

# Asymptotic Stability in the Lovász-Shapley Replicator Dynamic for Cooperative Games\*

André Casajus<sup>†</sup>    Michael Kramm<sup>‡</sup>    Harald Wiese<sup>§</sup>

November 20, 2019

## Abstract

We derive population dynamics from finite cooperative games with transferable utility, where the players are interpreted as types of individuals. We show that any asymptotically stable population profile is characterized by a coalition: while the types in the coalition have the same positive share, the other types vanish. The average productivity of such a stable coalition must be greater than the average productivity of any proper sub- or super-coalition. In simple monotonic games, this means that exactly the minimal winning coalitions are stable. Possible applications are the analysis of the organizational structure of businesses or the population constitution of eusocial species.

---

\*We would like to thank Markus Kunze, Pierfrancesco La Mura, Wolfgang Leininger, Lars Metzger, Heinrich Nax, Karl Friedrich Siburg, William Thomson, and Jörgen Weibull for valuable comments. This paper also benefited from discussions and remarks at the European Meeting on Game Theory 2015 (SING, St. Petersburg), the 2nd Workshop on Economics and Biology (Toulouse), and the 3rd Workshop on Cooperative Game Theory in Business Practice (Leipzig). We are grateful for the comments of two anonymous referees and an associate editor. André Casajus: Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - 388390901.

<sup>†</sup>HHL Leipzig Graduate School of Management, Germany and Dr. Hops Craft Beer Bar, Germany, Email: mail@casajus.de

<sup>‡</sup>Corresponding author, Technical University of Dortmund, Vogelpothsweg 87, 44221 Dortmund, Germany, Email: michael.kramm@tu-dortmund.de

<sup>§</sup>University of Leipzig, Email: wiese@wifa.uni-leipzig.de

**Keywords:** cooperative game theory, evolutionary game theory, replicator dynamics, asymptotic stability, Lovász-Shapley value, simple monotonic games

**JEL codes:** C71, C73

# 1 Introduction

For a long time, non-cooperative game theory has employed the evolutionary approach to model the behavior of boundedly rational agents and to analyze the relation between the stable outcomes of the dynamical process generated by their repeated interaction and the common static solution concept of Nash equilibrium (see, e.g., Mailath, 1998 or Samuelson, 2002 for discussions). In contrast to non-cooperative game theory, cooperative game theory shifts the focus from modeling strategic interactions and the agents' strategic reasoning to payoffs: it asks questions about which payoff distribution among the players of a given cooperative game can, for instance, be considered as fair or stable (see, e.g., McLean, 2015).

The non-cooperative concept of the Nash equilibrium has taken a central role in the study of game theory in a rather literary sense. On the one hand, the idea of the Nash program, starting with Nash (1953), is to provide non-cooperative foundations for cooperative solution concepts based on the Nash equilibrium. On the other hand, as mentioned above, Smith and Price (1973) and many other researchers started to think about evolutionary foundations of the Nash equilibrium about two decades later. Our modeling approach skips the Nash equilibrium in the chain of reasoning and directly connects cooperative and evolutionary game theory. We introduce the evolutionary methodology to cooperative game theory in order to answer questions on the relation between the properties of the underlying cooperative game and the stable outcomes of a dynamical evolutionary process.

In non-cooperative game theory, evolutionary pressures work against strategies, for instance, dominated strategies may die out (see, e.g., Hofbauer and Weibull, 1996). In our setup, evolution takes place over types of individuals and not over strategies. In this way, pressures work against

certain coalitions, for example, non-minimal winning coalitions (see Corollary 2). We can thus analyze the relation between asymptotically stable population profiles and the underlying transferable utility (TU) game.

The interpretation of the evolutionary dynamics in our model is illustrated with an economic example in Section 2 and can be outlined as follows. We consider a society that consists of infinitely many individuals of finitely many different types. The types correspond to the players of a cooperative game. They are (economic) entities characterized by distinct productive capabilities, such as specific business departments in an organization or specific casts of a eusocial species<sup>1</sup>. The individuals, e.g., members of a specific business department or cast, are programmed to perform certain tasks according to their type and contribute to the organization's or the society's total reproductive success or fitness, i.e., its growth in time. This is modelled by the Lovász extension of a finite TU game (see Lovász, 1983 and Algaba et al., 2004), which embodies a particular Leontief-type technology for the production of fitness. Eventually, the total fitness has to be distributed among the different types. This is modelled by the Lovász-Shapley value (see Casajus and Wiese, 2016) that rewards only the scarce types, which is in line with the Leontief-type technology. Given this distribution, the composition of the society evolves in time according to the replicator dynamic, which is the most commonly used imitative dynamic in the literature (see Sandholm, 2010, Section 5.3) and which is most suitable to capture the idea of individuals reproducing according to their fitness

---

<sup>1</sup>A common “criterion for eusociality is the presence of castes, which are groups of individuals that become irreversibly behaviorally distinct at some point prior to reproductive maturity. Eusocial societies are characterized by two traits: (1) helping by individuals of the less-reproductive caste, and (2) either behavioral totipotency of only the more reproductive caste (facultative eusociality) or totipotency of neither caste (obligate eusociality)” (Crespi and Yanega, 1995, p. 109). Examples are the European hornet (*Vespa crabro*) or the naked mole rat (*Heterocephalus glaber*).

derived from the cooperative game. Now, one can study the proliferation of the types and, in particular, stable compositions of the society for the resulting replicator dynamic under the Filippov solution (see Filippov, 1988).

To our knowledge, we are the first to derive an evolutionary dynamic from a cooperative game in a similar fashion as it is done in non-cooperative evolutionary game theory without incorporating non-cooperative aspects. This cooperative approach has particular appeal when types represent organizational departments of a collaborative enterprise, such as a large business corporation, or casts of a eusocial species.

Usually, it is considered that “[o]n general grounds of methodological individualism, noncooperative games are prior to cooperative games” (see Elster, 1982, p. 466). We do not contest this conventional approach of economic theory (or its value) to analyze aggregate phenomena by examining the actions and incentives of individual entities. Rather, our aim is to provide a complementary view on the “links between aggregate analysis and the study of individual behavior” (see Elster, 1982, p. 477). Some researchers have pushed forward that methodological individualism alone cannot explain all aspects of socio-economic interactions and thus that social context and institutions need to be included in the analysis (see Newton, 2019 and Hodgson, 2007). Furthermore, the debate on the evolution of eusociality<sup>2</sup> emphasizes that it might be insightful to provide theories

---

<sup>2</sup>The evolution of eusociality is debated controversially among evolutionary biologists. It has challenged the predominant idea of the inclusive fitness theory (in this case, kin selection; see, e.g., Hamilton, 1964), which postulates that evolution takes place at the level of individuals and thus is loosely related to the non-cooperative approach in game theory. In contrast, some researchers advance the idea of group selection to explain eusociality (see Wilson and Hölldobler, 2005 and Nowak et al., 2010), which is an approach similar to cooperative game theory, since it assumes that evolution takes place at the level of groups. In view of this debate, skipping the non-cooperative approach of



Pradelski (2019) apply evolutionary mechanisms to questions of cooperative game theory such as core selection and core stability in finite population matching models, mainly to transferable-utility assignment games. Newton (2012) shows that, in a finite-population evolutionary dynamic generated by repeated recontracting, i.e., a non-cooperative representation of the underlying cooperative game, stochastically stable states are characterized by maximizing the utility of the poorest individual in a Rawlsian manner. In a non-transferable utility setting and for a large class of perturbed dynamics, Newton and Sawa (2015, p. 1) show that “all stochastically stable matchings are [...] contained in the set of matchings which are most robust to one-shot deviation”. In this literature, the evolutionary dynamics involve randomly matching two individuals which then play a cooperative game. Furthermore, there is usually a non-cooperative flavor to these dynamics, since players can, for instance, individually make bids or break ties with former partners. This is fundamentally different from our model of the generation of worth, which occurs according to the Lovász extension in a purely cooperative game with an infinite population, and it is different from the evolutionary dynamic we have in mind, which describes how generated worth is translated into population growth of a certain type represented by a player.

Nash (2008) also addresses the idea of a “cooperative evolutionary game theory”. He models the formation of cooperative coalitions by allowing non-cooperative players to decide repeatedly whether or not to completely follow another players agency, i.e., to be cooperative.

Filar and Petrosjan (2000) analyze dynamic cooperative games and define a sequence of games so that one TU game is determined by the previous one and by the payoffs achieved under some solution concept. The players obtain the sum of payoffs for this sequence of coalition functions and the problem of whether the payoffs satisfy a consistency criterion is discussed.

The rest of this paper is structured as follows. The example in Section 2 illustrates the main ideas of our approach. In Section 3, we build up the framework in which coalitions of players generate worth, and in Section 4, we introduce the evolutionary model. Section 5 contains results on the asymptotic stability. Section 6 finally concludes.

## 2 Illustrative Example

An economic perspective on our approach is to interpret the types, i.e., the players of the cooperative game in Table 1, as business departments, for example, each having a certain endowment of workforce, which determines the amount of time (e.g., working hours per day) a business department can spend on specific projects, i.e., work in certain coalitions. For illustrative purposes, let the initial endowment with workers be given by the vector  $s = (\frac{2}{6}, \frac{1}{6}, \frac{3}{6})$ , where the first entry denotes the (absolute) amount of time of the first business department, and so on.

The Leontief-type technology for the production of fitness induced by the Lovász extension of the finite TU game implies that each business department can invest one sixth of a working hour per day into the project “grand coalition” to generate worth. Business units 1 and 3 have enough time left to form a coalition, i.e., to cooperate in a project consisting only of them both, and finally one sixth of a working hour is left for business department 2 to produce worth alone. Thus, for the initial population state  $s$  and the coalition function  $v$  defined by Table 1, the overall worth generated by the society is  $\bar{v}(s) = \frac{1}{6} \cdot 7 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 1$ .<sup>3</sup>

The Lovász-Shapley value, which satisfies desirable properties for the setting at hand, determines how the business departments distribute the

---

<sup>3</sup>The term  $\bar{v}$  stands for the Lovász-extension of the game  $v$ . Details on this will follow below.

worth among each other: each business department is paid according to its marginal contribution to the coalition in which players are ordered with respect to their population sizes, starting with the largest population, i.e., here they are ordered according to  $\rho(s) = (3, 1, 2)$ .<sup>4</sup> The calculated marginal productivity is then multiplied with the initial endowment of the department. Thus, we obtain Lovász-Shapley payoffs  $LS(v, s) = (\frac{8}{6}, \frac{2}{6}, \frac{3}{6})$ .

If we assume that, in a replicator dynamic, the Lovász-Shapley payoffs determine the growth of the business departments' disposable working hours in each state (e.g., they might hire new employees with the income generated by the executed projects), then we obtain the dynamic in Figure 1 depicting the evolution of the business departments' relative shares in the population of workers.

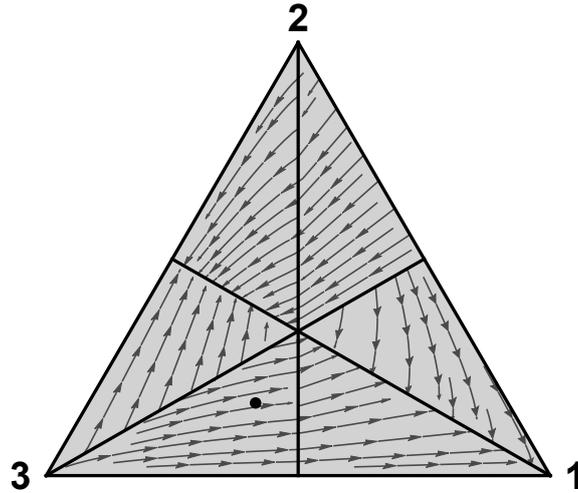


Figure 1: Vector field for the game described in Table 1. Each point in the simplex depicts the relative population shares of the three different business units.

---

<sup>4</sup>Intuitively, department 2's time is most scarce in the initial state  $s$  and thus its investment is in some way “most crucial” to form the grand coalition, so that it should obtain a payoff proportional to his marginal contribution to the grand coalition,  $MC_2^v(3, 1, 2) = 2$ . Next, department 1's time is most crucial to form the coalition of departments 1 and 3, so it should obtain a payoff proportional to  $MC_1^v(3, 1, 2) = 4$ .

In Figure 1, the solid dot indicates the initial population profile  $(\frac{2}{6}, \frac{1}{6}, \frac{3}{6})$ . The trajectory starting at this point converges to the asymptotically stable profile  $(1, 0, 0)$ . In contrast, the trajectories starting in the upper-left part of the simplex converge to the second asymptotically stable profile  $(0, \frac{1}{2}, \frac{1}{2})$ . In this paper, we focus on the analysis of asymptotically stable population profiles.

### 3 Populations of Players Generating Worth

In this section, we first provide the foundation of cooperative game theory needed for what follows. Then, we present the Lovász-Shapley value, which will be used to derive the fitness of type populations from cooperative games.

#### 3.1 Basic Definitions and Notation

A (finite) cooperative game with transferable utility (TU game) for a non-empty and finite set of players  $N$  is given by a coalition function

$$v \in \mathbb{V} := \{f \mid f : 2^N \rightarrow \mathbb{R}, f(\emptyset) = 0\}.$$

The latter describes the worths  $v(S)$  that can be generated by players who cooperate within coalitions  $S \subseteq N$ . A solution is a mapping  $\varphi : \mathbb{V} \rightarrow \mathbb{R}^N$ , which assigns a payoff  $\varphi_i(v)$  to any player  $i \in N$  for any game  $v \in \mathbb{V}$ .

For  $v, w \in \mathbb{V}$ ,  $\alpha \in \mathbb{R}$ , the coalition functions  $v + w \in \mathbb{V}$  and  $\alpha \cdot v \in \mathbb{V}$  are given by  $(v + w)(S) = v(S) + w(S)$  and  $(\alpha \cdot v)(S) = \alpha \cdot v(S)$  for all  $S \subseteq N$ . For  $T \subseteq N$ ,  $T \neq \emptyset$ , the game  $u_T \in \mathbb{V}$ ,  $u_T(S) = 1$  if  $T \subseteq S$  and  $u_T(S) = 0$  otherwise, is called an **unanimity game**. Any  $v \in \mathbb{V}$  can uniquely be represented by unanimity games,

$$v = \sum_{T \subseteq N: T \neq \emptyset} \lambda_T(v) \cdot u_T,$$

where the coefficients  $\lambda_T(v)$ , the so-called Harsanyi dividends (Harsanyi, 1959), can be determined recursively via

$$v(S) = \sum_{T \subseteq S: T \neq \emptyset} \lambda_T(v) \quad S \subseteq N. \quad (1)$$

A **rank order** of a set  $N$  is a bijection  $\rho : N \rightarrow \{1, \dots, |N|\}$  with the interpretation that  $i$  is the  $\rho(i)$ th player in  $\rho$ . The set of rank orders of  $N$  is denoted by  $R$ . The **marginal contribution** of  $i$  to  $S \subseteq N \setminus \{i\}$  is denoted by

$$MC_i^v(S) := v(S \cup \{i\}) - v(S) \stackrel{(1)}{=} \sum_{T \subseteq S \setminus \{i\}} \lambda_{T \cup \{i\}}(v),$$

and the marginal contribution of  $i$  under  $\rho$  is denoted by

$$MC_i^v(\rho) := MC_i^v(\{j \in N \mid \rho(j) < \rho(i)\}). \quad (2)$$

The **Shapley value** (Shapley, 1953),  $\text{Sh}$ , is given by

$$\text{Sh}_i(v) := |R|^{-1} \cdot \sum_{\rho \in R} MC_i^v(\rho) = \sum_{T \subseteq N: i \in T} |T|^{-1} \cdot \lambda_T(v), \quad v \in \mathbb{V}, i \in N.$$

Game  $v$  is a **simple game** if for all  $S \subseteq N$  we have  $v(S) \in \{0, 1\}$  and  $v(N) = 1$ , and it is a **monotonic game** if for all  $S, T \subseteq N$  such that  $S \subseteq T$ , we have  $v(S) \leq v(T)$ . Coalition  $T$  is a **minimal winning coalition** in a simple game  $v$  if we have  $v(T) = 1$  and  $v(S) = 0$  for all  $S \subsetneq T$ .

### 3.2 The Lovász-Shapley Value and the Fitness of Populations

Casajus and Wiese (2016) suggest to interpret the players in a TU game as **types of individuals**. The **population sizes** of the types are given by a vector  $s \in \mathbb{R}_+^N$  of non-negative weights, which need not sum up to one and where  $s_i$  denotes the size of the population (of individuals) of type  $i \in N$ .

We address pairs  $(v, s) \in \mathbb{V} \times \mathbb{R}_+^N$  as **population games**. A population solution is a mapping  $\varphi : \mathbb{V} \times \mathbb{R}_+^N \rightarrow \mathbb{R}^N$ , which assigns a payoff  $\varphi_i(v, s)$  to population of any type  $i$  for any population game  $(v, s)$ . Casajus and Wiese (2016, Equation 13) introduce a specific population solution, the **Lovász-Shapley value**, LS, given by

$$\text{LS}_i(v, s) := s_i \cdot \sum_{T \subseteq N: i \in \text{argmin}_T(s)} \frac{\lambda_T(v)}{|\text{argmin}_T(s)|} \quad (3)$$

for all  $v \in \mathbb{V}$ ,  $s \in \mathbb{R}_+^N$ , and  $i \in N$ , where

$$\min_T(x) := \min_{i \in T} x_i \quad \text{and} \quad \text{argmin}_T(x) := \{i \in T \mid x_i = \min_T(x)\}$$

for all  $x \in \mathbb{R}^N$ .

The Harsanyi dividends (Harsanyi, 1959)  $\lambda_T(v)$  can be interpreted as the gain from cooperation that can be ascribed to coalition  $T$  in a TU game  $v$ . The Shapley value distributes this gain equally among the players in  $T$ . Formula (3) indicates (i) how the population sizes affect the gains from cooperation and (ii) how these gains are distributed among players under the Lovász-Shapley value. For an axiomatic characterization and a detailed motivation of the Lovász-Shapley value, we refer the reader to Casajus and Wiese (2016).

The gain from cooperation generated by a coalition  $T$  in the population game  $(v, s)$  is given by  $s_i \cdot \lambda_T(v)$ , where  $s_i$  is the smallest population size of players from  $T$ . That is, the **scarcest type** in a coalition determines the gain from cooperation. This indicates that a Leontief-type “technology” underlies the Lovász-Shapley value. Moreover, the payoffs sum up to the worth generated by  $s$  in the Lovász extension  $\bar{v} : \mathbb{R}_+^N \rightarrow \mathbb{R}$  of  $v$  (Lovász, 1983; Algaba et al., 2004),

$$\bar{v}(s) = \sum_{T \subseteq N: T \neq \emptyset} \min_T(s) \cdot \lambda_T(v) = \sum_{i \in N} \text{LS}_i(v, s) \quad \text{for all } s \in \mathbb{R}_+^N,$$

where the first equation is due to Algaba et al. (2004, Theorem 5) and the second one is immediate from (3).

The gain from cooperation  $s_i \cdot \lambda_T(v)$  is distributed equally among the scarcest types from  $T$ . Since the scarcest types determine the gains from cooperation, this indicates that the Lovász-Shapley value assigns a **competitive remuneration** to the individuals in a population according to their marginal productivity.

Casajus and Wiese (2016, Equation 12) also provide a formula for the Lovász-Shapley value in terms of the players' marginal contributions. We have

$$\text{LS}_i(v, s) = |R(s)|^{-1} \cdot \sum_{\rho \in R(s)} s_i \cdot MC_i^v(\rho) \quad (4)$$

for all  $v \in \mathbb{V}$ ,  $s \in \mathbb{R}_+^N$ , and  $i \in N$ , where

$$R(s) := \{\rho \in R \mid \rho(i) < \rho(j) \text{ for all } i, j \in N \text{ with } s_i > s_j\}. \quad (5)$$

Compare this with the corresponding formula for the Shapley value. The marginal contributions are multiplied by the population sizes, indicating how the latter determine the generation of worth. Summation runs only over those rank orders where players with a larger population size precede those with a smaller population size.

If the population sizes are pairwise different,  $s_i \neq s_j$  for all  $i, j \in N$ ,  $i \neq j$ , then the set  $R(s)$  becomes a singleton and the Lovász-Shapley value can be expressed in a particularly simple way (Casajus and Wiese, 2016, Remark 5). For all  $\rho \in R$ ,  $v \in \mathbb{V}$ , and  $i \in N$ , we have

$$\text{LS}_i(v, s) = s_i \cdot MC_i^v(\rho) \quad \text{for all } s \in \mathbb{R}_+^N(\rho), \quad (6)$$

where

$$\mathbb{R}_+^N(\rho) := \{s \in \mathbb{R}_+^N \mid \text{for all } i, j \in N, \rho(i) < \rho(j) \text{ implies } s_i > s_j\}.$$

Note that  $s \in \mathbb{R}_+^N(\rho)$  implies  $R(s) = \{\rho\}$ . Equation (6) has important implications. We interpret the term  $s_i \cdot MC_i^v(\rho)$  as the fitness of the population of type  $i$ . It is obtained by multiplying the fitness of an individual,

$MC_i^v(\rho)$ , by the population size  $s_i$ . Now observe that the fitness of an individual does not depend on the populations sizes  $s \in \mathbb{R}_+^N(\rho)$  but only on the rank order  $\rho$ . This fact will prove to be very useful when we study the stability of population states.

## 4 A Framework for Evolutionary Cooperative Game Theory

The general idea is to establish an evolutionary setup in which the population dynamics are determined by an underlying TU game  $v$ . There is an (infinite) continuum of individuals (members) of a population (society). These individuals are of different types, where the types correspond to the players in the TU game. Whenever this does not cause confusion, we write “type  $i$ ” instead of “type corresponding to  $i$ ” for  $i \in N$ . The sizes of the (sub)populations of individuals of the different types are given by a population state  $s \in \mathbb{R}_+^N$ , where  $s_i$  stands for the size of the population of individuals of type  $i$ . As the population dynamic evolves in the continuous space of the types’ population sizes, we use the Lovász-Shapley value as a plausible concept to determine the fitness of a type’s population in each population state  $s \in \mathbb{R}_+^N$  from the underlying TU game.

### 4.1 The Lovász-Shapley Replicator Dynamics

If in total a society in population state  $s$  generates worth according to the Lovász extension  $\bar{v}$  of TU game  $v$ , one may view  $\bar{v}(s)$  as the total additional fitness of the society to be distributed among its members. We can thus interpret  $LS_i(s)$  as the growth of type  $i$ ’s (absolute) population size in state  $s$ , that is, the population dynamic is generally described by

$$\dot{s}_i = LS_i(v, s), \quad \text{for all } s \in \mathbb{R}_+^N, \quad (7)$$

where throughout we use the notation  $\dot{f} := \partial f / \partial t$  for the first derivative of function  $f$  with respect to (continuous) time  $t$ . We can reformulate these dynamics in relative terms, as the vector field of the population size dynamics of  $s$  is positively homogeneous in  $s$ :<sup>5</sup>

$$\dot{x}_i = g_i(x), \quad \text{for all } x \in \Delta_+^N, \quad (8)$$

where  $x_i := s_i / \sum_{j \in N} s_j$  is the population share of type  $i$ , that is  $x \in \Delta_+^N := \{x \in \mathbb{R}_+^N \mid \sum_{i \in N} x_i = 1\}$ . The real valued, measurable and bounded function  $g_i(x)$  in Equation (8) is implicitly defined by Equation (7) and will be discussed explicitly further below. Note that the right-hand side of Equation (7) is discontinuous (indicated by the minimum operator in Equation (3)), which implies that also  $g_i(x)$  is discontinuous.

Equation (7) implies that for generic population states, i.e.,  $s \in \mathbb{R}_+^N(\rho)$ , using Equation (6), we obtain the population dynamics in absolute terms for some given  $\rho \in R$  by

$$\dot{s}_i = s_i \cdot MC_i^v(\rho), \quad \text{for all } s \in \mathbb{R}_+^N(\rho), i \in N. \quad (9)$$

Therefore, the growth rate of type  $i$ 's population size is  $\dot{s}_i / s_i = MC_i^v(\rho)$  for  $\rho \in R(s)$ . Equation (9) defines an autonomous, linear and homogeneous dynamical system and the local solution  $\xi : I \rightarrow \mathbb{R}_+^N$  to the initial value problem  $\dot{s} = s \cdot MC^v(\rho)$  with the initial value condition  $s(0) = s_0 \in \mathbb{R}_+^N(\rho)$  is given by  $\xi(t) = \exp[MC^v(\rho) \cdot t] \cdot s_0$ , where  $\exp[\cdot]$  denotes the matrix exponential,  $MC^v(\rho)$  denotes the vector whose entries are the marginal contributions of all types under  $\rho$  in the game  $v$ , and  $I := [0, T]$  denotes the (time) interval of definition.

Looking at this setup in relative terms, yields the common replicator dynamics (see, e.g., Taylor and Jonker, 1978), which we focus on in

---

<sup>5</sup>See Remark 3 in Casajus and Wiese (2016) on the positive homogeneity of the Lovász-Shapley value in the population states.

the following. Define the weighted average of the population fitness by  $\overline{MC}^v(\rho, x) := \sum_{j \in N} x_j \cdot MC_j^v(\rho)$ , and let  $\Delta_+^N(\rho) := \Delta_+^N \cap \mathbb{R}_+^N(\rho)$ . For  $\rho \in R$  and  $x \in \Delta_+^N(\rho)$ , it follows from Equation (9) and (8) that

$$\dot{x}_i = g_i(x) := x_i \cdot (MC_i^v(\rho) - \overline{MC}^v(\rho, x)), \quad \text{for all } i \in N. \quad (10)$$

Hence, the growth rate of type  $i$ 's population share,  $\dot{x}_i/x_i$ , is given by the difference of its marginal contribution and the weighted average marginal contribution in population state  $x$ . Note that  $\Delta_+^N(\rho)$  is a region in  $\Delta_+^N$  in which population shares are *strictly* ordered according to  $\rho \in R$ .

## 4.2 The Dynamics as a Differential Inclusion and Its Filippov Solution

As mentioned above, Equation (8) generally describes a system with a discontinuous right-hand side, so that we use specific techniques for the analysis of such setups. Fixing  $\rho \in R$ , from the definition of  $MC_i^v(\rho)$  in Equation (2) we can see that this fitness indicator is constant on  $\Delta_+^N(\rho)$  for each type  $i \in N$ . This simplifies the behavior of the solution to Equation (8) for  $x \in \Delta_+^N(\rho), \rho \in R$ , which is described by Equation (10): the function  $g_i(x)$  is Lipschitz continuous in each region  $\Delta_+^N(\rho)$ , which guarantees the existence and uniqueness of the solution according to the Picard-Lindelöf theorem, when “looking at each region  $\Delta_+^N(\rho), \rho \in R$  separately”.

For  $J \subseteq N$  with  $|J| \geq 2$ , let

$$\Sigma_J := \{x \in \Delta_+^N \mid x_i = x_j, \text{ for all } i, j \in J \subseteq N\} \subset \Delta_+^N$$

be the set of population shares, where all players  $j \in J$  have the same population share. For  $|J| > 1$ , the set

$$D := \cup_{J \subseteq N} \Sigma_J,$$

which is closed relative to  $\Delta_+^N$ , then is a union of a finite number (of subsets) of hyperplanes between the (open) regions  $\Delta_+^N(\rho)$ ,  $\rho \in R$  and it is of measure zero relative to  $\Delta_+^N$ .<sup>6</sup> While the replicator dynamic is Lipschitz continuous on

$$C := \Delta_+^N \setminus D = \bigcup_{\rho \in R} \Delta_+^N(\rho),$$

the Lovász-Shapley value, which determines the behavior of Equation (8), has discontinuity points on  $D$ , so that below we describe the dynamics on the complete simplex  $\Delta_+^N = C \cup D$  as a differential inclusion.

A system of differential equations with a discontinuous right-hand side as in Equation (8) can be adequately dealt with using a differential inclusion, which in our case yields the set valued vector field  $G : \Delta_+^N \rightrightarrows \mathbb{R}^N$ , so that for all  $i \in N$ ,  $\rho \in R(x)$ ,

$$\dot{x}_i \in G_i(x) := \text{con} \left\{ \begin{array}{l} \lim_{\substack{(x_k) \in \Delta_+^N(\rho): \\ (x_k) \rightarrow x}} g_i(x_k) \end{array} \right\} = \begin{cases} \{g_i(x)\} & \text{if } x \in C, \\ \text{con} \left\{ \begin{array}{l} \lim_{\substack{(x_k) \in \Delta_+^N(\rho): \\ (x_k) \rightarrow x}} g_i(x_k) \end{array} \right\} & \text{if } x \in D, \end{cases} \quad (11)$$

where  $\text{con}\{p\}$  is defined as the convex hull of vectors in the set  $p$  and  $(x_k) \rightarrow x$  means that the sequence  $(x_k)$  converges to  $x$ .<sup>7</sup> The above differential inclusion depicts a crucial insight, which we further elaborate on below: the collection of hyperplanes,  $D$ , being of measure zero can be interpreted as being “too small to matter for the overall vector field”. This is reflected by the fact that in Equation (11) the vector field on  $D$  is determined by the points lying close to it in the adjacent regions. The Filippov solution

---

<sup>6</sup>For the case where  $|N| = 2$ , we define  $D := (\cup_{J \subseteq N} \Sigma_J) \cup \{x \in \Delta_+^N | x_i = 1 \text{ for some } i \in N\}$ .

<sup>7</sup>Note that all points  $x \in \Delta_+^N$  are regular, i.e.,  $\sum_i \dot{x}_i = 0$ , even if  $x \in D$ , as the vector field  $G_i(x)$  is determined by a convex combination of regular points.

(Filippov, 1988) uses this insight:<sup>8</sup>

**Definition 1** (Filippov Solution). *A Filippov solution at  $x^*$  of the differential equation (8) with a discontinuous right-hand side is an absolutely continuous function  $\xi : I \rightarrow \Delta_+^N$  for which  $\xi(0) = x^*$  and with respect to the differential inclusion (11) it satisfies  $\xi(t) \in G(x)$  almost everywhere on some interval  $I := [0, T] \subseteq \mathbb{R}$ .*

Note that, in the region  $C$ , the trajectory of the Filippov solution to (8) is equal to that of the “usual” solution, where the vector field is Lipschitz continuous.

## 5 Stability Analysis

Due to the possibility of the non-uniqueness of the Filippov solution, we need to use the accordingly adapted definition of asymptotic stability.

**Definition 2** (Asymptotic Stability). *A population profile  $x^* \in \Delta_+^N$  is asymptotically stable in (11), if both of the following properties hold:*

1. *For any  $\epsilon > 0$  there is some  $\delta > 0$  such that for any solution  $\xi$  to (11) and any  $x_0 \in \Delta_+^N$  such that  $\|x^* - x_0\| < \delta$  and  $\xi(x_0, 0) = x_0$ , we have  $\|\xi(x_0, t) - x^*\| < \epsilon$  for all  $t \in [0, \infty)$ , and*
2.  *$\lim_{t \rightarrow \infty} \|\xi(x_0, t) - x^*\| = 0$  for any solution  $\xi$  to (11) and any  $x_0 \in \Delta_+^N$  such that  $\|x^* - x_0\| < \delta$ .<sup>9</sup>*

---

<sup>8</sup>These techniques have been applied to economic theory, for instance, by Honkapohja and Ito (1983), Ito (1979) and Mohlin (2012), while other solution concepts such as the Caratheodory solution to differential inclusions have been applied in economics, for instance, by Lahkar and Sandholm (2008).

<sup>9</sup>In difference to the usual definition of stability, the properties have to hold for *any* solution to (DI) and the definition employs *absolute* continuity.

Equation (10) describes a system of non-linear, coupled first order ordinary differential equations in the region  $C$ . Therefore, an exact solution is difficult to determine for the general case and we use the insights from Lyapunov theory to analyze the stability of the system.<sup>10</sup> Note that  $x \in \Delta_+^N$  is an asymptotically stable point if and only if it is a strict local maximizer of a Lyapunov function. Thus, we introduce the following common definition.

**Definition 3** (Lyapunov function). *The  $C^1$ -function  $L : \mathbb{R}^N \rightarrow \mathbb{R}$  is an (increasing) strict Lyapunov function for the differential inclusion  $\dot{x} \in G(x)$  if  $\dot{L}(x) \geq 0$  for all  $x \in \Delta_+^N$ , with equality only at stationary points of  $g(x)$ .*

As it is commonly known, the definition implies that  $L$  increases along all solutions and is positive semi-definite (see Sydsaeter et al., 2008, pp. 273). Remember that  $D$  is a set of hyperplanes of measure zero, and thus the dynamics in this region are determined by the dynamics in the adjacent regions of continuity.

**Lemma 1.** *A Lyapunov function of the dynamical system (11) is given by*

$$L(x) := \sum_{i \in N} MC_i^v(\rho)x_i \quad \text{for } \rho \in R(x).$$

*Proof.* Using vector notation, due to the chain rule we have

$$\dot{L}(x) = \nabla L(x) \cdot \dot{x} = \sum_{i \in N} MC_i^v(\rho)\dot{x}_i.$$

Hence, we have  $\dot{L}(x) = 0$  if  $\dot{x} = 0$ , i.e., if  $x$  is a stationary point of  $g$ . Now, take some state  $x$ , where  $\dot{x} \neq 0$ . We can always divide the types into the

---

<sup>10</sup>For the case where  $|N| = 2$  the dynamics in a region  $\Delta_+^N(\rho)$  can be broken down to a Bernoulli differential equation and an exact solution can be found according to the common procedure used for such equations. For  $|N| = 3$  the system resembles a competitive Lotka-Volterra equation for two species (due to  $\dot{x}_1 + \dot{x}_2 + \dot{x}_3 = 0$ , we have one degree of freedom). Abdelkader (1974) shows that an exact solution can be determined for a model equivalent to the three-type-case at hand (his ‘‘Case 3’’).

two sets  $N_+(x) := \{i \in N | \dot{x}_i > 0\}$  and  $N_-(x) := \{i \in N | \dot{x}_i < 0\}$  according to their growth rate in state  $x$ . Rewriting the above expression yields

$$\begin{aligned} \sum_{i \in N} MC_i^v(\rho) \dot{x}_i &= \sum_{i \in N_+(x)} MC_i^v(\rho) \dot{x}_i + \sum_{i \in N_-(x)} MC_i^v(\rho) \dot{x}_i \\ &\geq \left[ \min_{i \in N_+(x)} MC_i^v(\rho) \right] \sum_{j \in N_+(x)} \dot{x}_j + \left[ \max_{i \in N_-(x)} MC_i^v(\rho) \right] \sum_{j \in N_-(x)} \dot{x}_j \\ &= \underbrace{\left( \min_{i \in N_+(x)} MC_i^v(\rho) - \max_{i \in N_-(x)} MC_i^v(\rho) \right)}_{>0} \underbrace{\sum_{j \in N_+(x)} \dot{x}_j}_{>0} > 0, \end{aligned}$$

using  $-\sum_{i \in N_+(x)} \dot{x}_i = \sum_{i \in N_-(x)} \dot{x}_i$  in the last step, which holds, as the population states have to lie in the simplex. This implies that  $L$  increases along all Filippov solutions to  $\dot{x} = g(x)$  with the starting condition  $x^0 \in \Delta_+^N(\rho)$ . Define  $\underline{0} \in \mathbb{R}^{|N|}$  to be a vector with all entries being zeros. Obviously,  $\dot{L}(x) = 0$  if  $\dot{x} = \underline{0}$ .

Note, that  $L$  is single-valued, even if  $R(x)$  is not a singleton. This can be seen in Equation (11) in Casajus and Wiese (2016).  $\square$

The Lyapunov function is a piecewise linear function on  $\Delta_+^N$ , it is non-differentiable (but continuous) at  $x \in D$ , and it is an indicator of the fitness of the whole population, as it coincides with the Lovász extension. Intuitively, the evolutionary process will drive the population to states with larger attainable fitness for a given starting point.

The following result shows that “inefficient” coalitions will not persist in the replicator dynamics derived from TU games. It allows us to relate the asymptotically stable profiles to the structure of the underlying TU game. The result says that a coalition of player types is stable if and only if it implies a higher average productivity than any of its super- or subcoalitions. This reflects the idea, that starting from a population profile, those types with the highest marginal productivity in that profile proliferate with the largest growth rate, and as the population as a whole strives for the

maximum fitness, the dynamics end up in asymptotically stable population profiles as characterized in the Theorem 1 below. For  $S \subseteq N$ ,  $S \neq \emptyset$ , let  $\mathbb{1}_S \in \Delta_+^N$  be given by

$$(\mathbb{1}_S)_i = \begin{cases} \frac{1}{|S|} & \text{if } i \in S, \\ 0 & \text{if } i \notin S, \end{cases}$$

that is, in population state  $\mathbb{1}_S$  only types  $i \in S$  have a positive (and equally split) population share.

**Theorem 1.** *If  $x$  is asymptotically stable, then there exists some  $S \subseteq N$ ,  $S \neq \emptyset$  such that  $x = \mathbb{1}_S$ . Furthermore, for  $S \subseteq N$ ,  $S \neq \emptyset$ ,  $\mathbb{1}_S$  is asymptotically stable if and only if  $\frac{v(S)}{|S|} > \frac{v(T)}{|T|}$  for all  $T \subseteq N$  such that  $T \subsetneq S$  or  $T \supsetneq S$ .*

*Proof.* The proof follows from the following three observations.

1) No population profile  $x \in C$  is asymptotically stable.

The linear functional form of  $L$  does not allow for strict local maxima lying in the region  $C$ .

2) If  $x$  is asymptotically stable, then there exists some  $S \subseteq N$ ,  $S \neq \emptyset$  such that  $x = \mathbb{1}_S \in D$ .

Let  $x$  be stable. We know that asymptotically stable profiles can only lie in areas of (potential) discontinuity, i.e.,  $x \in D$ . We know  $x \in D$  is an asymptotically stable point, if and only if the different parts of the piecewise defined Lyapunov function in each adjacent region attain a strict local maximum in  $x$ . Remember that we have shown that  $L(x)$  is single-valued even if  $x \in D$ . The set

$$\Delta_+^N(y) := \bigcup_{\rho \in R(y)} \Delta_+^N(\rho)$$

is the set of all regions  $\Delta_+^N(\rho)$  which are adjacent to some point  $y \in \Delta_+^N$ . Obviously,  $\Delta_+^N(y)$  is equivalent to  $\Delta_+^N(\rho)$  if  $y \in \Delta_+^N(\rho)$ , i.e., if  $R(y)$  is a singleton. Thus, by stability,

$$L(x) > L(y) \quad \text{for all } y \in \Delta_+^N(x).$$

As the Lyapunov function is piecewise linear and non-constant in the non-trivial cases, i.e., whenever  $\dot{L}$  vanishes on the whole domain, this implies for some  $S \subseteq N$ ,  $S \neq \emptyset$  that  $x = \mathbb{1}_S$ .

### 3) Characterization in terms of average contributions

We know that we only have to consider points  $x \in \Delta_+^N$  for which there exists a coalition  $S$  such that  $x = \mathbb{1}_S$ . A population profile  $\mathbb{1}_S$ ,  $S \subseteq N$ ,  $S \neq \emptyset$  is a strict local maximum of  $L$  if and only if for all  $T \subseteq N$  with  $T \subsetneq S$  or  $T \supsetneq S$  and  $\rho \in R$  such that  $\Delta_+^N(\rho) \subseteq \Delta_+^N(x)$  we have

$$\begin{aligned} L(\mathbb{1}_S) > L(x^T) &\Leftrightarrow \sum_{i \in S} MC_i(\rho)(\mathbb{1}_S)_i > \sum_{i \in T} MC_i(\rho)x_i^T \\ \Leftrightarrow \sum_{i \in S} \frac{MC_i(\rho)}{|S|} > \sum_{i \in T} \frac{MC_i(\rho)}{|T|} &\Leftrightarrow \frac{v(S)}{|S|} > \frac{v(T)}{|T|}. \end{aligned}$$

□

The next statements on the general properties of the stable population profiles follow immediately from Theorem 1.

**Corollary 1** (Existence and Robustness of Asymptotically Stable Population Profiles).

- *An asymptotically stable population profile generically<sup>11</sup> exists.*
- *Asymptotically stable population profiles are robust to small perturbations of the underlying TU game.*

---

<sup>11</sup>We employ the following notion of genericity: A property of games holds generically if there exists an open and dense subset  $\mathcal{V}$  of  $\mathbb{V}$  such that all games in  $\mathcal{V}$  show this property.

The interpretation of the second statement is that small “mutations” do not have an impact on the constellation of asymptotically stable profiles. The next result also follows immediately from Theorem 1.

**Corollary 2.** *A population profile  $x \in \Delta_+^N$  is asymptotically stable for a simple monotonic game  $v \in \mathbb{V}$ , if there exists a minimal winning coalition  $S$  for  $v$  such that  $x = \mathbb{1}_S$ .*

## 6 Conclusion

We have introduced an approach to derive an evolutionary dynamic from an underlying cooperative transferable utility game. This allows us to relate the survival of certain types of coalitions to the payoff structure of the cooperative game. For instance, we obtain the plausible result that, in simple monotonic games, only minimal winning coalitions, that is so to say efficient coalitions, can be asymptotically stable under the replicator dynamic.

Future research could explore whether our model allows to evaluate the “evolutionary plausibility” of cooperative solution concepts like the Shapley value or the Aumann-Dreze value (Aumann and Dreze, 1974) in the same way as non-cooperative evolutionary game theory allows to examine the “evolutionary plausibility” of Nash equilibria. Thus, it could also be viewed as an approach complementary to the Nash program, which tries to establish non-cooperative foundations of cooperative solution concepts.

Obviously, the dynamics of the model we discuss are driven by how worth is generated. This is captured here by the Lovász extension, telling the story of a society in which scarcity and competition are driving forces for the production of worth. A straightforward extension of the model would be to incorporate different and more general production functions, for example, the constant-elasticity-of-substitution production function.

## References

- Abdelkader, Mostafa A.**, Exact Solutions of Lotka-Volterra Equations, *Mathematical Biosciences*, 1974, *20*, 293–291.
- Algaba, Encarnacion, Jesus Mario Bilbao, Julio Rodrigo Fernandez, and Andres Jimenez**, The Lovász Extension of Market Games, *Theory and Decision*, 2004, *56*, 229–238.
- Aumann, Robert J. and Jaques H. Dreze**, Cooperative Games with Coalition Structures, *International Journal of Game Theory*, 1974, *3*, 217–237.
- Casajus, André and Harald Wiese**, Scarcity, Competition, and Value, *International Journal of Game Theory*, 2016, *45*, 1–16.
- Crespi, Bernard J. and Douglas Yanega**, The Definition of Eusociality, *Journal of Behavioral Ecology*, 1995, *6*, 109–115.
- Elster, Jon**, The Case for Methodological Individualism, *Theory and Society*, 1982, *11*, 453–482.
- Filar, Jerzy A. and Leon A. Petrosjan**, Dynamic Cooperative Games, *International Journal of Game Theory*, 2000, *2*, 47–65.
- Filippov, Aleksei F.**, *Differential Equations with Discontinuous Right-hand Sides*, Kluwer Academic Publishers, 1988.
- Gowdy, John and Irmi Seidl**, Economic man and selfish genes: the implications of group selection for economic valuation and policy, *Journal of Socio-Economics*, 2004, *33*, 343–358.
- Hamilton, William D.**, The genetical theory of social behaviour, *Journal of Theoretical Biology*, 1964, *7*, 1–52.

- Harsanyi, John C.**, A Bargaining Model for Cooperative  $n$ -Person Games., in A. W. Tucker and R. D. Luce, eds., *Contributions to the Theory of Games IV*, Vol. 2, Princeton University Press, 1959, pp. 325–355.
- Hodgson, Geoffrey M.**, Meanings of Methodological Individualism, *Journal of Economic Methodology*, 2007, 14, 211–226.
- Hofbauer, Josef and Jörgen Weibull**, Evolutionary Selection Against Dominated Strategies, *Journal of Economic Theory*, 1996, 71, 558–573.
- Honkapohja, Seppo and Takatoshi Ito**, Stability in Regime Switching, *Journal of Economic Theory*, 1983, 29, 22–48.
- Ito, Takatoshi**, A Filippov Solution of a System of Differential Equations with Discontinuous Right-Hand Sides, *Economics Letters*, 1979, 4, 349–354.
- Klaus, Bettina and Jonathan Newton**, Stochastic stability in assignment problems, *Journal of Mathematical Economics*, 2016, 62, 62–74.
- Lahkar, Ratul and William H. Sandholm**, The Projection Dynamic and the Geometry of Population Games, *Games and Economic Behavior*, 2008, 64, 565–590.
- Leshno, Jacob D. and Bary S. R. Pradelski**, The importance of memory for price discovery in decentralized markets, 2019. Available at SSRN: <http://dx.doi.org/10.2139/ssrn.3057554>.
- Lovász, László**, Submodular Functions and Convexity, in A. Bachem, M. Grötschel, and B. Korte, eds., *Mathematical Programming: The State of the Art*, Berlin: Springer, 1983, pp. 235–257.

- Mailath, George J.**, Do People Play Nash Equilibrium? Lessons from Evolutionary Game Theory, *Journal of Economic Literature*, 1998, *36*, 1347–1374.
- McLean, Richard P.**, Values of Non-Transferable Utility Games, in Robert J. Aumann and Sergiu Hart, eds., *Handbook of Game Theory with Economic Applications*, Vol. 3, North-Holland, 2015, pp. 2077–2120.
- Mohlin, Erik**, Evolution of Theories of Mind, *Games and Economic Behavior*, 2012, *75*, 299–318.
- Nash, John F.**, The Agencies Method for Modelling Coalitions and Cooperation in Games, *International Game Theory Review*, 2008, *10*, 539–564.
- Nash, John Forbes**, Two-person Cooperative Games, *Econometrica*, 1953, *21*, 128–140.
- Nax, Heinrich H. and Bary S. R. Pradelski**, Evolutionary Dynamics and Equitable Core Selection in Assignment Games, *International Journal of Game Theory*, 2015, *44*, 903–932.
- Nax, Heinrich H. and Bary S. R. Pradelski**, Core Stability and Core Selection in a Decentralized Labor Matching Market, *Games*, 2016, *7*, 1–16.
- Newton, Jonathan**, Recontracting and stochastic stability in cooperative games, *Journal of Economic Theory*, 2012, *147*, 364–381.
- Newton, Jonathan**, Evolutionary Game Theory: A Renaissance, *Games*, 2018, *9*, 1–67.
- Newton, Jonathan**, Agency Equilibrium, *Games*, 2019, *10*, 1–15.

- Newton, Jonathan and Ryoji Sawa**, A one-shot deviation principle for stability in matching problems, *Journal of Economic Theory*, 2015, *157*, 1–27.
- Nowak, Martin A., Corina E. Tarnita, and Edward O. Wilson**, The evolution of eusociality, *Nature*, 2010, *466*, 1057–1062.
- Samuelson, Larry**, Evolution and Game Theory, *The Journal of Economic Perspectives*, 2002, *16*, 47–66.
- Sandholm, William H.**, *Population Games and Evolutionary Dynamics*, The MIT Press, 2010.
- Shapley, Lloyd S.**, A Value for  $n$ -Person Games, in H.W. Kuhn and A.W. Tucker, eds., *Contributions to the Theory of Games*, Vol. II, Princeton University Press, 1953, pp. 307–317.
- Smith, John Maynard and George Robert Price**, The logic of animal conflict, *Nature*, 1973, *246*, 15–18.
- Sydsaeter, Knut, Peter Hammond, Alte Seierstad, and Arne Strom**, *Further Mathematics for Economic Analysis*, Pearson Education, 2008.
- Taylor, Peter and Leo Jonker**, Evolutionary Stable Strategies and Game Dynamics, *Mathematical Biosciences*, 1978, *40*, 145–156.
- Wilson, Edward O. and Bert Hölldobler**, Eusociality: Origin and consequences, *PNAS*, 2005, *102*, 13367–13371.