Article

Fractional Punishment of Free Riders to Improve Cooperation in Optional Public Good Games

Rocio Botta, Gerardo Blanco and Christian E. Schaerer *

Polytechnic School, National University of Asuncion, San Lorenzo 2111, Paraguay; rbotta@pol.una.py (R.B.); gblanco@pol.una.py (G.B.)
* Correspondence: cschaer@pol.una.py

Abstract: Improving and maintaining cooperation are fundamental issues for any project to be time-persistent, and sanctioning free riders may be the most applied method to achieve it. However, the application of sanctions differs from one group (project or institution) to another. We propose an optional, public good game model where a randomly selected set of the free riders is punished. To this end, we introduce a parameter that establishes the portion of free riders sanctioned with the purpose to control the population state evolution in the game. This parameter modifies the phase portrait of the system, and we show that, when the parameter surpasses a threshold, the full cooperation equilibrium point becomes a stable global attractor. Hence, we demonstrate that the fractional approach improves cooperation while reducing the sanctioning cost.

Keywords: evolutionary game theory; lyapunov stability; optional public good game; random punishment; stability perturbation

1. Introduction

Improving cooperation is a fundamental issue in human activities. The emergence and evolution of cooperation have been studied in life and social science; for instance, biologists have vastly studied altruism [1–3], and economists have studied social dilemmas [4,5]. In many cases, this problem has been modeled by using game theory [6,7] and, lately, with the development of evolutionary game theory [8] by using evolutionary dynamics models [9–11].

In the literature, two main mechanisms that affect cooperation are found: incentives and abstention. Incentives are known to improve the rate of cooperative behavior in a group [12–17], while abstention to participate in a project is useful to maintain cooperation [18,19]. However, incentives are expensive; for this reason, a punishment system is a public good itself [12].

In this paper, we consider improving cooperation by sanctioning only a fraction of the free riders, denominated from here onwards, as fractional punishment. This approach seeks to reduce the number of free riders while minimizing the cost of the sanctioning system. The punishment is built as an extension of the optional public good game model presented in [18].

Therefore, the optional public good game that will be analyzed in this paper consists of a set of randomly selected individuals from the population who shares an institutional service. Each of the individuals decides among three strategies: refuse to play, decide to play and contribute to the common pool (pay for the service), and decide to play but do not contribute (use the service but do not pay for it). In the public good game, the benefit from the common pool is divided equally among all the players. In this work, a fraction of the free riders are randomly selected, and their payoff is reduced to zero.

Sanctioning System

In public good games, the two well-studied forms of sanctioning free riders are peer punishment and pool punishment [20]. In peer punishment, after the game, an individual that contributed (a cooperator) can decide to use part of its payoff to sanction all the free riders in the population. On the contrary, pool punishment is defined before the game; each cooperator must decide to contribute or not to a second pool that will be later on used to sanction the free riders.

A cooperator that is willing to sanction free riders plays a punisher strategy. A punisher has to contribute to the game and to the sanctioning system. A cooperator that decides not to sanction free riders becomes themselves a free rider for the sanctioning system (called second-order free riders). With peer punishment, the cost for a small number of punishers in a group with mostly defectors prevents the punishers from prospering; conversely, in a group without defectors, cooperators and punishers have the same payoff. Therefore, punishers tend to disappear, allowing defectors space to invade the population. With pool punishment, this cannot occur since the punishers must contribute ex ante to the game. Pool punishment, however, is a fixed cost that exists even when there are no defectors in the population [20,21].

Both sanctioning methods were studied and compared in theoretical models and experiments [13,20,22,23]. For instance, it was found that pool punishment can stabilize cooperation if second-order free riders are also punished [21,22]. In addition, the second-order free-riding problem was analyzed in [23–25].

Peer and pool punishment are rarely applied in their pure form in practical situations. More recently, some authors have explored different ways to penalize free riders often inspired in real-life situations. In [26], the sanction that a defector has to pay is a fixed amount. Avoiding second-order free riding by introducing an entrance fee was analyzed in [15]. A set of the cooperators was randomly selected to perform as punishers in [27], the cost of punishing was shared between cooperators in [28], and a ceiling payoff for defectors was introduced to promote cooperation in [29,30].

In this work, we assume that the sanctions have to be implemented with the fee collected periodically in a fee pool. This is observed as a common practice in institutions that provide a service. Within this framework, second-order free-riding does not occur since those who pay for the service implicitly are paying for the sanctioning system. However, this also implies that the profit that each player receives is reduced, given that a part of the collected resources must be used to penalize free riders.

The institution decides the fee-pool portion that will be used for each expense: providing the service and sanctioning free riders. The sanctioning system may not be a fixed cost; when free-riding decreases, the fee-pool amount required to sanction free riders will also decrease. Besides, the institution may limit sanctioning expenses by penalizing only a fraction of the free riders.

The aim of this paper consists in modeling mathematically a fractional punishment mechanism commonly applied in real situations to reduce expenses. We prove that fractional punishment improves cooperation and reduces the cost of punishing, and full cooperation can be achieved when the fraction of sanctioned defectors surpasses a minimum. Moreover, depending on the size of the punished fraction, the cooperation can be attained faster. This means that, by adjusting the parameter, the fractional punishment modifies the nature of the equilibrium points of the system and the phase portrait itself. With this analysis, decision-makers of the institutions can produce scenarios to foresee the effect of the fraction sanctioned in the improvement of the cooperation.

This article is organized as follows: in Section 2, the proposed model and dynamic, as well as the definition of the payoff, are presented. In Section 3, the system is analyzed, and the equilibrium points are characterized as a function of the punishment parameter denoted by d. The implications of the model, along with the penalization of a limited set of the free riders, are discussed in Section 4, while some concluding remarks are finally presented in Section 5.

2. Materials and Methods

Consider a large population where from time to time n randomly selected individuals are offered to play in a public good game ($n = n_x + n_y + n_z$; $n \geq 2$). Those who refuse to play are known in the game as loners (n_z). The loners have a payoff σ that is constant and independent of the game. Those who decide to play the public good game will receive their payoff from the game ($s = n_x + n_y$). In the game, each player must decide if either contributes or not with a positive amount c to the common pool. Individuals who contribute are cooperators (n_x), and those who do not are defectors (n_y). The total contribution to the common pool (obtained from cooperators) is then multiplied by a factor r ($1 < r < n$); in a public good game, the amount in the common pool is distributed equally between all the players, regardless of whether they are cooperators or defectors. Instead, in this work, a randomly selected set of defectors will be sanctioned reducing their payoff to zero.

For modeling the population dynamics, in this paper and following [11], every individual in the population follows a strategy i (for $i = x, y, z$) from the three options available: cooperators x, defectors y, and loners z.

Let $0 \leq x(t), y(t)$, and $z(t) \leq 1$ be the frequency of each of the corresponding available strategies of the population at a specific time t. To simplify the notation, we drop henceforth the time dependency t and simply write x, y, and z. The frequency distribution of the whole population at a specific time t is defined by the state $[x, y, z]$, which belongs to the unit simplex $\mathcal{S}_3$ given by

$$\mathcal{S}_3 = \left\{ [x, y, z] \in \mathbb{R}^3 : x, y, z \geq 0 \text{ and } x + y + z = 1 \right\}. \tag{1}$$

The interior of $\mathcal{S}_3$ is defined as the set of points where all strategies are present (i.e., $x > 0$, $y > 0$, and $z > 0$) and the boundary of $\mathcal{S}_3$ is defined as the set of points that are not interior points. The boundary of the simplex comprises vertices and borders. The vertices x, y, and z represent a homogeneous population of cooperators, defectors, and loners, respectively. In the borders, xy, yz, and zx, one strategy is absent.

Assuming a large enough population in which generations blend continuously into each other, each strategy evolves with a rate that expresses its evolutionary success following the replicator equation

$$\begin{cases} \dot{x} = x(p_x - \bar{p}) \\ \dot{y} = y(p_y - \bar{p}) \\ \dot{z} = z(p_z - \bar{p}) \end{cases} \tag{2}$$

with $x(0)$, $y(0)$, and $z(0)$ as initial conditions in the simplex $\mathcal{S}_3$, with $\bar{p} = xp_x + yp_y + zp_z$, where p_i is the payoff of the ith strategy, and $\bar{p}$ is the average payoff of the population.

It is important to remark that the simplex $\mathcal{S}_3$ remains invariant under the flow of Equation (2) (see [11,31]); hence, if the state $[x, y, z] \in \mathcal{S}_3$ at $t = 0$, then $[x, y, z] \in \mathcal{S}_3 \ \forall t$, and it is endowed with the standard norm $\| \cdot \|_2$. Notice that, in Expression (2), an important behavior relies on the payoff p_i of the ith strategy with respect to the average population payoff $\bar{p}$ in the sense that, if $p_i > \bar{p}$, then the proportion of the i strategy increases, otherwise it decreases. Next, the payoff of each strategy is defined.

Considering a group of n_x cooperators, n_y defectors, and n_z loners and assuming that each contribution is equal ($c = 1$), the payoffs of each strategy in an optional public good game without punishment is defined by [18,32]:

$$p_x = r\frac{n_x}{s} - 1, \quad p_y = r\frac{n_x}{s}, \qquad p_z = \sigma, \tag{3}$$

where the parameters r, σ, and n are considered under the following assumptions:

Assumption 1. *The interest rate on the common pool r satisfies $1 < r < n$.*

If $1 < r$, i.e., if all individuals cooperate, they are better off than if all defects. If $r < n$, each individual is better off defecting than cooperating [18].

Assumption 2. *The payoff of the loner strategy σ satisfies $0 < \sigma < r - 1$.*

This means that the revenue of a cooperator in a group of cooperators (profiting $(r-1)$ each) is better off than loners that receive σ, but loners are better off than a defector in a group of defectors, where the payoff is equal to 0 [18].

To compute the payoff of a player in a group, consider that this group composition depends on the frequencies of all strategies in the population. In the sample group of size n, where s ($s \leq n$) is the number of those who will participate (cooperators and defectors), and $n - s$ is those who refuse (loners), if $s = 1$, the game cannot take place, and the payoff will be equal to a loner payoff regardless of the strategy of the player, and this happens with a probability z^{n-1}. If $s \geq 2$, an individual that agrees to play has $s - 1$ coplayers from the $n - 1$ sample given by the probability [18]

$$\binom{n-1}{s-1}(1-z)^{s-1}z^{n-s}.\tag{4}$$

Furthermore, the probability that j of these $s-1$ coplayers will be cooperators and the other $s-1-j$ defectors is [18]

$$\binom{s-1}{j}\left(\frac{x}{x+y}\right)^{j}\left(\frac{y}{x+y}\right)^{s-1-j}.\tag{5}$$

In this work, we consider that only a fraction d ($0 \leq d \leq 1$) of the defectors will be punished, modifying the payoff presented in (3). In particular, we assume that this set of randomly selected defectors will have their corresponding payoff reduced to 0, while the remaining free riders will obtain the normal payoff. Therefore, a defector not knowing if they will be punished or not, in the presence of j cooperators, will have a payoff equal to

$$p_y = (1-d)\left(\frac{rj}{s}\right)+d0.\tag{6}$$

The expected payoff for a defector in a group of s ($s = 2, ..., n$) players over all possible numbers of cooperators is

$$(1-d)\frac{r}{s}\sum_{j=0}^{s-1}j\binom{s-1}{j}\left(\frac{x}{x+y}\right)^{j}\left(\frac{y}{x+y}\right)^{s-1-j}=(1-d)\frac{r}{s}(s-1)\frac{x}{x+y}.$$

Taking into consideration that a defector can be alone ($s = 1$) or in a group ($s \geq 2$) and that both the number or participants (s) and cooperators (j) in the group are random variables, the average payoff for a defector over all possible numbers of participants is given by

$$p_y(x,z,d) = \sigma z^{n-1}+(1-d)\left(r\frac{x}{1-z}\sum_{s=1}^{n}\binom{n-1}{s-1}(1-z)^{s-1}z^{n-s}\left(1-\frac{1}{s}\right)\right).$$

Observing that $\binom{n-1}{s-1} = \binom{n}{s}\frac{s}{n}$, then the expression for p_y can be rewritten as [18]

$$p_y(x,z,d) = \sigma z^{n-1}+(1-d)\left(r\frac{x}{1-z}\left(1-\frac{1}{n}\sum_{s=1}^{n}\binom{n}{s}(1-z)^{s-1}z^{n-s}\right)\right)$$

$$= \sigma z^{n-1}+(1-d)\left(r\frac{x}{1-z}\left(1-\frac{1}{n}\frac{1-z^n}{(1-z)}\right)\right).$$

For the sake of simplicity, we define an auxiliary function $a := (1 - z^n)/(n(1 - z))$. The payoff then finally takes the following form:

$$p_y(x, z, d) = \sigma z^{n-1} + (1 - d)r\frac{x}{1-z}(1 - a). \tag{7}$$

A similar analysis can be performed to obtain the expected payoff of a cooperator in the population. To distinguish the contribution from one cooperator to the contribution of the other cooperators in the group, the payoff can be rewritten:

$$p_x = \left(\frac{r(n_x - 1)}{n} + \frac{r}{n}\right) - 1.$$

where $n_x - 1$ (the remaining cooperators in the game) is defined as j.

$$p_x = \left(\frac{rj}{n} + \frac{r}{n}\right) - 1 = \left(\frac{r(j+1)}{n}\right) - 1.$$

Therefore, the expected payoff for a cooperator in a group of s $(s = 2, ..., n)$ players over all possible numbers of cooperators is

$$\sum_{j=0}^{s-1} \binom{s-1}{j}\left(\frac{r}{s}(j+1) - 1\right)\left(\frac{x}{x+y}\right)^j\left(\frac{y}{x+y}\right)^{s-1-j} = \frac{r}{s}(s-1)\frac{x}{1-z} + \frac{r}{s} - 1.$$

If the cooperator is the only one that wants to play, that is, if $s = 1$, the payoff is σz^{n-1}. If $(s \geq 2)$, the payoff depends on the number or participants (s) in the group. Hence, the expected payoff for a cooperator is given by

$$p_x(x, z) = \sigma z^{n-1} + \sum_{s=2}^{n}\binom{n-1}{s-1}(1-z)^{s-1}z^{n-s}\left(\frac{r}{s}(s-1)\frac{x}{1-z} + \frac{r}{s} - 1\right)$$

$$p_x(x, z) = \sigma z^{n-1} + r\frac{x}{1-z}\left(\sum_{s=2}^{n}\binom{n-1}{s-1}(1-z)^{s-1}z^{n-s}\right) - r\frac{x}{1-z}\left(\sum_{s=2}^{n}\binom{n-1}{s-1}(1-z)^{s-1}z^{n-s}\frac{1}{s}\right)$$

$$+ r\left(\sum_{s=2}^{n}\binom{n-1}{s-1}(1-z)^{s-1}z^{n-s}\frac{1}{s}\right) - \sum_{s=2}^{n}\binom{n-1}{s-1}(1-z)^{s-1}z^{n-s}.$$

Observing that $\binom{n-1}{s-1} = \binom{n}{s}\frac{s}{n}$, we obtain

$$p_x(x, z) = \sigma z^{n-1} + r\frac{x}{1-z}(1 - z^{n-1}) - r\frac{x}{1-z}\left(\frac{1}{n}\frac{(1-z^n)}{(1-z)} - z^{n-1}\right)$$

$$+ r\left(\frac{1}{n}\frac{(1-z^n)}{(1-z)} - z^{n-1}\right) - 1 + z^{n-1}.$$

Rearranging and using, as before, the auxiliary function a, we obtain

$$p_x(x, z) = \sigma z^{n-1} + ra + r\frac{x}{1-z}(1 - a) + (1 - r)z^{n-1} - 1. \tag{8}$$

The difference in payoff between defector and cooperator strategies shows the relative benefit (or drawback) of defectors over cooperators, and it is essential to characterize the solution orbits. This difference is given by

$$g(x, z, d) := p_y - p_x = 1 + (r-1)z^{n-1} - \frac{r}{n}\frac{(1-z^n)}{(1-z)} - dr\frac{x}{(1-z)}\left(1 - \frac{1}{n}\frac{1-z^n}{(1-z)}\right)$$

$$:= 1 + (r - 1)z^{n-1} - ra - dr\frac{x}{(1-z)}(1 - a). \tag{9}$$

In order to highlight the effect of the parameter d, Expression (9) is rewritten as
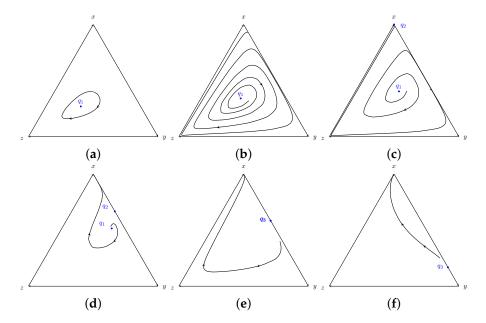
$$g(x,z,d) := m(z) - dxh(z),\tag{10}$$

where $m(z) = 1 + (r-1)z^{n-1} - ra$ and $h(z) = \frac{r}{1-z}(1-a)$. Observe that, when $d = 0$, then $g(x,z,d) = m(z)$, and the free system introduced in [18] is recovered.

3. Results

In this section, we analyze the system (2). To help the reader, and at the risk of oversimplification, before starting the analysis, we outline the effect of the parameter d on the phase portrait of the system (2). This is presented in Figure 1 for increasing values of d. Note that this system has three obvious equilibrium points, which are the vertices of the simplex: (i) $(\hat{x}, \hat{y}, \hat{z}) = (1, 0, 0)$ (ii) $(\hat{x}, \hat{y}, \hat{z}) = (0, 1, 0)$, and (iii) $(\hat{x}, \hat{y}, \hat{z}) = (0, 0, 1)$. Another equilibrium point exists when all strategies have the same payoff $p_x = p_y = p_z = \bar{p}$, which is an interior equilibrium point (denoted as q_1). Later, we show that another equilibrium point arises when incrementing the value of d (this point is denoted by q_2).

For $d = 0$ (no punishment), the vertices (x, y, z) are saddle points and the interior equilibrium (q_1) is a center (see Figure 1a), which is proved in [18]. The interesting effect of the punishment parameter d consists in the fact that the interior equilibrium point changes its stability (see Figure 1b). Furthermore, when d is large enough (d_1), an interior equilibrium point arises (q_2) in the border xy (see Figure 1c). As a consequence, the nature of the equilibrium in the vertex x also changes. As d increases, the equilibrium point q_1 shifts to the border xy, and for a value d_2, both equilibrium points (q_1 and q_2) merge into a new equilibrium point denoted by q_3 (see Figure 1e). The nature of each of the aforementioned fixed points is studied in the next subsections.

Figure 1. Phase diagram of the system when d increases. Parameters $N = 5$, $r = 3$, and $\sigma = 1$. For these values, $d_1 = 0.16667$ and $d_2 = 0.2857$. Notice that q_1, q_2, and q_3 are equilibrium points. (**a**) $d = 0$, (**b**) $d = 0.1$, (**c**) $d = d_1$, (**d**) $d = 0.25$, (**e**) $d = d_2$, and (**f**) $d = 0.99$.

For the analysis, we use the following sequence. In Section 3.1, the boundary of the simplex, particularly the border xy, where a new equilibrium point arises for a sufficiently large d, is analyzed. In Section 3.2, the interior of the simplex is analyzed. First, in Sections 3.2.1 and 3.2.2, the effect of d in the position of the interior equilibrium point is considered, while in Section 3.2.3, the stability of the interior equilibrium point is analyzed. The stability of the equilibrium in the vertex x is analyzed in Section 3.3. A summary is presented in Section 3.4.

3.1. Boundary of the Simplex

In the borders, System (2) reduces to two equations. Next, we present the analysis of System (2) restricted to the boundary of $\mathcal{S}_3$.

3.1.1. Border xy

This border represents a system with cooperators and defectors, where the game is compulsory. Given that, in this border, $x + y = 1$, the dynamics can be analyzed with the following equation:

$$\dot{x} = -x(1-x)g(x,d) = -x(1-x)(1 - (r/n) - drx(1 - 1/n)). \tag{11}$$

The equilibrium points are $\hat{x} = 1$ $(\hat{y} = 0)$, $\hat{x} = 0$ $(\hat{y} = 1)$, and $\hat{x} = (n-r)/(r(n-1)d)$. The qualitative characterization of the aforementioned equilibrium points were studied in [33] with the following lemma (see the proof in Appendix A.1).

Lemma 1. *Let us consider Equation (11) modeling the border xy where x denotes the cooperators, y the defectors, $x \geq 0$, $y \geq 0$, and $x + y = 1$. If the fraction of defectors being sanctioned is denoted by d, with $0 \leq d \leq 1$ and defining $d_1 := (n-r)/(r(n-1))$, then (1) for $0 \leq d \leq 1$, the equilibrium point $\hat{x} = 0$ is locally asymptotically stable, (2) $\hat{x} = 1$ is locally asymptotically stable for $d_1 < d \leq 1$, and (3) the point $\hat{x} = (n-r)/(r(n-1)d)$, which exists in the interior of the border for $d_1 < d$, is an unstable equilibrium point.*

The possible outcomes of the system presented in Lemma 3.1 are sketched in Figure 2. It is important to mention the strong dependence of the characterization of the equilibrium points on the function $g(x,d)$. If $0 \leq d < d_1$, then $g(x,d) \neq 0$ for all values of x, and the two equilibrium points are $\hat{x} = 0$ and $\hat{x} = 1$. In such a case, the $\hat{x} = 1$ is globally unstable, and $\hat{x} = 0$ is globally asymptotically stable (see Figure 2a). For $d_1 < d$, an unstable equilibrium point $\tilde{x}$ arises in the interior of the border (see Figure 2b). If $d = (n-r)/r(n-1)x$, then $g(x,d) = 0$, the corresponding state x is the equilibrium point $\tilde{x}$.

(a)   **(b)**

Figure 2. Stability of equilibrium points with different values of d: a black dot represents a stable point, and a white dot represents an unstable point. (**a**) With $d < d_1$, the $\hat{x} = 1$ is unstable and $\hat{y} = 1$ is asymptotically stable. (**b**) If $d_1 < d \leq 1$, an unstable equilibrium $\hat{x} = (n-r)/r(n-1)d$, denoted by $\tilde{x}$, appears, changing the stability of $\hat{x} = 1$, which becomes an stable equilibrium point.

3.1.2. Border zx

This border represents a population of cooperators and loners. When $x + z = 1$, it is sufficient to analyze the equation $\dot{x} = ((r-1) - \sigma)x(1-x)(1 - z^{n-1})$ to characterize the system. Since $(r-1) \neq \sigma$, the equilibrium points are $\hat{x} = 1$ $(\hat{z} = 0)$ and $\hat{x} = 0$ $(\hat{z} = 1)$. When $(r-1) = \sigma$, the system remains constant, and every point in the border is an equilibrium point, regardless of the value of z. The Jacobian of $\dot{x}$ has the following form:

$$J(x,z) = (1 - 2x)\big(((r-1) - \sigma)(1 - z^{n-1})\big) \tag{12}$$

Evaluating (12) shows that the equilibrium point $\hat{x} = 1$ $(\hat{x} = 0)$ is stable (unstable) if $(r-1) > \sigma$. The dynamic of the border goes from only loners in the population to full cooperation. On the contrary, if $(r-1) < \sigma$, the equilibrium point $\hat{x} = 1$ $(\hat{x} = 0)$ is unstable (stable) and the system reverses direction; it goes from full cooperation to a population of loners.

3.1.3. Border yz

This border represents a population composed of defectors and loners. Since $y + z = 1$, the system can be completely represented by $\dot{y} = y(1 - y)\sigma(z^{n-1} - 1)$. When $\sigma \neq 0$, the equilibrium points are $\hat{y} = 1$ ($\hat{z} = 0$) and $\hat{y} = 0$ ($\hat{z} = 1$). When $\sigma = 0$, the system remains constant, and every point in the border is an equilibrium point, regardless of the value of z. The Jacobian of $\dot{y}$ has the following form:

$$J(y, z) = (1 - 2y)\sigma(z^{n-1} - 1) \tag{13}$$

The equilibrium point $\hat{y} = 1$ ($\hat{y} = 0$) is unstable (stable) if $\sigma > 0$. The dynamic in the border goes from full defection to a population of loners. Conversely, if $\sigma < 0$, $\hat{y} = 1$ ($\hat{y} = 0$) is stable (unstable) and the system reverses direction, going from only loners to full defection.

Observe that the parameter d does not affect the borders zx and yz. Furthermore, since the game setup defines $0 < \sigma < r - 1$, the dynamic of the border yz goes from full defection to only loners in the population, while the dynamic of the border xz goes from only loners in the population to full cooperation.

3.2. Interior of the Simplex

To determine and analyze the interior equilibrium point, a new variable $f := x/(x + y)$ is defined [18]. This variable represents the fraction of cooperators among the players in the game and allows one to define a bijection $\mathcal{T} : (x, y, z) \rightarrow (f, z)$, provided that the restriction $x + y + z = 1$ is satisfied. Hence, the system (2) can be rewritten in a more convenient reduced form [18]:

$$\begin{cases} \dot{f} &= -f(1 - f)g(f, z, d) \\ \dot{z} &= z(\sigma - \bar{p}(f, z)) \end{cases} \tag{14}$$

where $(f, z) \in [0, 1] \times [0, 1]$. The function $g(f, z, d)$ is obtained replacing $x = f(1 - z)$ in Expression (10), so $g(f, z, d) = 1 + (r - 1)z^{n-1} - ra - drf(1 - a)$. To obtain a more explicit expression than (14), consider that $y = 1 - x - z$, and $f = x/(1 - z)$. The system of Equation (14) can then be rewritten in the following form:

$$\begin{cases} \dot{f} &= -f(1 - f)g(f, z, d) \\ \dot{z} &= z(1 - z)\left(\sigma + fg(f, z, d) - p_y(f, z, d)\right). \end{cases} \tag{15}$$

For each time t, d parametrizes the solution (f, z) of the system of Equation (15). An equilibrium point of System (15) exists in the interior of the simplex if the following expressions are satisfied:

$$g(f, z, d) = 0 \tag{16a}$$

$$\sigma - fg(f, z, d) - \sigma z^{n-1} - (1 - d)rf(1 - a) = 0, \tag{16b}$$

and since both equations have to be satisfied simultaneously, Equation (16b) is reduced to $\sigma(1 - z^{n-1}) - (1 - d)rf(1 - a) = 0$, so the value of f can be obtained by

$$f = \frac{\sigma}{(1 - d)r}\frac{(1 - z^{n-1})}{(1 - a)}. \tag{17}$$

By then introducing f in (16a), an equivalent definition for $g(\cdot)$ that depends only on z and d is obtained:

$$\tilde{g}(z, d) = m(z) + \tilde{h}(z, d) \tag{18}$$

where $\tilde{h}(z, d) = -\frac{d}{1-d}\sigma(1 - z^{n-1})$ and $m(z) = 1 + (r - 1)z^{n-1} - ra$ (as introduced in Expression (10)). Now, from (18), the value of z in the equilibrium, defined as $\hat{z}$, can be

numerically obtained as a function of d. Recalling that a depends only on z (as defined in Section 2) and denoting $\hat{a}$ as the expression of a with $\hat{z}$, then (16a) and (16b) become

$$\begin{cases} 1 + (r-1)\hat{z}^{n-1} - r\hat{a} - dr\hat{f}(1-\hat{a}) &= 0 \\ \sigma\hat{z}^{n-1} + (1-d)r\hat{f}(1-\hat{a}) &= \sigma \end{cases} \tag{19}$$

where $\hat{f}$ corresponds to the value of f at the equilibrium point associated with $\hat{z}$ for a d. Defining an auxiliary variable as $\tilde{b} := r\hat{a}$, the above equations take the form of a system of two equations with two unknowns ($\hat{f}$ and $\hat{b}$), as follows:

$$\begin{cases} 1 + (r-1)\hat{z}^{n-1} - \tilde{b} - d\hat{f}(r-\tilde{b}) &= 0 \\ \sigma\hat{z}^{n-1} + (1-d)\hat{f}(r-\tilde{b}) &= \sigma \end{cases} \tag{20}$$

Extracting $\hat{f} = (\sigma(1-\hat{z}^{n-1})/((1-d)(r-\tilde{b}))$ from the second equation and introducing it in the first equation, $\tilde{b}$ is obtained. Replacing $\tilde{b}$ in the second equation, a new form of $\hat{f}$ in the equilibrium is then obtained:

$$\hat{f} = \frac{\sigma}{(r-1) + d(\sigma - (r-1))}. \tag{21}$$

Notice that the interior equilibrium point $(\hat{f}, \hat{z})$ corresponds in the system $(\hat{x}, \hat{y}, \hat{z})$ to the equilibrium point in the interior of the simplex $\mathcal{S}_3$ given by

$$(\hat{x}, \hat{y}, \hat{z}) = \left(\frac{\sigma}{\alpha}(1-\hat{z}), (1-\frac{\sigma}{\alpha})(1-\hat{z}), \hat{z}\right)$$

where $\alpha = (r-1) + d(\sigma - (r-1))$ and the dependence on the parameter d is given explicitly. To relate both systems, a summary of the equilibrium points with respect to the variables (x, y, z) and its corresponding (f, z) are given in Table 1.

Table 1. Equilibrium points relationship between $(\hat{x}, \hat{y}, \hat{z})$ and $(\hat{f}, \hat{z})$. $\alpha = (r-1) + d(\sigma - (r-1))$, $\beta = (N-r)/(r(N-1)d)$.

$(\hat{x}, \hat{y}, \hat{z})$	$(\hat{f}, \hat{z})$
$(\hat{x}, \hat{y}, \hat{z}) = (1, 0, 0)$	$(\hat{f}, \hat{z}) = (1, 0)$
$(\hat{x}, \hat{y}, \hat{z}) = (0, 1, 0)$	$(\hat{f}, \hat{z}) = (0, 0)$
$(\hat{x}, \hat{y}, \hat{z}) = (0, 0, 1)$	$(\hat{f}, \hat{z}) = (0, 1)$
$(\hat{x}, \hat{y}, \hat{z}) = (\frac{\sigma}{\alpha}(1-\hat{z}), (1-\frac{\sigma}{\alpha})(1-\hat{z}), \hat{z})$	$(\hat{f}, \hat{z}) = (\frac{\sigma}{\alpha}, \hat{z})$
$(\hat{x} = f, \hat{y} = 1 - f, \hat{z} = 0)$	$(\hat{f}, \hat{z}) = (\beta, 0)$

The first three equilibrium points in Table 1 are the equilibrium points corresponding to the vertex of $\mathcal{S}_3$, the fourth in the interior of the simplex $\mathcal{S}_3$, and the fifth in the border xy. In the next subsections, the effect of d in the equilibrium point will be analyzed. For that reason, $(\hat{f}, \hat{z})$ will henceforth be redefined as $(\hat{f}_d, \hat{z}_d)$ (for $d \neq 0$) to distinguish from the equilibrium when $d = 0$, which will henceforth be defined as $(\hat{f}_0, \hat{z}_0)$.

Before further analysis, let us define another threshold for the parameter d that will be used in the following sections. Notice that Expression (18) when $z = 0$ is reduced to $(1 - r/n) - \sigma(d/1 - d)$, from which $d_2 = (n-r)/(n\sigma + n - r)$ can be obtained.
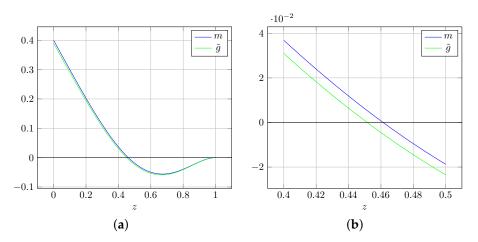
3.2.1. Effect of d over the Value of z in the Equilibrium

Here, the effect of parameter d on functions $g(f, z, d)$, $\tilde{g}(z, d)$, and in $\hat{z}$ itself is analyzed. To this end, consider that, for $d_0 = 0$, there is a value of $\hat{z}_0 \in (0, 1)$ such that $\tilde{g}(\hat{z}_0, 0) = m(\hat{z}_0) = 0$ (see Expressions (10) and (18)). We claim that the punishment parameter d, when $d > d_0 = 0$, displaces the value $\hat{z}_0$ to another $\hat{z}_d$ such that $\hat{z}_d \leq \hat{z}_0$ with $\tilde{g}(\hat{z}_d, d) = 0$. To corroborate this affirmation about the effect of d, we introduce the following two lemmas. In all cases, Assumptions 1 and 2 are fulfilled. The prime symbol denotes the derivative of

the function with respect to its independent variable, for instance, $m'(z) = \frac{dm(z)}{dz}$. We are interested in the relationships between $\tilde{g}(z, d)$ and $\tilde{g}(z, 0) = m(z)$ and between $\tilde{g}'(z, d)$ and $m'(z)$ for a small $d > 0$ and for a z in the neighborhood of $\hat{z}_0$ where $z > \hat{z}_0$.

Lemma 2. *Consider $\hat{z}_0 \in (0, 1)$ associated to $d_0 = 0$ such that $\tilde{g}(\hat{z}_0, 0) = m(\hat{z}_0) = 0$. Let $0 < \epsilon << 1$ such that $\mu = \hat{z}_0 + \epsilon$, and $m(\mu)$, $\tilde{g}(\mu, d)$, $m'(\mu)$, and $\tilde{g}'(\mu, d)$ are negative. Thus, a) $|\tilde{g}(\mu, d)| \geq |m(\mu)|$ and b) $|m'(\mu)| \geq |\tilde{g}'(\mu, d)|$.*

The effect of the parameter d on the solution $\hat{z}_d$ of $\tilde{g}(z, d) = 0$ with respect to the solution $\hat{z}_0$ of $\tilde{g}(\hat{z}_0, d_0) = m(\hat{z}_0) = 0$ of the free system model is analyzed in the following lemma (see the proof in Appendix A.3).
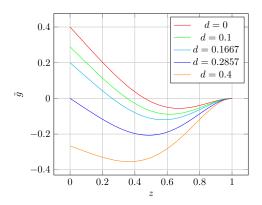
Lemma 3. *Considering the values of $\hat{z}_0$ and $\hat{z}_d$ associated to $d_0 = 0$ and $d > d_0$ such that $\tilde{g}(\hat{z}_0, 0) = m(\hat{z}_0) = 0$ and $\tilde{g}(\hat{z}_d, d) = 0$; hence, $\hat{z}_d \leq \hat{z}_0$.*

It is important to mention that, due to the value of d considered being small and the smoothness of $m(z)$ and $\tilde{g}(z, d)$, the high order terms in the expansion in (A4) do not significantly affect the analysis and can be neglected. The results of Lemmas 2 and 3 can be observed in Figure 3, where the functions $m(z)$ and $\tilde{g}(z, d)$ for a $d = 0.01$ with respect to z in the interval $(0, 1)$ are presented. The effect of displacing $\hat{z}$ to the left is observed; i.e., $\hat{z}_d$ with $d \neq 0$ is smaller than $\hat{z}$ with $d = 0$.

Figure 3. (**a**) $m(z)$ and $\tilde{g}(z, d)$ with $d = 0.01$. In the whole interval $z \in (0, 1)$, (**b**) Zoom in the neighborhood of $m(z) = 0$ and $\tilde{g}(z, d) = 0$. It can be observed that d decreases the value of $\hat{z}$ in the equilibria. It shows how d shifts the equilibrium point $\hat{z}$ to the border of $z = 0$. Parameters: $n = 5$, $r = 3$, and $\sigma = 1$.

Figure 4 shows the evolution of $\hat{z}_d$ with respect to several values of the parameter d (increasing values) remaining constant n, r, and σ. It can be observed that the main effect of the parameter d is precisely to displace $\hat{z}$ to the left from its initial position. This property is exploited for the remainder of the article.

Figure 4. Graph of $\tilde{g}$ with several values of d. Observe that, if $z = 0$, $\tilde{g} = 0$ when $d = d_2(0.2857)$. Parameters: $n = 5$, $r = 3$, and $\sigma = 1$.

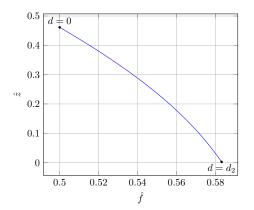### 3.2.2. Effect of d over the Value of f in the Equilibrium

Here, the effect of the parameter d on the value of f in the equilibrium is analyzed. Recall that, at the interior equilibrium point, $\hat{f}_d = \frac{\sigma}{(r-1)+d(\sigma-(r-1))}$. Furthermore, due to game setup, $\sigma < (r-1)$ (see [18]). As a consequence, we have the expression

$$(r-1) + d(\sigma - (r-1)) < (r-1) \tag{22}$$

for all $0 < d < 1$. This means that, for a $0 \leq d < \tilde{d}$ and its corresponding $\hat{f}_0$, $\hat{f}_d$, and $\hat{f}_{\tilde{d}}$, respectively, the following relationship holds:

$$\hat{f}_0 = \frac{\sigma}{(r-1)} \leq \hat{f}_d = \frac{\sigma}{(r-1)+d(\sigma-(r-1))} < \hat{f}_{\tilde{d}} = \frac{\sigma}{(r-1)+\tilde{d}(\sigma-(r-1))}. \tag{23}$$

Therefore, the parameter d increments the equilibrium value $\hat{f}$ or, equivalently, improves the level of cooperation. Figure 5 shows the changes in the value of $\hat{f}$ and $\hat{z}$ when d increases its value from 0 to d_2. It can be seen how $\hat{f}$ increases while $\hat{z}$ decreases.

Figure 5. Interior equilibrium point ($\hat{f}$ and $\hat{z}$) for increasing values of d. Observe that, when d increases, the value of $\hat{f}$ increases, and the value of $\hat{z}$ decreases. Parameters: $n = 5$, $r = 3$, $\sigma = 1$, and $d_2 = (n-r)/(n\sigma + n - r)$.

3.2.3. Effect of d in the Interior Equilibrium Point

Using the definition for a (introduced in (7)), Equation (14) (or equivalently Equation (15)) can be rewritten as follows:

$$\begin{cases} \dot{f} &= -f(1-f)\,g(f,z,d) \\ \dot{z} &= z(1-z)\left(\left(1-z^{n-1}\right)\left(\sigma - f(r-1)\right) - dfr(1-a)(f-1)\right). \end{cases} \tag{24}$$

For simplicity, we will drop the arguments in $g(f, z, d)$ and write simply g. To characterize the equilibrium point at the interior of the simplex, we consider the Jacobian of System (24), which has the following form:

$$J = \begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix}$$

where

$$J_{11} = d\dot{f}/df = -((1-2f)g - f(1-f)dr(1-n)), \tag{25}$$

$$J_{12} = d\dot{f}/dz = -f(1-f)((n-1)(r-1)z^{n-2} - rn' + drfn'), \tag{26}$$

$$J_{21} = d\dot{z}/df = z(1-z)((1-r)(1-z^{n-1}) - (1-a)dr(2f-1)) \tag{27}$$

$$J_{22} = d\dot{z}/dz = (1-2z)((1-z^{n-1})b - e(1-a)) + z(1-z)(-(n-1)z^{n-2}b - e(-a')) \tag{28}$$

with $a' = (1 - z^n - nz^{n-1}(1-z))/(n(1-z)^2)$, $b = \sigma - f(r-1)$ and $e = drf(f-1)$.

In the interior equilibrium equilibrium point, $g = 0$ and $(1 - z^{n-1})b - e(1-a) = 0$. Evaluating the Jacobian then yields

$$J = \begin{pmatrix} f(1-f)dr(1-a) & \begin{array}{c} -f(1-f)((n-1)(r-1)z^{n-2} \\ -ra' + drfa') \end{array} \\ \begin{array}{c} z(1-z)((1-r)(1-z^{n-1}) \\ -(1-a)(2drf - dr)) \end{array} & \begin{array}{c} z(1-z)(-(n-1)z^{n-2}b \\ -e(-a')) \end{array} \end{pmatrix}. \tag{29}$$

To simplify the analysis, considering the influence of the parameter d, we separate the Jacobian in two matrices J_R and J_T of dimension 2×2 as $J = J_R + d J_T$ with

$$J_R = \begin{pmatrix} 0 & -f(1-f)((n-1)(r-1)z^{n-2} - ra') \\ z(1-z)((1-r)(1-z^{n-1})) & -z(1-z)(n-1)z^{n-2}(\sigma - f(r-1)) \end{pmatrix} \tag{30}$$
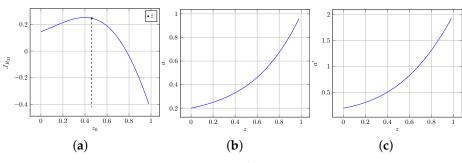
and

$$J_T = \begin{pmatrix} f(1-f)r(1-a)) & -f(1-f)(rfa') \\ -z(1-z)((1-a)r(2f-1)) & z(1-z)rf(1-a)(f-1)a' \end{pmatrix}. \tag{31}$$

It is important to mention that all the entries of the matrices J_R and J_T have bounded derivatives. If $d = 0$, the Jacobian matrix is reduced to J_R. When (30) is evaluated at $(\hat{f}_0, \hat{z}_0)$, it becomes

$$J_R = \begin{pmatrix} 0 & J_{R_{12}} \\ J_{R_{21}} & 0 \end{pmatrix}. \tag{32}$$

Since $\hat{f}_0 = \sigma/(r-1)$ (see (21)), $J_{R_{22}} = 0$. The term $J_{R_{21}}$ is negative because $r > 1$. In addition, the term $J_{R_{12}}$ is positive, for $\hat{z}_0 \in (0,1)$ and for the parameters n, σ, and r with values normally used in the literature. Figure 6a shows the value of $J_{R_{12}}$ for $0 \le z_0 \le 1$ and particularly the value in the equilibrium $(\hat{f}_0, \hat{z}_0)$.

Figure 6. (a) Values of $J_{R_{12}}$ for $0 \leq z_0 \leq 1$ and $\hat{f}_0 = \sigma/(r-1)$; for $\hat{z}_0$, the term is positive. (b) Values of a for $0 \leq z < 1$. (c) Values of a' for $0 \leq z < 1$. Parameters: $n = 5$, $r = 3$, and $\sigma = 1$.

The eigenvalues of J_R are given by $\lambda^2 = J_{R_{21}} J_{R_{12}}$, observing that $J_{R_{21}}$ is negative and $J_{R_{12}}$ is positive. The eigenvalues of J_R are then pure imaginary. This result is important for the rest of the analysis of the interior equilibrium point when $d \neq 0$. The fact of the eigenvalues of J being pure imaginary is not conclusive when the interior equilibrium point is a center. However, we are assuming here that $d = 0$, so the system is reduced to the model in [18], which proves, using a Hamiltonian technique, that the interior equilibrium point is a center.

Next, we analyze the case of the Jacobian matrix J perturbed by the parameter d. To this end, we consider that $d \approx 0$. It is important to remark the implicit dependence of the matrices J_R and J_T on the parameter d.
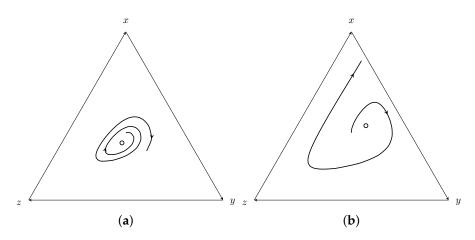
Theorem 1. *Suppose a Jacobian matrix of System (24) denoted by J and decomposed as $J = J_R + dJ_T$ with entries $J_{R_{ij}}$ and $J_{T_{ij}}$ for $i, j \in \{1, 2\}$ with bounded derivatives. Suppose also an initial equilibrium point $(\hat{f}_0, \hat{z}_0)$ of System (24) that corresponds to the parameter $d = 0$, and another equilibrium point for $0 < d$ denoted by $(\hat{f}_d, \hat{z}_d)$. Thus, the Jacobian J, evaluated at $(\hat{f}_d, \hat{z}_d)$ with $d \approx 0$, has complex eigenvalues with positive real parts.*

Theorem 1 (see the proof in Appendix A.4) shows that, for a small parameter d (considered positive), the eigenvalues of the Jacobian J of System (24) are complex with positive real parts when evaluated at the interior equilibrium point $(\hat{f}_d, \hat{z}_d)$. This means that this equilibrium point is an unstable focus; i.e., all solution orbits of System (24) are repelled from the equilibrium point to the boundaries. This can be observed in Table 2, where some values of d, the corresponding equilibrium points $(\hat{f}_d, \hat{z}_d)$, the entries of the Jacobian J, and its eigenvalues are shown. For $d = 0$, the eigenvalues of J are pure imaginary corresponding to the equilibrium point being a center. This result is consistent with the literature [18]. With small positive values of the parameter d ($d = 0.001$ and $d = 0.01$), the eigenvalues have positive real parts, implying that the corresponding equilibrium point is an unstable focus in accordance with Theorem 1 (see also Figure 7a). It is important to mention that, for larger values of the parameter d ($0 \leq d < d_2$), the equilibrium point remains an unstable focus. Moreover, it is observed that $J_{R_{22}}$ is positive for $0 < d$ and $J_{T_{22}} < J_{T_{11}}$, as mentioned in Theorem 1.

As d increases, the equilibrium point $(\hat{f}_d, \hat{z}_d)$ changes its position from the equilibrium point $(\hat{f}_0, \hat{z}_0)$ with $d = 0$ in the interior of the simplex towards the border xy by decreasing $\hat{z}_d$ and increasing $\hat{f}_d$ (see Sections 3.2.1 and 3.2.2).

Table 2. Effect of the parameter d on the entries of the Jacobian J evaluated at equilibrium point $(\hat{f}_d, \hat{z}_d)$ and its corresponding eigenvalues. For $d = 0$, $(\hat{f}_d, \hat{z}_d) = (\hat{f}_0, \hat{z}_0)$. Parameters $N = 5$, $r = 3$, and $\sigma = 1$; $\psi = J_{R_{22}} - \varepsilon_{22} + \varepsilon_{11}$ and $\Delta = (J_{R_{22}} - \varepsilon_{22} + \varepsilon_{11})^2 - 4(J_{R_{21}} \pm \varepsilon_{21})(J_{R_{12}} + \varepsilon_{12})$.

d	$J = J_R + d\,J_T$	$(\hat{f}_d, \hat{z}_d)$	$\lambda = 0.5\,(\psi) \pm 0.5\,\sqrt{\Delta}$
$d = 0.01$	$\begin{bmatrix} 0 & 0.24831 \\ -0.47471 & 0.00046 \end{bmatrix} + d\begin{bmatrix} 0.48162 & -0.21732 \\ -0.00437 & -0.06878 \end{bmatrix}$	$(0.5025, 0.4516)$	$\lambda_1 = 0.00210 - 0.34182i$ $\lambda_2 = 0.00210 + 0.34182i$
$d = 0.001$	$\begin{bmatrix} 0 & 0.24688 \\ -0.47454 & 0.00005 \end{bmatrix} + d\begin{bmatrix} 0.47778 & -0.22112 \\ -0.00044 & -0.06996 \end{bmatrix}$	$(0.5003, 0.4604)$	$\lambda_1 = 0.00021 - 0.34213i$ $\lambda_2 = 0.00021 + 0.34213i$
$d = 0$	$\begin{bmatrix} 0 & 0.24671 \\ -0.47450 & 0 \end{bmatrix} + d\begin{bmatrix} 0.47735 & -0.22155 \\ 0 & -0.07008 \end{bmatrix}$	$(0.5, 0.4613)$	$\lambda_1 = -0.34215i$ $\lambda_2 = +0.34215i$

Figure 7. Effect of d on the interior equilibrium point. Parameters $n = 5$, $r = 3$, and $\sigma = 1$, and values of d as (**a**) $d = 0.1$ and (**b**) $d = 0.2$. The white dot represents an unstable point.

3.3. The Equilibrium Point $(x, y, z) = (1, 0, 0)$

This equilibrium represents a homogeneous population composed exclusively of co-operators. When $d = 0$, this equilibrium point is a saddle point [19]. Now, we characterize this equilibrium point for $0 < d$ and particularly for $d_1 < d$. For simplicity, the system (f, z) (see (14)) will be used. As can be seen in Table 1, the equilibrium point $(\hat{x}, \hat{y}, \hat{z}) = (1, 0, 0)$ corresponds to $(\hat{f}, \hat{z}) = (1, 0)$. Evaluating the Jacobian at this equilibrium point is obtained:

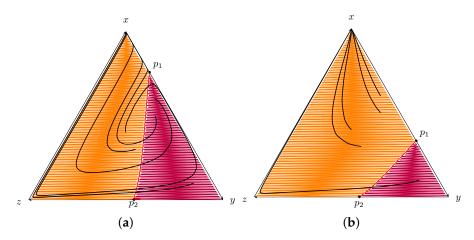$$J = \begin{bmatrix} 1 - r/n - dr(1 - 1/n) & 0 \\ 0 & \sigma - (r - 1) \end{bmatrix}. \tag{33}$$

The eigenvalues of (33) are as follows: $\lambda_1 = 1 - r/n - dr(1 - 1/n)$ and $\lambda_2 = (\sigma - (r - 1))$. Notice that λ_2 is negative given the condition in the game that $0 < \sigma < (r - 1)$ (Assumption 2) and the choice of d determines the sign of λ_1. Recall that $d_1 = (n - r)/r(n - 1)$. Thus, for $d < d_1$, λ_1 is positive, and the equilibrium point is unstable (a saddle); however, if $d_1 < d$, λ_1 is negative and the equilibrium point is locally asymptotically stable.

To better understand this equilibrium point and its basin of attraction, let us consider the relative-entropy function used in [10,34] as a candidate Lyapunov function.

Theorem 2. *Suppose that* $(\hat{f}, \hat{z}) = (1, 0)$ *and the candidate Lyapunov function* $V_{(\hat{f},0)} = \hat{f} \ln\left(\hat{f}/f\right)$ *such that* $V_{(\hat{f},0)} > 0$ *for all* $(f,z) \neq (\hat{f},0)$ *and* $V_{(\hat{f},0)} = 0$ *for all* $(f,z) = (\hat{f},0)$. *If*

$$df > \frac{n(1-z)(1+(r-1)z^{n-1}) - r(1-z^n)}{n(1-z)r - r(1-z^n)} \tag{34}$$

then $\dot{V}_{(\hat{f},0)} < 0$, *and the equilibrium point is locally asymptotically stable (see proof in Appendix A.5).*

In Figure 8, the condition (34) is analyzed for increasing values of d, with the restriction $d > d_1$, since this is a requisite for the equilibrium point to be stable. There are two main regions in the simplex. The region colored in orange contains all the states where $\dot{V}_{\hat{f}} < 0$, while the region colored in purple contains all the states where $\dot{V}_{\hat{f}} > 0$. Both regions are separated by a red line that divides the simplex from a point in the border xy (where $z = 0$) defined as p_1 to another in the border yz (where $x = 0$) defined as p_2 (see Figure 8). This division corresponds to the set of points where $\dot{V}_{\hat{f}}$ changes sign.

Figure 8. Values of $\dot{V}_{\hat{f}}$ for each possible state of the system. The orange area represent states $\dot{V}_{\hat{f}} < 0$; the purple area represent states $\dot{V}_{\hat{f}} > 0$. The red line between both regions represent the set of points where $\dot{V}_{\hat{f}}$ changes sign. The black lines are trajectories from several initial values in the simplex. Parameters $n = 5$, $r = 3$, and $\sigma = 1$, and increasing values of d: (**a**) $d = 0.25$ and (**b**) $d = 0.5$.

As d increases the region where $\dot{V}_{\hat{f}} < 0$ also increases. This happens because the point p_1 moves to the vertex $y = 1$. Notice that the value of p_1 obtained from $\dot{V}_{\hat{f}} = 0$ is $x = (n-r)/r(n-1)d$. This is the position of the unstable equilibrium in the border xy analyzed in Section 3.1.1. The point p_2, however, does not depend on d. When $x = 0$, the expression is reduced to $1 + (r-1)z^{n-1} - ra = 0$ (see (10)). Therefore, the value of z in the point p_2 is fixed and corresponds to the value of $\hat{z}$ when $d = 0$ as in [18].

Additionally, in Figure 8, some paths with initial value points from both regions— $\dot{V}_{\hat{f}} < 0$ and $\dot{V}_{\hat{f}} > 0$—are shown. Due to the structure of the solution, even trajectories with initial values outside the basin of attraction are drawn inside and eventually end in the equilibrium point. In such a case, the equilibrium point $(1,0,0)$ is globally asymptotically stable.
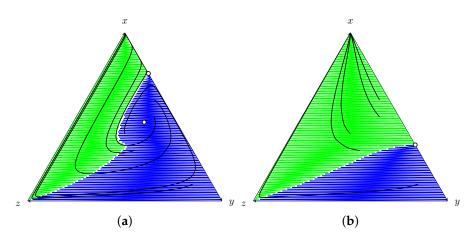
Notice in Figure 8a that some trajectories oscillate before reaching the equilibrium. This is caused by the unstable equilibrium point in the interior of the simplex analyzed in Section 3.2. Recall that the interior equilibrium shifts to the border xy as d increases, reaching it when $d = d_2$, which, for the parameter in Figure 8, is $d_2 = 0.2857$.

From a practical point of view, we are interested in trajectories where the frequency of cooperators does not decrease and, as a consequence, in the set of initial values from which the number of cooperators increases along all the trajectory to the equilibrium point. To observe this behavior, two regions are defined in the simplex $\mathcal{S}_3$; we denote by Ω a

set of initial values where the frequency of the cooperator strategy in the solution orbits continuously grows until full cooperation is reached, i.e.,

$$\Omega := \{(x(t_0), y(t_0), z(t_0)) : x(\tilde{t}) \le x(t) \quad \forall t_0 \le \tilde{t} \le t. \tag{35}$$

In Figure 9, the region Ω is in green. It can also be observed (and this is an important consideration) that Ω is enlarged as the parameter d increases. In addition, the set of initial values in the simplex $\mathcal{S}_3$, where the frequency of cooperators decreases before reaching the equilibrium point, i.e., $\mathcal{S}_3 - \Omega$, is colored in blue.

(a)   **(b)**

Figure 9. The behavior of x in the trajectories for each possible initial value. The green area represents initial values where the frequency of cooperators x increases along the orbit; the blue area represents initial values where at some point of the trajectories x decreases. The black lines are trajectories from several initial values in the simplex. $n = 5$, $r = 3$, and $\sigma = 1$, and values of d are (**a**) $d = 0.22$ and (**b**) $d = 0.5$.

3.4. Summary of the Fractional Punishment Effect in the System

In this section, we complete the outline presented in Figure 1 in Section 3. Notice that the whole effect of the parameter d in the dynamic and the equilibrium points of System (2) can be described using the thresholds d_1 and d_2 as introduced in Section 3.1.1 (Lemma 1) and Section 3.2, respectively.

Recall that, when $d = 0$, the interior equilibrium point q_1 is a center, and the vertices are saddle points [18]. Figure 1a shows the periodic orbits around the equilibrium.

Positive values of d cause two transformations in the system. First, the equilibrium point q_1 becomes unstable; second, the position of the equilibrium changes. It moves towards the border xy. Figure 1b shows how, with small values of d, $d < d_1$, the trajectories approaching the border of the simplex oscillate.

As d increases, i.e., $d_1 < d < d_2$, a new equilibrium point q_2 arises in the border xy. This is an unstable equilibrium point that moves to the vertex y as d increments its value. Furthermore, when q_2 appears, the vertex x changes its nature from unstable (saddle point) to stable (see Figure 1c,d). Notice that the number of oscillations needed to reach the vertex x decreases as d grows.

When $d = d_2$, the equilibrium point q_1 reaches the border xy and merges with the equilibrium point q_2 in the border. This new equilibrium q_3 remains unstable (see Figure 1e).

From $d_2 < d$ to $d = 1$, there is no equilibrium in the interior of the simplex. The point q_3 continues to move toward the vertex y, but it does not reach the vertex; when $d = 1$, the equilibrium point on the xy border takes the value $\hat{x} = (n - r)/r(n - 1)$.

4. Discussion

In this work, we present a mathematical model in which a fractional punishment is incorporated in an optional public good game under the replicator dynamic. To design the

model, we consider two common situations regarding the administration of institutions: (1) not all those who did not pay their fees can be punished due to a lack of resources; (2) both the service provided and the cost of the sanctioning system come from the fee paid by the users (fee pool).

Our proposal seeks to handle this situation by limiting the sanction to a random set of defectors, warranting at the same time both the economic feasibility and effective punishment to at least a reduced number of free riders.

To this end, we assume in the game that a randomly selected portion of free riders would be penalized, but we do not include a separate pool to be used to sanction defectors. With these assumptions, we eliminate the possibility of second-order free-riding and the cost of second-order punishment, since cooperators, when paying the fee, are also unintentionally contributing to penalizing free riders. Furthermore, the cost of the sanctioning system is mainly dependent on the size of the portion punished, and it may be adapted in accordance with the resources available to the institution.

Small values of d produce solutions that oscillate approaching the border. The effect of the parameter d on the trajectory towards the equilibrium point is an important issue. A wide excursion could jeopardize the survival of any institution because of the increment of the loners and the possibility that this triggers the disappearance of the active population.

As d increases, the participation in the game is improved by reducing the share of loners within the population. This can be understood by the convergence of the equilibrium point to the xy axis. For an appropriate minimum value of d, a population composed exclusively of cooperators can be achieved. If d surpasses a first threshold value d_1, a new unstable equilibrium point arises in the border xy while the equilibrium point in the vertex x becomes stable. Then, when d achieves a second threshold value d_2, the interior equilibrium point reaches the border, and both equilibria merge. Notice that, with fractional punishment, the equilibrium reaches the border with $r < n$, whereas in the model introduced in [18], it occurs when $r = n$.

Even with a value of d adequate to achieve full cooperation, the trajectories can oscillate depending on the initial composition of the population. In this context, the set of initial values from which full cooperation is attained through trajectories with an ever-increasing frequency of cooperators is of special interest. This set of points (denoted by Ω in the results) increases as d increases.

The fractional punishment is a useful mechanism when resources are scarce; it can improve cooperation at a lower cost. However, with this approach, the pressure on the correct use of the contribution (fee pool) is higher, reaching a balance between the expenditure dedicated to the operation with the expenditure on sanction free-riding is more important. Sanctioning free riders improves cooperation but decreases the profit from the game. We consider that an adequate balance between both expenses is perhaps one of the characteristics of a successful institution.

5. Conclusions

We have analyzed the effect of randomly punishing a percentage of defectors in an optional public good game modeled using replicator dynamics. To this end, the payoff of the strategies is redefined to introduce a punishment parameter. This parametrizes the simplex phase diagram, modifying the trajectories of the strategies in the population as well as the nature of some important equilibrium points. A description of the main characteristics variations of the trajectories in the simplex due to the punishment parameter parametrization was presented, and for the equilibrium point that represents full cooperation, the basin of attraction was described. Future works include modeling how the institution decides to set the value of the parameter d (probably adaptive) and its optimal choice, as well as the redistribution to cooperators of the benefits taken from the punished free riders.

Author Contributions: Conceptualization, R.B., C.E.S. and G.B.; methodology, R.B., C.E.S.; formal analysis, R.B., C.E.S. and G.B.; writing—original draft preparation, R.B., C.E.S. and G.B.; writing—

Appendix A

Appendix A.1. Proof of the Lemma 1

Proof. The Jacobian of the Equation (11), defined by $J(x,d) = d\dot{x}/dx$ has the following form:

$$J(x,d) = -(1-2x)(1-(r/n) - drx(1-1/n)) + x(1-x)dr(1-1/n). \quad (A1)$$

When Expression (A1) is evaluated at equilibrium point $\hat{x} = 0$, it takes the form $J(0,d) = -(1-r/N)$. As a consequence, since $r < n$, the eigenvalue is negative, and the equilibrium point is locally asymptotically stable. This also means that the reciprocal equilibrium point $\hat{y} = 1$ is locally asymptotically stable, since $x + y = 1$.

If Expression (A1) is evaluated at equilibrium point $\hat{x} = 1$, the Jacobian depends on the value of the parameter d, i.e., $J(1,d) = (1-r/n) - dr(1-1/n)$. The eigenvalue is negative for $d > d_1$, and $\hat{x} = 1$ is consequently locally asymptotically stable; however, if $d < d_1$, then the eigenvalue is positive and $\hat{x} = 1$ is locally unstable.

In order to analyze the equilibrium point $\hat{x} = (n-r)/(r(n-1)d)$, it should be noticed that d and x are bounded between 0 and 1. If $d = 1$, then $\hat{x} = (n-r)/(r(n-1))$, which is the minimum value of x for which the equilibrium point exists. Notice also that the maximum value of x is 1, so $d = (n-r)/(r(n-1))$, which is d_1, the minimum value of d for which the equilibrium point exists. For $\hat{x} = (n-r)/(r(n-1)d)$, Expression (A1) takes the following form:

$$J(\hat{x},d) = \hat{x}(\hat{x}-1)(-dr(1-1/n)). \quad (A2)$$

Given that $0 < (1-1/n)$ and considering $(n-r)/r(n-1) \le \hat{x} < 1$, the eigenvalue of (A2) is positive for positive values of d; therefore, for $d_1 < d$, the equilibrium point is unstable. $\square$

Appendix A.2. Proof of the Lemma 2

Proof. Part 1. Consider $\mu = \hat{z}_0 + \epsilon$ with $\epsilon << 1$ and compare the expression $\tilde{g}(z,d)$ (introduced in (18)) in two situations: (1) $z = \mu$ and $d_0 = 0$ and (2) $z = \mu$ and $d \ne 0$. Notice that $m(\mu)$ is negative by the hypothesis. In addition, $\tilde{h}(\mu) := -d\sigma(1-\mu^{n-1})/(1-d)$ is negative since $\mu \in (0,1)$ and d is positive. Hence, $|m(\mu)| < |\tilde{g}(\mu,d)|$ for $d \ne 0$ and $|m(\mu)| = |\tilde{g}(\mu,d)|$ for $d = 0$.

Part 2. Consider $\mu \in (0,1)$ with $m(\mu) < 0$ and $m'(\mu) < 0$. Thus,

$$\tilde{g}'(\mu,d) = m'(\mu) + \frac{d}{1-d}\sigma\mu^{n-2} \quad (A3)$$

Notice that $-\frac{d}{1-d}\sigma(1-\mu^{n-1})'|_\mu = +\frac{d}{1-d}\sigma\mu^{n-2}$ is positive, and since $m'(\mu)$ and $g'(\mu,d)$ are negative, then $|\tilde{g}'(\mu,d)| \le |m'(\mu)|$. $\square$

Appendix A.3. Proof of the Lemma 3

Proof. Consider $\eta = \hat{z}_0 + \epsilon$ with $0 < \epsilon << 1$. Thus, $\eta \approx \hat{z}_0$. Considering the Taylor series expansion truncated at the first order for $m(\eta)$ as follows, i.e., high-order terms of the series are neglected, we have

$$m(\eta - \epsilon) = m(\eta) - m'(\eta)\epsilon. \quad (A4)$$

Since $\hat{z}_0 = \eta - \epsilon$ and $m(\hat{z}_0) = 0$ (because $\hat{z}_0$ satisfies $\tilde{g}(\hat{z}_0, 0) = m(\hat{z}_0)$ and $\tilde{g}(\hat{z}_0, 0) = 0$), $\epsilon = m(\eta)/m'(\eta)$.

Consider $\eta = \hat{z}_d + \zeta$, performing a similar analysis for $m(\eta - \epsilon)$ but now for $\tilde{g}(\eta - \zeta, d)$. Thus,

$$\tilde{g}(\hat{z}_d, d) = \tilde{g}(\eta, d) - \tilde{g}'(\eta, d)\zeta = 0 \tag{A5}$$

and $\zeta = \tilde{g}(\eta, d)/\tilde{g}'(\eta, d)$.

Using the results of Lemma 2, we obtain

$$|\epsilon| = \frac{|m(\eta)|}{|m'(\eta)|} \leq \frac{|\tilde{g}(\eta, d)|}{|m'(\eta)|} \leq \frac{|\tilde{g}(\eta, d)|}{|\tilde{g}'(\eta, d)|} = |\zeta| \tag{A6}$$

Since $\hat{z}_0 = \eta - \epsilon$ and $\hat{z}_d = \eta - \zeta$, it is possible to conclude that $\hat{z}_d \leq \hat{z}$. $\square$

Appendix A.4. Proof of the Theorem 1

Proof. Consider

$$J = \begin{pmatrix} 0 & J_{R_{12}} \\ J_{R_{21}} & J_{R_{22}} \end{pmatrix} + d \begin{pmatrix} J_{T_{11}} & J_{T_{12}} \\ J_{T_{21}} & J_{T_{22}} \end{pmatrix} \tag{A7}$$

Thus, for a positive d and $d = \epsilon \approx 0$, we obtain

$$J = \begin{pmatrix} \varepsilon_{11} & J_{R_{12}} - \varepsilon_{12} \\ J_{R_{21}} \pm \varepsilon_{21} & J_{R_{22}} - \varepsilon_{22} \end{pmatrix} \tag{A8}$$

where the sign of ε_{ij} is explicitly specified due to the structure of the entries of the matrix J_T. It can be observed that $\varepsilon_{11} > 0$, ε_{12} and ε_{22} are both negative (see Expression (31), and the values of a and a' in Figure 6b,c) and ε_{21} can be positive or negative depending on the value of f on the trajectory of the solution of System (24).

The eigenvalues of (A8) are obtained by

$$(\lambda - \varepsilon_{11})(\lambda - J_{R_{22}} + \varepsilon_{22}) - ((J_{R_{21}} \pm \varepsilon_{21})(J_{R_{12}} - \varepsilon_{12})) = 0. \tag{A9}$$

Rearranging the expression above, we can obtain the second-order equation in λ

$$\lambda^2 + (-J_{R_{22}} + \varepsilon_{22} - \varepsilon_{11})\lambda - (\varepsilon_{11}(\varepsilon_{22} - J_{R_{22}}) + (J_{R_{21}} \pm \varepsilon_{21})(J_{R_{12}} - \varepsilon_{12})) = 0 \tag{A10}$$

with a solution

$$\lambda = \frac{1}{2}(J_{R_{22}} - \varepsilon_{22} + \varepsilon_{11}) \pm \frac{1}{2}\sqrt{\Delta} \tag{A11}$$

where the discriminant $\Delta = (J_{R_{22}} - \varepsilon_{22} + \varepsilon_{11})^2 - 4(J_{R_{21}} \pm \varepsilon_{21})(J_{R_{12}} - \varepsilon_{12}) + 4\varepsilon_{11}(\varepsilon_{22} - J_{R_{22}})$.

Analysis of $\psi = J_{R_{22}} - \varepsilon_{22} + \varepsilon_{11}$. First observe that, from Expression (23), $\hat{f}_0 < \hat{f}_d$. By Assumptions 1 and 2, and recalling that $J_{R_{22}} = -z(1-z)(n-1)z^{n-2}(\sigma - f(r-1))$, $J_{R_{22}}$ is positive. In addition, comparing ε_{11} and ε_{22} (see that $z(1-z)rf(1-a)(f-1)a' < f(1-f)r(1-a)$ in Expression (31)), and considering $d = \varepsilon$, then it is possible to observe that $\varepsilon_{22} < \varepsilon_{11}$, which is controlled by $d = \varepsilon \approx 0$. Hence, the term $J_{R_{22}} - \varepsilon_{22} + \varepsilon_{11}$ is positive.

Analysis of discriminant Δ. Observe that $J_{R_{22}}$ is an analytic function with respect to d (since it depends on $\hat{z}_d$), and $J_{R_{22}} = 0$ for $d = 0$. Considering the Taylor series expansion of $J_{R_{22}} = J_{R_{22}}(0 + d)$ with respect to the parameter d, $J_{R_{22}}$ is given by the expansion

$$J_{R_{22}}(0 + d) = dJ'_{R_{22}}|_{d=0} + \mathcal{O}(d^2)$$

where $J'_{R_{22}}$ denotes the derivative with respect to the parameter d. Observe that d controls $J'_{R_{22}}|_{d=0}$, since choosing arbitrarily d can make the term $dJ'_{R_{22}}|_{d=0}$ as small as is necessary.

In addition, the expression $(J_{R_{22}} - \varepsilon_{22} + \varepsilon_{11})^2$ inside the discriminant takes the form

$$(J_{R_{22}} - \varepsilon_{22} + \varepsilon_{11})^2 = \left((dJ'_{R_{22}}|_{d=0} + \mathcal{O}(d^2) - \varepsilon_{22} + \varepsilon_{11} \right)^2$$

with all terms controlled arbitrarily by d. As a consequence, $(J_{R_{22}} - \varepsilon_{22} + \varepsilon_{11})^2 = \mathcal{O}(d^2)$. In this case, the discriminant Δ is

$$4J_{R_{21}}J_{R_{12}} \pm 4\varepsilon_{21}J_{R_{12}} - 4\varepsilon_{12}J_{R_{21}} \pm 4\varepsilon_{21}\varepsilon_{12} + 4\varepsilon_{11}\varepsilon_{22} - 4\varepsilon_{11}J_{R_{22}} + \mathcal{O}(d^2)$$

where the dominant term is $4J_{R_{21}}J_{R_{12}}$, since the remaining terms are arbitrarily small. Recalling that $J_{R_{21}}$ and $J_{R_{12}}$ have opposite signs and for a sufficiently small d, the discriminant Δ is negative. As a consequence, the eigenvalues of (A8) are complex with a positive real part. □

Appendix A.5. Proof of the Theorem 2

For the proof of Theorem 2, and following [10], we consider the carrier of a given state as the set of pure strategies with the positive probability in that state. Similarly, we consider the domain of a given state as the set of states in the simplex where the pure strategies that are present in the carrier of the given state have positive probabilities.

Proof. In the equilibrium point $(\hat{f}, \hat{z}) = (1, 0)$, only f has a positive probability; therefore, by using the concept of a carrier introduced above, the candidate Lyapunov function takes the following form:

$$V_{\hat{f}}(f) = \hat{f} \ln \left(\frac{\hat{f}}{f} \right), \tag{A12}$$

with $V_{\hat{f}}(f) > 0$ for all $(f, z) \neq (\hat{f}, \hat{z})$ and $V_{\hat{f}}(f) = 0$ for $(f, z) = (\hat{f}, \hat{z})$.

The derivative with respect of time of this candidate Lyapunov function is given by

$$\dot{V}_{\hat{f}}(f) = -\frac{1}{f} \dot{f}. \tag{A13}$$

Substituting $\dot{f}$ given by Expression (15) in Expression (A13), we obtain

$$\dot{V}_{\hat{f}}(f) = (1 - f)g(f, z, d), \tag{A14}$$

where $\dot{V}_{\hat{f}}(f) = 0$ in the equilibrium point $(\hat{f}, \hat{z}) = (1, 0)$ and is negative when $g(f, z, d) < 0$. To obtain a condition for the parameter d, the explicit form of $g(f, z, d)$ (see Equation (14)) is replaced in Equation (A14). Thus,

$$\dot{V}_{\hat{f}}(f) = (1 - f) \left(1 + (r - 1)z^{n-1} - ra - drf(1 - a) \right). \tag{A15}$$

Observe that $\dot{V}_{\hat{f}}(f)$ is negative if $d > \left(1 + (r - 1)z^{n-1} - ra \right) / (r(1 - a)f)$, or, equivalently,

$$df > \frac{n(1 - z)(1 + (r - 1)z^{n-1}) - r(1 - z^n)}{n(1 - z)r - r(1 - z^n)} \tag{A16}$$

In addition $\dot{V}_{\hat{f}}(f)$ is negative along the trajectories specified by System (15). When (A16) holds, the equilibrium point $(\hat{f}, \hat{z}) = (1, 0)$ is locally asymptotically stable. This means that $(\hat{f}, \hat{z}) = (1, 0)$, and the equivalent point in $\mathcal{S}_3$, $(\hat{x}, \hat{y}, \hat{z}) = (1, 0, 0)$ is locally asymptotically stable. □

References

1. Hamilton, W.D. The genetic theory of social behavior. I and II. *J. Theor. Biol.* **1964**, *7*, 1–52. [CrossRef]
2. Haldane, J.B. Population genetics. *New Biol.* **1955**, *18*, 34–51.
3. Nowak, M.A. Five Rules for the Evolution of Cooperation. *Science* **2006**, *314*, 1560–1563 [CrossRef] [PubMed]
4. Hardin, G. The Tragedy of the Commons. *Science* **1968**, *162*, 1243–1248.
5. Ostrom, E. *Governing the Commons: The Evolution of Institutions for Collective Action*; Cambridge University Press: Cambridge, UK, 1990.
6. Trivers, R.L. The Evolution of Reciprocal Altruism Published by : The University of Chicago Press Stable URL. *Q. Rev. Biol.* **1971**, *46*, 35–57. [CrossRef]
7. Axelrod, R.; Hamilton, W. The evolution of cooperation. *Science* **1981**, *211*, 1390–1396. [CrossRef]
8. Smith, J.M. Evolutionary game theory. *Phys. D Nonlinear Phenom.* **1986**, *22*, 43–49. [CrossRef]
9. Taylor, P.D.; Jonker, L.B. Evolutionary stable strategies and game dynamics. *Math. Biosci.* **1978**, *40*, 145–156. [CrossRef]
10. Weibull, J. *Evolutionary Game Theory*; MIT Press: Cambridge, MA, USA, 1997.
11. Hofbauer, J.; Sigmund, K. *Evolutionary Games and Population Dynamics*; Cambridge University Press: Cambridge, UK, 1998; p. 323. [CrossRef]
12. Yamagishi, T. The provision of a sanctioning system as a public good. *J. Personal. Soc. Psychol.* **1986**, *51*, 110–116. [CrossRef]
13. Fehr, E.; Gächter, S. Altruistic punishment in humans. *Nature* **2002**, *415*, 137–140. [CrossRef] [PubMed]
14. Sigmund, K. Punish or perish? Retaliation and collaboration among humans. *Trends Ecol. Evol.* **2007**, *22*, 593–600. [CrossRef] [PubMed]
15. Sasaki, T.; Brännström, A.; Dieckmann, U.; Sigmund, K. The take-it-or-leave-it option allows small penalties to overcome social dilemmas. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 1165–1169. [CrossRef] [PubMed]
16. Hauert, C.; Traulsen, A.; Brandt, H.; Nowak, M.A.; Sigmund, K. Public Goods With Punishment and Abstaining in Finite and Infinite Populations. *Biol. Theory* **2008**, *3*, 114–122. [CrossRef] [PubMed]
17. Hilbe, C.; Sigmund, K. Incentives and opportunism: From the carrot to the stick. *Proc. Biol. Sci. R. Soc.* **2010**, *277*, 2427–2433. [CrossRef] [PubMed]
18. Hauert, C.; De Monte, S.; Hofbauer, J.; Sigmund, K. Replicator dynamics for optional public good games. *J. Theor. Biol.* **2002**, *218*, 187–194. [CrossRef] [PubMed]
19. Hauert, C.; De Monte, S.; Hofbauer, J.; Sigmund, K. Volunteering as Red Queen mechanism for cooperation in public goods games. *Science* **2002**, *296*, 1129–1132. [CrossRef] [PubMed]
20. Sigmund, K.; Hauert, C.; Traulsen, A.; Silva, H. Social Control and the Social Contract: The Emergence of Sanctioning Systems for Collective Action. *Dyn. Games Appl.* **2010**, *1*, 149–171. [CrossRef]
21. Traulsen, A.; Rohl, T.; Milinski, M. An economic experiment reveals that humans prefer pool punishment to maintain the commons. *Proc. R. Soc. B Biol. Sci.* **2012**, *279*, 3716–3721. [CrossRef]
22. Sigmund, K.; De Silva, H.; Traulsen, A.; Hauert, C. Social learning promotes institutions for governing the commons. *Nature* **2010**, *466*, 861–863. [CrossRef]
23. Hilbe, C.; Traulsen, A.; Rohl, T.; Milinski, M. Democratic decisions establish stable authorities that overcome the paradox of second-order punishment. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 752–756. [CrossRef] [PubMed]
24. Ozono, H.; Kamijo, Y.; Shimizu, K. Punishing second-order free riders before first-order free riders: The effect of pool punishment priority on cooperation. *Sci. Rep.* **2017**, *7*, 1–6. [CrossRef] [PubMed]
25. Sasaki, T.; Unemi, T. Replicator dynamics in public goods games with reward funds. *J. Theor. Biol.* **2011**, *287*, 109–114. [CrossRef]
26. Dercole, F.; De Carli, M.; Della Rossa, F.; Papadopoulos, A.V. Overpunishing is not necessary to fix cooperation in voluntary public goods games. *J. Theor. Biol.* **2013**, *326*, 70–81. [CrossRef]
27. Chen, X.; Szolnoki, A.; Perc, M. Probabilistic sharing solves the problem of costly punishment. *New J. Phys.* **2014**, *16*. [CrossRef]
28. Zhang, C.; Zhu, Y.; Chen, Z.; Zhang, J. Punishment in the form of shared cost promotes altruism in the cooperative dilemma games. *J. Theor. Biol.* **2017**, *420*, 128–134. [CrossRef] [PubMed]
29. Zhang, J.; Zhu, Y.; Chen, Z.; Cao, M. Evolutionary Game Dynamics Driven by Setting a Ceiling in Payoffs of Defectors. In Proceedings of the 36th Chinese Control Conference, Dalian, China, 26–28 July 2017; pp. 11289–11295.
30. Zhang, J.; Zhu, Y.; Li, Q.; Chen, Z. Promoting cooperation by setting a ceiling payoff for defectors under three-strategy public good games. *Int. J. Syst. Sci.* **2018**, *49*, 2267–2286. [CrossRef]
31. Zeeman, E.C. Population dynamics from game theory. In *Global Theory of Dynamical Systems, Lecture Notes in Mathematics*; Springer: Berlin/ Heidelberg, Germany, 1980; Volume 819, pp. 471–497. [CrossRef]
32. Sigmund, K.; Haiden, N.; Hauert, C. The dynamics of public goods. *Discret. Contin. Dyn. Syst. Ser. B* **2004**, *4*, 575–587. [CrossRef]
33. Botta, R.; Schaerer, C.; Blanco, G. Cooperation and punishment in community managed water supply system. In Proceedings of the Conference of Computational Interdisciplinary Science CCIS-2019, Atlanta, GA, USA, 19–22 March 2019.
34. Sandholm, W. *Population Games and Evolutionary Dynamics*; Economic Learning and Social Evolution; MIT Press: Cambridge, MA, USA, 2010.