Altruistic Punishment Can Help Resolve Tragedy of the Commons Social Dilemmas

Garrison W. Greenwood
Portland State University

Altruistic Punishment Can Help Resolve Tragedy of the Commons Social Dilemmas

Garrison W. Greenwood
Department of Electrical & Computer Engineering
Portland State University
Portland, OR 97207–0751 USA
Email: greenwd@pdx.edu

Abstract—Social dilemmas force individuals to choose between cooperation, which benefits a group, and defection which benefits the individual. The unfortunate outcome in most social dilemmas is mutual defection where nobody benefits. Researchers frequently use mathematical games such as public goods games to help identify circumstances that might improve cooperation levels within a population. Altruistic punishment has shown promise in these games. Many real-world social dilemmas are expressed via a tragedy of the commons metaphor. This paper describes an investigation designed to see if altruistic punishment might work in tragedy of the commons social dilemmas. Simulation results indicate not only does it help resolve a tragedy of the commons but it also effectively deals with the associated first-order and second-order free rider problems.

I. INTRODUCTION

Social dilemmas arise whenever individuals must choose between self-interests and collective interests. The study of social dilemmas has dramatically increased over recent years because they describe many real-world pressing problems such as overfishing, use of public lands and energy consumption. In such social dilemmas individuals can choose to "cooperate" by putting their self-interests aside for the collective good or "defect" by putting their self-interests first even if it means the collective benefits are reduced. Social dilemmas have two conflicting characteristics: (1) individuals who defect get higher rewards regardless of what others decide, but (2) mutual cooperation yields greater returns for all individuals than if everyone defects. The unfortunate outcome of most social dilemmas is everyone defects and the group suffers.

Mathematical games provide an ideal framework for studying social dilemmas. These N-player games ($N > 2$) start with a random allocation of individuals in a population choosing to cooperate or defect and then observing how those population choices evolve over time. The objective is to gain insight into why humans make particular choices. The most widely studied such game is a *Public Goods Game* (PGG). Another less studied game (but the author would argue a more realistic social dilemma model) is the *Tragedy of the Commons* (TOC). The difference between these two social dilemmas will be discussed in the next section.

Nowak [1] posited five rules that could promote cooperation in populations: direct reciprocity, indirect reciprocity, network reciprocity, kin selection, and group selection. These rules possibly explain the emergence of cooperation in pairwise interactions but they are practically too simplistic to explain it in large groups. A more holistic approach is needed. For example, indirect reciprocity, which is based on reputation, probably has little effect after 2 or 3 interaction rounds. Direct reciprocity, where individuals make decisions in reaction to other players' previous interactions, are unlikely to explain how an individual makes decisions in a group where say half of the other individuals cooperated while the other half defected. Moreover, Nowak's rules only deal with personal perspectives. In other words, they are responses to past decisions other players made; they do not consider deliberate actions players might take to influence the future choices other players make. One such deliberate action is *altruistic punishment* where players punish others who did not cooperate in previous rounds, while incurring a personal cost to inflict this punishment.

Altruistic punishment has been studied in PGGs, but little work has explored how it affects TOC social dilemmas. TOC is fundamentally different than a PGG so it is interesting to see how altruistic punishment might help resolve a TOC. This issue was investigated and our results are reported in this paper. These results indicate that altruistic punishment can help resolve a TOC.

The paper is organized as follows, In Section II the difference between a PGG and TOC are explained and some past work on altruistic punishment in PGGs is presented. Section III describes our model while Section IV presents experimental results. Section V discusses how altruistic punishment should be implemented to help resolve TOC social dilemmas and provides recommendations for future work.

II. BACKGROUND

Altruistic punishment has been applied to PGGs but not in a TOC. These two social dilemmas have a prisoner's dilemma schema, but they are not the same so it is important to explain the difference. The evolution of player strategies in a finite population is governed by discrete replicator equations which are also described in this section.

A. PGG and TOC

In each round of a PGG N players independently decide whether to cooperate by contributing a fixed amount y to a common pool or defect by contributing nothing. An external benefactor multiplies the pool by a factor $r < N$ and

then equally distributes this increased amount to every player whether or not they contributed. Defectors are therefore free riders because they benefit from the contributions of others.

In each round of a TOC N players use or consume a fixed resource. This resource is periodically renewed. Cooperators limit their consumption to help preserve the resource whereas defectors maximize their consumption regardless of how the resource is affected. But by limiting their consumption cooperators make more of the resource available for the defectors. Thus defectors free-ride by exploiting the goodwill of others. A TOC is considered "resolved" if the resource remains viable—i.e., it is preserved for future use. A population of all cooperators resolves a TOC.

Definition 1. *A good is non-excludable if everyone can use it.*

An example of non-excludability is national defense. Everyone benefits from national defense whether or not they pay the taxes needed to finance it. Cooperators pay taxes whereas defectors are free riders who pay little or no taxes.

Definition 2. *Diminishable goods are finite and can be completely depleted unless replenished.*

One characteristic of diminishable goods is use by one individual denies use by another individual. A good example is fisherman in the Tasman Sea. The amount of fish is finite so any fish caught by one fisherman cannot by caught by another fisherman. Cooperators voluntarily limit the amount they catch to keep the fish population viable. Defectors are free riders because they exploit the benevolence of cooperators because there are now more fish available to take.

The inevitable outcome for a PGG or a TOC is everybody defects. Both a PGG and a TOC are non-excludable because everyone is allowed to participate. However, a PGG is not diminishable because regardless of the number of rounds and regardless of the amount contributed, it is assumed the external benefactor always has sufficient funds to increase the pool amount and distribute it. Conversely, in a TOC the resource is diminishable because it is finite and, unless replenished, will eventually become depleted. That difference may appear subtle, but the distinction is important: players are rivals only in the TOC. In a PGG a payoff given to one player does not reduce the payoff to another player because the pool is distributed equally. However, in a TOC the shared resource is finite so any amount consumed by one player is not available to another player. TOC is thus a zero-sum game while a PGG is not.

B. Altruistic Punishment

Diminishable goods and the associated rivalry are important distinctions between a PGG and a TOC social dilemma. Those distinctions lead to different motivations for punishment.

Definition 3. *Altruistic punishment is punishment inflicted on free riders even if costly to the punisher and even if the punisher receives no material benefits from it.*

The motivation behind punishing defectors is to convince them it is better to cooperate. In a PGG the motivation is purely self-serving: the more others cooperate the larger the pool and the higher the payoff to existing cooperators. Conversely, the motivation for punishment in a TOC is for the collective good: the more others cooperate the higher the likelihood the public good will be preserved. This raises an important question. If the motivation behind the punishment is different in the two social dilemmas, will punishment have a similar affect on cooperation levels in both social dilemmas?

Altruistic punishment is easily added to a PGG. There are now three strategies: *cooperators* (C) who contribute an amount y to a common pool; *defectors* (D) who contribute nothing; and *punishers* (P) who also contribute y but, at some small cost $\alpha > 0$ to impose a fine on defectors. Let $\pi(z)$ be the return to a player $z \in \{C, D, P\}$. Then with a population size N and multiplication factor r the payoffs each round are

$$\pi(z) = \begin{cases} rky/N - m\beta & \text{defector} \\ rky/N & \text{cooperator} \\ rky/N - \alpha & \text{punisher} \end{cases} \quad (1)$$

where k is the number of cooperators, m is the number of punishers, β is the imposed punishment (e.g., a fine) per punisher and α is the fixed cost paid for inflicting a punishment.. Notice the punisher acts like a cooperator, but receives a slighter lower return because of the cost paid to inflict punishment on the defectors. The returns are identical if there are no defectors in the population because the cost is incurred if and only if punishment is administered.

There are a number possible PGG variations. For example, in Eq. (1) the punishment for a defector depends on the number of punishers but the cost α to the punisher is the same regardless of how many defectors are punished. This cost could be changed to say $\ell\alpha$ where ℓ is the number of defectors. Also, punishment could be added to cooperators to address the so-called 2nd order free rider problem with an associated additional cost to the punisher. (These additional punishment forms and costs are incorporated into our TOC experiments.)

Human experiments have shown altruistic punishment has a positive affect on cooperation levels in PGGs. Fehr and Gächter [2] had 240 university students participate in two sets of 6-round PGGs: one set with punishment and one set without. They found that 72% of above average contributors punished below average contributors. They found cooperation levels were considerably higher ($\approx 40\%$) with punishment regardless of whether the set of games with punishment preceded the set without punishment or vice versa. Furthermore, they found contribution levels either remained constant or slightly increased over time with punishment but tended to decrease over time with no punishment[1].

---

[1]Fehr and Gächter also believed Nowak's 5 rules were too simplistic to explain group behavior.
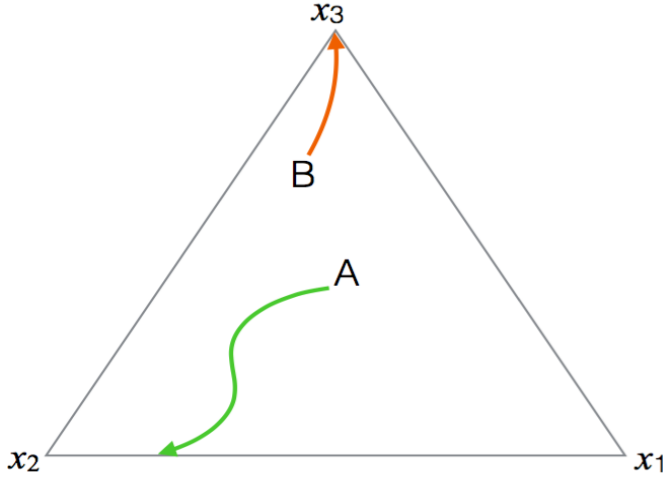
Fig. 1. A 2-simplex showing two trajectories representing the evolution of an infinite population of three strategies. The green trajectory starts at point A and terminates a fixed point on the x_1–x_2 boundary. The final population mixture contains only type 1 and type 2 strategies with more type 2 strategies. The red trajectory starts at point B and ends at the x_3 vertex. This final population has been taken over by type 3 strategies.
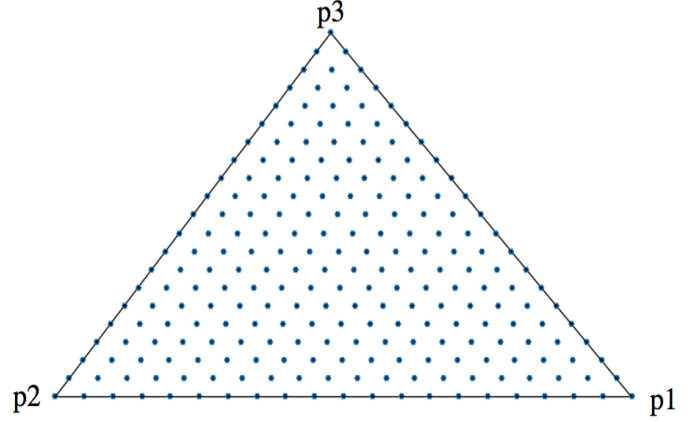

Fig. 2. A 2-simplex for a finite population with $N = 20$. Only the points shown represent an integer number of strategies where $\sum_i k_i = N$. A trajectory can only move between these points.

C. Replicator Equations

How do we represent the evolution of the population over time? A convenient representation is a simplex, Figure 1 shows a 2-D simplex, which is suitable for three strategies. Let $x_i \in [0, 1]$ be the frequency of strategy i where $\sum_i x_i = 1.0$. Each point in the simplex represents a mixture of three strategies in an infinite size population—i.e., every point has coordinates $[x_1\ x_2\ x_3]$. Trajectories in the simplex reflect how the population evolves over time. This evolution is governed by *replicator equations* which are 1st-order differential equations of the form

$$\dot{x}_i = x_i \left(F_i - \bar{F} \right) \qquad (2)$$

where F_i is the fitness of strategy i and $\bar{F}$ is the mean fitness of the population. If the term in parenthesis is positive, then strategy i increases; if negative it decreases; and if zero it does not change. The population size is infinite so a trajectory can pass through any point in the simplex.

Figure 1 shows two trajectories. Trajectories are smooth because their path is described by a differential equation. The green trajectory begins at point A and terminates on the x_1–x_2 boundary. This is a fixed point where the final population consists of (mostly) x_2 strategies but some x_1 strategies. The red trajectory begins at point B and terminates at the x_3 vertex. That final population consists solely of x_3 strategies.

Human populations, however, are always finite. Let k_i, $i \in \{C, P, D\}$ be the number of players choosing strategy i. In a finite population of size N the frequency of strategy i at time t is $p_i^t = k_i/N$. Then the population evolution over time is now given by the *discrete replicator equation*

$$p_i^{t+1} = p_i^t \left(\frac{\pi_i^t}{\bar{\pi}^t} \right) \qquad (3)$$

where where π_i^t is the return for strategy i at time t. The term inside the parenthesis is the ratio of the return of strategy i to the average population return $\bar{\pi}^t$. If the term is greater than 1.0 then strategy i grows in the population; if equal to 1.0 it remains the same; and if less than 1.0 it decreases.

But there is a problem. Eq. (3) can be rewritten (after multiplying both sides by N) as

$$k_i^{t+1} = k_i^t \left(\frac{\pi_i^t}{\bar{\pi}^t} \right) \qquad (4)$$

The problem is there is no guarantee the term in parenthesis is an integer which means the left hand side of Eq. (4) may not be an integer. Clearly $k_1 + k_2 + k_3 = N$ is required. To overcome this problem the quantization algorithm below (from [3]) was used. The algorithm takes $\{p_i^{t+1}\}$ and N and returns k_i' where $\sum k_i' = N$.

1) Compute

$$k_i' = \left\lfloor N p_i + \tfrac{1}{2} \right\rfloor \quad , \quad N' = \sum_i k_i'$$

2) Let $d = N' - N$. If $d = 0$, then go to step 4. Otherwise, compute the errors $\delta_i = k_i' - N p_i$.
3) If $d > 0$, decrement the d k_i''s with the largest δ_i values. If $d < 0$, increment the $|d|$ k_i''s with the smallest δ_i values.
4) Return $[k_1'\ \ k_2'\ \ k_3']$ and exit.

For this investigation we use a 2-D simplex to represent the population evolution. But in finite populations only certain points are valid because there are only a finite number of points where $\sum_i k_i = N$. Figure 2 shows a 2-D simplex for a finite population with $N = 20$. The trajectories showing the population evolution are piecewise linear since only transitions between these points are allowed.

III. The TOC Model

The model has $N = 20$ players. We adopt the same notion used in the previous section where k_i is the number of players

in the population using strategy $i \in \{C, P, D\}$ and $p_i^t = k_i/N$ is the corresponding frequency at time t. (For convenience the t will be dropped when there is no ambiguity.) Each round each player consumes an amount of a finite resource, which is assumed to have an initial capacity 50000 "units" (an arbitrary unit of measure). C and D players consume 40 units while D players consume 50 units. The slight higher consumption rate for D players reflects a greedier, self-serving approach. After 10 rounds the remaining resource capacity is increase by 20%. If the population is composed entirely of C or P players the total consumption over 10 rounds would be $20 \cdot 10 \cdot 40 = 8000$ units. A 20% increase boosts the capacity to slightly over 50000 before starting the next 10 rounds. This slight increase over 50000 keeps the resource capacity viable even if a small number of defectors are present. The game continues for a fixed number of rounds or until the resource is depleted, whichever comes first.

The return is 39 units per round for C or P players and 82 units per round for D players. There is a strong incentive to consume more since the return for defecting is more than twice that of not defecting.

IV. EXPERIMENTAL RESULTS

All simulations were conducted with $N = 20$. The rewards for each strategy i were

$$
\pi(i) = \begin{cases} 82 - k_2\beta & \text{defectors} \\ 39 & \text{cooperators} \\ 39 - ck_3\alpha & \text{punishers} \end{cases} \quad (5)
$$

where β is the defector's punishment and $\alpha = 1.0$ is the cost a punishers pays. The constant c equals 1 if there are defectors in the population or 0 otherwise. This constant ensures punishers pay costs only when defectors are present. Notice punishments and costs are additive since each punisher imposes a punishment β on a defector and pays a cost α for every defector punished.

The returns in Eq. (5) are slightly different from Eq. (1) because the former is for a TOC game and the latter for a PGG. The returns in a PGG are based on a contribution amount y, a multiplication factor r and the frequency of cooperators. In the TOC the constants 82 and 39 represent player consumption rates which only depend on the strategy played. The punishment for defection is the same in both games. Eq. (5) has an additional penalty for cooperator free riding and a corresponding additional cost for the punisher.

Figure 3 shows the population evolution for an initial strategy distribution of $[p_1\ p_2\ p_3] = [0.35\ 0.35\ 0.30]$ corresponding to $k_1 = k_2 = 7$ and $k_3 = 6$. The replicator equations show for low β values the punishment isn't high enough to prevent defectors from taking over the population. The punishment becomes high enough to coerce defectors to switch strategy as β approaches 7. The trajectory with $\beta = 7.1$ terminates at an interior fixed point. (Fixed points are discussed in the next section.) The trajectory with $\beta = 10.0$ terminates at the p_1–p_2 boundary which means there are no defectors left
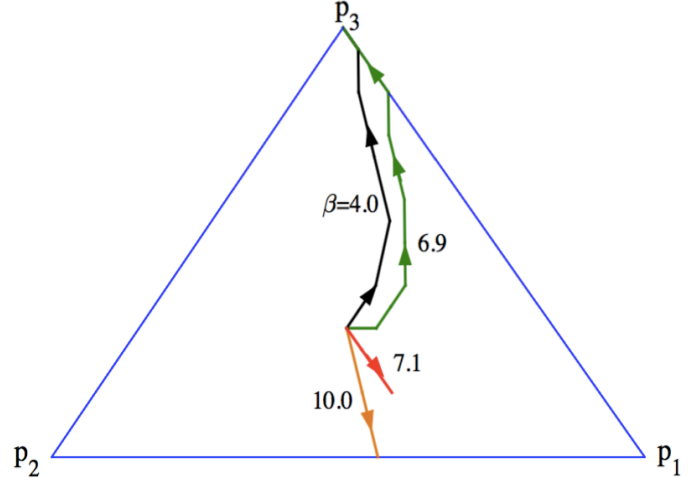

Fig. 3. Evolution of a finite populations ($N = 20$) for various punishment values (β). Punisher cost is $\alpha = 1$. Initial distribution for all trajectories is $[p_1\ p_2\ p_3] = [0.35\ 0.35\ 0.30]$ which corresponds to $k_1 = k_2 = 7$ and $k_3 = 6$. Payoffs are 39, 39 and 82 for C, P and D players respectively. For $\beta \approx 7$ or less defectors eventually take over the population.

in the population. Consequently, $c = 0$ yielding the same payoffs for a punisher and a cooperator. In that case there is no distinction between the two players and the TOC was completely resolved.

Figure 4 shows just two trajectories from Figure 3. The black dashed line is parallel to the p_2–p_3 boundary. Every point on this dashed line has $p_1 = 0.35$ ($k_1 = 7$). Similarly the red dashed line is parallel to the p_1–p_3 boundary and every point on it has $p_2 = 0.35$. Notice the trajectory with $\beta = 7.1$ exactly overlaps the red dashed line. This means the number of punishers did not change from the initial value of $k_2 = 7$. Thus all defectors who switched strategies became cooperators. The trajectory with $\beta = 10.0$ intersects the $p1$–$p2$ boundary to the right of where the dashed black line intersects that boundary indicating the cooperator frequency increased. But it intersects to the left of where the red dashed line intersects indicating the punisher frequency also increased. Thus in this case some defectors switched to cooperators while others switched to punishers[2].

There are actually two ways of inflicting more punishment on defectors: fix k_2 and increase β or fix β and increase k_2. To explore this latter case more thoroughly we conducted an experiment with various initial values of k_2 while fixing the initial value of k_3 at 7. For all runs $\beta = 6.9$ which Figure 3 indicates was too small to prevent the population from being taken over by defectors. Figure 5 shows 7 and even 8 punishers were not enough to prevent the take over by defectors. However, for $k_2 \geq 9$ the population contains enough punishers to coerce defectors to switch strategies.

---

[2]The initial population had $k_1 = 7, k_2 = 7$ and $k_3 = 6$ whereas the final population had $k_1 = 11, k_2 = 9$ and $k_3 = 0$.
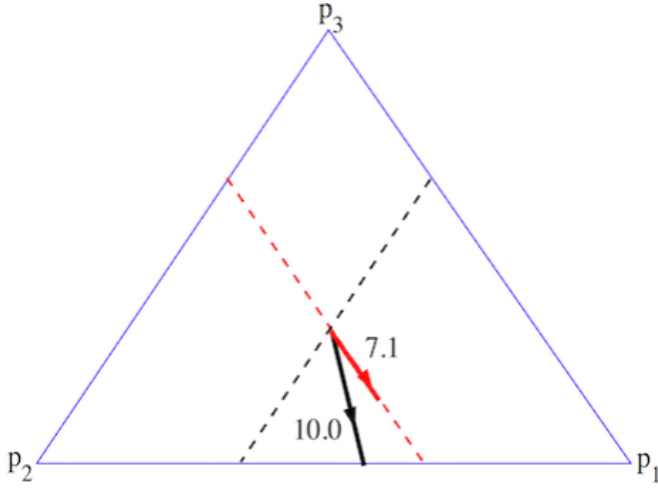
Fig. 4. Evolution of a finite populations ($N = 20$) for $\beta = 7.1$ and 10.0. Dashed lines represent constant strategy frequencies for p_1 (black) and p_2 (red). The initial distribution is $[p_1\ p_2\ p_3] = [0.35\ 0.35\ 0.30]$. These trajectories are the same as those in Figure 3.
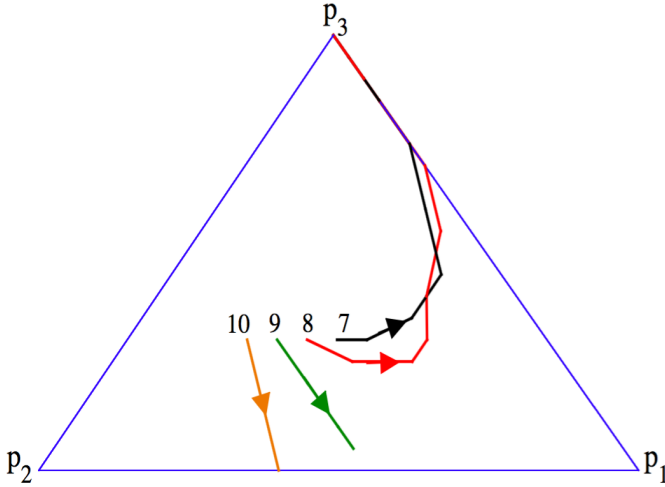

Fig. 5. Effect of varying the frequency of punishers with fixed punishment $\beta = 6.9$ and $k_3 = 6$ (cf., Figure 3). Numbers indicate the initial number of punishers. Punisher cost is $\alpha = 2$.

V. DISCUSSION

The 2-D simplex was used for trajectories showing how strategies evolve in the population. These trajectories were generated from the discrete replicator equations. Strategy frequencies will not change any further if the trajectory hits a *fixed point*. This simplex contains several fixed points. For instance, it is easy to prove every vertex is a fixed point. All interior fixed points are caused by the quantization process. Consider the red trajectory in Fig 3. That trajectory hits a fixed point at $p_i = [0.528\ 0.341\ 0.131]$ which corresponds to $k_1 = 10, k_2 = 7$ and $k_3 = 3$. Plugging those p_i values into the quantization algorithm returns the same population mixture. The location of these interior fixed points will change as N increases and completely disappear as $N \rightarrow \infty$. However, there are some natural fixed points on one simplex boundary

(other than at the vertices) which do not disappear and the exact number grows with N. For infinite population sizes every point on this boundary is a fixed point.

Theorem. Every point on the p_1–p_2 boundary of the 2-D simplex is a fixed point.

Proof. Suppose at time τ a trajectory intersects a point on the p_1–p_2 boundary. Then $k_3 = 0$ and by Eq. (4) remains so for all $t \geq \tau$. This makes $c = 0$ in Eq. (5) so the return to a punisher or a cooperator is the same. The proof follows because with identical returns there is no incentive to change strategies. ☐

Referring back to Figure 3, the initial population mixture was $[p_1\ p_2\ p_3] = [0.35\ 0.35\ 0.30]$. $\beta = 6.9$ wasn't a sufficient punishment to prevent defectors from eventually taking over the population. This poses an interesting question: if the punishment were subsequently increased would this induce defectors to switch strategies?

To investigate this issue further an experiment was conducted where punishers doubled the punishment if p_3 exceeded 0.35. The results are shown in Figure 6. The replicators equations predict after an initial decrease in defectors their number increased again until they take over the population (green trajectory). A second experiment (red trajectory) investigated what would happen if the punishers immediately doubled the punishment if there was *any* increase in defector frequency. In this case, after an initial increase in defectors, they rapidly decreased until a fixed point was reached. This suggests if punishers want to prevent defectors from taking over a population, then they shouldn't wait too long before increasing the punishment.

In this last experiment β was doubled to stop the growth of defectors. That may seem excessive since Figure 3 shows $\beta = 6.9$ wasn't sufficient to stop defector growth but $\beta = 7.1$ was sufficient. Notice the trajectories where defectors prevail move towards the $p2$–$p3$ simplex boundary and then up to the p_3 vertex. In other words, the number of punishers is decreasing. Consequently, the punishment *per punisher* must increase sharply to stop defector growth because there are fewer punishers in the population.

Cooperators limit their consumption of the shared resource for the good of the group. This allows the resource to replenish itself, thereby keeping it viable. Defectors are free riders. They may consume more of the shared resource by free riding on cooperators who want to preserve the resource. The experimental results show punishment can effectively reduce the number of defectors in a finite population, but if and only if the punishment is sufficiently large enough. This can be achieved two ways: either keep the same number of punishers in the population and increase the punishment level or keep the punishment level the same but increase the number of punishers. In both cases the defectors are reduced.

But there is another type of free rider. Altruistic punishment comes with a cost to the punisher, which reduces his return. Cooperators who do not punish free ride on punishers. That is, they reap the benefits of fewer defectors in the population,
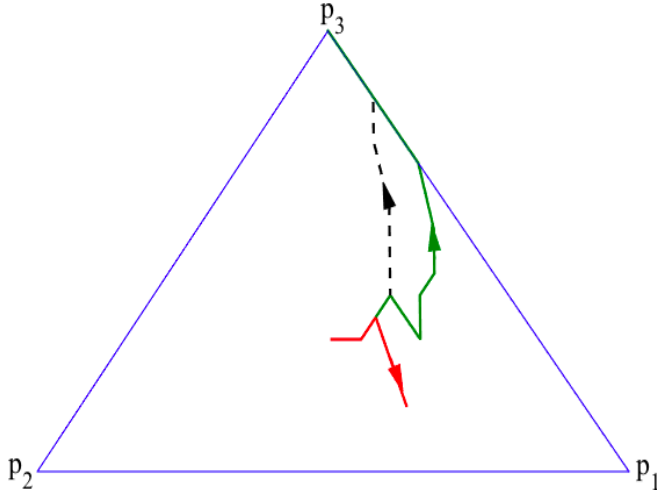
Fig. 6. Effect of adapting β. The dashed trajectory is the same trajectory for $\beta = 6.9$ from Figure 3 included for reference. Initial distribution is $[p_1 \ p_2 \ p_3] = [0.35 \ 0.35 \ 0.30]$. β is doubled when $p_3 \geq 0.35$ (green trajectory) or $p_3 > 0.30$ (red trajectory).
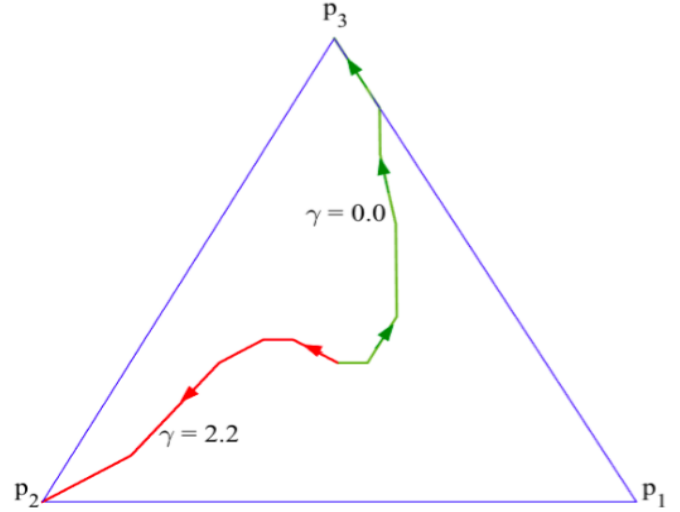

Fig. 7. Effect of punishing cooperators with punisher cost $\eta = 0.1$. γ is punishment to cooperator. Defector punishment is $\beta = 6.9$. $\gamma = 0.0$ trajectory is same trajectory as the green trajectory from Figure 3.

but they let punishers pay the associated cost. This is referred to as the *second-order free rider problem*.

Punishment effectively handles the first-order free rider problem—i.e., defector free riding—so it is reasonable to assume it would also handle the second-order free rider problem. Now cooperators are punished for not punishing defectors. Each punisher pays a cost η to reduce a cooperator's return by γ. The new returns for a strategy i are

$$\pi(i) = \begin{cases} 82 - k_2\beta & \text{defectors} \\ 39 - ck_2\gamma & \text{cooperators} \\ 39 - c[(k_3\alpha) + (k_1\eta)] & \text{punishers} \end{cases} \quad (6)$$

The punishments and costs can be summarized as follows:

1) defectors get punished β by each punisher
2) cooperators get punished by γ by each punisher
3) punishers pay a cost α for each defector punished plus a cost η for each cooperator punished.
4) if there are no defectors $c = 0$ to remove all costs and punishments

Figure 7 shows how punishing cooperators who don't punish defectors helps resolve the second-order free rider problem. The defector punishment was $\beta = 6.9$ with a cost $\alpha = 1.0$. The green trajectory has $\gamma = 0.0$ so there is no cooperator punishment. (This is the same as the green trajectory depicted in Figure 3). The red trajectory shows a cooperator punishment of $\gamma = 2.2$, with a cost to the punisher of $\eta = 0.1$ The defectors are rapidly switching to punishers who eventually take over the population—i.e., $p_2 = 1.0$ and $p_1 = p_3 = 0.0$. But a population of all punishers has the same affect on the shared resource as does a population of all cooperators; no defectors, no punishment costs so the returns are identical. Punishing defectors resolves the first-order free rider problem and punishing cooperators who don't punish resolves the

second-order free rider problem. More importantly, this dual punishment approach resolves the TOC at the same time.

In finite populations discrete replicator equations predict strategy evolution but some form of quantization is necessary. Quantization can produce fixed points in the simplex interior. Interior fixed points always have $k_3 \geq 1$. The TOC is only partially resolved in this case, which can prove problematic. The defectors that remain will continue to over-consume the shared resource and, unless the replenishment rate (20% after 10 rounds in these experiments) is sufficient, the resource will eventually become depleted. Under those circumstances β should be increased to force a complete resolving. An example is shown in Figure 3. $\beta = 7.1$ resulted in the population reaching an interior fixed point whereas $\beta = 10.0$ completely purged defectors from the population.

Defectors always do better in social dilemmas regardless of what others do. The simulation results show altruistic punishment can help resolve a TOC. These results are consistent with previous studies where it was shown punishment is most effective in iterated scenarios where group memberships don't change [4]. But there still is an open question: why do individuals choose particular strategies in a TOC? Defectors may be motivated purely by self interest. Other individuals may choose cooperation because they feel morally obligated to preserve a finite resource. But why would individuals choose to be a punisher, especially if punishment is costly?

Unfortunately replicator dynamics will probably *not* be very helpful in getting an answer. Replicator equations only describe how strategies in a population evolve over time. There is no mutation involved so new strategies cannot emerge and lost strategies cannot reappear. Evolution is dictated strictly in terms of fitness relative to the average population fitness. Put another way, replicator equations only provide proximate causes, not underlying reasons. For instance, Figure 7 shows

punisher growth. But to get this growth punishers had to address both the first-order and the second-order free rider problem and with high enough punishment levels. Punisher growth occurs if and only if the cost incurred is less than the punishment inflicted—i.e., $\beta/\alpha > 1$ and $\gamma/\eta > 1$. Note these punishment-to-cost ratios are necessary but not sufficient conditions. Recall the punishment inflicted depends on the number of available punishers (see Figure 5). Natural selection will still favor defectors or cooperators if punishers are rare. Replicator equations also do not explain why an individual decided to be a punisher in the first place.

The human experiments conducted by Fehr and Gächter [2] may provide some proximate explanations. Subjects participated in a PGG and afterwards recorded their anger and annoyance at free riders. They found subjects who contributed more were far more likely to be angry at free riders. They were also more likely to retaliate by inflicting punishment. Selp et al. [5] found that just witnessing non-cooperative behaviors provided the "extra fuel" needed for people to engage in costly punishment. Prior work by this author [6] showed that both anger and guilt are present in TOC social dilemmas.

These results regarding the role emotions play in social dilemmas have profound implications. In the past researchers have focused on Nowak's five rules as the genesis of cooperative behavior—e.g., kin selection or reciprocity are essential mechanisms. Those mechanisms may provide insight for pairwise interactions, but not group dynamics. For example, directory reciprocity is the underlying mechanism in the tit-for-tat strategy in a 2-player iterated prisoner's dilemma game. But now consider a 5-player PGG where 2 players cooperate and the other 2 defect. It is not obvious in this case how direct reciprocity helps the 5th player develop an effective strategy for the next round.

Rational individuals weigh the costs and benefits of new strategies before making any change but emotions are the triggering events that cause individuals to reassess their current strategy. Under what circumstances would a cooperator decide to become a punisher? A cooperator who is merely irritated by defectors may decide no strategy change is warranted; an outraged cooperator may decide differently. Emotions may also provide some insight into how high punishment levels are set and what costs punishers are willing to pay. Future efforts should focus on explaining the origins of altruistic punishment rather than mechanisms.

REFERENCES

[1] M. Nowak, "Five rules for the evolution of cooperation," *Science*, vol. 314, no. 5805, pp. 1560–1563, 2006.
[2] E. Fehr and S. Gächter, "Altruistic punishment in humans," *Nature*, vol. 413, pp. 137–140, 2002.
[3] V. Chandrasekhar, Y. Reznik, G. Takacs, D. Chen, S. Tsai, R. Grzeszczuk, and R. Girod, "Quantization schemes for low bitrate compressed histogram of gradients descriptors," in *2010 IEEE Computer Vision and Pattern Recognition Workshops*, 2010, pp. 33–40.
[4] D. Bailiet, L. Mulder, and P. V. Lange, "Reward, punishment and cooperation: a meta-analysis," *Psych. Bull.*, vol. 137, no. 4, pp. 594–615, 2011.
[5] E. Selp, W. van Dijk, and M. Rotteveel, "On hotheads and dirty harries," *Annals N. Y. Acad. Sci.*, vol. 1167, pp. 190–196, 2009.
[6] G. Greenwood, "Evolving strategies to help resolve tragedy of the commons social dilemmas," in *Proc. 2015 IEEE Conf. Comput. Intell. & Games*, 2015, pp. 383–390.