

Article

Population Games, Stable Games, and Passivity

Michael J. Fox and Jeff S. Shamma *

School of Electrical and Computer Engineering, Georgia Institute of Technology, 777 Atlantic Drive NW, Atlanta, GA 30332, USA

* Author to whom correspondence should be addressed; E-mail: shamma@gatech.edu.

Received: 4 April 2013; in revised form: 3 September 2013 / Accepted: 26 September 2013 /

Published: 7 October 2013

Abstract: The class of “stable games”, introduced by Hofbauer and Sandholm in 2009, has the attractive property of admitting global convergence to equilibria under many evolutionary dynamics. We show that stable games can be identified as a special case of the feedback-system-theoretic notion of a “passive” dynamical system. Motivated by this observation, we develop a notion of passivity for evolutionary dynamics that complements the definition of the class of stable games. Since interconnections of passive dynamical systems exhibit stable behavior, we can make conclusions about passive evolutionary dynamics coupled with stable games. We show how established evolutionary dynamics qualify as passive dynamical systems. Moreover, we exploit the flexibility of the definition of passive dynamical systems to analyze generalizations of stable games and evolutionary dynamics that include forecasting heuristics as well as certain games with memory.

Keywords: population games; evolutionary games; passivity theory

1. Introduction

Evolutionary game theory (e.g., [1]), along with the related topic of learning in games, explores the dynamics of interacting players or populations. The framework entails introducing an “evolutionary dynamic” (or behavioral rule) and proceeding to analyze the resulting trajectories that a game induces. One of the goals of evolutionary game theory is to understand how rational solution concepts, such as Nash equilibrium, can emerge and be selected through simplistic evolutionary interactions. Oftentimes, an evolutionary dynamic for a specific game need not induce any stable equilibrium points. Indeed, different evolutionary dynamics can produce outcomes ranging from chaos to convergence (cf., [2])

and [3]). Furthermore, games have been constructed that can be shown to never exhibit stable Nash equilibria under only very mild conditions on the dynamics themselves [4]. In contrast to such specific examples, researchers also have sought to identify broad classes of games for which correspondingly broad classes of dynamics converge to equilibrium, one example of which is potential games [5]. For a broader discussion of these and related issues, see [6–8].

We focus here on the recently proposed class of *stable games* [9]—a generalization of a number of earlier classes of games ranging from concave potential games to symmetric normal form games with an interior evolutionarily stable strategy (ESS). The appealing property of stable games is that their Nash equilibria comprise a convex set that many dynamics are guaranteed to reach [9].

In this paper, we show that stable games can be identified as a special case of the feedback-system-theoretic notion of passive input–output systems. Passivity is an abstraction of energy conservation and dissipation in mechanical and electrical systems [10] that has become a standard tool in the design and analysis of nonlinear feedback systems [10–13]. It provides conditions under which particular system interconnections will be stable.

The connection to evolutionary games is that an evolutionary dynamic can be viewed as a dynamical system in feedback with a game. Accordingly, after we identify stable games as passive systems, we are guaranteed that play by any admissible passive evolutionary dynamic will result in stability and convergence—provided that one indeed can define an analogous notion of passivity for evolutionary dynamics. It turns out that the various dynamics that guarantee global convergence in stable games do indeed satisfy a natural notion of passivity. In particular, passivity of an evolutionary dynamic can be interpreted as *long run correlation* between the time derivative of payoffs and the direction of motion (see Equation (54)). While passivity techniques have been used in analysis of game theoretic learning dynamics employed in certain specific engineering models [14,15], the notion of passivity capturing a class of dynamics or games is novel.

The contributions of this paper are summarized as follows:

- We show that stable games can be formulated as dynamical systems exhibiting an appropriate form of passivity. We develop a complementary notion of passivity for learning dynamics, resulting in sufficient conditions for stability.
- We show that learning dynamics known to exhibit convergence in stable games do indeed satisfy a relevant passivity condition.
- The passivity conditions we introduce also apply to more general forms of games and dynamics. We provide stability results for certain games with memory as well as for learning dynamics that employ forecasting heuristics.
- Finally, we extend our methods to games and dynamics that depend on finite histories of strategy change and proceed to analyze learning techniques that can achieve convergence in certain games with lags or time delays.

In pursuing a broad class of evolutionary dynamics that converges for the class of stable games, the analysis in [9] considers a particular evolutionary dynamic called “excess payoff/target” (EPT) dynamics that generalizes various other evolutionary dynamics. Certain EPT dynamics exhibit a

property known as “positive correlation” (cf., its variant, termed virtual positive correlation [16]). In particular, there is positive correlation between the instantaneous payoffs and the instantaneous direction of motion. It turns out that positive correlation alone need not imply convergence for stable games [9]. This realization motivated the introduction of additional properties such as “integrability” or the stronger “separability” specific to EPT dynamics. We will see that, rather than positive correlation, the aforementioned long run correlations provide the essential characteristic for convergence under stable games. Furthermore, these results suggest an interpretation of passive learning that complements the interpretation of stable games as strategic environments exhibiting self-defeating externalities.

An immediate benefit of our characterization, beyond providing a unifying framework to assess stability, is the novel generalizations it enables. Evolutionary game theory has historically placed particular emphasis on the study of memoryless games, in which payoffs are a function of current population levels rather than histories of population levels. Likewise, evolutionary game theory has historically only considered evolutionary dynamics of restricted dimension, in particular with order equal to the total number of strategies across all players. While our definitions include such settings, they are not restricted to them. Dynamic learning schemes that utilize additional, auxiliary states in reckoning strategy changes also can be analyzed using passivity. In particular, we identify games that preserve the convergence properties of passive learning dynamics when they are combined with prevalent forecasting heuristics like smoothing and trend following. Alternatively, certain dynamic games, that is, strategic environments where payoffs can depend on the entire action trajectory, can be shown to exhibit passivity.

Lastly, we probe the limits of the class of passive learning dynamics by suggesting an evolutionary dynamic in which players attempt to update strategies in a contrarian manner. Specifically, they discount payoffs to actions that have seen a rise in popularity over a defined lookback period. This scheme leads to an infinite-dimensional system. We find that this predisposition has no consequences for convergence of passive dynamics in stable games and all other passive strategic environments.

The remainder of this paper is organized as follows. Section 2 presents background material on population games and stable games. Section 3 presents a brief tutorial on passivity analysis for feedback systems. Section 4 contains the main results that define the notion of passivity for evolutionary dynamics and establish stability when coupled with a stable game. This section goes on to present various generalizations. Finally, Section 5 contains concluding remarks.

This paper expands and develops on the work reported by the authors in [17].

2. Background

2.1. Population Games

We begin with a description of single population games. The main results herein are extendable to multi-population games considered in [9] in a straightforward manner, and so the restriction to single populations simplifies notation.

A single population has a set of available strategies $S = \{1, 2, \dots, n\}$. The set of strategy distributions is $X = \{x \in \mathbb{R}_+^n : \sum_{i \in S} x_i = 1\}$. Since strategies lie in the simplex, admissible changes in strategy are restricted to the *tangent space* $TX = \{z \in \mathbb{R}^n : \sum_{i \in S} z_i = 0\}$.

The payoff function $F : X \rightarrow \mathbb{R}^n$ is a continuous map associating each strategy distribution with a payoff vector so that $F_i : X \rightarrow \mathbb{R}$ is the payoff to strategy $i \in S$.

A state $x \in X$ is a *Nash equilibrium*, denoted $x \in NE(F)$, if each strategy in the support of x receives the maximum payoff available to the population, i.e.,

$$x \in NE(F) \Leftrightarrow [x_i > 0 \Rightarrow F_i(x) \geq F_j(x)], \quad \forall i, j \in S. \quad (1)$$

2.2. Stable Games

We say that $F : X \rightarrow \mathbb{R}^n$ is a *stable game* if

$$(y - x)^T (F(y) - F(x)) \leq 0, \quad \forall x, y \in X. \quad (2)$$

For a detailed discussion with examples of stable games, see [9]. Many evolutionary dynamics are well-behaved when restricted to the stable games. The above definition has the following equivalence when F is continuously differentiable (i.e., C^1).

Theorem 2.1 [9] *Suppose the population game F is continuously differentiable. Then F is a stable game if and only if the Jacobian matrix $DF(x)$ is negative semidefinite with respect to TX for all $x \in X$, i.e., $z^T DF(x)z \leq 0$ for all $z \in TX$ and $x \in X$.*

If $DF(x)$ is negative definite for all $x \in X$, then we say that F is *strictly stable*.

Reference [9] gives the definition of stable games the interpretation of *self-defeating externalities*. That is, the payoff improvements to strategies being adopted are dominated by the payoff improvements to strategies being abandoned. This is easy to see by letting $z = e_j - e_i \in TX$, the difference between two unit vectors, and noting that (by definition) $z^T DF(x)z \leq 0$.

Many games are known to be stable games including zero sum games, symmetric multi/network zero sum games, and concave potential games. For a thorough discussion see [9].

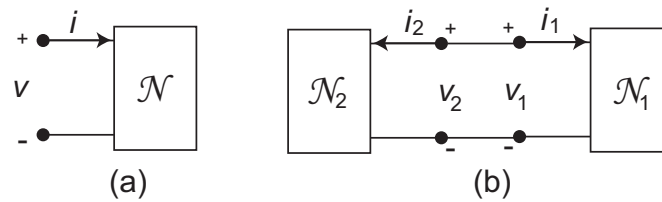
3. A Primer on Passivity in Feedback Systems

This section presents a brief tutorial on passivity analysis for feedback systems. For further discussion, see [10–13].

3.1. Circuit Origins

The term “passivity” is derived from circuit theory. Figure 1a illustrates a circuit network with two terminals. Let $v(t)$ denote the time-varying voltage across the terminals and $i(t)$ denote the corresponding time-varying current into the network. The relationship between $v(\cdot)$ and $i(\cdot)$ is determined by the specifics of the network. Let $E(t)$ denote the energy stored in the network. The instantaneous power delivered to the network at time t is $v(t) \cdot i(t)$. If the network is *passive*, it does not have any internal “active” or “power generating” components. Accordingly, over any time interval $[t_0, t]$,

$$E(t) \leq E(t_0) + \int_{t_0}^t v(\tau) i(\tau) d\tau. \quad (3)$$

Figure 1. (a) One-port circuit network. (b) Interconnection of two circuit networks.

That is, the change in stored energy is less than the supplied energy (“less than” because energy may have dissipated, e.g., through resistors).

Figure 1b illustrates an *interconnection* of two networks. If both networks are passive, i.e., have no power generation, then one would expect that the stored energy, which can be transferred from network to network, does not increase overall.

Mathematically, passivity of the networks individually means

$$E_1(t) \leq E_1(t_0) + \int_{t_0}^t v_1(\tau) i_1(\tau) d\tau \quad (4)$$

$$E_2(t) \leq E_2(t_0) + \int_{t_0}^t v_2(\tau) i_2(\tau) d\tau \quad (5)$$

where $v_j(\cdot)$, $i_j(\cdot)$, and $E_j(\cdot)$, $j \in \{1, 2\}$, are the voltage, current, and energy of the two networks. Furthermore, the interconnection implies that

$$v_1(t) = v_2(t) \quad \text{and} \quad i_1(t) = -i_2(t). \quad (6)$$

Combining these relations with the above passivity conditions implies that

$$E_1(t) + E_2(t) \leq E_1(t_0) + E_2(t_0), \quad (7)$$

as expected. Passivity analysis extends these ideas from circuits to feedback interconnections of dynamical systems.

3.2. Input–Output Operators and Feedback Interconnections

A dynamical system can be viewed as a (nonlinear) operator defined on function spaces. Let $\mathbb{L}_2$ denote the usual Hilbert space of square integrable functions¹ mapping $[0, \infty) := \mathbb{R}_+$ to $\mathbb{R}^n$ with inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$, i.e., for $f, g \in \mathbb{L}_2$,

$$\langle f, g \rangle = \int_0^\infty f(t)^T g(t) dt, \quad (8)$$

$$\|f\| = \left(\int_0^\infty f(t)^T f(t) dt \right)^{1/2} = \langle f, f \rangle^{1/2}. \quad (9)$$

Let $\mathbb{L}_{2,e}$ denote the “extended” space of functions that are square integrable over finite intervals, i.e.,

$$\mathbb{L}_{2,e} = \left\{ f : \mathbb{R}_+ \rightarrow \mathbb{R}^n : \int_0^T f(t)^T f(t) dt < \infty \text{ for all } T \in \mathbb{R}_+ \right\}. \quad (10)$$

For $\mathbb{U}, \mathbb{Y} \subset \mathbb{L}_{2,e}$ subsets of functions, an input–output operator is a mapping $S : \mathbb{U} \rightarrow \mathbb{Y}$.

¹For notational simplicity, we suppress the dimension n in $\mathbb{L}_2$.

Figure 2. (a) Feedback interconnection with negative feedback. (b) Feedback interconnection with positive feedback.

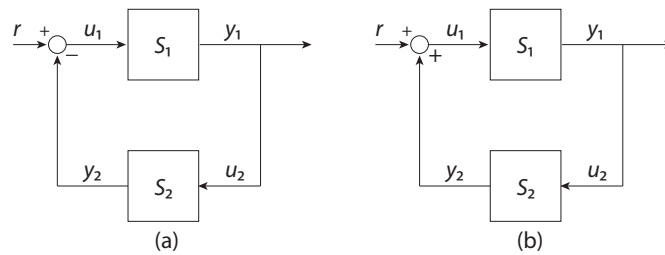


Figure 2a represents a feedback interconnection of two input–output operators, $S_1 : \mathbb{U} \rightarrow \mathbb{Y}$ and $S_2 : \mathbb{Y} \rightarrow \mathbb{U}$. The feedback interconnection makes the “output” of S_1 becomes the “input” of S_2 . Likewise, the negative output of S_2 becomes the input of S_1 after summation with the external $r \in \mathbb{U}$. Traditionally in control theory, feedback interconnections are with negative feedback as in Figure 2a rather than positive feedback as in Figure 2b. However, we will use positive feedback interconnections for games in the following analysis.

Figure 2a graphically represents the equations

$$\begin{aligned} u_1 &= r - y_2 \\ y_1 &= S_1 u_1 \\ y_2 &= S_2 u_2 \\ u_2 &= y_1, \end{aligned} \quad (11)$$

which imply

$$u_1 = r - S_2 S_1 u_1. \quad (12)$$

We assume that for any $r \in \mathbb{U}$, there exists a solution $u_1 \in \mathbb{U}$. Given this solution, u_1 , one can go on to construct y_1 , u_2 , and y_2 accordingly.

3.3. Passivity with Input–Output Operators

We now define passivity for input–output operators and present a basic theorem on the stability of feedback interconnections of passive operators. First, for $T \in \mathbb{R}_+$, define the truncated inner-product

$$\langle f, g \rangle_T = \int_0^T f(t)^T g(t) dt \quad (13)$$

and truncated norm

$$\|f\|_T = \langle f, f \rangle_T^{1/2}. \quad (14)$$

The input–output operator $S : \mathbb{U} \rightarrow \mathbb{Y}$ is *passive* if there exists a constant α such that

$$\langle Su, u \rangle_T \geq \alpha, \quad \text{for all } u \in \mathbb{U}, T \in \mathbb{R}_+, \quad (15)$$

and *input strictly passive* if there exist $\beta > 0$ and α such that

$$\langle Su, u \rangle_T \geq \alpha + \beta \langle u, u \rangle_T, \quad \text{for all } u \in \mathbb{U}, T \in \mathbb{R}_+. \quad (16)$$

Theorem 3.1 Consider the feedback interconnection of S_1 and S_2 defined by (11). Assume S_1 is passive and S_2 is input strictly passive. Then $r \in \mathbb{L}_2$ implies $y_1 \in \mathbb{L}_2$.

In words, Theorem 3.1 can be interpreted as the stability of a feedback interconnection between a passive and strictly passive system. The proof is straightforward. Using Equation (11),

$$\langle r, y_1 \rangle_T = \langle u_1 + y_2, y_1 \rangle_T = \langle u_1, S_1 u_1 \rangle_T + \langle S_2 u_2, u_2 \rangle_T. \quad (17)$$

Therefore

$$\langle r, y_1 \rangle_T \geq \alpha_1 + \alpha_2 + \beta_2 \langle u_2, u_2 \rangle_T = \alpha_1 + \alpha_2 + \beta_2 \langle y_1, y_1 \rangle_T, \quad (18)$$

where α_1 , α_2 , and β_2 are the passivity constants associated with S_1 and S_2 . By the Schwarz inequality,

$$\|r\|_T \|y_1\|_T \geq \alpha_1 + \alpha_2 + \beta_2 \|y_1\|_T^2. \quad (19)$$

By assumption, $r \in \mathbb{L}_2$, and so $\|r\|_T$ is uniformly bounded over $T \geq 0$. Since $\beta_2 > 0$, it follows that $\|y_1\|_T$ is also uniformly bounded over $T \geq 0$.

3.4. Passivity with State Space Systems

The previous discussion was in terms of general input–output operators and illustrates that passivity methods need not be restricted to dynamical systems described by differential equations. We now present specialized results when the input–output operators are, in fact, represented by a system of differential equations. Consider the state space system

$$\begin{aligned} \dot{x} &= f(x, u), & x(0) &= x_0, \\ y &= g(x, u), \end{aligned} \quad (20)$$

where $u(t) \in \mathcal{U} \subset \mathbb{R}^m$, $y(t) \in \mathbb{R}^m$, and $x_0 \in \mathcal{M} \subset \mathbb{R}^n$. Let us assume that for some classes of functions, $\mathbb{U}$ and $\mathbb{Y}$, and for any initial condition in $\mathcal{M}$, there exists a solution for all $u \in \mathbb{U}$ resulting in $y \in \mathbb{Y}$ and $x(t) \in \mathcal{M}$ for all $t \geq 0$ (i.e., $\mathcal{M}$ is invariant). These equations define a family of input–output operators (indexed by the initial condition, x_0), mapping $u(\cdot)$ to $y(\cdot)$. Accordingly, one can apply input–output operator notions of passivity for each initial condition.

For state space systems, however, the traditional approach is to define passivity irrespective of initial conditions as follows. A state-space system is *passive* if there exists a storage function $L : \mathcal{M} \rightarrow \mathbb{R}_+$ such that for all $x_0 \in \mathcal{M}$, $u \in \mathbb{U}$, and $t \in \mathbb{R}_+$,

$$L(x(t)) \leq L(x_0) + \int_0^t u(\tau)^T y(\tau) d\tau. \quad (21)$$

The parallels with the discussion of circuits are that the “input” is identified as the voltage and “output” is identified as the current (or *vice versa*), which in turns motivate the terminology of a “storage” function.

We will focus on C^1 storage functions, in which case the above definition may be written as

$$\dot{L} := DL(x)f(x, u) \leq u^T y = u^T g(x, u). \quad (22)$$

for all $x \in \mathcal{M}$ and $u \in \mathcal{U}$. Inequalities such as (22) are referred to as “dissipation inequalities”.

We now consider a feedback interconnection of two state space systems as in Figure 2a, in which each system, S_i , is represented by state space equations

$$\begin{aligned}\dot{x}_i &= f_i(x, u) & x_i(0) &= x_{i0} \in \mathcal{M}_i, & i &\in \{1, 2\} \\ y_i &= g_i(x_i, u_i), & i &\in \{1, 2\} \\ u_1 &= r - y_2 \\ u_2 &= y_1.\end{aligned}\tag{23}$$

Again, we assume the existence of solutions.

Theorem 3.2 *Consider the feedback interconnection of S_1 and S_2 defined by Equation (23). Assume each S_i is passive with storage function L_i . Then for all $x_1(0) \in \mathcal{M}_1$, $x_2(0) \in \mathcal{M}_2$, and $t \in \mathbb{R}_+$,*

$$L_1(x_1(t)) + L_2(x_2(t)) \leq L_1(x_1(0)) + L_2(x_2(0)) + \int_0^t r(\tau)^T y_1(\tau) d\tau.\tag{24}$$

An interpretation is that the *closed-loop* system mapping r to y_1 is passive with storage function $L_1 + L_2$. The derivation is an immediate consequence of the definitions.

In case $r = 0$, Theorem 3.2 implies that $L_1(x_1(t)) + L_2(x_2(t))$ is non-increasing along solutions of state space (23). Compare to the previous discussion on circuits, in which energy was non-increasing for interconnected circuits. This monotonicity can have stability implications depending on the underlying specifics. For example, if both L_1 and L_2 are positive definite with $L_i(0) = 0$, and if the passivity inequalities of (22) are strict for $x_i \neq 0$, then the origin is locally asymptotically stable. These details are omitted here. The main point for now is that, as before, there are stability implications associated with the feedback interconnection of two passive systems.

4. Main Results

4.1. Preview

Traditionally, we think of games as memoryless mappings from strategy $x \in X$ to payoff $F(x)$. In a dynamic setting, we can extend this viewpoint to a mapping of strategy *trajectories* $x(\cdot)$ to payoff *trajectories* $\pi(\cdot) = F(x(\cdot))$. Likewise, evolutionary dynamics (see below) can be viewed as mappings from payoff trajectories to strategy trajectories. Accordingly, we can view evolutionary games as a feedback interconnection. In terms of the positive feedback digram Figure 2b, we will set $r = 0$ and associate S_1 with an evolutionary dynamic process and S_2 with a game. Accordingly, $u_1 = y_2$ represent payoff trajectories and $y_1 = u_2$ represent strategy trajectories.

Here we find the first departure from standard passivity analysis. Namely, we are dealing with positive feedback as in Figure 2b instead of the traditional negative feedback of Figure 2a, and so we will need to make suitable (simple) modifications to the analysis.

There is a more significant departure from standard passivity analysis that is at the heart of associating stable games with passive systems. As mentioned above, let us think of a passive game as an input–output operator mapping strategy trajectories to payoff trajectories. Then for C^1 games,

$$\dot{\pi}(t) = DF(x(t))\dot{x}(t). \quad (25)$$

By the definition of stable games,

$$\dot{x}(t)^T \dot{\pi}(t) = \dot{x}(t)^T DF(x(t))\dot{x}(t) \leq 0 \quad (26)$$

pointwise in time. Consequently,

$$\int_0^T \dot{x}(t)^T \dot{\pi}(t) dt = \langle \dot{\pi}, \dot{x} \rangle_T \leq 0. \quad (27)$$

While this inequality resembles a passivity condition, we find two differences. First, the direction of the inequality has changed. This reversal suits the shift to positive feedback interconnections and will result in a definition of “anti-passivity”.

Second, and more significantly, the inner product involves the *derivatives* of the input and output. The approach taken in the precursor paper [17] was to represent a C^1 game as an input–output system through these derivatives as follows:

$$\dot{x} = u \quad (28)$$

$$y = \dot{\pi} = DF(x)u. \quad (29)$$

Here, we see that the “input” u is actually $\dot{x}$, and the “output” y is $\dot{\pi}$. Using such associations enables the application of standard passivity theory at the cost of associating time derivatives as inputs and outputs.

In this paper, we take an alternative and more direct (but essentially equivalent) approach. We will continue to view evolutionary dynamics as mappings from payoff trajectories (not their derivatives) to strategy trajectories. However, we will introduce a notion of “differential” or simply δ -passivity (see also [18]) which mimics standard passivity, but uses derivatives of inputs and outputs as the integrand for passivity conditions.

Our notion of differential passivity can be interpreted as a local version of “incremental passivity”, defined in [11] as

$$\langle Su - Su', u - u' \rangle_T \geq 0, \quad (30)$$

which closely resembles the original definition of Equation (2).

4.2. δ -Passivity and δ -Anti-Passivity

We begin with definitions for input–output operators. Let $S : \mathbb{U} \rightarrow \mathbb{Y}$ be an input–output operator. We assume both $\mathbb{U}$ and $\mathbb{Y}$ are subsets of locally Lipschitz functions over $[0, \infty)$. This assumption implies a requisite differentiability (for almost all $t \in \mathbb{R}_+$) as well as certain required boundedness properties as well.

The input–output operator $S : \mathbb{U} \rightarrow \mathbb{Y}$ is

- δ -passive if there exists a constant α such that

$$\langle (\dot{S}u), \dot{u} \rangle_T \geq \alpha, \quad \text{for all } u \in \mathbb{U}, T \in \mathbb{R}_+. \quad (31)$$

- input strictly δ -passive if there exist $\beta > 0$ and α such that

$$\langle (\dot{S}u), \dot{u} \rangle_T \geq \alpha + \beta \langle \dot{u}, \dot{u} \rangle_T, \quad \text{for all } u \in \mathbb{U}, T \in \mathbb{R}_+. \quad (32)$$

- δ -anti-passive if $-S$ is passive.
- input strictly δ -anti-passive if $-S$ is input strictly δ -passive.

We now present a stability theorem for δ -passivity analogous to Theorem 3.1 for the positive feedback interconnection Figure 2b, where $S_1 : \mathbb{U} \rightarrow \mathbb{Y}$ and $S_2 : \mathbb{Y} \rightarrow \mathbb{U}$. The feedback equations are:

$$\begin{aligned} u_1 &= r + y_2 \\ y_1 &= S_1 u_1 \\ y_2 &= S_2 u_2 \\ u_2 &= y_1. \end{aligned} \quad (33)$$

We assume existence of solutions and that r is such that $u_1 \in \mathbb{U}$ (which is satisfied, in particular, for $r = 0$).

Theorem 4.1 Consider the feedback interconnection of S_1 and S_2 defined by Equation (33). Assume S_1 is δ -passive and S_2 is input strictly δ -anti-passive. Then $\dot{r} \in \mathbb{L}_2$ implies $\dot{y}_1 \in \mathbb{L}_2$.

The proof parallels that of Theorem 3.1 and is omitted.

We continue with a definition of δ -passivity for state space systems as in Equation (20). We assume that $g(\cdot, \cdot)$ is C^1 and input functions in $\mathbb{U}$ are locally Lipschitz continuous. We assume further that functions in $\mathbb{U}$ are $\mathcal{U}$ valued, with $\mathcal{U} \subset \mathbb{R}^m$, and denote $T\mathcal{U}$ to be the associated tangent space. Throughout this paper, we will have $T\mathcal{U}$ as either $\mathbb{R}^m$ or TX .

A state space system is δ -passive if there exists a C^1 storage function $L : \mathcal{M} \times \mathbb{R}^m \rightarrow \mathbb{R}_+$ such that for all $x \in \mathcal{M}$, $u \in \mathcal{U}$, and $\dot{u} \in T\mathcal{U}$,

$$\nabla_x L(x, u) f(x, u) + \nabla_u L(x, u) \dot{u} \leq \dot{u}^T \left(\nabla_x g(x, u) f(x, u) + \nabla_u g(x, u) \dot{u} \right), \quad (34)$$

or more succinctly,

$$\dot{L} \leq \dot{u}^T \dot{y}. \quad (35)$$

It is δ -anti-passive if

$$\dot{L} \leq -\dot{u}^T \dot{y}. \quad (36)$$

In case the state space system has no state, as in $y = g(u)$, then we take $L = 0$ in the above definitions.

We now state a stability theorem that is tailored for the forthcoming discussion. We are concerned with a special case of positive feedback depicted in Figure 2b. Each system, S_i , is represented by state space equations, as in

$$\begin{aligned}\dot{x}_i &= f_i(x, u) & x_i(0) &= x_{i0} \in \mathcal{M}_i, & i &\in \{1, 2\} \\ y_1 &= g_1(x_1), & y_2 &= g_2(x_2, u_2), & i &\in \{1, 2\} \\ u_1 &= y_2 \\ u_2 &= y_1.\end{aligned}\tag{37}$$

In particular, we have set $r = 0$ and eliminated dependence of y_1 on u_1 to avoid any pathological issues with algebraic loops. Again, we assume the existence of solutions.

Theorem 4.2 *Consider the feedback interconnection of S_1 and S_2 defined by Equation (37). Assume S_1 is δ -passive and S_2 is δ -anti-passive, with storage functions L_1 and L_2 , respectively.*

1. *For all $x_1(0) \in \mathcal{M}_1$, $x_2(0) \in \mathcal{M}_2$, and for $t \in \mathbb{R}_+$,*

$$\dot{L}_1 + \dot{L}_2 \leq 0.\tag{38}$$

2. *Furthermore, if the level set*

$$\{(x_1, x_2) : L_1(x_1, g_2(x_2, g_1(x_1))) + L_2(x_2, g_1(x_1)) \leq \ell\}\tag{39}$$

with

$$\ell = L_1(x_1(0), g_2(x_2(0), g_1(x_1(0)))) + L_2(x_2(0), g_1(x_1(0)))\tag{40}$$

is compact, and

$$\dot{L}_1 + \dot{L}_2 \leq -\psi(\dot{h}),\tag{41}$$

for some positive definite $\psi(\cdot)$ and C^1 function $h : \mathcal{M}_1 \times \mathcal{M}_2 \rightarrow \mathbb{R}^k$, then $\lim_{t \rightarrow \infty} \dot{h}(t) = 0$.

Proof. Statement 1 is a direct consequence of the definitions of δ -passive and δ -anti-passive. Expanding the definition of L_1 and L_2 , we see that

$$L_1(x_1, u_1) + L_2(x_2, u_2) = L_1(x_1, g_2(x_2, g_1(x_1))) + L_2(x_2, g_1(x_1))\tag{42}$$

is non-increasing. Statement 2 assumes some “strictness” in passivity expressed through $\psi(\dot{h})$. The conclusion follows from an application of LaSalle’s invariance theorem. $\square$

We conclude this section with a discussion relating passivity and δ -passivity. Starting from an original state space system:

$$\begin{aligned}\dot{x} &= f(x, u), & x(0) &= x_0 \\ y &= g(x, u),\end{aligned}\tag{43}$$

we can construct an extended system defined by

$$\begin{aligned}\dot{u} &= u', & u(0) &= u_0 \\ \dot{x} &= f(x, u), & x(0) &= x_0 \\ y' &= \nabla_x g(x, u)f(x, u) + \nabla_u g(x, u)u'.\end{aligned}\quad (44)$$

The extended system has as an “input” u' which equals $\dot{u}$ of the original system. Similarly, the extended system has as an “output” y' which equals $\dot{y}$ of the original. Also, note that the state of the expanded system, x' is (x, u) .

The condition for a storage function, $L(x, u)$, for δ -passivity for the original system is

$$\dot{L} \leq \dot{u}^T \dot{y} \quad (45)$$

which is expanded as

$$\nabla_x L(x, u)f(x, u) + \nabla_u L(x, u)\dot{u} \leq \dot{u}^T (\nabla_x g(x, u)f(x, u) + \nabla_u g(x, u)\dot{u}). \quad (46)$$

The condition for a storage function, L' , for *standard* passivity of the extended system is

$$\dot{L}' \leq u'^T y'. \quad (47)$$

Since the state of the extended system is (x, u) , the resulting expansion is

$$\nabla_x L'(x, u)f(x, u) + \nabla_u L'(x, u)u' \leq u'^T (\nabla_x g(x, u)f(x, u) + \nabla_u g(x, u)u'). \quad (48)$$

Identifying u' with $\dot{u}$ we see that δ -passivity for the original system corresponds to standard passivity for the extended system.

This correspondence was used in the precursor paper [17] by defining an extended system for both the evolutionary dynamic and population. That approach resulted in a duplication of states in the two extended systems. The present approach, by defining δ -passivity, avoids such difficulties and is more direct.

4.3. Application to Stable Games and Evolutionary Dynamics

4.3.1. Stable Games

We now state formally the motivating connection between stable games and passivity. Let $\mathbb{X}$ denote locally Lipschitz X -valued functions over $\mathbb{R}_+$, and $\mathbb{P}$ denote locally Lipschitz $\mathbb{R}^n$ -valued functions over $\mathbb{R}_+$.

Theorem 4.3 *A C^1 stable game mapping $\mathbb{X}$ to $\mathbb{P}$ is δ -anti-passive, i.e.,*

$$\langle (F(\dot{x})), \dot{x} \rangle_T \leq 0, \quad \text{for all } x(\cdot) \in \mathbb{X}, T \in \mathbb{R}_+. \quad (49)$$

Furthermore, if F is strictly stable, then the mapping is input strictly δ -anti-passive, i.e., for some $\beta > 0$,

$$\langle (F(\dot{x})), \dot{x} \rangle_T \leq -\beta \langle \dot{x}, \dot{x} \rangle_T, \quad \text{for all } x(\cdot) \in \mathbb{X}, T \in \mathbb{R}_+. \quad (50)$$

Proof. By definition,

$$\langle (F(\dot{x})), \dot{x} \rangle_T = \int_0^T \dot{x}(\tau)^T DF(x(\tau)) \dot{x}(\tau) d\tau. \quad (51)$$

δ -anti-passivity follows from negative definiteness of the integrand. Furthermore, if F is strictly stable, there exists a $\beta > 0$ such that

$$\dot{x}(\tau)^T DF(x(\tau)) \dot{x}(\tau) < -\beta \dot{x}(\tau)^T \dot{x}(\tau), \quad (52)$$

which implies the desired result. $\square$

In identifying stable games as δ -anti-passive systems, we see that evolutionary dynamics that are δ -passive are the complement of stable games in that the stability Theorems 4.1–4.2 are applicable for any δ -passive dynamic. Inspecting the definition for δ -passivity, we will require that for some α ,

$$\int_0^T \dot{x}(t)^T \dot{p}(t) dt \geq \alpha. \quad (53)$$

Since this equation holds for all $T \geq 0$, it implies a long run correlation between the *flow* of population state with the flow of payoffs, namely,

$$\liminf_{T \rightarrow \infty} \frac{1}{T} \int_0^T \dot{x}(t)^T \dot{p}(t) dt \geq 0. \quad (54)$$

4.3.2. Passive Evolutionary Dynamics

We now examine evolutionary dynamics from the perspective of δ -passivity. A general form for evolutionary dynamics is

$$\dot{x} = V(x, F(x)) \quad (55)$$

which describes the evolution of the strategy state, $x \in X$, for the population game F . From a feedback interconnection perspective, an evolutionary dynamic describes how strategy trajectories evolve in response to payoff trajectories. Accordingly, we will remove any explicit game description and write

$$\dot{x} = V(x, p). \quad (56)$$

In terms of previous discussions, the payoff vector p is an “input” and the strategy x is the output. Again, in establishing that an evolutionary dynamic is δ -passive, we do not assert that $p = F(x)$. Rather, $p(\cdot)$ is drawn from a class of trajectories. Since stable games are δ -anti-passive as mappings from $\mathbb{X}$ to $\mathbb{P}$, we are interested in conditions for an evolutionary dynamic to be δ -passive as a mapping from $\mathbb{P}$ to $\mathbb{X}$. Specializing the definition of δ -passivity for state space systems to the current setup, we seek to find a storage function $L : X \times \mathbb{R}^n \rightarrow \mathbb{R}_+$ such that

$$\nabla_x L(x, p) V(x, p) + \nabla_p L(x, p) \dot{p} \leq \dot{p}^T V(x, p) \quad (57)$$

for all admissible $\dot{p}$. Once the above equality is established, then one can employ Theorems 4.1 and 4.2 to make conclusions about stability.

We will focus specifically on so-called excess payoff target (EPT) dynamics [9,19]. These dynamics form a class of evolutionary dynamics that contain several well studied cases.

First, define the *excess payoff* function $\xi : X \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ by

$$\xi(x, p) = p - (x^T p) \cdot \mathbf{1}, \quad (58)$$

where $\mathbf{1}$ is a vector of ones. EPT dynamics take the form

$$\dot{x} = V_{\text{EPT}}(x, p) := \tau(\xi(x, p)) - (\mathbf{1}^T \tau(\xi(x, p))) \cdot x, \quad (59)$$

where $\tau : \mathbb{R}^n \rightarrow \mathbb{R}_+^n$ is called the *revision protocol* (see [9] for a thorough discussion.). Following [9], we make the following assumptions:

- positive correlation: $V_{\text{EPT}}(x, p) \neq 0 \Rightarrow p^T V_{\text{EPT}}(x, p) > 0$.
- integrability: $\tau = \nabla \gamma$ for some C^1 function $\gamma : \mathbb{R}^n \rightarrow \mathbb{R}$.

Finally, let $\mathbb{P}_\rho$ denote the subset

$$\{p \in \mathbb{P} : |p(t)| \leq \rho, \text{ for all } t \in \mathbb{R}_+\}, \quad (60)$$

where $|\cdot|$ denotes the euclidean norm on $\mathbb{R}^n$.

Theorem 4.4 For any $\rho > 0$, EPT dynamics are δ -passive as a mapping from $\mathbb{P}_\rho$ to $\mathbb{X}$ with storage function $\gamma(\xi(x, p)) + C$ for some constant C .

Proof. Following [9], take as a candidate storage function,

$$L(x, p) = \gamma(\xi(x, p)). \quad (61)$$

From integrability of τ ,

$$\begin{aligned} \nabla_x L(x, p) \dot{x} + \nabla_p L(x, p) \dot{p} &= \tau(\xi(x, p))^T \nabla_x \xi(x, p) \dot{x} + \tau(\xi(x, p))^T \nabla_p \xi(x, p) \dot{p} \\ &= \tau(\xi(x, p))^T \left(-\mathbf{1} \cdot p^T V_{\text{EPT}}(x, p) \right) + \tau(\xi(x, p))^T \left((I - \mathbf{1} \cdot x^T) \dot{p} \right) \\ &= -(\mathbf{1}^T \tau(\xi(x, p))) (p^T V_{\text{EPT}}(x, p)) + (\tau(\xi(x, p))^T - ((\tau(\xi(x, p))^T \mathbf{1}) \cdot x^T) \dot{p} \\ &= -(\mathbf{1}^T \tau(\xi(x, p))) (p^T V_{\text{EPT}}(x, p)) + \dot{p}^T V_{\text{EPT}}(x, p) \\ &\leq \dot{p}^T V_{\text{EPT}}(x, p) \\ &= \dot{p}^T \dot{x}, \end{aligned} \quad (62)$$

where the last inequality is due to non-negativity of τ and positive correlation.

The remainder of the proof resolves a technicality that defines storage functions to be positive. Set

$$C = \min \{ \gamma(\xi(x, p)) : x \in X, |p| \leq \rho \}. \quad (63)$$

Then $\gamma(\xi(x, p)) + C$ is non-negative for all $x \in X$ and $|p| \leq \rho$. $\square$

The proof closely follows the stability proof in [9] that establishes $\gamma(\xi(x, F(x)))$ as a Lyapunov

function. This resemblance should not be surprising, in light of Theorem 4.2, which establishes that the sum of the storage functions in a feedback interconnection is non-increasing.

An important difference here is that the proof does not specify the origins of the payoff trajectory. That is, we do not presume that $p(t) = F(x(t))$ for some stable game. Accordingly, the established δ -passivity will have stability implications for “generalized” stable games (see forthcoming sections).

The above proposition also establishes that EPT dynamics are δ -passive as an input–output operator. In particular, since p is the “input” and x is the “output”,

$$\gamma(\xi(x(T), p(T))) - \gamma(\xi(x(0), p(0))) \leq \langle \dot{x}, \dot{p} \rangle_T. \quad (64)$$

One can construct a suitable passivity constant α by maximizing over $x(0)$, $x(T)$, $p(0)$, and $p(t)$, as in the construction of C .

It also is possible to establish δ -passivity for other learning dynamics considered in [9]. As in the case with EPT dynamics, the proofs parallel previous stability proofs in [9], but with new interpretations with broader implications. The only subtlety is that, as before, we do not associate $p = F(x)$ in the process of establishing passivity.

We illustrate this argument for so-called impartial pairwise comparison dynamics, which are not of the EPT form [9]. First, define Lipschitz continuous switch rates

$$\phi_j : \mathbb{R} \rightarrow \mathbb{R}_+ \quad (65)$$

with the property that

$$\phi_j(\delta) > 0 \Leftrightarrow \delta > 0. \quad (66)$$

Impartial pairwise comparison dynamics are defined as

$$\dot{x}_i = \sum_{j=1}^n x_j \phi_i(p_i - p_j) - \sum_{j=1}^n x_i \phi_j(p_j - p_i). \quad (67)$$

The interpretation from [9] is that the flow from strategy i to j depends on the relative payoffs, p_i and p_j . Furthermore, impartiality means that the flow rate, $\phi_j(\cdot)$, only depends on the destination (and not origin) strategy.

Theorem 4.5 *Impartial pairwise comparison dynamics defined by Equation (67) are δ -passive as a mapping from $\mathbb{P}$ to $\mathbb{X}$.*

Proof. Following [9], take as a candidate storage function

$$L(x, p) = \sum_{i=1}^n x_i \left(\sum_{j=1}^n \int_0^{p_j - p_i} \phi_j(s) ds \right). \quad (68)$$

The derivative is

$$\dot{L} = \sum_{i=1}^n \left(\dot{x}_i \left(\sum_{j=1}^n \int_0^{p_j - p_i} \phi_j(s) ds \right) + x_i \left(\sum_{j=1}^n \phi_j(p_j - p_i) (\dot{p}_j - \dot{p}_i) \right) \right). \quad (69)$$

Note that we do not take $\dot{p} = \nabla F(x)\dot{x}$.

Arguments in [9] establish that the summation of the first term satisfies

$$\sum_{i=1}^n \dot{x}_i \left(\sum_{j=1}^n \int_0^{p_j - p_i} \phi_j(s) ds \right) \leq 0. \quad (70)$$

Rearranging the summation of the second term,

$$\begin{aligned} & \sum_{i=1}^n x_i \left(\sum_{j=1}^n \phi_j(p_j - p_i)(\dot{p}_j - \dot{p}_i) \right) \\ &= \left(\sum_{j=1}^n \dot{p}_j \sum_{i=1}^n x_i \phi_j(p_j - p_i) \right) - \left(\sum_{i=1}^n \dot{p}_i \sum_{j=1}^n x_i \phi_j(p_j - p_i) \right) \end{aligned} \quad (71)$$

$$= \sum_{i=1}^n \dot{p}_i \left(\sum_{j=1}^n x_j \phi_i(p_i - p_j) - \sum_{j=1}^n x_i \phi_j(p_j - p_i) \right) \quad (72)$$

$$= \sum_{i=1}^n \dot{p}_i \dot{x}_i. \quad (73)$$

Therefore,

$$\dot{L} \leq \dot{p}^T \dot{x}, \quad (74)$$

as desired. $\square$

4.3.3. Dynamically Modified Payoffs

We now illustrate how passivity methods can be used for the analysis of evolutionary dynamics with auxiliary states. In this section, we consider evolutionary dynamics acting on dynamically modified payoffs. These modifications can be interpreted in two ways: (i) dynamic modification as part of an evolutionary process coupled with a static game, or (ii) a game with dynamic dependencies coupled with a standard evolutionary dynamic. Figure 3 illustrates these two perspectives. In either case, the interconnection, and hence analysis, remains the same.

A consequence of dynamic modifications is the introduction of auxiliary states other than the strategy states. This setting is a departure from much of the literature on evolutionary games, which, almost exclusively, considers evolutionary dynamics whose dimension equals the number of strategies. Likewise, game payoffs typically are static functions of strategies.

Throughout this section, the stable games we consider are affine functions the state, i.e.,

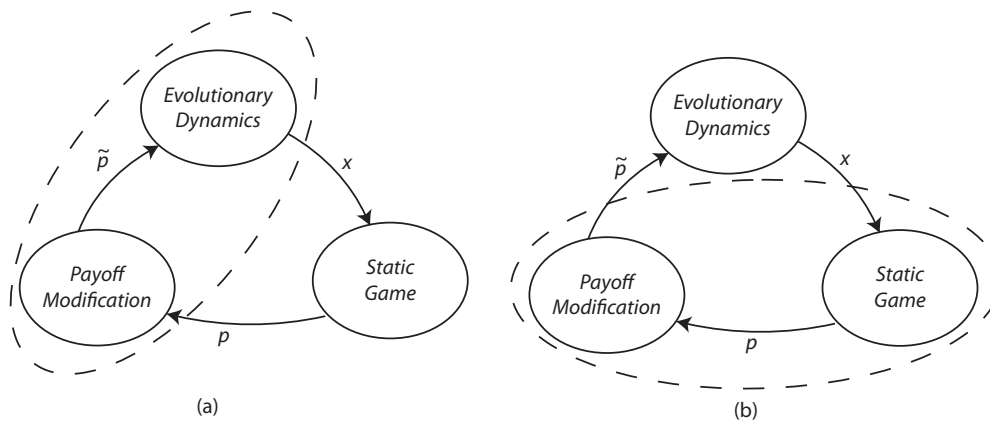
$$p = Ax + b. \quad (75)$$

where A is symmetric negative definite.

Smoothed payoff modification: Payoffs are subject to an exponentially weighted moving average. Given a payoff stream $p(t)$, the smoothed payoffs are

$$\tilde{p}(t) = e^{-\lambda t} p(0) + \int_0^t e^{-\lambda(t-\tau)} p(\tau) d\tau. \quad (76)$$

Figure 3. (a) Static game with dynamically modified evolution. (b) Standard evolution with dynamically modified payoffs.



An effect of the averaging is to smooth out short term fluctuations in order to isolate longer term trends.

In state space form, the mapping from strategies to modified payoffs is described by

$$\dot{\tilde{p}} = \lambda(\underbrace{Ax + b}_p - \tilde{p}). \quad (77)$$

Theorem 4.6 The state space system (77) is δ -anti-passive as a mapping from $\mathbb{X}$ to $\mathbb{P}$.

Proof. Take as a candidate storage function

$$L(\tilde{p}, x) = -\frac{\lambda}{2}(Ax + b - \tilde{p})^T A^{-1}(Ax + b - \tilde{p}). \quad (78)$$

Then

$$\dot{L} = \nabla_{\tilde{p}} L(\tilde{p}, x) \dot{\tilde{p}} + \nabla_x L(\tilde{p}, x) \dot{x} \quad (79)$$

$$= -\lambda(Ax + b - \tilde{p})^T A^{-1}(A\dot{x} - \dot{\tilde{p}}) \quad (80)$$

$$= -\dot{\tilde{p}}^T \dot{x} + \dot{\tilde{p}}^T A^{-1} \dot{\tilde{p}} \quad (81)$$

$$\leq -\dot{\tilde{p}}^T \dot{x}, \quad (82)$$

where the last inequality is due to the negative definiteness of A . $\square$

An implication of Theorem 4.6 is that Theorems 4.1 and 4.2 are now applicable for any δ -passive evolutionary dynamic coupled with smoothed payoffs of an affine stable game.

Anticipatory payoff modification: In anticipatory payoff modification, payoff streams are used to construct myopic forecasts of payoffs. Evolution (or learning) then acts on these myopic forecasts rather than the instantaneous payoffs. The concept is inspired by classical methods in feedback control as well as the psychological tendency to extrapolate from past trends. Anticipatory learning was utilized in [3,20,21], where it was shown how anticipatory learning can alter the convergence to both mixed and pure equilibria.

The state space equations for anticipatory payoff modification are

$$\begin{aligned} \dot{q} &= \lambda(Ax + b - q) \\ \dot{\tilde{p}} &= (Ax + b) + k\lambda(Ax + b - q). \end{aligned} \quad (83)$$

Here, the modified payoff is a combination of the original payoff and an estimate of its derivative, i.e.,

$$\tilde{p} \approx p + k\dot{p}^{\text{est}}. \quad (84)$$

The specific estimate of $\dot{p}$ here is $\dot{q}$ (see the discussion in [20]), which can be constructed from payoff measurements. The scalar k reflects the weighting on the derivative estimate.

Theorem 4.7 *The state space system (83) is δ -anti-passive as a mapping from $\mathbb{X}$ to $\mathbb{P}$.*

Proof. Take as a candidate storage function

$$L(q, x) = -\frac{\lambda}{2}(Ax + b - q)^T A^{-1}(Ax + b - q). \quad (85)$$

Then

$$\dot{L} = -\lambda(Ax + b - q)^T A^{-1}(A\dot{x} - \dot{q}). \quad (86)$$

By definition,

$$\dot{\tilde{p}} = A\dot{x} + k\lambda(A\dot{x} - \dot{q}). \quad (87)$$

Therefore,

$$\dot{L} = -\lambda(Ax + b - q)^T A^{-1}(\dot{\tilde{p}} - A\dot{x}) \frac{1}{k\lambda} \quad (88)$$

$$= -\dot{q}^T A^{-1}(\dot{\tilde{p}} - A\dot{x}) \frac{1}{k\lambda}. \quad (89)$$

Using the above equation for $\dot{\tilde{p}}$ results in

$$\dot{L} = \left(\frac{1}{k\lambda}(\dot{\tilde{p}} - A\dot{x}) - A\dot{x} \right)^T A^{-1}(\dot{\tilde{p}} - A\dot{x}) \frac{1}{k\lambda} \quad (90)$$

$$= \frac{1}{(k\lambda)^2}(\dot{\tilde{p}} - A\dot{x})^T A^{-1}(\dot{\tilde{p}} - A\dot{x}) + \frac{1}{k\lambda}\dot{x}^T A\dot{x} - \frac{1}{k\lambda}\dot{x}^T \dot{\tilde{p}} \quad (91)$$

$$\leq -\frac{1}{k\lambda}\dot{x}^T \dot{\tilde{p}}, \quad (92)$$

where the last inequality is due to the negative definiteness of A . We see that rescaling $L_{\text{new}}(q, x) = k\lambda L(q, x)$ leads to the desired result. $\square$

The proof of Theorem 4.7 reveals that the associated dissipation inequality is satisfied *strictly* because of the two terms involving A . In particular, the anticipatory payoff modification of Equation (83) defines an input strictly δ -anti-passive system. The following representative theorem is then an immediate consequence of Theorem 4.1.

Theorem 4.8 *In the positive feedback interconnection defined by Equation (33) (as in Figure 2b), let S_1 be any δ -passive evolutionary dynamic mapping payoff trajectories $p(\cdot) \in \mathbb{P}$ to state trajectories $x(\cdot) \in \mathbb{X}$, and let S_2 be anticipatory payoff modification defined by Equation (83). Then $\dot{x}(\cdot) \in \mathbb{L}_2$.*

In case the evolutionary dynamic has a state space description (e.g., EPT), one can use Theorem 4.2 to conclude $\lim_{t \rightarrow \infty} \dot{x}(t) = 0$.

Implicit in the above discussion is that convergence of $\dot{x}$ has implications about convergence to Nash equilibrium. Any such conclusions are specific to the underlying evolutionary dynamic.

4.3.4. Contrarian Effect Payoffs

In this section, we illustrate the use of passivity methods in the presence of lags or time delays. In this model, players perceive advantages in avoiding strategies that have seen net increase in recent usage. In particular, for a fixed lag $\ell > 0$, payoffs are given by “contrarian effect” payoffs, defined by

$$\tilde{p}(t) = F(x(t)) - \Lambda(x(t) - x(t - \ell)), \quad (93)$$

where $\Lambda > 0$ is a diagonal scaling matrix. We assume that strategies are initialized by some Lipschitz continuous $x_0 : [-\ell, 0] \rightarrow X$, so that for $t - \ell \leq 0$,

$$x(t - \ell) := x_0(t - \ell). \quad (94)$$

As intended, an increase in the usage of a strategy diminishes the perceived payoff derived from that strategy. While such a contrarian effect defined here may seem simplistic, our main interest is to illustrate the analysis of delays using passivity methods.

In this section, we will deal exclusively with the input–output operator formulation of passivity. We begin by establishing the following *variant* of δ -anti-passivity of contrarian effect payoffs. First, define $\mathbb{X}_{x_0}$ to be the restriction of $\mathbb{X}$ to functions with $x(0) = x_0$.

Proposition 4.1 *Let F be a strictly stable game with strict passivity constant, $\beta_2 > 0$, i.e.,*

$$\dot{z}^T DF(x) \dot{z} \leq -\beta_2 \dot{z}^T \dot{z}, \quad \forall z \in TX, x \in X. \quad (95)$$

Let $x_0 : [-\ell, 0] \rightarrow X$ be Lipschitz continuous. Let $S : \mathbb{X}_{x_0(0)} \rightarrow \mathbb{P}$ be the (contrarian effect payoff) input–output operator defined by Equation (93). Then

$$-\langle \dot{\tilde{p}}, \dot{x} \rangle_T \geq \beta_2 \langle \dot{x}, \dot{x} \rangle_T - \alpha'_2 \|\dot{x}\|_\ell, \quad (96)$$

with $\alpha'_2 = (\int_{-\ell}^0 \dot{x}_0(t)^T \Lambda^2 \dot{x}_0(t) dt)^{1/2}$.

Proof. See appendix. $\square$

Proposition 4.1 states that contrarian effect payoffs satisfy a version of passivity that deviates slightly from the usual definition of δ -anti-passivity, since the associated passivity lower bound (associated with α) is $-\alpha'_2 \|\dot{x}\|_\ell$ and depends on the input signal, but in a bounded manner.

The following theorem now follows from arguments similar to those for Theorem 4.1.

Theorem 4.9 *In the positive feedback interconnection defined by Equation (33) (as in Figure 2b), let S_1 be any δ -passive evolutionary dynamic mapping payoff trajectories $p(\cdot) \in \mathbb{P}$ to state trajectories $x(\cdot) \in \mathbb{X}$, and let S_2 be contrarian effect payoffs defined by Equation (93) with F strictly stable. Then $\dot{x}(\cdot) \in \mathbb{L}_2$.*

Proof. Let α_1 be the passivity constant of the δ -passive evolutionary dynamic, so that

$$\langle \dot{\tilde{p}}, \dot{x} \rangle_T \geq \alpha_1. \quad (97)$$

By Proposition 4.1, contrarian effect payoffs satisfy

$$-\langle \dot{\tilde{p}}, \dot{x} \rangle_T \geq \beta_2 \langle \dot{x}, \dot{x} \rangle_T - \alpha'_2 \|\dot{x}\|_\ell. \quad (98)$$

Summing these inequalities leads to

$$-\alpha_1 + \alpha'_2 \|\dot{x}\|_\ell \geq \beta_2 \langle \dot{x}, \dot{x} \rangle_T, \quad (99)$$

which then implies that $\dot{x}(\cdot) \in \mathbb{L}_2$, as desired. $\square$

5. Concluding Remarks

This paper has proposed passivity theory as a unifying and extending framework for the study of evolutionary games. In particular, the passivity condition property of long run correlated payoff flows and population flows in Equation (54) appears to be a natural complement to the class of stable games. The methods are applicable to generalizations of stable games with dynamic payoffs or evolutionary dynamics that include auxiliary states. We believe that there is significant potential in bringing in related methods of feedback control theory, such as generalized dissipativity, multipliers, and loop transformations, to complement the more traditional analytical approaches to evolutionary games. A lingering question is to understand the extent to which passive evolutionary dynamics and stable games are complementary. We conjecture that if an evolutionary dynamic is *not* passive then one can construct a (generalized) stable game that results in instability. Stated differently, an evolutionary dynamic results in stability for all generalized stable games if and only if it is passive.

Acknowledgements

Research supported by ONR project #N00014-09-1-0751. We thank William H. Sandholm, Georgios Piliouras, Nikhil Chopra, Jason Marden, Georgios Chasparis, Ozan Candogan, and Georgios Kotsalis for helpful discussions.

A. Proof of Proposition 4.1

Contrarian effect payoffs in Equation (93) can be written as the sum of two terms,

$$\tilde{p}(t) = \tilde{p}_{\text{SG}}(t) - \tilde{p}_{\text{CE}}(t), \quad (100)$$

i.e., the “stable game” portion,

$$\tilde{p}_{\text{SG}}(t) = F(x(t)) \quad (101)$$

and the “contrarian effect” portion,

$$\tilde{p}_{\text{CE}}(t) = \Lambda(x(t) - x(t - \ell)). \quad (102)$$

Accordingly,

$$-\langle \dot{\tilde{p}}, \dot{x} \rangle_T = -\langle \dot{\tilde{p}}_{\text{SG}}, \dot{x} \rangle_T + \langle \dot{\tilde{p}}_{\text{CE}}, \dot{x} \rangle_T \quad (103)$$

$$= -\int_0^T \dot{x}(\tau)^T DF(x(\tau)) \dot{x}(\tau) d\tau + \langle \dot{\tilde{p}}_{\text{CE}}, \dot{x} \rangle_T \quad (104)$$

$$\geq \beta_2 \langle \dot{x}, \dot{x} \rangle_T + \langle \dot{\tilde{p}}_{\text{CE}}, \dot{x} \rangle_T. \quad (105)$$

It then remains to show that

$$\langle \dot{\tilde{p}}_{\text{CE}}, \dot{x} \rangle_T \geq -\alpha'_2 \|\dot{x}\|_\ell. \quad (106)$$

We will need the following lemma, whose proof relies on standard arguments from systems theory (e.g., [11]).

Lemma A.1 For any $v \in \mathbb{L}_2$ and $\ell \geq 0$,

$$\int_0^\infty (v(t) - v(t - \ell))v(t) dt \geq 0, \quad (107)$$

where $v(t - \ell) := 0$ for $t - \ell < 0$.

Proof. Define the extensions $\bar{v} : (-\infty, \infty) \rightarrow \mathbb{R}^n$

$$\bar{v}(t) = \begin{cases} 0 & t < 0; \\ v(t) & t \geq 0. \end{cases} \quad (108)$$

Let $\hat{v}$ be the Fourier transform of $\bar{v}$, i.e.,

$$\hat{v}(j\omega) := \int_{-\infty}^\infty \bar{v}(t) e^{-j\omega t} dt. \quad (109)$$

Likewise, let

$$\overline{w}(t) = \bar{v}(t) - \bar{v}(t - \ell) \quad (110)$$

and let $\hat{w}$ be the Fourier transform of $\overline{w}$. Parseval's Theorem states that

$$\int_{-\infty}^\infty \overline{w}(t)^T \bar{v}(t) dt = \frac{1}{2\pi} \int_{-\infty}^\infty \hat{w}(j\omega)^* \hat{v}(j\omega) d\omega, \quad (111)$$

where superscript “*” denotes complex conjugate transform. Using that

$$\hat{w}(j\omega) = \hat{v}(j\omega) - e^{-j\omega\ell} \hat{v}(j\omega) \quad (112)$$

results in

$$\int_{-\infty}^\infty \overline{w}(t)^T \bar{v}(t) dt = \frac{1}{2\pi} \int_{-\infty}^\infty \hat{v}(j\omega)^* (I - e^{j\omega\ell} I) \hat{v}(j\omega) d\omega, \quad (113)$$

$$= \frac{1}{2\pi} \int_{-\infty}^\infty \hat{v}(j\omega)^* (I - \frac{1}{2}(e^{j\omega\ell} + e^{-j\omega\ell}) I) \hat{v}(j\omega) d\omega, \quad (114)$$

$$= \frac{1}{2\pi} \int_{-\infty}^\infty (1 - \cos(\omega\ell)) \hat{v}(j\omega)^* \hat{v}(j\omega) d\omega \quad (115)$$

$$\geq 0, \quad (116)$$

as desired. $\square$

Lemma A.1 almost provides the desired conclusion, except that

$$\langle \dot{p}_{\text{CE}}, \dot{x} \rangle_T = \langle \Lambda^{1/2}(\dot{x}(t) - \dot{x}(t - \ell)), \Lambda^{1/2} \dot{x} \rangle_T \quad (117)$$

involves terms due to the initialization $x_0(\cdot)$ over the interval $[-\ell, 0]$. These terms can be bounded as follows. For any $T > 0$, define

$$v(t) = \begin{cases} 0, & t < 0; \\ \dot{x}(t), & 0 \leq t \leq T; \\ 0 & t \geq T. \end{cases} \quad (118)$$

and

$$v_0(t) = \begin{cases} 0, & t < -\ell; \\ \dot{x}_0(t), & -\ell \leq t \leq 0; \\ 0, & t \geq 0. \end{cases} \quad (119)$$

and set

$$w(t) = \Lambda(v(t) - v(t - \ell) - v_0(t - \ell)) \quad (120)$$

for $t \in (-\infty, \infty)$. Then

$$\begin{aligned} \langle \dot{p}_{\text{CE}}, \dot{x} \rangle_T &= \langle w, v \rangle_T \\ &= \int_0^T w(t)^T v(t) dt \\ &= \int_0^T (v(t) - v(t - \ell))^T \Lambda v(t) dt - \int_0^T v_0(t - \ell)^T \Lambda v(t) dt \\ &= \int_{-\infty}^T (v(t) - v(t - \ell))^T \Lambda v(t) dt - \int_0^T v_0(t - \ell)^T \Lambda v(t) dt, \quad (\text{since } v(t) = 0 \text{ for } t < 0) \\ &= \int_{-\infty}^{\infty} (v(t) - v(t - \ell))^T \Lambda v(t) dt - \int_0^T v_0(t - \ell)^T \Lambda v(t) dt, \quad (\text{since } v(t) = 0 \text{ for } t > T) \end{aligned} \quad (121)$$

Using Lemma A.1, the first term above is positive, and the second term is bounded from below via

$$- \int_0^T v_0(t - \ell)^T \Lambda v(t) dt \geq - \|v\|_{\ell} \alpha'_2. \quad (122)$$

References

1. Sandholm, W.H. *Population Games and Evolutionary Dynamics (Economic Learning and Social Evolution)*; The MIT Press: Cambridge, MA, USA, 2011.
2. Sato, Y.; Akiyama, E.; Farmer, J.D. Chaos in learning a simple two-person game. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4748–4751.

3. Arslan, G.; Shamma, J.S. Anticipatory Learning in General Evolutionary Games. Proceedings of the 45th IEEE Conference on Decision and Control, San Diego, CA, USA, 13–15 December, 2006; pp. 6289–6294.
4. Hart, S.; Colell, A.M. Uncoupled dynamics do not lead to Nash equilibrium. *Am. Econ. Rev.* **2003**, *93*, 1830–1836.
5. Monderer, D.; Shapley, L. Potential games. *Games Econ. Behav.* **1996**, *14*, 124–143.
6. Hart, S. Adaptive Heuristics. *Econometrica* **2005**, *73*, 1401–1430.
7. Young, H.P. *Strategic Learning and its Limits*; Oxford University Press: Oxford, UK, 2005.
8. Fudenberg, D.; Levine, D. Learning and equilibrium. *Annu. Rev. Econom.* **2009**, *1*, 385–420.
9. Sandholm, W.H.; Hofbauer, J. Stable games and their dynamics. *J. Econ. Theory* **2009**, *144*, 1665–1693.
10. Willems, J.C. Dissipative dynamical systems — Part I, Part II. *Arch. Ration. Mech. Anal.* **1972**, *45*, 321–393.
11. Desoer, C.; Vidyasagar, M. *Feedback Systems: Input-Output Properties*; Academic Press, Inc.: New York, NY, USA, 1975.
12. Khalil, H. *Nonlinear Systems*, 3rd ed.; Prentice Hall: Upper Saddle River, New Jersey, 2002.
13. van der Schaft, A. *L2-Gain and Passivity Techniques in Nonlinear Control*; Springer Verlag: Berlin/Heidelberg, Germany, 2000.
14. Ramirez-Llanos, E.; Quijano, N. A population dynamics approach for the water distribution problem. *Int. J. Control* **2010**, *83*, 1947–1964.
15. Fan, X.; Alpcan, T.; Arcak, M.; Wen, T.J.; Basar, T. A passivity approach to game-theoretic CDMA power control. *Automatica* **2006**, *42*, 1837–1847.
16. Hofbauer, J.; Sandholm, W.H. Evolution in games with randomly disturbed payoffs. *J. Econ. Theory* **2007**, *132*, 47–69.
17. Fox, M.J.; Shamma, J.S. Population Games, Stable Games, and Passivity. Proceedings of the 51st IEEE Conference on Decision and Control, Maui, Hawaii, 10–13 December, 2012, pp. 7445–7450.
18. Wang, H. Differential-Passivity based Controlled Synchronization of Networked Robots with Additive Disturbances. Proceedings of the 31st Chinese Control Conference (CCC'2), Hefei, China, 25–27 July, 2012, pp. 5838–5843.
19. Sandholm, W. Excess payoff dynamics and other well-behaved evolutionary dynamics. *J. Econ. Theory* **2005**, *124*, 149–170.
20. Shamma, J.S.; Arslan, G. Dynamic fictitious play, dynamic gradient play, and distributed convergence to Nash equilibria. *IEEE Trans. Automat. Contr.* **2005**, *50*, 312–327.
21. Chasparis, G.; Shamma, J. Distributed dynamic reinforcement of efficient outcomes in multiagent coordination and network formation. *Dynamic Games and Applications* **2012**, *2*, 18–50.