)-(4.7). Then there exists an open dense set of functions $\mathcal{F} \subset \mathcal{C}^2(\mathbb{R}^2)$ such that, for every $F \in \mathcal{F}$, the following holds.*

The global minimum of F on $\mathcal{A} = \text{graph}(R)$ is attained at a point $(\bar{x}, \bar{y})$ in generic position, as defined in (2.2). Moreover there exists constants $C, \delta > 0$ such that, if

$$\|\widetilde{F} - F\|_{\mathcal{C}^2} \leq \delta, \qquad \|\widetilde{G} - G\|_{\mathcal{C}^3} \leq \delta, \tag{5.4}$$

then the corresponding perturbed optimization problem

$$\min_{(x,y)\in\widetilde{\mathcal{A}}} \widetilde{F}(x,y) \tag{5.5}$$

has a unique minimizer $(\tilde{x}, \tilde{y})$, also in generic position. Here $\widetilde{\mathcal{A}} = graph(\widetilde{R})$ is the best reply map corresponding to the cost function $\widetilde{G}$. In addition

$$|\tilde{x} - \bar{x}| + |\tilde{y} - \bar{y}| \leq C \cdot \Big(\|\widetilde{F} - F\|_{\mathcal{C}^2} + \|\widetilde{G} - G\|_{\mathcal{C}^3}\Big). \tag{5.6}$$

Proof. Consider a cost function $G \in \mathcal{C}^3(\mathbb{R}^2)$ in the open dense set where none of the conditions (i)–(xiii) in the proof of Theorem 4.1 are satisfied. Then the corresponding best reply map $R(\cdot)$ has the generic structure described at in (4.5)-(4.7). In particular (see Fig. 2), there is a finite covering of the square $[0,1] \times [0,1]$ with open sets V_q, $q = 1, \ldots, \nu$, such that the restriction of the graph of $R(\cdot)$ to each set V_q can be described by a finite set of equalities and inequalities involving

- functions of the form $\phi(x,y) = y - \varphi_k(x)$, where φ_k is implicitly defined by the identity $\nabla_y G(x,, \varphi_k(x)) = 0$.
- the two functions $\phi(x,y) = y$ and $\phi(x,y) = y - 1$,
- functions of the form $\phi(x,y) = x - x_k$, where x_k is a point where the best reply map has a jump.

If now $\widetilde{G}$ is a another cost function with $\|\widetilde{G} - G\|_{\mathcal{C}^3}$ sufficiently small, then by repeated applications of the implicit function theorem we check that the best reply map $\widetilde{\mathcal{A}}$ determined by $\widetilde{G}$ also admits the structure (4.5)-(4.7). Furthermore, all the corresponding functions $\tilde{\phi}_i$ in the characterization **(C)** satisfy

$$\|\phi_i - \tilde{\phi}_i\|_{\mathcal{C}^2} \leq C\,\|G - \widetilde{G}\|_{\mathcal{C}^3} \qquad \text{for all } \; i \in I,$$

for some constant $C > 0$. The estimate (5.6) now follows from Corollaries 2.2 and 2.3. □

6 Multi-dimensional strategies for the follower

We now extend the analysis to the case where the follower chooses his strategy within a multi-dimensional space. To avoid technicalities associated with the boundary of the sets X, Y, we will simply assume that $x \in X = \mathbb{R}$, $y \in Y = \mathbb{R}^n$. The follower will thus solve a minimization problem on $\mathbb{R}^n$, depending on a scalar parameter. The generic structure of the best reply map, shown in Fig. 4, can be readily described. As in Section 3, we denote by $\mathcal{G}$ the set of all functions $G : \mathbb{R} \times \mathbb{R}^m \mapsto \mathbb{R}$ which satisfy **(A1)-(A2)**.

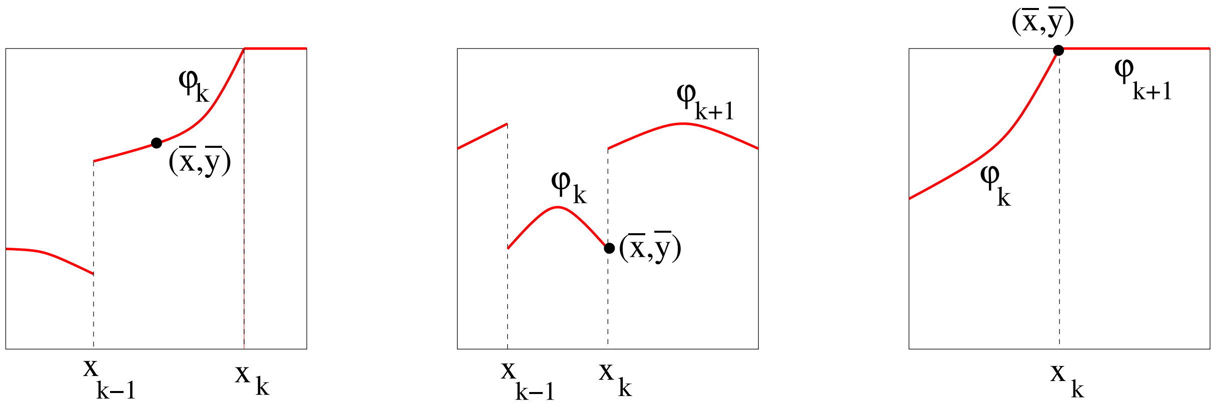

Figure 3: Examples of a best reply map $R(\cdot)$ with generic structure, according to Theorem 4.1. Three different cases where the function F can attain a global minimum at a point $(\bar{x}, \bar{y})$ on the graph of $R(\cdot)$, in generic position.

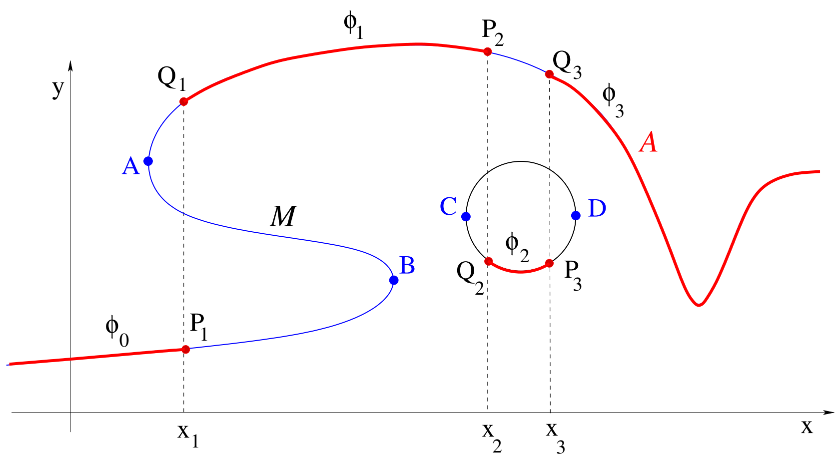

Figure 4: The generic structure of the best reply map $x \mapsto R(x) \subset \mathbb{R}^n$, depending on a scalar parameter $x \in \mathbb{R}$. As x varies, at the points A, B, C, D, where a new pair of local minima of $G(x, \cdot)$ is created, the value of G must be strictly larger than the global minimum.

Theorem 6.1 *Given $\alpha > 0$, there exists an open dense subset $\mathcal{G}^\sharp \subset \mathcal{G}$ of cost functions such that, for any $G \in \mathcal{G}^\sharp$, restricted to the interval $[-\alpha, \alpha]$, the best reply map has the following structure.*

There exists finitely many points $-\alpha = x_0 < x_1 < \cdots < x_\nu = \alpha$, and functions $\varphi_k \in \mathcal{C}^2(\mathbb{R};\ \mathbb{R}^n)$ such that

$$\{(x,y)\,;\quad y \in R(x)\,,\quad |x| \le \alpha\} \;=\; \bigcup_{k=1}^{\nu} \Big\{(x, \varphi_k(x))\,;\quad x \in [x_k, x_{k+1}]\Big\}. \tag{6.1}$$

Moreover, $\varphi_k(x_k) \neq \varphi_{k-1}(x_k)$ for all $k = 1, \ldots, \nu - 1$.

Proof. 1. Let a function $G_0 \in \mathcal{G}$ be given, together with constants $\kappa, \rho > 0$. Using Lemma 3.1, for any $\varepsilon > 0$ we can find a function $g \in \mathcal{G} \cap \mathcal{C}^\infty$ with $\|g - G_0\|_{\mathcal{C}^3} < \varepsilon$ for which the following implication holds:

$$|x| \le \kappa, \quad |y| \le \rho, \quad \nabla_y g(x,y) \;=\; 0, \qquad \Longrightarrow \qquad \mathrm{rank}\big(D(\nabla_y g)(x,y)\big) \;=\; n. \tag{6.2}$$

As a consequence, for every sufficiently small $\mathcal{C}^3$ perturbation of g, the implication (6.2) still holds.

For a given function $G \in \mathcal{C}^3$, we consider the $n \times (n+1)$ matrix of partial derivatives

$$A \;=\; \Big(A_0\Big|A_1\Big|\cdots\Big|A_n\Big) \;\doteq\; \begin{pmatrix} \partial_x\partial_{y_1}G & \partial_{y_1}\partial_{y_1}G & \ldots & \partial_{y_n}\partial_{y_1}G \\ \vdots & \vdots & & \vdots \\ \partial_x\partial_{y_n}G & \partial_{y_1}\partial_{y_n}G & \ldots & \partial_{y_n}\partial_{y_n}G \end{pmatrix}. \tag{6.3}$$

The condition $\mathrm{rank}(A) = n$ at every point in the set

$$\mathcal{M} \;\doteq\; \big\{(x,y) \in \mathbb{R}^{1+n}\,;\;\; \nabla_y G(x,y) = 0\big\}, \tag{6.4}$$

guarantees that $\mathcal{M}$ is a 1-dimensional embedded manifold in $\mathbb{R}^{1+n}$. By the implicit function theorem, near points where

$$\mathrm{rank}(A_1|\cdots|A_n) \;=\; n, \tag{6.5}$$

this manifold $\mathcal{M}$ can be represented as the graph of a function $x \mapsto (y_1(x), \ldots, y_n(x))$.

We observe that, by removing any column from the matrix A in (6.3), we obtain an $n \times n$ matrix whose rank can be either n or $n - 1$.

2. Let g be a smooth function for which (6.1) holds. Given any $\delta > 0$, we claim that there exists $G \in \mathcal{C}^\infty$ with $\|G - g\|_{\mathcal{C}^3} < \delta$ and such that, for $|x| \le \alpha$ and $|y| \le \rho$, none of the following statements holds true.

(i) *There exist three distinct points (x, y^1), (x, y^2), (x, y^3) such that*

$$G(x,y^1) \;=\; G(x,y^2) \;=\; G(x,y^3), \qquad \nabla_y G(x,y^1) \;=\; \nabla_y G(x,y^2) \;=\; \nabla_y G(x,y^3) \;=\; 0. \tag{6.6}$$

Notice that, by requiring that (i) fails, we rule out the possibility that, for some x_1, the global minimum of $G(x_1,\cdot)$ is attained at three or more distinct points (see Fig. 5, left).

(ii) *There exist two distinct points* (x, y^1), (x, y^2) *such that*

$$G(x,y^1) \;=\; G(x,y^2), \qquad \nabla_y G(x,y^1) \;=\; \nabla_y G(x,y^2) \;=\; 0, \tag{6.7}$$

$$\mathrm{rank}\big(A_1|\cdots|A_n\big)(x,y_1) \;<\; n. \tag{6.8}$$

By requiring that (ii) fails, we preclude the existence of x_2 such that the minimum of $G(x_2,\cdot)$ is attained at two distinct points, and at one of these points the tangent vector to $\mathcal{M}$ is vertical (see again Fig. 5).

(iii) *There exist two distinct points* (x, y^1), (x, y^2) *such that*

$$rank\big(A_1|\cdots|A_n\big)(x,y^1) \;=\; rank\big(A_1|\cdots|A_n\big)(x,y^2) \;=\; n, \tag{6.9}$$

and moreover

$$\begin{aligned} G(x,y^1) \;&=\; G(x,y^2), \qquad \nabla_y G(x,y^1) \;=\; \nabla_y G(x,y^2) \;=\; 0,\\ \frac{d}{dx}G(x,\phi_1(x)) \;&=\; \frac{d}{dx}G(x,\phi_2(x)). \end{aligned} \tag{6.10}$$

Here ϕ_1,ϕ_2 *are the functions implicitly defined by the system of* n *equations*

$$\nabla_y G(x,\phi(x)) \;=\; 0, \tag{6.11}$$

with $\phi_1(x) = y^1$ *and* $\phi_2(x) = y^2$.

Notice that, by (6.9), both functions ϕ_1,ϕ_2 are well defined in a neighborhood of x, and have $\mathcal{C}^2$ regularity. To understand the meaning of (iii), assume that the function $G(x_3,\cdot)$ attains its global minimum at two distinct points (x_3, y^1) and (x_3, y^2). By (6.9), on a neighborhood of x_3 the equation (6.11) implicitly defines two functions $y = \phi_1(x)$, $y = \phi_2(x)$. If (iii) fails, at every point $x \neq x_3$ with $|x - x_3|$ small enough, we have

$$G(x,\phi_1(x)) \;\neq\; G(x,\phi_2(x)).$$

In particular, the minimum of $G(x,\cdot)$ cannot be simultaneously attained at more then one point, for all x in a neighborhood of x_3 (see Fig. 5).

Our last condition implies that there cannot exist $x_4 \in \mathbb{R}$ such that the 1-dimensional manifold $\mathcal{M}$ has a second order tangency with the vertical hyperplane $\{(x,y)\,;\; x = x_4\}$, as shown in Fig. 5, right.

(iv) *There exists a point* $(\bar x,\bar y) \in \mathbb{R}^{1+n}$ *such that*

$$\nabla_y G(\bar x,\bar y) \;=\; 0, \quad \mathrm{rank}(A_1|A_2|\cdots|A_n)(\bar x,\bar y) \;=\; n-1. \tag{6.12}$$

Moreover, if $s \mapsto (x(s), y(s))$ *is a local arc-length parameterization of* $\mathcal{M}$ *with* $(x(0), y(0)) = (\bar x,\bar y)$*, then*

$$\frac{d^2}{ds^2}x(s)\bigg|_{s=0} \;=\; 0. \tag{6.13}$$

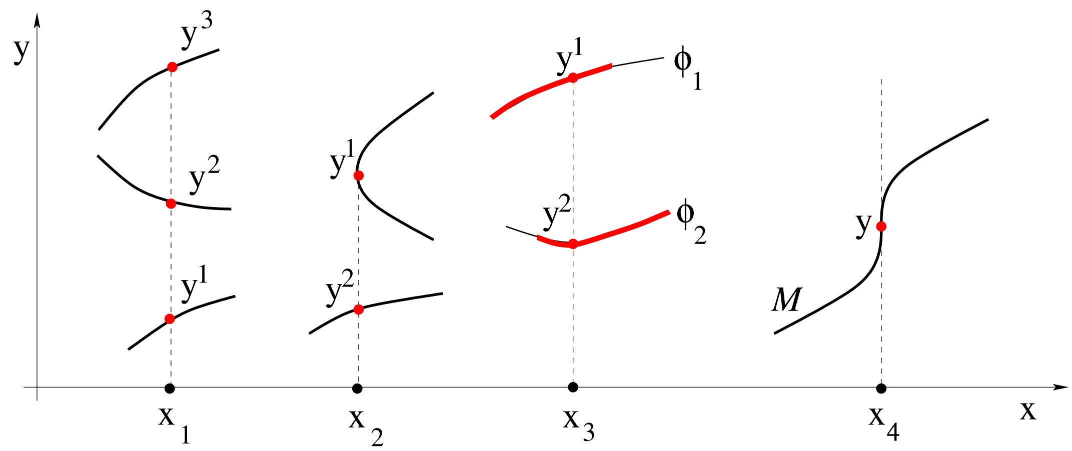

Figure 5: Four non-generic configurations, described at (i)–(iv) of the proof of Theorem 6.1.

Our claims will be proved in the forthcoming steps. Notice that in all cases (i)–(iv), we are considering sets of points (x, y^i) defined by a number of equations which is strictly larger than the dimension of the spaces they live in. Hence, for a generic function G, these sets are empty.

3. In this step we prove that the set of functions G which do not satisfy (i) at any point $x \in \mathbb{R}$ is dense. Indeed, consider the open domain

$$X^{(3)} \doteq \Big\{(x, y^1, y^2, y^3) \in \mathbb{R}^{1+3n}\,; \quad y^i \neq y^j \text{ for } 1 \leq i < j \leq 3\Big\}. \tag{6.14}$$

Call $J^1(X^{(3)}, \mathbb{R})$ the bundle of all 1-jets of functions $f : X^{(3)} \mapsto \mathbb{R}$. Given a smooth map $G : \mathbb{R} \times \mathbb{R}^n \mapsto \mathbb{R}$, this yields a map

$$j^1G(x, y^1, y^2, y^3) \doteq \Big(G(x, y^1), G(x, y^2), G(x, y^3), \nabla_y G(x, y^1), \nabla_y G(x, y^2), \nabla_y G(x, y^3)\Big).$$

Consider the set

$$\mathcal{M}_1 \doteq \Big\{(z_1, z_2, z_3, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \in \mathbb{R}^3 \times \mathbb{R}^{3n}\,; \quad z_1 - z_2 = z_2 - z_3 = 0, \quad \mathbf{v}_1 = \mathbf{v}_2 = \mathbf{v}_3 = 0\Big\}. \tag{6.15}$$

It is immediate to check that $\mathcal{M}_1$ is a smooth submanifold of $J^1(X^{(3)}, \mathbb{R})$. By the multi-jet version of Thom's transversality theorem (proved in the Appendix), the set of all $G \in \mathcal{C}^\infty(\mathbb{R}^{1+n}, \mathbb{R})$ such that j^1G is transversal to $\mathcal{M}_1$ is residual in the $\mathcal{C}^\infty$ topology. Observing that $\dim(X^{(3)}) = 1+3n$ while $\text{codim}(\mathcal{M}_1) = 2+3n$, transversality implies that the intersection is empty:

$$\Big\{j^1G(x, y^1, y^2, y^3)\,; \quad (x, y^1, y^2, y^3) \in X^{(3)}\Big\} \cap \mathcal{M}_1 = \emptyset. \tag{6.16}$$

4. Next, we show that the set of functions G which do not satisfy (ii) at any point $(x, y) \in \mathbb{R} \times \mathbb{R}^n$ is dense. For this purpose consider the open set

$$X^{(2)} \doteq \Big\{(x, y^1, y^2) \in \mathbb{R}^{1+2n}\,; \quad y^1 \neq y^2\Big\}. \tag{6.17}$$

Call $J^2(X^{(2)}, \mathbb{R})$ the bundle of all 2-jets of functions $f : X^{(2)} \mapsto \mathbb{R}$. Given a smooth map $G : \mathbb{R} \times \mathbb{R}^n \mapsto \mathbb{R}$, this yields a map

$$j^2G(x, y^1, y^2) \doteq \Big(G(x, y^1), G(x, y^2), \nabla_y G(x, y^1), \nabla_y G(x, y^2), D^2_y G(x, y^1), D^2_y G(x, y^2)\Big).$$

where $D_y^2 G \doteq (\partial^2_{y_i y_j} G)$ denotes the $n \times n$ Hessian matrix of second derivatives of G w.r.t. $y_1, \ldots, y_n$. For each $k = 1, 2, \ldots, n-1$, consider the set

$$\mathcal{M}_{2,k} \doteq \Big\{(z_1, z_2, \mathbf{v}_1, \mathbf{v}_2, A_1, A_2)\,;\quad z_1 - z_2 = 0,\quad \mathbf{v}_1 = \mathbf{v}_2 = 0,\ \mathrm{rank}(A_1) = k\Big\}. \qquad (6.18)$$

Using Proposition 3.2.6 in [3], p. 33, we check that $\mathcal{M}_2$ is a smooth submanifold of $J^2(X^{(2)}, \mathbb{R})$. By a multi-jet version of Thom's transversality theorem, the set of all $G \in \mathcal{C}^\infty(\mathbb{R}^{1+n}, \mathbb{R})$ such that j^2G is transversal to $\mathcal{M}_{2,k}$ is residual in the $\mathcal{C}^\infty$ topology. We now observe that $\dim(X^{(2)}) = 1+2n$ while $\mathrm{codim}(\mathcal{M}_{2,k}) = 1+2n+(n-k)^2$. Hence, for every $k = 0, 1, \ldots, n-1$, transversality implies that the intersection is empty:

$$\Big\{j^2G(x, y^1, y^2)\,;\quad (x, y^1, y^2) \in X^{(2)}\Big\} \cap \mathcal{M}_{2,k} = \emptyset. \qquad (6.19)$$

5. We now prove that the set of all functions G which do not satisfy (iii) at any point $(x, y^1, y^2) \in X^{(2)}$ is dense.

Call $J^2(X^{(2)}, \mathbb{R})$ the bundle of all 1-jets of functions $f : X^{(2)} \mapsto \mathbb{R}$. Given a smooth map $G : \mathbb{R} \times \mathbb{R}^n \mapsto \mathbb{R}$, consider the map

$$j^1G(x, y^1, y^2) \doteq \Big(G(x, y^1),\ G(x, y^2),\ G_x(x, y^1),\ G_x(x, y^2),\ \nabla_y G(x, y^1),\ \nabla_y G(x, y^2)\Big).$$

Define the set

$$\mathcal{M}_3 \doteq \Big\{(z_1, z_2, w_1, w_2, \mathbf{v}_1, \mathbf{v}_2) \in \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^{2n}\,;\quad z_1 - z_2 = 0,\quad w_1 - w_2 = 0,\quad \mathbf{v}_1 = \mathbf{v}_2 = 0\Big\}.$$

It is clear that $\mathcal{M}_3$ a smooth submanifold of $J^1(X^{(2)}, \mathbb{R})$. By a multi-jet version of Thom's transversality theorem (proved in the Appendix), the set of all $G \in \mathcal{C}^\infty(\mathbb{R}^{1+n}, \mathbb{R})$ such that j^1G is transversal to $\mathcal{M}_3$ is residual in the $\mathcal{C}^\infty$ topology. Since $\dim(X^{(2)}) = 1 + 2n$ while $\mathrm{codim}(\mathcal{M}_3) = 2 + 2n$, transversality implies that the intersection is empty:

$$\Big\{j^1G(x, y^1, y^2);\ (x, y^1, y^2) \in X^{(2)}\Big\} \cap \mathcal{M}_3 = \emptyset. \qquad (6.20)$$

Notice that the assumption (6.9) guarantees that ϕ_1, ϕ_2 are well defined. Since $\nabla_y G(x, \phi_1(x)) = \nabla_y G(x, \phi_2(x)) = 0$, one has the equivalence

$$\frac{d}{dx} G(x, \phi_1(x)) = \frac{d}{dx} G(x, \phi_2(x)) \qquad \Longleftrightarrow \qquad G_x(x, \phi_1(x)) = G_x(x, \phi_2(x)).$$

Therefore, the set of all functions $G \in \mathcal{C}^\infty(\mathbb{R}^{1+n}, \mathbb{R})$ which do not satisfy (iii) at any point $(x, y^1, y^2) \in X^{(2)}$ is the intersection of all the set of functions G such that j^2G is transversal to $\mathcal{M}_{2,k}$ for $k = 0, \ldots, n-1$, and moreover j^1G is transversal to $\mathcal{M}_3$. We thus conclude that the set of functions $G \in \mathcal{C}^\infty(\mathbb{R}^{1+n}, \mathbb{R})$ which do not satisfy (iii) at any point $(x, y^1, y^2) \in X^{(2)}$ is dense in $\mathcal{C}^3(\mathbb{R}^{1+n}, \mathbb{R})$.

6. Finally, we show that the set of functions G which do not satisfy (iv) at any point $(\bar{x}, \bar{y}) \in \mathbb{R}^{1+n}$ is dense. For this purpose, we need to show that the family of third order jets of functions $G : \mathbb{R}^{1+n} \mapsto \mathbb{R}$, which satisfy all the conditions in (iv), is a smooth manifold with codimension $n - 2$.

We recall that the assumption $\mathrm{rank}(A) = n$ at every point $(x, y) \in \mathcal{M}$ implies that the set $\mathcal{M}$ in (4.3) is a 1-dimensional manifold, embedded in $\mathbb{R}^{1+n}$. The condition $\mathrm{rank}(A_1|\cdots|A_n)(\bar{x}, \bar{y}) = n-1$ implies that the tangent vector to $\mathcal{M}$ at $(\bar{x}, \bar{y})$ is vertical, i.e.:

$$\frac{d}{ds}x(s)\Big|_{s=0} \;=\; 0. \tag{6.21}$$

To obtain an expression for the second derivative in (6.13), we first differentiate the identity $\nabla_y G(x, y) = 0$ and obtain the linear system

$$\Big(A_0\Big|A_1\Big|\cdots\Big|A_n\Big)\begin{pmatrix} dx/ds \\ dy_1/ds \\ \vdots \\ dy_n/ds \end{pmatrix} \;=\; \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{6.22}$$

Calling A_{-i} the $n \times n$ matrix obtained from $A = (A_0|A_1|\cdots|A_n)$ by deleting the i-th column, the solution to (6.22) can be written as

$$\frac{dx}{ds} \;=\; \kappa \cdot \det(A_1|A_2|\cdots|A_n), \qquad\qquad \frac{dy_i}{ds} \;=\; \kappa \cdot \det(A_{-i}), \qquad i = 1, \ldots, n, \tag{6.23}$$

where $\kappa = \kappa(A)$ is a normalizing factor, chosen in order to achieve

$$\left(\frac{dx}{ds}\right)^2 + \sum_{i=1}^{n}\left(\frac{dy_i}{ds}\right)^2 \;=\; 1.$$

At a point $(\bar{x}, \bar{y})$ where $\det(A_1|A_2|\cdots|A_n) = 0$, using (6.23) the second derivative (6.13) is computed by

$$\begin{aligned} \frac{d^2}{ds^2}x(s) \;&=\; \kappa\,\nabla_y\Big[\det(A_1|A_2|\cdots|A_n)\Big]\cdot\frac{d}{ds}y(s) \\ &=\; \kappa\sum_{i=1}^{n}\partial_{y_i}\Big[\det(A_1|A_2|\cdots|A_n)\Big]\cdot\frac{dy_i}{ds} \\ &=\; \kappa^2\sum_{i=1}^{n}\sum_{j=1}^{n}\det(H_{i,j})\,\det(A_{-i}). \end{aligned} \tag{6.24}$$

Here

$$H_{i,j} \;\doteq\; \begin{pmatrix} \partial_{y_1}\partial_{y_1}G & \cdots & \partial_{y_{j-1}}\partial_{y_1}G & \partial_{y_i}\partial_{y_j}\partial_{y_1}G & \partial_{y_{j+1}}\partial_{y_1}G & \ldots & \partial_{y_n}\partial_{y_1}G \\ \vdots & & & \cdots & & & \vdots \\ \partial_{y_1}\partial_{y_n}G & \cdots & \partial_{y_{j-1}}\partial_{y_n}G & \partial_{y_i}\partial_{y_j}\partial_{y_n}G & \partial_{y_{j+1}}\partial_{y_n}G & \ldots & \partial_{y_n}\partial_{y_n}G \end{pmatrix} \tag{6.25}$$

is the matrix obtained by differentiating the j-th column of $(A_1|\cdots|A_n)$ by y_i.

Call $J^3(\mathbb{R}^{1+n}; \mathbb{R})$ the bundle of all 1-jets of functions $f : \mathbb{R} \times \mathbb{R}^n \mapsto \mathbb{R}$. Given a smooth map $G : \mathbb{R} \times \mathbb{R}^n \mapsto \mathbb{R}$, its third-order jet at a point (x, y) is

$$j^3G(x, y) \;=\; \Big(G(x, y),\, \nabla_y G(x, y),\, D_y^2 G(x, y),\, D_y^3 G(x, y)\Big).$$

For any $(z, \mathbf{v}, A, T) \in J^3(\mathbb{R}^{1+n}; \mathbb{R})$, denote by $\widetilde{H}_{i,j}(A, T)$ the $n \times n$ matrix obtained from (6.25) by replacing the second and third derivatives of G with the corresponding elements in A and T.

With this notation, the family of third order jets of functions satisfying (iv) can thus be expressed by

$$\mathcal{M}_4 \doteq \Big\{(z, \mathbf{v}, A, T) \in J^3(\mathbb{R}^{1+n}; \mathbb{R}) \ ; \quad \mathbf{v} = 0, \qquad \mathrm{rank}(A_1|A_2|\cdots|A_n) = n-1, \\ \sum_{i=1}^n \sum_{j=1}^n \det(\widetilde{H}_{i,j}(A,T)) \det(A_{-i}) = 0\Big\}. \tag{6.26}$$

In order to apply Thom's transversality theorem, we need to show that $\mathcal{M}_4$ is a smooth manifold with $\mathrm{codim}(\mathcal{M}_4) = n+2$. Indeed, the vector equation $\mathbf{v} = 0$ yields n scalar equations, while the condition

$$\mathrm{rank}(A_1|A_2|\cdots|A_n) = n-1, \tag{6.27}$$

provides one more, independent scalar equation. The last condition in (6.26) can be written as

$$\Psi(A,T) \doteq \sum_{i=1}^n \sum_{j=1}^n \det\big(\widetilde{H}_{i,j}(A,T)\big) \det(A_{-i}). \tag{6.28}$$

We claim that the above equation determines a smooth manifold of codimension 1, transversal to the manifolds determined by the previous equations in (6.26). This will be true if at least one of the partial derivatives of Ψ does not vanish. Namely

$$\frac{\partial}{\partial G_{y_1 y_i y_j}} \Psi(A,T) \neq 0 \tag{6.29}$$

for some indices i, j.

To simplify the computations, we observe that the property (iv) which defines $\mathcal{M}_4$ is invariant w.r.t. rotations of the coordinates $y = (y_1, \ldots, y_n)$. Without loss of generality, we can thus assume that the tangent vector to the 1-dimensional manifold $\mathcal{M}$ at the point $(\bar{x}, \bar{y})$ is

$$\left(\frac{dx}{ds}, \frac{dy_1}{ds}, \frac{dy_2}{ds}, \ \ldots \ , \frac{dy_n}{ds}\right) = (0, 1, 0, \ldots, 0). \tag{6.30}$$

In this case, one has

$$\det(A_{-2}) = \cdots = \det(A_{-n}) = 0. \tag{6.31}$$

For notational convenience, we shall here denote by

$$\widetilde{A} = (a_{ij})_{1 \le i,j \le n} = (A_1|\cdots|A_n)$$

is the $n \times n$ symmetric matrix of second derivatives $G_{y_i y_j}$. Using (6.31), from (6.28) one obtains

$$\frac{\partial}{\partial G_{y_1 y_k y_\ell}} \Psi(A,T) = \sum_{\ell=1}^n \frac{\partial}{\partial G_{y_1 y_k y_\ell}} \det\big(\widetilde{H}_{1,\ell}(A,T)\big) = \begin{cases} \frac{\partial}{\partial a_{k\ell}} \det(\widetilde{A}) & \text{if } i = j, \\ 2\frac{\partial}{\partial a_{k\ell}} \det(\widetilde{A}) & \text{if } i \neq j. \end{cases}$$

By assumption, $\widetilde{A}$ has rank $n-1$. Hence there exists a minor $M^{k\ell}$ obtained from $\widetilde{A}$ by removing the k-th row and the ℓ-th column, which has rank $n-1$.

If $k=\ell$, then

$$\frac{\partial}{\partial G_{y_1y_ky_k}}\,\Psi(A,T) \;=\; \frac{\partial}{\partial a_{kk}}\det(\widetilde{A}) \;=\; \det(M^{kk}) \;\neq\; 0.$$

If $k\neq\ell$, recalling that $\widetilde{A}$ is a symmetric matrix, then by symmetry

$$\frac{\partial}{\partial G_{y_1y_ky_\ell}}\Psi(A,T) \;=\; \left(\frac{\partial}{\partial a_{k\ell}}+\frac{\partial}{\partial a_{\ell k}}\right)\det(\widetilde{A}) \;=\; 2\cdot(-1)^{k+\ell}\det(M^{k\ell}) \;\neq\; 0.$$

In both cases, we find a partial derivative which does not vanish. As a result, $\mathcal{M}_4$ is a smooth manifold with $\mathrm{codim}(\mathcal{M}_4)=n+2$.

We can now use Thom's transversality theorem and conclude that the set of all $G\in\mathcal{C}^\infty(\mathbb{R}^{1+n},\mathbb{R})$ such that j^3G is transversal to $\mathcal{M}_4$ is residual in the $\mathcal{C}^\infty$ topology. Since $\mathrm{codim}(\mathcal{M}_4)=n+2$, transversality implies that the intersection is empty:

$$\{j^3G(x,y);\quad (x,y)\in\mathbb{R}^{1+n}\}\cap\mathcal{M}_4 \;=\; \emptyset.$$

7. In this step we show that the family $\mathcal{G}^\sharp\subset\mathcal{C}^3(\mathbb{R}^{1+n},\mathbb{R})$ of all cost functions G that do not satisfy any of the conditions (i)–(iv), at any point (x,y) where $G(x,\cdot)$ attains its global minimum, is open in the topology induced by $\mathcal{C}^3$.

Assume, on the contrary, that there exists a convergent sequence of functions $G_k\to G$ in $\mathcal{C}^3$, with $G_k\notin\mathcal{G}^\sharp$ for every $k\geq 1$. We claim that $G\notin\mathcal{G}^\sharp$ as well. Since there are four conditions that each G_k can satisfy, we consider them separately.

CASE 1: Suppose that all functions G_k satisfy condition (i). Hence there exists a sequence of points (x_k,y_k^1,y_k^2,y_k^3) satisfying (6.6) for G_k, with

$$|x_k|\leq\kappa,\quad |y_k^i|\leq\rho \text{ for } i=1,2,3,\quad y_k^i\neq y_k^j \text{ for all } 1\leq i<j\leq 3.$$

By possibly taking a subsequence, one has

$$x_k\to x_*,\; y_k^1\to y_*^1,\; y_k^2\to y_*^2,\; y_k^3\to y_*^3\quad\text{with}\quad |x_*|\leq\kappa, |y_*^1|\leq\rho, |y_*^2|\leq\rho, |y_*^3|\leq\rho.$$

Three sub-cases can arise.

Case 1a: $y_*^i\neq y_*^j$ for all $1\leq i<j\leq 3$. In this case, by continuity, G satisfies (i) at (x_*,y_*^1,y_*^2,y_*^3).

Case 1b: Two of the points y_*^1,y_*^2,y_*^3 coincide with each other while the third point is different. Without loss of generality, suppose $y_*^1=y_*^2\neq y_*^3$. In this case, since $\nabla_yG_k(x_k,y_k^1)=\nabla_yG_k(x_k,y_k^2)=0$, in a neighborhood of (x_*,y_*) for every k large enough there must be a point $(x_k^\sharp,y_k^\sharp)\in\mathcal{M}_k$ where the tangent vector to $\mathcal{M}_k$ is vertical. This means

$$\nabla_yG(x_k^\sharp,y_k^\sharp) \;=\; 0,\qquad\qquad \det\big(\partial_{y_iy_j}G\big)(x_k^\sharp,y_k^\sharp) \;=\; 0.$$

Moreover, $(x_k^\sharp,y_k^\sharp)\to(x_*,y_*^1)$ as $k\to\infty$. By continuity, we conclude that G satisfies (ii) at (x_*,y_*^1,y_*^3).

Case 1c: $y_*^1 = y_*^2 = y_*^3$. As in the previous case, in a neighborhood of (x_*, y_*) for every k large enough there must be a point $(x_k^\sharp, y_k^\sharp) \in \mathcal{M}_k$ where the tangent vector to $\mathcal{M}_k$ is vertical. Letting $k \to \infty$ we again conclude

$$\det\big(\partial_{y_i y_j} G\big)(x_*, y_*) \;=\; 0. \tag{6.32}$$

By (6.32), the limit manifold $\mathcal{M}$ has a vertical tangent at (x_*, y_*). Let $s \mapsto (x(s), y(s))$ be a local arc-length parameterization of $\mathcal{M}$ with $(x(0), y(0)) = (x_*, y_*)$. If (6.13) holds, then the condition (iv) is verified and we are done (Fig. 6, left). Otherwise, to fix the idea assume

$$\frac{d^2}{ds^2}x(s)\bigg|_{s=0} \;>\; 0.$$

We claim that (x_*, y_*) cannot be a point of global minimum for $G(x_*, \cdot)$. Indeed, consider a strictly increasing sequence $x_\nu \to x_*$. Let y_ν be a value where the function $G(x_\nu, \cdot)$ attains its global minimum. By taking a subsequence we can assume the convergence $(x_\nu, y_\nu) \;\to\; (x_*, y_\sharp)$ for some $y_\sharp$ (see Fig. 6, right). Notice that we must have $y_\sharp \neq y_*$, because the manifold $\mathcal{M}$ does not contain any point (x, y) with $x < x_*$, in a neighborhood of (x_*, y_*). By continuity, $(x_*, y_\sharp)$ must provide a global minimum to $G(x_*, \cdot)$. We conclude that all conditions in (ii) are now satisfied, with

$$(x, y^1) \;=\; (x_*, y_*), \qquad\qquad (x, y^2) \;=\; (x_*, y_\sharp).$$

CASE 2: Suppose that all G_k satisfy condition (ii). This implies that there exists a sequence of points $(x_k, y_{1,k}, y_{2,k})$ satisfying (ii) for G_k with

$$|x_k| \leq \kappa, \quad |y_k^i| \leq \rho \text{ for } i = 1, 2, \quad y_k^1 \neq y_k^2\,.$$

By possibly taking a subsequence, one has

$$x_k \to x_*,\; y_k^1 \to y_*^1,\; y_k^2 \to y_*^2, \quad \text{with} \quad |x_*| \leq \kappa,\; |y_*^1| \leq \rho,\; |y_*^2| \leq \rho.$$

Two sub-cases must be considered.

Case 2a: $y_*^1 \neq y_*^2$. In this case, by continuity the limit function G satisfies (ii) at (x_*, y_*^1, y_*^2).

Case 2b: $y_*^1 = y_*^2$. Since G_k satisfies (ii) at (x_k, y_k^1, y_k^2), this case can be handled by the same arguments as case 1c.

CASE 3: Suppose that all G_k satisfy condition (iii), say at points (x_k, y_k^1, y_k^2), with

$$|x_k| \leq \kappa, \quad |y_k^i| \leq \rho \text{ for } i = 1, 2, \quad y_k^1 \neq y_k^2.$$

By possibly taking a subsequence, one has

$$x_k \to x_*, \quad y_k^1 \to y_*^1, \quad y_k^2 \to y_*^2, \quad \text{with} \quad |x_*| \leq \kappa,\; |y_*^1| \leq \rho,\; |y_*^2| \leq \rho.$$

Two sub-cases must be considered.

Case 3a: $y_*^1 \neq y_*^2$. In this case, if both ϕ_1 and ϕ_2 are still well defined at (x_*, y_*^1) and (x_*, y_*^2), then by continuity G satisfies (iii) at (x_*, y_*^1, y_*^2). The remaining possibility is that at least one of the functions ϕ_1 or ϕ_2 that is not well defined. Without loss of generality, assume that

ϕ_1 is not well defined. Namely, $\det(\partial^2_{y_i,y_j} G(x, y^1_*)) = 0$ and hence G satisfies (ii). In all cases we conclude that $G \notin \mathcal{G}^\sharp$.

Case 3b: $y^1_* = y^2_* = y_*$. In this case, for all k large enough $(x_k, \phi_1(x_k))$ and $(x_k, \phi_2(x_k))$ should lie on the same component of a smooth one-dimensional manifold $\mathcal{M}_k$ defined by the equation $\nabla_y G_k(x, y) = 0$. In a neighborhood of (x_*, y_*), for every k large enough there must be a point $(x^\sharp_k, y^\sharp_k) \in \mathcal{M}_k$ where the tangent vector to $\mathcal{M}_k$ is vertical. Letting $k \to \infty$ we again conclude that (6.32) holds. Assuming that $G(x_*, \cdot)$ attains its global minimum at (x_*, y_*), the same arguments used in case 1c again imply that either (ii) or (iv) hold.

CASE 4: Suppose that every function G_k satisfies condition (iv) at some point (x_k, y_k) with

$$|x_k| \le \kappa, \quad |y_k| \le \rho.$$

By possibly taking a subsequence, we can assume $(x_k, y_k) \to (x_*, y_*)$. By the convergence $G_k \to G$ in $\mathcal{C}^3$, we conclude that G satisfies the condition (iv) at the point (x_*, y_*).

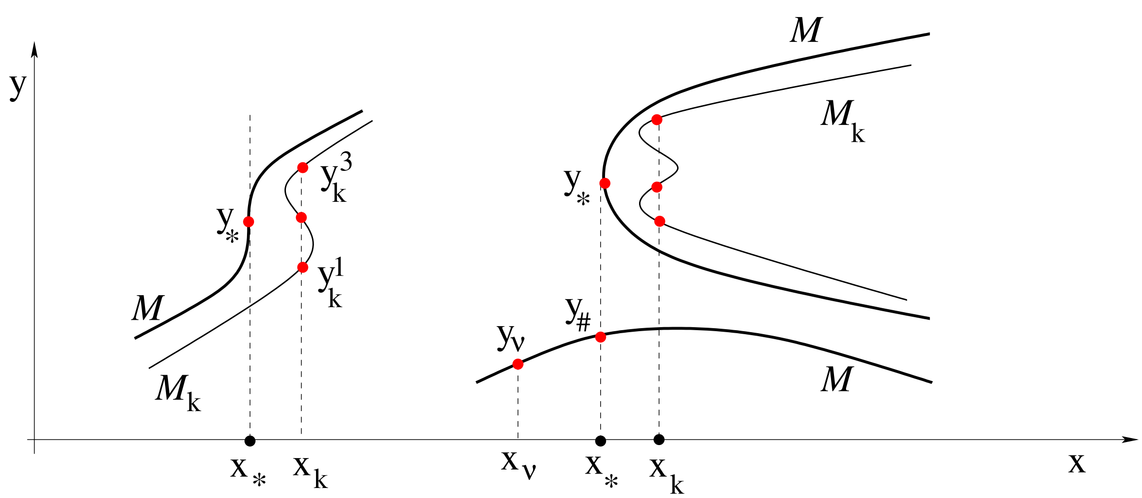

Figure 6: The two main cases considered in step **7** of the proof of Theorem 6.1.

8. In this step we claim that, if none of the conditions (i)–(iv) holds true at any point where $G(x, \cdot)$ attains its global minimum, then the graph of the best reply map satisfies the conclusion of the theorem.

Indeed, consider any point $\bar{x}$. By (i) the minimum of the function $y \mapsto G(\bar{x}, y)$ is attained at most at two distinct points.

CASE 1: The global minimum is attained at the two points $y_1 \neq y_2$. Then, since (ii) fails, near (x, y_1) and (x, y_2) the manifold $\mathcal{M}$ coincides with the graph of two functions $y = \phi_1(x)$ and $y = \phi_2(x)$. By the previous steps, we can assume that (iii) fails. To fix the ideas let

$$\frac{d}{dx} G(x, \phi_1(x)) \;>\; \frac{d}{dx} G(x, \phi_2(x)).$$

This implies that, for some $\delta > 0$ small enough, the global minimum is attained at $(x, \phi_1(x))$ for $x \in [\bar{x} - \delta, \bar{x}]$ and at $(x, \phi_2(x))$ for $x \in [\bar{x}, \bar{x} + \delta]$.

CASE 2: The global minimum is attained at a single point $\bar{y}$. We claim that

$$\det\left(\partial^2_{y_i y_j} G(\bar{x}, \bar{y})\right) \neq 0, \tag{6.33}$$

hence in a neighborhood of $\bar{x}$ the best reply map is single-valued: $R(x) = \{\phi(x)\}$, for some $\mathcal{C}^2$ function $\phi(\cdot)$.

Indeed, if (6.33) fails, since G does not satisfy (iv) at $(\bar{x}, \bar{y})$, we can assume that $\frac{d^2 x}{ds^2}|_{s=0} > 0$, the other case is entirely similar. By the same arguments as case 1c, $G(\bar{x}, \cdot)$ achieves the global minimum at two points $\bar{y} \neq \hat{y}$, reaching a contradiction.

Combining the above two cases, the proof of the theorem is completed. □

An application of Theorems 2.1 and 2.2 now yields the generic stability of Stackelberg equilibria. As in Section 2, we denote by $\mathcal{F}^\infty$ the family of functions satisfying **(B1)**, with the distance (2.6).

Theorem 6.2 *Consider a generic function $G \in \mathcal{G}^\sharp \subset \mathcal{C}^3(\mathbb{R} \times \mathbb{R}^N)$, satisfying the conclusion of Theorem 6.1. Then there exists an open dense set of functions $\mathcal{F}^\sharp \subset \mathcal{F}^\infty$ such that, for every $F \in \mathcal{F}^\sharp$, the following holds.*

The global minimum of F on $\mathcal{A} = \mathrm{graph}(R)$ is attained in generic position, as defined in (2.2). Moreover there exists constants $C, \delta > 0$ such that, if

$$\|\widetilde{F} - F\|_{\mathcal{C}^2} \leq \delta, \qquad \|\widetilde{G} - G\|_{\mathcal{C}^3} \leq \delta, \tag{6.34}$$

then the corresponding perturbed optimization problem

$$\min_{(x,y)\in\widetilde{\mathcal{A}}} \widetilde{F}(x, y) \tag{6.35}$$

has a unique minimizer $(\tilde{x}, \tilde{y})$, also in generic position. Here $\widetilde{\mathcal{A}} = graph(\widetilde{R})$ is the best reply map corresponding to the cost function $\widetilde{G}$. In addition

$$|\tilde{x} - \bar{x}| + |\tilde{y} - \bar{y}| \leq C \cdot \left(\|\widetilde{F} - F\|_{\mathcal{C}^2} + \|\widetilde{G} - G\|_{\mathcal{C}^3}\right). \tag{6.36}$$

Proof. The same arguments used in the proof of Theorem 5.1 apply here as well. We remark that, in the present case, the strategy of the follower is not constrained to a closed set, hence the best reply map has the simpler structure described in Theorem 6.1 (compare Fig. 4 with Fig. 2). The conclusion of Theorem 6.2 thus follows already from Theorems 2.1 and 2.2, without using Corollaries 2.2 and 2.3. □

7 Concluding remarks

In this paper we used techniques from differential geometry to analyze the generic structure of solutions to noncooperative games. In particular, we described the structure of the best reply map and proved the uniqueness and stability of the Stackelberg equilibrium, for an open dense set of cost functions to the leader and to the follower.

While our results only cover some specific settings, it is clear that these techniques can have broader applications. To keep the discussion simple, we considered Stackelberg games where the leader's strategy lies in a one-dimensional space. We expect that a similar analysis can be performed also in the case where the leader chooses his strategy within a two-dimensional manifold, and hence the follower solves a minimization problem depending on two parameters. In this case, the manifold $\mathcal{M}$ defined at (1.6) will then contain not only "fold" but also "cusp" singularities [1, 7]. On the other hand, when the leader's strategy ranges in a high dimensional space, the generic structure of the best reply map will be more difficult to describe.

The main motivation for the present work came from the analysis of infinite-horizon stochastic games with discrete state space [5]. In that paper, a key role is played by a stability assumption on the Stackelberg equilibrium solution, when the follower adopts a myopic strategy. The arguments used in the proof of Theorem 4.1 show that this assumption is not very restrictive. Indeed, it is satisfied by "almost all" cost functionals.

8 Appendix: working tools from differential geometry

Assume $f : \mathbb{R}^m \mapsto \mathbb{R}^n$

- $x \in \mathbb{R}^m$ is a **critical point** of f if the Jacobian matrix $Df(x)$ has rank $< n$. Equivalently: if the differential $Df(x)$ is not surjective.
- $y \in \mathbb{R}^n$ is a **critical value** of f if $y = f(x)$ for some critical point x.

Sard's Lemma [16].

For any $f \in \mathcal{C}^\infty(\mathbb{R}^m; \mathbb{R}^n)$, the set of critical values has n-dimensional measure zero.

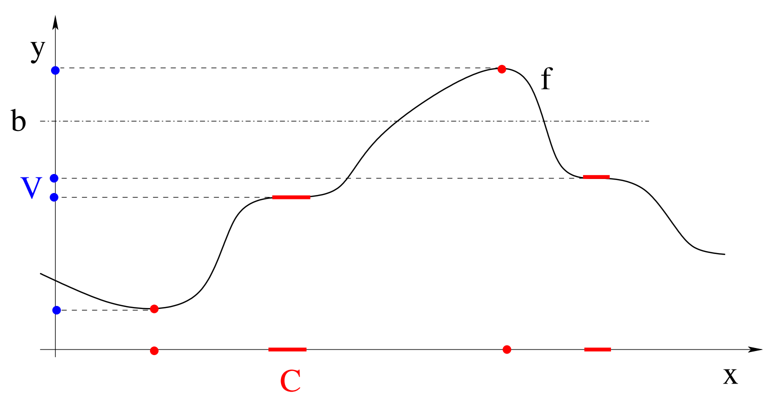

Figure 7: An illustration of Sard's theorem. Here the se of critical points of f is large, but its image has measure zero.

8.1 Transversality.

Let $f : X \mapsto Y$ be a smooth map of manifolds and let W be a submanifold of Y.

We say that f is **transverse** to W at a point p, and write $f \pitchfork_p W$, if

- either $f(p) \notin W$,
- or else $f(p) \in W$ and $(df)_p(T_p X) + T_{f(p)}W = T_{f(p)}Y$.

We say that f is **transverse** to W, and write $f \pitchfork W$, if $f \pitchfork_p W$ for every $p \in X$.

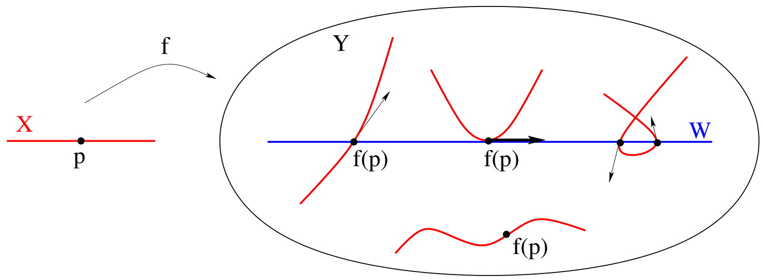

Figure 8: The map shown in the center, whose graph is tangent to W, is not transversal. All other functions are transversal.

Example. Take $X = \mathbb{R}^m$, $Y = \mathbb{R}^n$ with $m < n$. Assume $W = \{y_0\}$ for some $y_0 \in \mathbb{R}^n$. In this case, transversality implies

- either $f(x) \neq y_0$,
- or else $f(x) = y_0$ and rank $(Df(x)) = n$.

Since the second alternative is impossible,

$$f \pitchfork W \qquad \Longleftrightarrow \qquad f(x) \neq y_0 \qquad \text{for all } x \in \mathbb{R}^2$$

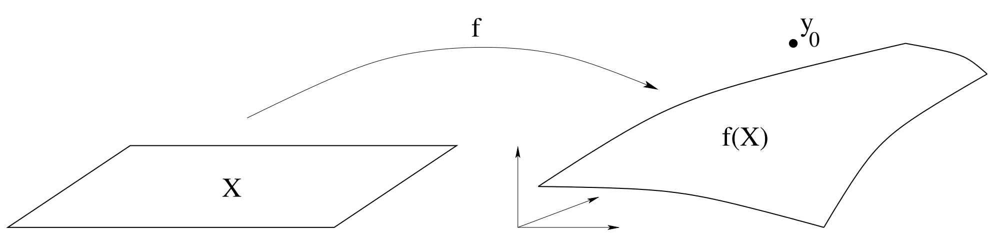

Figure 9: A map $f : \mathbb{R}^2 \mapsto \mathbb{R}^3$ is transversal to the manifold $W = \{y_0\}$ if and only if its image does not contain the point y_0.

Transversality lemma. *Let X, Θ, and Y be smooth manifolds, W a submanifold of Y. Let $\theta \mapsto \phi^\theta$ be a smooth map which to each $\theta \in \Theta$ associates a function $\phi^\theta \in \mathcal{C}^\infty(X, Y)$, and define $\Phi : X \times \Theta \mapsto Y$ by setting $\Phi(x, \theta) = \phi^\theta(x)$.*

If $\Phi \pitchfork W$, then the set $\{\theta \in \Theta, ; \quad \phi^\theta \pitchfork W\}$ is dense in Θ.

8.2 A multi-jet transversality theorem.

We state here a version of the multi-jet transversality theorem which is used several times in the paper. In the following, for given integers $m, n \geq 1$ we consider maps $f : \mathbb{R}^{m+n} \mapsto \mathbb{R}$. Let

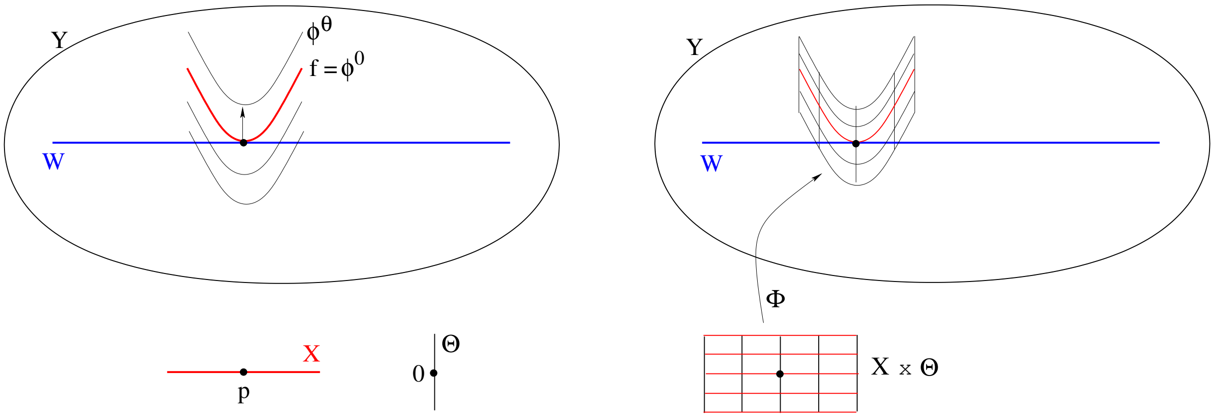

Figure 10: A family of maps $\theta \mapsto \phi^\theta$ depending on a parameter $\theta \in \Theta$ can also be seen as a single map Φ from the product space: $X \times \Theta \mapsto Y$. If Φ is transversal, then for a.e. θ the map ϕ^θ is transversal.

$\mathcal{P}^k$ the space of polynomials of degree $\leq k$ in the $m+n$ variables $(x, y) \in \mathbb{R}^m \times \mathbb{R}^n$. Identifying a polynomial with its coefficients, it is clear that $\mathcal{P}^k$ is a finite dimensional vector space. The product space

$$J^k(\mathbb{R}^{m+n}; \mathbb{R}) \doteq \mathbb{R}^{m+n} \times \mathcal{P}^k$$

is a jet bundle over the space R^{m+n}. Any function $f : \mathbb{R}^{m+n} \mapsto \mathbb{R}$ determines a section of this bundle, defined as $j^k f(x, y) = P^{(x,y)}$, where the polynomial $P^{(x,y)}(\cdot)$ is the k-th order Taylor approximation of the function f at the point (x, y).

Next, for $s \geq 1$ we call

$$Z^{(s)} \doteq \{(x, y_1, \ldots, y_s) \in \mathbb{R}^m \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n \,; \quad y_i \neq y_j \ \text{ for all } \ 1 \leq i < j \leq s\}. \tag{8.1}$$

Notice that $Z^{(s)}$ is an open subset of a vector space of dimension $m+sn$; hence it is a manifold. The set

$$J^k_s(\mathbb{R}^{m+n}; \mathbb{R}) = \Big\{(x, y_1, \ldots, y_s, P_1, \ldots, P_s) \in \mathbb{R}^m \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n \times \mathcal{P}^k \times \cdots \times \mathcal{P}^k, \quad y_i \neq y_j \ \text{ for all } \ 1 \leq i < j \leq s\Big\} \tag{8.2}$$

is a k-th order jet bundle over $Z^{(s)}$. Any smooth function $f : \mathbb{R}^{m+n} \mapsto \mathbb{R}$ determines a section of this bundle defined as

$$j^k_s f(x, y_1, \ldots, y_s) = (Q^{(x,y_1)}, \ \ldots \ , Q^{(x,y_s)}), \tag{8.3}$$

where $Q^{(x,y_i)}$ is the polynomial of degree $\leq k$ determined by the k-th order Taylor approximation to f at the point (x, y_i). We can now state a version of the multi-jet transversality theorem which is used in our paper. The proof is similar to the one on p. 57–59 of [7], with some simplifications due to the fact that our maps are defined on Euclidean spaces, rather than on general manifolds.

Theorem 8.1 *Let W be a smooth submanifold of $J^k_s(\mathbb{R}^{m+n}; \mathbb{R})$. Then the set of functions $f \in \mathcal{C}^\infty(\mathbb{R}^{m+n}; \mathbb{R})$ which are transversal to W is dense, in the $\mathcal{C}^\infty$ topology.*

Proof. 1. Cover the open set $Z^{(s)}$ with countably many open sets V_ν, $\nu \geq 1$, such that

- If $(x, y_1, \ldots, y_s) \in V_\nu$ and $(\tilde{x}, \tilde{y}_1, \ldots, \tilde{y}_s) \in V_\nu$, then $y_i \neq \tilde{y}_j$ for all $1 \leq i < j \leq s$.

Construct $\mathcal{C}^\infty$ functions $\phi_\nu : \mathbb{R}^{m+n} \mapsto [0,1]$, $\nu \geq 1$, with the following properties.

- $\mathrm{Supp}(\phi_\nu) \subset V_\nu$.
- For each $\nu \geq 1$, call $V'_\nu \subset V_\nu$ the interior of the set where $\phi_\nu = 1$. Then $\bigcup_{\nu \geq 1} V'_\nu = Z^{(s)}$.

2. For each $\nu \geq 1$, and any polynomials $P_1, \ldots, P_s$ of degree $\leq k$, define the function

$$f^{(P)}(x,y) \;=\; f(x,y) + \sum_{\ell=1}^{s} \phi_k(x,y) P_\ell(x, y_\ell). \tag{8.4}$$

We now consider the map

$$(x, y_1, \ldots, y_s, P_1, \ldots, P_s) \;\mapsto\; \Big(j^k f^{(P)}(x, y_1), \;\ldots\; j^k f^{(P)}(x, y_s) \Big). \tag{8.5}$$

The right hand sides are the coefficients of the k-th order Taylor approximations to the maps $(x,y) \mapsto f^{(P)}(x,y)$ at the points (x, y_ℓ). Since $\phi_\nu \equiv 1$ on V'_ν, it is clear that the differential of the map (8.5) has full rank. Hence this map is transversal to any manifold W, restricted to V'_ν. By the transversality theorem, there is a residual set $\mathcal{S}_\nu \subset \mathcal{C}^\infty(\mathbb{R}^{m+n}; \mathbb{R})$ such that, for every $f \in \mathcal{S}_\nu$, the map $j^k_s f$ in (8.3) is transversal to W at every point $(x, y_1, \ldots, y_s, P_1, \ldots, P_s) \in W$ such that $(x, y_1, \ldots, y_s) \in V'_\nu$.

3. Repeating the same argument for every $\nu \geq 1$, we obtain a sequence of residual subsets $\mathcal{S}_\nu$. The intersection $\mathcal{S} \doteq \bigcap_{\nu \geq 1} \mathcal{S}_\nu$ is still residual in $\mathcal{C}^\infty(\mathbb{R}^{m+n}; \mathbb{R})$. By construction, for every $f \in \mathcal{S}$ the map $j^k_s f$ is transversal to W. □

References

[1] V. Arnold, *Catastrophe Theory*, Third edition. Springer-Verlag, Berlin-Heidelberg, 1992.

[2] T. Basar and G. J. Olsder, *Dynamic Noncooperative Games*, Academic Press, New York,1982.

[3] J. M. Bloom, The local structure of smooth maps of manifolds, B.A. Thesis, Harvard U., 2004.

[4] A. Bressan, Noncooperative differential games. *Milan J. of Mathematics*, **79** (2011), 357-427.

[5] A. Bressan and Y. Jiang, Self-consistent feedback Stackelberg solutions for infinite horizon stochastic games, preprint, 2019.

[6] J. Dieudonné, *Foundations of Modern Analysis.* Academic Press, New York - London, 1969.

[7] M. Golubitsky and V. Guillemin, *Stable Mappings and Their Singularities.* Springer-Verlag, New York, 1973.

[8] L. C. Evans, *Partial Differential Equations.* Second edition. American Mathematical Society, Providence, RI, 2010.

[9] J. Harsanyi, Oddness of the number of equilibrium points: a new proof. *Internat. J. Game Theory* **2** (1973), 235–250.

[10] M. B. Lignola and J. Morgan, Topological existence and stability for Stackelberg problems. *J. Optim. Theory Appl.* **84** (1995), 145–169.

[11] M. B. Lignola and J. Morgan, α-well-posedness for Nash equilibria and for optimization problems with Nash equilibrium constraints. *J. Global Optim.* **36** (2006), 439–459.

[12] L. Mallozzi and J. Morgan, Weak Stackelberg problem and mixed solutions under data perturbations, *Optimization* **32** (1995), 269–290.

[13] A. Marhfour, Mixed solutions for weak Stackelberg problems: existence and stability results. *J. Optim. Theory Appl.* **105** (2000), 417–440.

[14] C. Meroni and C. Pimienta, The structure of Nash equilibria in Poisson games. *J. Economic Theory* **169** (2017), 128–144.

[15] P. Olver, *Applications of Lie Groups to Differential Equations*, Second Edition. Springer-Verlag, New York, 1993.

[16] A. Sard, The measure of critical values of differentiable maps. *Bull. Amer. Math. Soc.* **48** (1942), 883–890.

[17] M. Simaan and J. Cruz, On the Stackelberg strategies in nonzero-sum games, *J. Optim. Theory Appl.* **11** (1973), 533–555.

[18] H. von Stackelberg, *The Theory of the Market Economy.* Oxford Univ. Press, 1952.