The target projection dynamic*

ELIAS TSAKAS†

Department of Economics, Göteborg University, Sweden

Department of Economics, Maastricht University, The Netherlands

MARK VOORNEVELD‡

Department of Econometrics and Operations Research, Tilburg University, The Netherlands

Department of Economics, Stockholm School of Economics, Sweden

Abstract

We study the target projection dynamic, a model of learning in normal form games. The dynamic is given a microeconomic foundation in terms of myopic optimization under control costs due to a certain status-quo bias. We establish a number of desirable properties of the dynamic: existence, uniqueness, and continuity of solution trajectories, Nash stationarity, positive correlation with payoffs, and innovation. Sufficient conditions are provided under which strictly dominated strategies are wiped out. Finally, some stability results are provided for special classes of games.

1. Introduction

The most well-known and extensively used solution concept in noncooperative game theory is the Nash equilibrium. The question how players may reach such equilibria is studied in a branch of game theory employing dynamic models of learning and strategic adjustment. The main dynamic processes in the theory of strategic form games include the replicator dynamic (Taylor and Jonker,

*We are indebted to Bill Sandholm, Jörgen Weibull, and an anonymous referee for their very useful comments. We thank Geir Asheim, Martin Dufwenberg, Friederike Mengel, Andrés Perea, Jeff Steif, and Jeroen Swinkels for fruitful discussions. Tsakas thanks the Stockholm School of Economics and the University of California, Berkeley, for their hospitality while working on this paper. Financial support from the Netherlands Organization for Scientific Research (NWO), the Wallander/Hedelius Foundation, and the Adlerbertska Foundation is gratefully acknowledged.

†Tel:+31-43-38 83649; Fax:+31-43-38 84878; E-mail: e.tsakas@algec.unimaas.nl

‡Tel:+46-8-736 9217; Fax:+46-8-31 32 07; E-mail: mark.voorneveld@hhs.se (corresponding author)

1978), the best-response dynamic (Gilboa and Matsui, 1991), and the Brown-Nash-von Neumann (BNN) dynamic (Brown and von Neumann, 1950). Sandholm (2005) introduced a definition for well-behaved evolutionary dynamics through a number of desiderata (see Theorem 3.2 for precise definitions):

> EXISTENCE, UNIQUENESS, AND CONTINUITY OF SOLUTIONS to the specified dynamic process,
>
> NASH STATIONARITY: the stationary points of the process coincide with the game's Nash equilibria,
>
> POSITIVE CORRELATION: roughly speaking, the probability of "good" strategies increases, that of "bad" strategies decreases.

He showed that — unlike the replicator and the best-response dynamics — the family of BNN or excess-payoff dynamics is well-behaved.

In the present paper we analyze the *target projection dynamic* that was mentioned only briefly in the same paper by Sandholm (2005, pp. 166–167). Our main results include the following:

Although the dynamic has a certain geometric appeal, Sandholm (2005, p. 167) wrote: "Unfortunately, we do not know of an appealing way of deriving this dynamic from a model of individual choice." This is remedied in Theorem 3.1, which provides a microeconomic foundation for the target projection dynamic. Following the control cost approach (Van Damme, 1991; Mattsson and Weibull, 2002; Voorneveld, 2006), we show that it models rational behavior in a setting where the players have to exert some effort/incur costs to deviate from incumbent strategies. In other words: the target projection dynamic is a best-response dynamic under a certain status-quo bias.

The fact that the players face control costs makes their adjustment process slower. This makes sense, since they are averse — to some extent — to deviations from their current behavior. However, despite the fact that their learning mechanism is quite conservative, the target projection dynamic is well-behaved in the sense described above. It also satisfies an additional property:

> INNOVATION: if some player is not at a stationary state and has an unused best response, then a positive probability is assigned to this best response.

This is established in Theorem 3.2. These properties imply (Hofbauer and Sandholm, 2007) that there are games where strictly dominated strategies survive under the target projection dynamic. Nevertheless, we show that strictly dominated strategies are wiped out if the "gap" between the dominated and dominant strategy is sufficiently large (Proposition 3.1) or if there are only two pure strategies (Proposition 3.2).

Like most other dynamics, the target projection dynamic belongs to family of uncoupled dynamics, where the behavior of one player is independent of payoffs to other players. Therefore, the process cannot converge to Nash equilibrium in all games (Hart and Mas-Colell, 2003). Nevertheless, some special cases can be established:

- sufficiently close to interior Nash equilibria of zero-sum games, the (standard Euclidean) distance to such an equilibrium remains constant (Corollary 4.1),
- strict Nash equilibria are asymptotically stable (Proposition 4.2),
- as are evolutionary stable strategies (Maynard Smith, 1982) if they are interior (Corollary 4.2) or the game is 2×2 (Proposition 4.3).

The first two points rely on the analysis of stable games in Hofbauer and Sandholm (2008). The analysis is facilitated by the fact that in certain cases (see Proposition 3.3) the target projection dynamic coincides with the projection dynamic extensively studied by Lahkar and Sandholm (2008) and Sandholm et al. (2008).

The paper is structured as follows. Section 2 specifies notation, recalls the framework of learning in normal form games, and provides some useful results on projections. Section 3 presents the target projection dynamic, including its control-cost motivation, establishes that it is well-behaved, and contains our results on strictly dominated strategies. Section 4 studies the target projection dynamic in special classes of games.

2. Notation and preliminaries

2.1. Learning in normal form games

Consider a finite normal form game $G = (N, (A_i)_{i \in N}, (u_i)_{i \in N})$ defined as usual: $N = \{1, \ldots, n\}$ is the set of players, $A_i = \{a_i^1, \ldots, a_i^{J_i}\}$ is player i's finite set of actions (pure strategies) with a_i being the typical element of A_i, and $u_i : A \to \mathbb{R}$ is i's payoff function, where $A = \times_{i \in N} A_i$ is the game's action space. Consider i's set of mixed strategies $\Delta_i := \{\alpha_i \in \mathbb{R}_+^{J_i} : \sum_{j=1}^{J_i} \alpha_i^j = 1\}$, with typical element $\alpha_i = (\alpha_i^1, \ldots, \alpha_i^{J_i})$, and let α_i^j denote the probability that α_i assigns to a_i^j. We say that α_i is completely mixed if $\alpha_i^j > 0$ for all $j = 1, \ldots, J_i$. With slight abuse of notation we write $\alpha = (\alpha_i, \alpha_{-i})$, where $\alpha_{-i} = (\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_n)$. We define the expected payoff $u_i : \Delta \to \mathbb{R}$ as usual, where $\Delta := \times_{i \in N} \Delta_i$. Then, we rewrite i's expected payoff as follows: $u_i(\alpha) = \langle \alpha_i, U_i(\alpha) \rangle$, where $\langle x, y \rangle := \sum_{i=1}^{m} x_i y_i$ denotes the usual inner product of two vectors

$x, y \in \mathbb{R}^m$, and $U_i(\alpha) = (U_i^1(\alpha), \ldots, U_i^{J_i}(\alpha))$ is the vector of expected payoffs to player i's pure strategies given that everybody else plays according to α, i.e., $U_i^j(\alpha) = u_i(a_i^j, \alpha_{-i})$.

We say that α_i is a best response to α if $\langle \alpha_i, U_i(\alpha) \rangle \geq \langle \beta_i, U_i(\alpha) \rangle$ for all $\beta_i \in \Delta_i$. If α_i is a best response to α for every $i \in N$ then α is a Nash equilibrium. If the inequality is strict for every $i \in N$ then α is a strict Nash equilibrium. Obviously, strict equilibria appear only in pure strategies. Action a_i^j strictly dominates a_i^k if $U_i^j(\alpha) > U_i^k(\alpha)$ for all $\alpha \in \Delta$. Then a_i^k is referred to as a dominated action.

Consider the standard framework of learning in normal form games (Börgers and Sarin, 1997; Fudenberg and Levine, 1998; Hopkins, 2002). There is a population of individuals. Every individual repeatedly participates in G against other individuals, always holding the same role, i.e., always being the same $i \in N$. The way the individuals are matched can vary according to the model: they play against the same opponents at every period (single-matching) or they are randomly matched with individuals playing the other roles (random-matching). At all periods every player i chooses a mixed strategy $\alpha_i \in \Delta_i$. After having observed the own payoff and (usually) the strategy profile $\alpha \in \Delta$, every $i \in N$ adjusts the own strategy according to some updating rule $\dot{\alpha}_i$, which typically puts higher weight on more profitable strategies and less weight on the ones that entail lower payoff. The adjustment rule describes the mechanism through which players learn how to play in the future, given their current observations, i.e., $\dot{\alpha} = f(\alpha)$. Dynamic processes that do not depend on the payoff earned by the other players are called uncoupled dynamics (Hart and Mas-Colell, 2003).

As we will see in the text following Theorem 3.1, the dynamic in this paper is most naturally interpreted in terms of learning, i.e., with one player per role and the profile of mixed strategies as state variable. Other dynamics come from models of evolution (some — like the best-response, logit, and replicator dynamic — have a natural interpretation in both contexts), where there is a population of players for each role and the state variable reflects the empirical distributions of strategies; see Fudenberg and Levine (1998), Sandholm (2006), and Weibull (1995). In somewhat more detail, there are n populations, each with a continuum of agents: one for each player role. Agents from the different populations are randomly matched to play the normal form game; agents from population i choose actions from A_i. The distribution $\alpha_i \in \Delta_i$ of actions is the population state. After having observed all population states and the earned payoff, individuals get the opportunity to revise their strategies, typically in such a way that populations evolve towards states where more profitable strategies are more popular. The expected motion of the populations defines the corresponding evolutionary dynamic.

2.2. *Projections*

This subsection contains some results on projections. In particular, Proposition 2.1 establishes a link between maximizing a linear function and certain projection problems. Proposition 2.2 gives a simple expression for projection onto the unit simplex. The proofs are in the Appendix.

As we have already mentioned, let $\langle x, y \rangle = \sum_{i=1}^{m} x_i y_i$ denote the usual inner product of two vectors $x, y \in \mathbb{R}^m$. Let $\|\cdot\|$ denote the standard Euclidean norm, i.e., $\|x\| = \langle x, x \rangle^{1/2}$. For $z \in \mathbb{R}$, let $[z]_+ := \max\{z, 0\}$.

Proposition 2.1. *Let $C \subseteq \mathbb{R}^n$ be nonempty and convex, $a \in \mathbb{R}^n$, $c \in C$ and $k > 0$.*

(i) The following two claims are equivalent:

(a) c solves $\max_{x \in C} \langle a, x \rangle$

(b) c solves $\max_{x \in C} \langle a, x \rangle - \frac{1}{k}\|x - c\|^2$.

(ii) The problem in (b) reduces to projecting $c + \frac{k}{2}a$ onto C:

$$\arg\max_{x \in C} \langle a, x \rangle - \frac{1}{k}\|x - c\|^2 = \arg\min_{x \in C} \|x - c - \frac{k}{2}a\|^2. \tag{1}$$

The following proposition characterizes projection onto a unit simplex.

Proposition 2.2. *Let $P : \mathbb{R}^n \to \Delta_n$ denote the projection on the $(n-1)$-dimensional unit simplex Δ_n: for $x \in \mathbb{R}^n$, $P(x) := \arg\min_{y \in \Delta_n} \|y - x\|$.*

(i) P is Lipschitz continuous.

(ii) For every $x \in \mathbb{R}^n$ the projection on Δ_n can be rewritten as follows

$$P(x) = ([x_1 + \lambda(x)]_+, \ldots, [x_n + \lambda(x)]_+), \tag{2}$$

where $\lambda(x) \in \mathbb{R}$ is the unique solution to $\sum_{i=1}^{n} [x_i + \lambda(x)]_+ = 1$.

Remark 2.1. Proposition 2.2*(ii)* immediately implies that for all $x \in \mathbb{R}^n$ and $i, j \in \{1, \ldots, n\}$: if $x_i - x_j \geq 1$, then $P_j(x) = 0$. $\triangleleft$

3. The target projection dynamic

The target projection dynamic, the dynamic process governing the learning mechanism that we study in this paper, was mentioned briefly in the concluding section of Sandholm (2005, pp. 166-167). It was originally defined in the framework of congestion networks by Friesz et al. (1994).

Definition 3.1. Consider a normal form game $(N, (A_i)_{i\in N}, (U_i)_{i\in N})$. The ***target projection dynamic (TPD)*** is defined, for each $i \in N$, by the differential equation

$$\dot{\alpha}_i = P_{\Delta_i}[\alpha_i + U_i(\alpha)] - \alpha_i, \tag{3}$$

where P_{Δ_i} denotes the projection on Δ_i with respect to the usual Euclidean distance.

The basic idea is simple and standard for most dynamic processes in game theory: the payoffs associated with the different actions determine the direction in which their weights are changed by reinforcing the better actions and decreasing the weight of worse ones.

In particular, under the TPD, player $i \in N$ starts from current mixed strategy α_i and moves in the direction of the payoff vector $U_i(\alpha)$ to obtain a preliminary target point for the revised strategy. Of course, simply running in the direction of the payoff vector $U_i(\alpha)$ might take you outside the strategy simplex, but Proposition 2.2*(ii)* assures that projection onto the strategy simplex — i.e., choosing the nearest feasible strategy in terms of Euclidean distance — does not affect the order of the coordinates. See Figure 1 for illustrations of the dynamic for a player with just two actions. The left panel involves a "standard" orthogonal projection; in the right panel, the feasible strategy nearest the preliminary target point $\alpha_i + U_i(\alpha)$ involves setting the probability of playing the second action to zero.

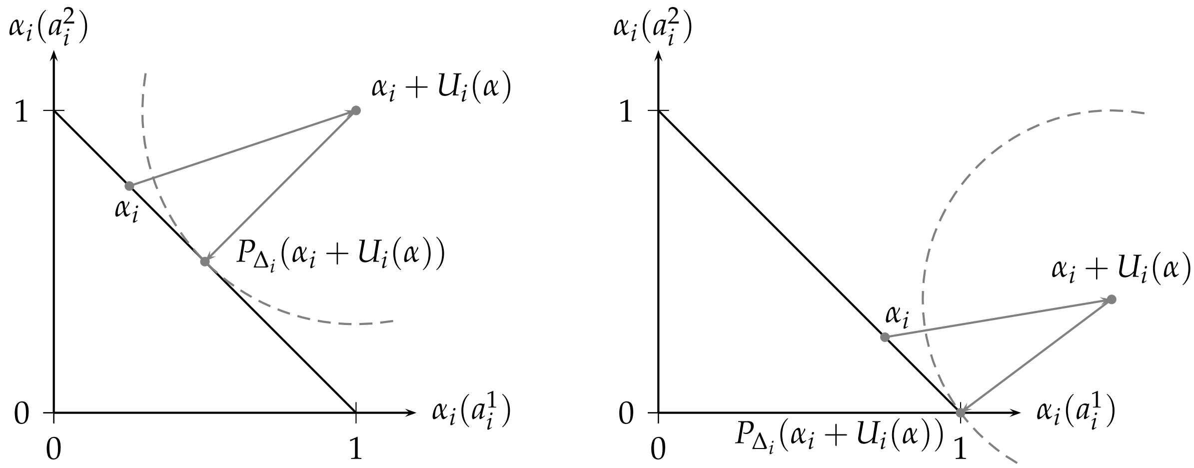

Figure 1: Two examples of the target projection dynamic for a player i with two actions.

The simplest dynamic embodying the basic idea of reinforcing actions with high payoffs at the expense of those with low ones is the best-response dynamic (Gilboa and Matsui, 1991): players place higher weight to their best responses according to

$$\dot{\alpha}_i = BR_i(\alpha) - \alpha_i, \tag{4}$$

where $BR_i(\alpha) = \arg\max_{\beta_i \in \Delta_i} \langle \beta_i, U_i(\alpha) \rangle$. In fact, Equation (4) is a differential inclusion, rather than a differential equation, since $BR_i(\alpha)$ need not be a singleton. In any case, as it becomes obvious from Equation (4), players do not switch to — but towards — their best responses. That is, the best-response dynamic involves inertia which precludes them from switching directly to their best response.

The next theorem indicates that the TPD is essentially a best-response dynamic, albeit under a bounded rationality assumption, involving the introduction of a certain status-quo bias. We follow the control cost approach, which since its introduction by Van Damme (1991) in the study of equilibrium refinements has proved to be a versatile way of providing microeconomic foundations for a variety of models of strategic behavior (Hofbauer and Sandholm, 2002; Mattsson and Weibull, 2002; Voorneveld, 2006). It does so by showing that such behavior is rational for decision makers who have to make some effort (incur costs) to implement their strategic choices. One intuitive way of modeling status-quo bias by player i could be as follows. Suppose that deviation from the current α_i is costly/requires effort in the sense that by switching to a strategy β_i, player i incurs a cost of $\frac{1}{2}||\beta_i - \alpha_i||^2$: staying at the current mixed strategy is costless, whereas large deviations, i.e., strategies further away from the current one in terms of Euclidean distance, incur larger costs. Taking such costs into account changes the optimization problem to

$$\max_{\beta_i \in \Delta_i} \langle \beta_i, U_i(\alpha) \rangle - \frac{1}{2}||\beta_i - \alpha_i||^2. \tag{5}$$

Let $B_i(\alpha) \in \Delta_i$ denote player i's (unique due to strict concavity of the goal function) best response against α, i.e., the unique solution to problem (5). Subject to these assumptions, we can now formulate the TPD as a best response dynamic:

Theorem 3.1. *Let $(N, (A_i)_{i \in N}, (U_i)_{i \in N})$ be a normal form game. The TPD is the best response dynamic for the control cost problem in (5), i.e., for each $i \in N$:*

$$\dot{\alpha}_i = P_{\Delta_i}[\alpha_i + U_i(\alpha_{-i})] - \alpha_i = B_i(\alpha) - \alpha_i. \tag{6}$$

Proof. By definition,

$$P_{\Delta_i}[\alpha_i + U_i(\alpha)] = \arg\min_{\beta_i \in \Delta_i} \|\beta_i - \alpha_i - U_i(\alpha)\|^2, \tag{7}$$

so we need to establish that

$$\arg\min_{\beta_i \in \Delta_i} \|\beta_i - \alpha_i - U_i(\alpha)\|^2 = \arg\max_{\beta_i \in \Delta_i} \langle \beta_i, U_i(\alpha) \rangle - \frac{1}{2}\|\beta_i - \alpha_i\|^2,$$

which follows from Proposition 2.1(ii) for $C = \Delta_i$, $x = \beta_i$, $c = \alpha_i$, $a = U_i(\alpha)$ and $k = 2$. $\square$

Theorem 3.1 shows that the TPD essentially is a best-response dynamic in a model of learning subject to two types of inertia. Firstly, it preserves the inertia that all best-response dynamics exhibit, i.e., the players do not switch to any of their best responses, but they place higher weight to them, thus shifting their behavior towards them. Secondly, players dislike moving away from their current strategies, the state variables in dynamic approaches to learning. The fact that they are averse to changing their current behavior makes the TPD a conservative rule of adjustment, and learning becomes slow.

It is much less natural to interpret the dynamic in terms of evolution in population games: there does not seem to be a convincing explanation — perhaps conformity issues might play a role — for why an agent for why an agent would care about the connection between his own strategy and the distribution of strategies over his population.

As pointed out to us by Bill Sandholm, the TPD belongs to a larger family of dynamics, characterized by the degree of aversion towards shifting away from the current strategy. Consider the parametric dynamic

$$\dot{\alpha}_i = P_{\Delta_i}[\alpha_i + \frac{k}{2}U_i(\alpha)] - \alpha_i, \tag{8}$$

where $k > 0$. Comparing (3) and (8) shows that the latter is the TPD after rescaling the game's payoffs with a factor $k/2$. Such rescaling affects neither the best response correspondence, nor the direction of the payoff vector (important for positive correlation). Hence, the parametric family of TPD in (8), no matter how conservative rules of adjustment it imposes, has a series of nice properties as shown in the following sections. By Proposition 2.1(ii), increasing k leads to lower control costs: the weight placed on these costs by the perturbed payoff function is $\frac{1}{k}$. The TPD arises when $k = 2$. When $k \to \infty$, the control cost becomes arbitrarily small and players will — roughly speaking — put more effort into choosing a best response in the original game, with less concern about the costs from deviating from their current strategy. On the other hand, the lower k the more conservative the players become, since they start placing more weight on their aversion towards change. In the limit ($k \to 0$), players never change their current strategy.

The fact that the TPD is interpreted as a best-response dynamic with control costs does not imply that the dynamic is susceptible to the extensive literature on (perturbed) best response dynamics (Gilboa and Matsui, 1991; Hofbauer and Sandholm, 2002; Fudenberg and Levine, 1998): Our status-quo bias models control costs arising due to deviations from the *current* state, while the usual approaches use control cost functions which:

- are independent of the current state: they define costs in terms of deviations from a fixed strategy, often uniform randomization (close your eyes and pick an action) as in Mattsson and Weibull (2002), and Voorneveld (2006),

- are often required to be steep near the boundary of the strategy space, as in Hofbauer and Sandholm (2002).

The logit dynamic (Fudenberg and Levine, 1998), for instance, takes the form

$$\forall i \in N, \forall j \in \{1, \ldots, J_i\}: \qquad \dot{\alpha}_i^j = \frac{\exp \eta^{-1} U_i^j(\alpha)}{\sum_{k=1}^{J_i} \exp \eta^{-1} U_i^k(\alpha)} - \alpha_i^j,$$

for some noise parameter $\eta > 0$ and is obtained as a perturbed best-response dynamic where each player i uses as control cost the Kullback-Leibler divergence (or relative entropy) from a given mixed strategy β_i to the mixed strategy $(1/J_i, \ldots, 1/J_i)$ assigning equal probability to all actions. Formally, this relative entropy is defined as

$$\sum_{j=1}^{J_i} \beta_i^j \ln \frac{\beta_i^j}{1/J_i} = \sum_{j=1}^{J_i} \beta_i^j \ln \beta_i^j + c,$$

where $c = \ln J_i$ is a constant and we use the convention that $0 \ln 0 = 0$. This transforms, in contrast with (5), player i's optimization problem to

$$\max_{\beta_i \in \Delta_i} \langle \beta_i, U_i(\alpha) \rangle - \eta \sum_{j=1}^{J_i} \beta_i^j \ln \beta_i^j,$$

where $\eta > 0$ is the weight assigned to the control cost term.

3.1. General properties

Theorem 3.2 states that the target projection dynamic satisfies a number of desirable properties of "nice" evolutionary dynamics. Indeed, Sandholm (2005) calls a dynamic ***well-behaved*** if it satisfies the first three properties of Theorem 3.2; he states, but does not prove, that the TPD is well-behaved.

Theorem 3.2. *Let $(N, (A_i)_{i \in N}, (U_i)_{i \in N})$ be a normal form game. The TPD satisfies the following properties:*

NASH STATIONARITY: *The stationary points of the TPD and the game's Nash equilibria coincide.*

BASIC SOLVABILITY: *For every initial state, a solution to the TPD exists, is unique, Lipschitz continuous in the initial state, and remains inside Δ at all times.*

POSITIVE CORRELATION: *Growth rates are positively correlated with payoffs: for each $i \in N$, if $\dot{\alpha}_i \neq 0$, then $\langle \dot{\alpha}_i, U_i(\alpha) \rangle > 0$.*

INNOVATION: *If some player is not at a stationary state and has an unused best response, then a positive probability is assigned to this best response. Formally, for each* $\alpha \in \Delta$ *and* $i \in N$, *if* $\dot{\alpha}_i \neq 0$, *but there is an action* $a_i^j \in A_i$ *with* $U_i^j(\alpha) = \max_{k \in \{1,\ldots,J_i\}} U_i^k(\alpha)$ *and* $\alpha_i^j = 0$, *then* $\dot{\alpha}_i^j > 0$.

Proof. NASH STATIONARITY: Let $\alpha \in \Delta$. By Proposition 2.1, Theorem 3.1, and (3), the following chain of equivalences holds:

$$
\begin{aligned}
\alpha \text{ is a Nash equilibrium} \quad &\Leftrightarrow \quad \forall i \in N : \alpha_i \in \arg\max_{\beta_i \in \Delta_i} \langle \beta_i, U_i(\alpha) \rangle \\
&\Leftrightarrow \quad \forall i \in N : \alpha_i \in \arg\max_{\beta_i \in \Delta_i} \langle \beta_i, U_i(\alpha) \rangle - \frac{1}{2}\|\beta_i - \alpha_i\|^2 \\
&\Leftrightarrow \quad \forall i \in N : \alpha_i = P_{\Delta_i}[\alpha_i + U_i(\alpha)] \\
&\Leftrightarrow \quad \forall i \in N : \dot{\alpha}_i = 0.
\end{aligned}
$$

BASIC SOLVABILITY: The target projection dynamic (3) is Lipschitz continuous. Let $i \in N$. By assumption, the payoff U_i is Lipschitz continuous, say with expansion factor $C > 0$. By Proposition 2.2, the projection is Lipschitz continuous with expansion factor 1. Using the triangle inequality, it follows for each $\alpha, \beta \in \Delta$ that

$$
\begin{aligned}
\|P[\alpha_i + U_i(\alpha)] - \alpha_i - P[\beta_i + U_i(\beta)] + \beta_i\| &\leq \|P[\alpha_i + U_i(\alpha)] - P[\beta_i + U_i(\beta)]\| + \|\alpha_i - \beta_i\| \\
&\leq \|\alpha_i + U_i(\alpha) - \beta_i - U_i(\beta)\| + \|\alpha_i - \beta_i\| \\
&\leq \|U_i(\alpha) - U_i(\beta)\| + 2\|\alpha_i - \beta_i\| \\
&\leq (C + 2)\|\alpha_i - \beta_i\|,
\end{aligned}
$$

establishing Lipschitz continuity of the vector field in (3). Since P_{Δ_i} maps onto Δ_i, it follows that $\sum_{j=1}^{J_i} \dot{\alpha}_i^j = 0$. Moreover, if $\alpha_i^j = 0$, then $\dot{\alpha}_i^j \geq 0$. This makes Δ forward-invariant. Together, these properties imply (Hirsch and Smale, 1974, Ch. 8) that for every initial state, a solution exists, is unique, Lipschitz continuous in the initial state, and remains in Δ at all times.

POSITIVE CORRELATION: Let $\alpha \in \Delta$ and $i \in N$. Suppose $\dot{\alpha}_i = P_{\Delta_i}[\alpha_i + U_i(\alpha)] - \alpha_i \neq 0$. Let $\beta_i = P_{\Delta_i}[\alpha_i + U_i(\alpha)] \neq \alpha_i$. Then, using Theorem 3.1, one obtains:

$$
\begin{aligned}
\langle \beta_i, U_i(\alpha) \rangle &> \langle \beta_i, U_i(\alpha) \rangle - \frac{1}{2}\|\alpha_i - \beta_i\|^2 \\
&\geq \langle \alpha_i, U_i(\alpha) \rangle - \frac{1}{2}\|\alpha_i - \alpha_i\|^2 \\
&= \langle \alpha_i, U_i(\alpha) \rangle,
\end{aligned}
$$

So $\langle \dot{\alpha}_i, U_i(\alpha) \rangle = \langle \beta_i - \alpha_i, U_i(\alpha) \rangle > 0$.

INNOVATION: Assume that the premises of the innovation property hold, but that $\dot{\alpha}_i^j \leq 0$. We derive a contradiction. By Proposition 2.2 there is a $\lambda \in \mathbb{R}$ such that

$$
P_{\Delta_i}[\alpha_i + U_i(\alpha)] = ([\alpha_i^1 + U_i^1(\alpha) + \lambda]_+, \ldots, [\alpha_i^{J_i} + U_i^{J_i}(\alpha) + \lambda]_+).
$$

By assumption, action j is unused ($\alpha_i^j = 0$) and $\dot{\alpha}_i^j \leq 0$, so

$$0 \geq \dot{\alpha}_i^j = [\alpha_i^j + U_i^j(\alpha) + \lambda]_+ - \alpha_i^j = [U_i^j(\alpha) + \lambda]_+ \geq 0,$$

i.e., $\dot{\alpha}_i^j = 0$ and $U_i^j(\alpha) + \lambda \leq 0$. But action j is a best response: $U_i^j(\alpha) = \max_{k \in \{1,\ldots,J_i\}} U_i^k(\alpha)$. Consequently, for every action $k \in \{1, \ldots, J_i\}$:

$$\dot{\alpha}_i^k = [\alpha_i^k + U_i^k(\alpha) + \lambda]_+ - \alpha_i^k \leq [\alpha_i^k + U_i^j(\alpha) + \lambda]_+ - \alpha_i^k \leq [\alpha_i^k + 0]_+ - \alpha_i^k = 0.$$

Since $\sum_{k=1}^{J_i} \dot{\alpha}_i^k = 0$, this implies that $\dot{\alpha}_i^k = 0$ for all $k \in \{1, \ldots, J_i\}$, in contradiction with the assumption that $\dot{\alpha}_i \neq 0$. □

3.2. *Strict domination: mind the gap*

Berger and Hofbauer (2006) show that under the Brown-von Neumann-Nash (BNN) dynamic, introduced in Brown and von Neumann (1950), there are games where a strictly dominated strategy survives. Hofbauer and Sandholm (2007) generalize this example: for each evolutionary dynamic satisfying the properties in Theorem 3.2 — actually, they restrict attention to single-population games — it is possible to construct a game with a strictly dominated strategy that survives along solutions of most initial states.

As their result applies to our TPD, it is of interest to investigate whether there are additional conditions under which such "bad" actions *are* wiped out. The next result shows that this is the case if one action strictly dominates another and the "gap" between them is sufficiently large.

Proposition 3.1. *Let $(N, (A_i)_{i \in N}, (U_i)_{i \in N})$ be a normal form game and let $i \in N$. Suppose there are actions $a_i^\ell, a_i^m \in A_i$ such that $U_i^\ell - U_i^m \geq 2$, i.e., action a_i^ℓ strictly dominates action a_i^m, and the gap between the payoffs is at least equal to 2. Then, the probability α_i^m converges to zero in the TPD.*

Proof. We show that the differential equation for the probability α_i^m of action a_i^m is given by $\dot{\alpha}_i^m = -\alpha_i^m$, because then $\alpha_i^m(t) = \alpha_i^m(0)e^{-t} \to 0$ as $t \to \infty$. Let $\alpha \in \Delta$. By (3), it suffices to show that the m-th coordinate of the projection $P_{\Delta_i}[\alpha_i + U_i(\alpha)]$ is zero. By Proposition 2.2(*ii*), there is a $\lambda \in \mathbb{R}$ such that its m-th and ℓ-th coordinate can be written as $[\alpha_i^m + U_i^m(\alpha) + \lambda]_+$ and $[\alpha_i^\ell + U_i^\ell(\alpha) + \lambda]_+$. Suppose, contrary to what we want to prove, that $[\alpha_i^m + U_i^m(\alpha) + \lambda]_+ > 0$. Then

$$\begin{aligned}
[\alpha_i^\ell + U_i^\ell(\alpha) + \lambda]_+ - [\alpha_i^m + U_i^m(\alpha) + \lambda]_+ &\geq \alpha_i^\ell - \alpha_i^m + U_i^\ell(\alpha) - U_i^m(\alpha) \\
&\geq \alpha_i^\ell - \alpha_i^m + 2 \\
&\geq 1,
\end{aligned} \tag{9}$$

since the difference between probabilities is bounded in absolute value by one. However, since $[\alpha_i^m + U_i^m(\alpha) + \lambda]_+ > 0$, the left-hand side of (9) is smaller than 1, a contradiction. □

Recall from the discussion on the parametric family of target projection dynamics (8) that they coincide with the TPD after rescaling payoffs. In particular, if one action strictly dominates another, the "gap" between them can be made arbitrarily large. This proves:

Corollary 3.1. *Let $(N, (A_i)_{i\in N}, (U_i)_{i\in N})$ be a normal form game and let $i \in N$. Suppose there are actions $a_i^\ell, a_i^m \in A_i$ such that a_i^ℓ strictly dominates a_i^m. For $k > 0$ sufficiently large, the probability α_i^m converges to zero in the parametric TPD (8).*

Also if a player has only two actions and one of them is strictly dominated, then it is eventually eliminated:

Proposition 3.2. *Let $(N, (A_i)_{i\in N}, (U_i)_{i\in N})$ be a normal form game and let $i \in N$. If $A_i = \{a_i^1, a_i^2\}$, and a_i^1 strictly dominates a_i^2, the probability assigned to a_i^2 converges to zero in the TPD.*

Proof. Let $\alpha \in \Delta$. By Proposition 2.2, there is a $\lambda(\alpha) \in \mathbb{R}$ such that the target projection dynamic for each of the two actions $j = 1, 2$ of population i can be rewritten as

$$\dot{\alpha}_i^j = [\alpha_i^j + U_i^j(\alpha) + \lambda(\alpha)]_+ - \alpha_i^j \tag{10}$$

Then $[\alpha_i^1 + U_i^1(\alpha) + \lambda(\alpha)]_+ > 0$. Suppose, to the contrary, that

$$\alpha_i^1 + U_i^1(\alpha) + \lambda(\alpha) \leq 0.$$

Since we project the two-dimensional vector onto the simplex, this implies

$$[\alpha_i^2 + U_i^2(\alpha) + \lambda(\alpha)]_+ = \alpha_i^2 + U_i^2(\alpha) + \lambda(\alpha) = 1.$$

Combining these two expressions gives

$$\alpha_i^2 - \alpha_i^1 \geq 1 + U_i^1(\alpha) - U_i^2(\alpha) > 1,$$

a contradiction, since the left-hand side is at most one. By continuity of the payoffs on the compact set Δ and strict domination, there is an $\varepsilon > 0$ such that $U_i^1(\alpha) - U_i^2(\alpha) > \varepsilon$ for each $\alpha \in \Delta$.

Distinguish two cases. First, if $[\alpha_i^2 + U_i^2(\alpha) + \lambda(\alpha)]_+ = 0$, then $\dot{\alpha}_i^2 = -\alpha_i^2$, so the probability α_i^2 decreases at an exponential rate. Second, if $[\alpha_i^2 + U_i^2(\alpha) + \lambda(\alpha)]_+ > 0$, combine this with the facts that $[\alpha_i^1 + U_i^1(\alpha) + \lambda(\alpha)]_+ > 0$ and that these two numbers add up to one, to deduce that $\lambda(\alpha) = -\frac{1}{2}(U_i^1(\alpha) + U_i^2(\alpha))$. So $\dot{\alpha}_i^2 = \frac{1}{2}(U_i^2(\alpha) - U_i^1(\alpha)) < -\frac{1}{2}\varepsilon$, i.e., the probability α_i^2 decreases at a rate bounded away from zero. Hence, along any solution trajectory, the probability α_i^2 of the dominated action converges to zero. □

3.3. The projection dynamic and the target projection dynamic

The projection dynamic was first developed by Nagurney and Zhang (1997) as part of the transportation literature, and was later introduced to game theory by Sandholm (2006), Lahkar and Sandholm (2008), and Sandholm et al. (2008). Let $T_i = \{\beta_i \in \mathbb{R}^{J_i} : \sum_{j=1}^{J_i} \beta_i^j = 0\}$ be the tangent space of Δ_i. Every $z_i \in T_i$ describes a motion between two points in Δ_i. The tangent cone of Δ_i at some strategy $\alpha_i \in \Delta_i$ is the set of feasible motions from α_i towards some other strategy in Δ_i, i.e., $T_i(\alpha) = \{\beta_i \in T_i : \alpha_i^j = 0 \Rightarrow \beta_i^j \geq 0\}$. Then, the projection dynamic is defined as follows

$$\dot{\alpha}_i^P = P_{T_i(\alpha)}[U_i(\alpha)], \tag{11}$$

The following proposition shows that the projection dynamic and the target projection dynamic coincide *(i)* in the interior of the simplex if the target projection is orthogonal to the motion it causes and *(ii)* close to completely mixed Nash equilibria.

Proposition 3.3. *Let $(N, (A_i)_{i\in N}, (U_i)_{i\in N})$ be a normal form game.*

(i) If $\alpha \in int(\Delta)$ and $\langle \dot{\alpha}_i, U_i(\alpha) - \dot{\alpha}_i \rangle = 0$, with $\dot{\alpha}$ as in (3), the TPD coincides with the projection dynamic.

(ii) If α is a completely mixed Nash equilibrium, then there is a neighborhood $\mathcal{O}$ of α such that the projection dynamic and the target projection dynamic coincide for all $\beta \in \mathcal{O}$.

Proof. *(i)* By definition the TPD solves, for each player $i \in N$, the following optimization problem:

$$\min_{\beta_i \in \Delta_i} \|\beta_i - \alpha_i - U_i(\alpha)\|^2, \quad \text{s.t.} \ \sum_{j=1}^{J_i} \beta_i^j = 1 \text{ and } \ \beta_i^j \geq 0 \text{ for all } j = 1, \ldots, J_i.$$

It follows from the Karush-Kuhn-Tucker conditions that there are $\nu \in \mathbb{R}$ and $\mu_j \geq 0$ such that for all $j = 1, \ldots, J_i$

$$\begin{aligned} \beta_i^j - \alpha_i^j - U_i^j(\alpha) + \nu - \mu_j &= 0, \\ \mu_j \beta_i^j &= 0. \end{aligned} \tag{12}$$

Summing (12) for all j, solving with respect to ν and substituting back yields

$$\beta_i^j - \alpha_i^j - U_i^j(\alpha) + \frac{1}{J_i}\sum_{j=1}^{J_i} U_i^j(\alpha) - \mu_j + \frac{1}{J_i}\sum_{j=1}^{J_i} \mu_j = 0. \tag{13}$$

We multiply by β_i^j, sum over j, and use complementary slackness ($\mu_j \beta_i^j = 0$):

$$\sum_{j=1}^{J_i} \Big({\beta_i^j}^2 - \alpha_i^j \beta_i^j - \beta_i^j U_i^j(\alpha) \Big) + \frac{1}{J_i}\sum_{j=1}^{J_i} U_i^j(\alpha) + \frac{1}{J_i}\sum_{j=1}^{J_i} \mu_j = 0. \tag{14}$$

We multiply now (13) by α_i^j and sum over j:

$$\sum_{j=1}^{J_i}\left(\alpha_i^j\beta_i^j - \alpha_i^{j2} - \alpha_i^j U_i^j(\alpha)\right) + \frac{1}{J_i}\sum_{j=1}^{J_i} U_i^j(\alpha) - \sum_{j=1}^{J_i}\mu_j\alpha_i^j + \frac{1}{J_i}\sum_{j=1}^{J_i}\mu_j = 0. \tag{15}$$

Now subtract (14) from (15) to find

$$\langle \beta_i - \alpha_i, \beta_i - \alpha_i - U_i(\alpha)\rangle = \sum_{j=1}^{J_i}\mu_j\alpha_i^j.$$

As $\beta_i = P_{\Delta_i}[\alpha_i + U_i(\alpha)]$, the orthogonality assumption implies that $\langle \beta_i - \alpha_i, \beta_i - \alpha_i - U_i(\alpha)\rangle = 0$. As $\mu_j\alpha_i^j \geq 0$ for all $j = 1, \ldots, J_i$ and $\alpha \in int(\Delta)$, it follows that $\mu_j = 0$ for all pure strategies $j = 1, \ldots, J_i$. Then, it follows from Proposition 2.2(ii) that the projection can be rewritten as

$$P_{\Delta_i}[\alpha_i + U_i(\alpha)] = \left([\alpha_i^1 + U_i^1(\alpha) + \lambda]_+, \ldots, [\alpha_i^{J_i} + U_i^{J_i}(\alpha) + \lambda]_+\right),$$

with $\alpha_i^j + U_i^j(\alpha) + \lambda \geq 0$ for all j, since $\mu_j = 0$. Hence, $\lambda = -\frac{1}{J_i}\sum_{j=1}^{J_i} U_i^j(\alpha)$, which implies that for every $i \in N$ and every $j \in J_i$ the TPD becomes

$$\dot{\alpha}_i^j = U_i^j(\alpha) - \frac{1}{J_i}\sum_{k=1}^{J_i} U_i^k(\alpha).$$

The previous formula is the projection dynamic for all completely mixed strategies (Sandholm et al., 2008), which proves the proposition.

(ii) By Nash stationarity, $\dot{\alpha}_i = 0$. Since $\alpha \in \mathrm{int}(\Delta)$, it follows that $\alpha_i^j > 0$ for every $i \in N$ and $j = 1, \ldots, J_i$. Let $i \in N$. By Proposition 2.2*(ii)*, there is a unique $\lambda(\alpha)$ such that $\alpha_i^j + U_i^j(\alpha) + \lambda(\alpha) > 0$ for all $j = 1, \ldots, J_i$. By continuity of the projection, there is a neighborhood $\mathcal{O}$ of α such that $\beta_i^j + U_i^j(\beta) + \lambda(\beta) > 0$, for every j and $\beta \in \mathcal{O}$. As above, it follows that

$$\lambda(\beta) = -\frac{1}{J_i}\sum_{j=1}^{J_i} U_i^j(\beta),$$

reducing the TPD to the projection dynamic on $\mathcal{O}$. □

4. Special classes of games

In this section we study the properties of the TPD in some special classes of games.

4.1. *Stable games*

Sandholm et al. (2008) prove a number of stability results for potential and stable games under the projection dynamic. Stable games (Hofbauer and Sandholm, 2008) are a family of normal form games characterized by the following condition:

$$\langle \alpha_i - \beta_i, U_i(\alpha) - U_i(\beta) \rangle \leq 0, \tag{16}$$

for every $\alpha_i, \beta_i \in \Delta_i$ and for all $i \in N$. The game is null (strictly) stable if (16) holds with equality (strict inequality).

Proposition 4.1. *Let $(N, (A_i)_{i\in N}, (U_i)_{i\in N})$ be a normal form game and let α be a completely mixed Nash equilibrium.*

(i) If the game is stable then α is Lyapunov stable under the TPD.

(ii) If the game is strictly stable then α is asymptotically stable under the TPD.

(iii) If the game is null stable then there is a neighborhood of α where the squared Euclidean distance to α, i.e., the function $\beta \mapsto \|\beta - \alpha\|^2$, defines a constant of motion under the TPD.

Proof. By Proposition 3.3*(ii)*, there is a neighborhood $\mathcal{O}$ of α where the projection dynamic coincides with the target projection dynamic for all players. Let $\varepsilon > 0$ be such that

$$\mathcal{O}_\varepsilon := \{\beta \in \Delta : \|\beta - \alpha\|^2 \leq \varepsilon\} \subseteq \mathcal{O}.$$

Sandholm et al. (2008) show that in the three cases of our proposition the function $L : \Delta \to \mathbb{R}$ with $L(\beta) := \|\beta - \alpha\|^2$ is (i) a Lyapunov function, (ii) a strict Lyapunov function, and (iii) defines a constant motion around α under the projection dynamic. Since $\mathcal{O}_\varepsilon \subseteq \mathcal{O}$, every trajectory starting in $\mathcal{O}_\epsilon$ will remain in it forever under the projection dynamic, and therefore under the target projection dynamic, completing the proof. □

4.2. Zero-sum games

A very interesting and widely explored class of games is the zero-sum games. We say that a normal form game is zero-sum if $\sum_{i\in N} u_i(a) = 0$ for every $a \in A$. Hofbauer and Sandholm (2008) establish that zero-sum games are null stable. By Proposition 4.1(iii), every trajectory of the TPD that gets sufficiently close to a completely mixed equilibrium in two-players zero-sum games forms a closed cyclical orbit around it.

Corollary 4.1. *Let $(N, (A_i)_{i\in N}, (U_i)_{i\in N})$ be a two-player normal form zero-sum game. If α is a completely mixed Nash equilibrium, there is a neighborhood $\mathcal{O}$ of α on which the the TPD forms a constant of motion around α.*

That is, if $\beta_0 \in \mathcal{O}$ and β belongs to the trajectory of the (unique) solution of (3) with initial value β_0, then $\|\beta - \alpha\| = \|\beta_0 - \alpha\|$. Typical examples include the matching pennies and the rock-paper-scissors games.

4.3. *Games with strict Nash equilibria*

Recall that in finite strategic games a Nash equilibrium is strict if each player chooses the unique best reply, i.e., $\alpha \in \Delta$ is a strict Nash equilibrium if $\langle \alpha_i, U_i(\alpha) \rangle > \langle \beta_i, U_i(\alpha) \rangle$ for all $\beta \in \Delta$ and all $i \in N$. Consequently, strict Nash equilibria are equilibria in pure strategies. This follows from the fact that a mixed strategy α_i is a best response to α if and only if all actions assigned positive probability by α_i are best responses to α, implying that i is indifferent between these actions, and therefore the necessary and sufficient condition for the strict equilibrium is violated.

Proposition 4.2. *Let $(N, (A_i)_{i\in N}, (U_i)_{i\in N})$ be a normal form game. If α is a strict Nash equilibrium, it is asymptotically stable under the TPD.*

Proof. Let α be a strict Nash equilibrium. Since α must be in pure strategies, without loss of generality, each $i \in N$ plays his first action: $\alpha_i = a_i^1$. By definition, for each $i \in N$ and $j \in \{2, \ldots, J_i\}$: $U_i^1(\alpha) > U_i^j(\alpha)$, so that $(\alpha_i^1 + U_i^1(\alpha)) - (\alpha_i^j + U_i^j(\alpha)) = 1 + U_i^1(\alpha) - U_i^j(\alpha) > 1$. By continuity, there is a neighborhood $\mathcal{O}$ of α such that for all $\beta \in \mathcal{O}$, $i \in N$, and $j \in \{2, \ldots, J_i\}$:

$$(\beta_i^1 + U_i^1(\beta)) - (\beta_i^j + U_i^j(\beta)) \geq 1.$$

For all $\beta \in \mathcal{O}$ and $i \in N$, Remark 2.1 implies that $P_{\Delta_i}(\beta_i + U_i(\beta)) = \alpha_i$; so $\dot{\beta}_i = \alpha_i - \beta_i$. Hence, the function $L : \mathcal{O} \to \mathbb{R}$ with $L(\beta) := \sum_{i\in N} \|\beta_i - \alpha_i\|^2$ is a Lyapunov function: It is non-negative, zero only at α, and if $\beta \in \mathcal{O} \setminus \{\alpha\}$:

$$\dot{L} = 2\sum_{i\in N} \langle \beta_i - \alpha_i, \dot{\beta}_i \rangle = 2\sum_{i\in N} \langle \beta_i - \alpha_i, \alpha_i - \beta_i \rangle = -2\sum_{i\in N} \|\beta_i - \alpha_i\|^2 < 0.$$

Given the existence of the Lyapunov function L, the equilibrium α is asymptotically stable (Hirsch and Smale, 1974, Ch. 9). $\square$

4.4. *Games with evolutionarily stable strategies*

We focus on symmetric two-player normal form games. A two player game is called symmetric if $A_1 = A_2 = A$ and $u_1(a_1, a_2) = u_2(a_1, a_2) = u(a_1, a_2)$ for all $a_1, a_2 \in A$. We say that (α, α) is an evolutionarily stable strategy (ESS) of a symmetric two-player normal form game (Maynard Smith, 1982; Weibull, 1995; Fudenberg and Levine, 1998) if for all $\beta \neq \alpha$:

(a) $\langle \alpha, U(\alpha) \rangle > \langle \beta, U(\alpha) \rangle$, or

(b) $\langle \alpha, U(\alpha) \rangle = \langle \beta, U(\alpha) \rangle$ and $\langle \alpha, U(\beta) \rangle > \langle \beta, U(\beta) \rangle$.

That is, a strategy is evolutionarily stable if it is robust to behavioral mutations: small mutations receive a strictly lower post-entry payoff than the incumbent strategy. ESS is a refinement of Nash equilibrium, so all ESS are rest points of every dynamic that satisfies Nash stationarity. Taylor and Jonker (1978), and Hofbauer et al. (1979) show that every ESS is asymptotically stable under the replicator dynamic. However, a similar result cannot be established for the TPD. Instead we provide some partial stability results.

Games with a completely mixed evolutionary stable strategy are strictly stable (Hofbauer and Sandholm, 2008). Proposition 4.1(ii) thus implies:

Corollary 4.2. *Let $(N, (A_i)_{i \in N}, (U_i)_{i \in N})$ be a symmetric two-player normal form game and let α be a completely mixed ESS. Then, it is asymptotically stable in the TPD.*

If in addition we restrict players to choose between only two actions we can extend the previous result to all ESS:

Proposition 4.3. *Let $(N, (A_i)_{i \in N}, (U_i)_{i \in N})$ be a symmetric 2×2 normal form game and let α be an ESS. Then, it is asymptotically stable in the TPD.*

Proof. For convenience, denote the strategy space by $\Delta = \{\beta \in \mathbb{R}^2_+ : \beta_1 + \beta_2 = 1\}$. Let $A = \{a_1, a_2\}$ and let $\alpha \in \Delta$ be an ESS. If $\alpha = (\alpha_1, \alpha_2)$ is completely mixed, apply Corollary 4.2. If not, assume without loss of generality that $\alpha = e_1$, where $e_1 = (1, 0)$, i.e., the pure strategy a_1 is an ESS.

If $U_1(e_1) > U_2(e_1)$ then it follows from convexity of the payoffs that

$$\langle e_1, U(e_1) \rangle = U_1(e_1) > \beta_1 U_1(e_1) + \beta_2 U_2(e_1) = \langle \beta, U(e_1) \rangle,$$

for all $\beta = (\beta_1, \beta_2) \in \Delta$. Hence, from Proposition 4.2 it follows that e_1 is asymptotically stable in the TPD.

Suppose now that $U_1(e_1) = U_2(e_1)$. Since e_1 is an ESS, it is also a Nash equilibrium, and therefore due to Theorem 3.2 it is a rest point. This implies that $P_\Delta[e_1^i + U_i(e_1)] = e_1^i$, which yields $P_\Delta[1 + U_1(e_1)] = 1 > 0$ for $i = 1$. It follows from continuity that there is a neighborhood $\mathcal{O}$ of e_1 such that $P_\Delta[\beta_1 + U_1(\beta)] > 0$ for all $\beta \in \mathcal{O}$. From Proposition 2.2(ii) it follows that there is $\lambda(\beta) \in \mathbb{R}$ such that $P_\Delta[\beta_1 + U_1(\beta)] = [\beta_1 + U_1(\beta) + \lambda(\beta)]_+ > 0$, implying that $P_\Delta[\beta_1 + U_1(\beta)] = \beta_1 + U_1(\beta) + \lambda(\beta)$, for all $\beta \in \mathcal{O}$. Hence,

$$\dot{\beta}_1 = U_1(\beta) + \lambda(\beta), \tag{17}$$

for all $\beta \in \mathcal{O}$. Now, we consider two cases:

CASE 1: Let $[\beta_2 + U_2(\beta) + \lambda(\beta)]_+ > 0$, which implies again due to Proposition 2.2(ii) that $\lambda(\beta) =$

$-\frac{1}{2}(U_1(\beta) + U_2(\beta))$. Hence, it follows from Equation (17) that

$$\begin{aligned}\dot{\beta}_1 &= U_1(\beta) - \frac{1}{2}(U_1(\beta) + U_2(\beta)) \\ &= \frac{1}{2}(U_1(\beta) - U_2(\beta)) \\ &= \frac{1}{2}\beta_1\Big(\langle e_1, U_1(e_1)\rangle - \langle e_2, U_2(e_1)\rangle\Big) + \frac{1}{2}\beta_2\Big(\langle e_1, U_1(e_2)\rangle - \langle e_2, U_2(e_2)\rangle\Big) \\ &= \frac{1}{2}\beta_2\Big(\langle e_1, U_1(e_2)\rangle - \langle e_2, U_2(e_2)\rangle\Big) > 0, \end{aligned} \tag{18}$$

where $e_2 = (0, 1)$ denotes the pure strategy a_2.

CASE 2: Let $[\beta_2 + U_2(\beta) + \lambda(\beta)]_+ = 0$, which implies again due to Proposition 2.2(ii) that $\lambda(\beta) = 1 - \beta_1 - U_1(\beta)$. Substituting into Equation (17) yields

$$\dot{\beta}_1 = 1 - \beta_1 > 0. \tag{19}$$

Consider now $L(\beta) = 1 - \beta_1$ for $\beta \in \mathcal{O}$, which is a Lyapunov function: it is non-negative, equal to zero only at e_1, and for all $\beta \in \mathcal{O} \setminus \{e_1\}$

$$\dot{L} = -\dot{\beta}_1 < 0,$$

which follows from (18) and (19) and completes the proof. □

Appendix

Proof of Proposition 2.1. (i) $[(a) \Rightarrow (b)]$ Assume (a) holds. Since $\|c - c\| = 0$, it follows, for each $x \in C$, that

$$\langle a, c\rangle - \frac{1}{k}\|c - c\|^2 = \langle a, c\rangle \geq \langle a, x\rangle \geq \langle a, x\rangle - \frac{1}{k}\|x - c\|^2,$$

so (b) holds.

$[(b) \Rightarrow (a)]$ Assume (b) holds. Let $x \in C$ and $\lambda \in (0, 1)$. By convexity, $\lambda x + (1 - \lambda)c \in C$. Since $\|c - c\| = 0$, it follows that

$$\begin{aligned}\langle a, c\rangle &\geq \langle a, \lambda x + (1-\lambda)c\rangle - \frac{1}{k}\|(\lambda x + (1-\lambda)c) - c\|^2 \\ &= \lambda\langle a, x\rangle + (1-\lambda)\langle a, c\rangle - \frac{\lambda^2}{k}\|x - c\|^2.\end{aligned}$$

Rearrange terms and divide by $\lambda > 0$ to obtain that $\langle a, c\rangle \geq \langle a, x\rangle - \frac{\lambda}{k}\|x - c\|^2$. Since $\lambda \in (0, 1)$ is arbitrary, let λ approach zero to establish (a).

(ii) Maximizing the function $x \mapsto \langle a, x\rangle - \frac{1}{k}\|x - c\|^2$ is equivalent with minimizing $x \mapsto \frac{1}{k}\|x - c\|^2 - \langle a, x\rangle$. It therefore suffices to show that the latter function is a positive affine transformation of $x \mapsto \|x - c - \frac{k}{2}a\|^2$.

Using the linearity and symmetry properties of the inner product, we find

$$\begin{aligned}\|x-c-\frac{k}{2}a\|^2 &= \langle x-c-\frac{k}{2}a, x-c-\frac{k}{2}a\rangle \\ &= \langle x-c, x-c\rangle - 2\langle\frac{k}{2}a, x\rangle + 2\langle c, \frac{k}{2}a\rangle + \langle\frac{k}{2}a, \frac{k}{2}a\rangle \\ &= \|x-c\|^2 - k\langle a, x\rangle + k\langle c, a\rangle + \frac{k^2}{4}\|a\|,\end{aligned}$$

which completes the proof, since the final two terms are independent of x, and $\|x-c\|^2 - k\langle a, x\rangle$ is a simple rescaling of $\frac{1}{k}\|x-c\|^2 - \langle a, x\rangle$. □

Proof of Proposition 2.2. (i) Recall from the Projection Theorem (see, for instance, Luenberger, 1969, p. 69) that for every $z \in \mathbb{R}^n$, $P(z)$ is characterized by $\langle z - P(z), w - P(z)\rangle \leq 0$ for all $w \in \Delta_n$. In particular, for all $x, y \in \mathbb{R}^n$:

$$\langle x - P(x), P(y) - P(x)\rangle \leq 0 \text{ and } \langle y - P(y), P(x) - P(y)\rangle \leq 0.$$

Write $\langle y - P(y), P(x) - P(y)\rangle = \langle P(y) - y, P(y) - P(x)\rangle$, add the two inequalities, and use Cauchy-Schwarz to establish

$$\begin{aligned} 0 &\geq \langle x - P(x) + P(y) - y, P(y) - P(x)\rangle \\ &= \|P(y) - P(x)\|^2 - \langle y - x, P(y) - P(x)\rangle \\ &\geq \|P(y) - P(x)\|^2 - \|y - x\| \, \|P(y) - P(x)\|.\end{aligned}$$

Conclude that $\|P(y) - P(x)\| \leq \|y - x\|$, i.e., P is Lipschitz continuous with expansion factor 1.
(ii) Let $x \in \mathbb{R}^n$. The function $T : \mathbb{R} \to \mathbb{R}$ defined for each $\lambda \in \mathbb{R}$ by $T(\lambda) = \sum_{i=1}^n [x_i + \lambda]_+$ is the composition of continuous functions and therefore continuous itself. Let $m = \max\{x_1, \ldots, x_n\}$. Then $T(\lambda) = 0$ for all $\lambda \in (-\infty, -m]$ and T is strictly increasing on $[-m, \infty)$, with $T(\lambda) \to \infty$ as $\lambda \to \infty$. By the Intermediate Value Theorem, there is a unique $\lambda(x) \in [-m, \infty)$ such that $T(\lambda(x)) = 1$.

By definition, $P(x)$ is the unique solution to $\min_{y \in \Delta_n} \frac{1}{2}\sum_{i=1}^n (y_i - x_i)^2$. This is a convex quadratic optimization problem with linear constraints, so the Karush-Kuhn-Tucker conditions are necessary and sufficient to characterize the minimum location: $y^* \in \Delta_n$ solves the problem if and only if there exist Lagrange multipliers $\mu_i \geq 0$ associated with the inequality constraints $y_i \geq 0$ and $\nu \in \mathbb{R}$ associated with the equality constraint $\sum_{i=1}^n y_i = 1$ such that for each $i = 1, \ldots, n$:

$$y_i^* - x_i - \mu_i + \nu = 0, \tag{20}$$

$$\mu_i y_i^* = 0. \tag{21}$$

Condition (20) is the first order condition obtained from differentiating the Lagrangian

$$(y, \mu_1, \ldots, \mu_n, \nu) \mapsto \frac{1}{2}\sum_{i=1}^n (y_i - x_i)^2 - \sum_{i=1}^n \mu_i y_i + \nu\left(\sum_{i=1}^n y_i - 1\right)$$

with respect to y_i and condition (21) is the complementary slackness condition. It is now easy to see that $y^* := ([x_1 + \lambda(x)]_+, \ldots, [x_n + \lambda(x)]_+)$ solves the minimization problem: set $\mu_i = 0$ if $[x_i + \lambda(x)]_+ > 0$,

$\mu_i = -x_i - \lambda(x) \geq 0$ if $[x_i + \lambda(x)]_+ \leq 0$, and $\nu = -\lambda(x)$. Substitution in (20) and (21) shows that these necessary and sufficient conditions are satisfied. $\square$

References

BERGER, U., HOFBAUER, J., (2006). Irrational behavior in the Brown-von Neumann-Nash dynamics. *Games and Economic Behavior* 56, 1–6.

BROWN, G., VON NEUMANN, J., (1950). Solutions of games by differential equations. *Annals of Mathematical Studies* 24, 73–79.

BÖRGERS, T., SARIN, R., (1997). Learning through reinforcement and replicator dynamics. *Journal of Economic Theory* 77, 1–14.

FRIESZ, T., BERNSTEIN, D., MEHTA, N., TOBIN, R., GANJALIZADEH, S., (1994). Day-to-day dynamic network disequilibria and idealized traveler information systems. *Operations Research* 46, 1120–1136.

FUDENBERG, D., LEVINE, D., (1998). *The Theory of Learning in Games*, MIT Press, Cambridge, Massachusetts.

GILBOA, I., MATSUI, A., (1991). Social stability and equilibrium. *Econometrica* 59, 859–867.

HART, S., MAS-COLELL, A., (2003). Uncoupled dynamics do not lead to Nash equilibrium, *American Economic Review* 93, 1830–1836.

HIRSCH, M., SMALE, S., (1974). *Differential equations, dynamical systems, and linear algebra*. New York: Academic Press.

HOFBAUER, J., SCHUSTER, P., SIGMUND, K., (1979). A note on evolutionarily stable strategies and game dynamics. *Journal of Theoretical Biology* 81, 609–612.

HOFBAUER, J., SANDHOLM, W., (2002). On the global convergence of stochastic fictitious play. *Econometrica* 70, 2265-2294.

——— (2007). Survival of dominated strategies under evolutionary dynamics, Working paper, University of Wisconsin.

——— (2008). Stable population games and integrability for evolutionary dynamics, Working paper, University of Wisconsin.

HOPKINS, E., (2002). Two competing models of how people learn in games. *Econometrica* 70, 2141–2166.

LAHKAR, R., SANDHOLM, W., (2008). The projection dynamic and the geometry of population games. *Games and Economic Behavior*, forthcoming.

LUENBERGER, D., (1969). *Optimization by vector space methods*. New York: John Wiley & Sons.

MAYNARD SMITH, J., (1982). *Evolution and the theory of games*, Cambridge University Press.

MATTSSON, L.-G., WEIBULL, J., (2002). Probabilistic choice and procedurally bounded rationality. *Games and Economic Behavior* 41, 61–78.

NAGURNEY, A., ZHANG, D., (1997). Projected dynamical systems in the formulation, stability analysis, and computation of fixed demand traffic network equilibria. *Transportation Science* 31, 147–158.

SANDHOLM, W., (2005). Excess payoff dynamics and other well-behaved evolutionary dynamics. *Journal of Economic Theory* 124, 149–170.

——— (2006). *Population games and evolutionary dynamics*. MIT Press, forthcoming.

SANDHOLM, W., DOKUMACI, E., LAHKAR, R., (2008). The projection dynamic and the replicator dynamic. *Games and Economic Behavior*, forthcoming.

TAYLOR, P., JONKER, L., (1978). Evolutionarily stable strategies and game dynamics. *Mathematical Biosciences* 16, 76–83.

VAN DAMME, E., (1991). *Stability and perfection of Nash equilibria*, 2nd ed. Springer, Berlin, Heidelberg, New York.

VOORNEVELD, M., (2006). Probabilistic choice in games: Properties of Rosenthal's *t*-solutions. *International Journal of Game Theory* 34, 105–121.

WEIBULL, J., (1995). *Evolutionary Game Theory*, MIT Press, Cambridge, Massachusetts.