TI 2006-115/1
Tinbergen Institute Discussion Paper

# Evolution of Strategies in Repeated Games with Discounting

*Matthijs van Veelen*

*Universiteit van Amsterdam.*

**Tinbergen Institute**
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

**Tinbergen Institute Amsterdam**
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

**Tinbergen Institute Rotterdam**
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Most TI discussion papers can be downloaded at http://www.tinbergen.nl.

# Evolution of strategies in repeated games with discounting

Matthijs van Veelen
Department of Economis and Econometrics
Universiteit van Amsterdam
Roetersstraat 11
1018 WB Amsterdam
the Netherlands
C.M.vanVeelen@uva.nl

December 18, 2006

**Abstract**

In repeated games there is in general a large set of equilibria. We also know that in the repeated prisoners dilemma there is a profusion of neutrally stable strategies, but no strategy that is evolutionarily stable. This paper investigates whether and how neutrally stable strategies can be upset in a process of mutation and selection. While neutral stability excludes that mutants have a selective advantage themselves, it does not rule out the possibility that mutants that are neutral can enter a population and create a selective advantage for a second mutant. This will be called an indirect invasion and the central results show that, for high enough continuation probability, there is no strategy that is robust against indirect invasions. Such stepping stone paths out of equilibrium generally exist both in the direction of more and in the direction of less cooperation.

**Keywords:** Repeated games, evolutionary stability, robust against indirect invasions.

**JEL Classification Number:** C73

## 1 Introduction: joy in repetition

Repeating a game generally opens up a variety of possibilities for equilibrium behaviour that the one-shot version does not possess. Repeated games therefore have been studied extensively; see for instance Friedman (1971), Aumann & Shapley (1976), Rubinstein (1979, 1980), Fudenberg & Maskin (1986), Abreu (1988) and van Damme (1989). The standard example under study is the prisoners dilemma, where the Nash equilibrium in the one-shot game is not Pareto-efficient and where repetition offers a possible escape from inefficiency. An interesting and natural follow up question is if evolution found an escape route too, and if it did, if it is the same escape route as the one that game theorists found. Again the literature is quite substantial, with for instance Axelrod (1984), Boyd & Lorberbaum (1987), Farrell & Ware (1989), Fudenberg & Maskin (1990), Binmore & Samuelson (1992, 1997), Bendor & Swistak (1995, 1997, 1998) and Lorberbaum, Bohning, Shastri & Sine (2002). The main problem these papers face is that in general there is no evolutionarily stable strategy in repeated games, while evolutionary stability is the main and usually also the most promising tool from the evolutionary game theory toolbox (see Weibull, 1995).

This paper examines how unfortunate that is. Helped by the careful distinctions between different definitions of stability from Bendor & Swistak (1995) and using arguments that are similar (but not identical) to those in Selten & Hammerstein (1984) and Farrell & Ware (1989) we begin with a general theorem concerning the non-existence of a finite mixture of strategies that is evolutionarily stable in the classical definition (Maynard Smith & Price, 1973, Maynard Smith, 1974). One way of dealing with such a negative result is to try out less demanding equilibrium refinements in order to overcome non-existence. Although positive results have been achieved with this approach (see Bendor & Swistak (1995, 1997, 1998) and, in a slightly different setting, Binmore & Samuelson (1992, 1997)), we will argue that there is a fundamental instability of all equilibria in interesting, non-trivial repeated games. In order to show why that is, we define robustness against indirect invasions. This definition rules out indirect invasions throwing an equilibrium off balance and thereby it formally acknowledges the possibility that one at first harmless mutant can serve as a stepping stone, or a springboard, for the invasion of a second mutant. It is shown that for repeated games such stepping stone paths out of equilibrium generally exist both in the direction of more and in the direction of less cooperation - that is, if a higher resp. lower level of cooperation is possible and, for increases in cooperation, if the probability of breakdown is small enough. This indicates that there is no population state that, once it is reached, cannot be overturned by a succession of mutants.

After the main results in Section 3, the dynamics that follow typical indirect invasions will be discussed in Section 4. Stability properties of the concept of robustness against indirect invasions are discussed in Section 5, including application to games other than repeated ones.

# 2 No ESS

The literature concerning evolutionary stability and repeated games can at first sight be a bit confusing. The reason, as Bendor & Swistak (1995) show, is that different authors have used different definitions of evolutionary stability. They also convincingly argue that Maynard Smith's (1974) definition of an evolutionarily stable strategy (ESS), or a weaker version, that Maynard Smith (1982) calls a neutrally stable strategy (NSS) is dynamically much more interesting and meaningful. We will therefore adopt the more standard definition of an evolutionarily stable strategy. Here $\mathcal{S}$ is a space of pure strategies for the repeated game and $\Pi : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ is the payoff function, where $\Pi(S, T)$ is the payoff of a player playing strategy $S$ against a player playing strategy $T$. The payoff of the opponent $T$ in this encounter is given by $\Pi(T, S)$, thereby assuming that the game is symmetric.

The pure strategy version of the definition - **[2]** in Bendor & Swistak (1995) - is as follows.

**Definition 1** ***(Pure ESS)*** *A strategy $S$ is evolutionarily stable if both*
$\Pi(S,S) \geq \Pi(T,S)$ *for all $T$ and*
*if* $\Pi(S,S) = \Pi(T,S)$ *then* $\Pi(S,T) > \Pi(T,T)$ *for all $T \neq S$*

As the standard definition of an ESS also allows for mixed strategies, we would like to do the same. We therefore equate mixed strategies with probability distributions over the strategy space $\mathcal{S}$ to obtain the following definition (Section 3 and Appendix A show how $\mathcal{S}$ can be endowed with a metric to make it a separable metric space).

**Definition 2** ***(Mixed ESS)*** *A strategy $P$ is evolutionarily stable if both*
$\Pi(P,P) \geq \Pi(Q,P)$ *for all $Q$ and*
*if* $\Pi(P,P) = \Pi(Q,P)$ *then* $\Pi(P,Q) > \Pi(Q,Q)$ *for all $Q \neq P$*

Using Definition 1 - the pure strategy definition of an ESS - Selten & Hammerstein (1984) argue that every pure strategy in every non-trivial repeated game has neutral mutants (where a trivial game would be one in which the stage game has a singleton strategy set). The reason is that for every strategy $S$ playing against itself, there is always an off-equilibrium path. On the off-equilibrium path a strategy can be changed without consequences for payoffs. This creates a mutant strategy $T$ for which $\Pi(T,S) = \Pi(S,S) = \Pi(T,T) = \Pi(S,T)$ and hence no strategy $S$ can be ESS.

The following theorem states that finite mixtures of strategies can also not be evolutionary stable. In terms of Definition 2, that is, no strategy $P$ can be ESS if $P$ is a probability distribution which puts probability 1 on a finite number of strategies. The proof is a simple generalization of Selten & Hammerstein's argument; in a finite mixture there is only a finite number of equilibrium paths and hence there is always an infinite number of off-equilibrium paths left on which behaviour can be changed without affecting payoffs. Please note that Farrell & Ware (1989) make the same claim - and prove it - but they use a different definition of evolutionary stability. Furthermore we will focus on games with discouning, but this theorem holds for undiscounted repeated games too.

**Theorem 3** *In a non-trivial repeated game there is no finite mixture of strategies that is evolutionarily stable*

**Proof.** Assume that $P$ is a finite mixture of strategies. Let $P_1, ..., P_n$ denote the composing pure strategies of $P$ and let $p_1, ..., p_n$ with $\sum_{i=1}^{n} p_i = 1$ be the probabilities with which they are played in $P$. It is safe to assume that $P$ is a Nash equilibrium, as being ESS implies being a Nash equilibrium.

There can be no more than $n^2$ paths that are followed by combinations of two pure strategies from this mixture. There is, however, an (uncountably) infinite number of possible paths; if $k$ represents the number of possible actions of each player in the stage game, then there are $k^2$ possible action profiles per repetition, and there is an infinite number of repetitions. (Note that a game is non-trivial if $k > 1$). For every finite mixture of strategies, we can create a new strategy that performs exactly as well as the other strategies in the mixture. Take one of the strategies present in the mixture, say strategy $n$, and mutate it into strategy $n+1$ by only changing its behaviour for a history that does not occur along any of the at most $n^2$ paths followed by duo's of strategies from this mixture interacting. Some such changes could turn it into one of the other $n - 1$ strategies, but there is a (countably) infinite number of possible histories to chose from (see also Section 3) and only a finite number of strategies in the mixture, so there always exists one such mutant that really is a new strategy. This new strategy does not cause any changes; when paired with any of the $n$ strategies both strategies $n$ and $n + 1$ follow the same paths and also the path of $n$ with itself is the same as $n + 1$ with itself. Hence $n + 1$ receives exactly the same payoff as the other strategies from the mixture and we have a mutant that is not driven out. Therefore the finite mixture is not evolutionarily stable. ■

One reaction to a non-existence result like this is to be less demanding. Bendor & Swistak (1995, 1997, 1998) did this and chose to look at strategies that satisfy a weaker condition - **[3]** in their paper. This condition equals Definition 1, but then with all inequalities non-strict. They chose to name strategies that satisfy this relaxed condition evolutionarily stable too, but clarity might be served with following Maynard Smith (1982) and Weibull (1995) in terming such strategies neutrally stable (NSS). In the current paper the definition also includes mixed strategies, as opposed to Bendor & Swistak (1995, 1997, 1998)

**Definition 4** ***(Mixed NSS)*** *A strategy $P$ is neutrally stable if both*

$\Pi(P,P) \geq \Pi(Q,P)$ *for all $Q$ and*

*if* $\Pi(P,P) = \Pi(Q,P)$ *then* $\Pi(P,Q) \geq \Pi(Q,Q)$

While there is no *ESS*, Bendor & Swistak (1995, 1997, 1998) do find a profusion of (pure) *NSS*'es. They also find that nice and retaliatory strategies have larger basins of 'non-repulsion'.[1]

A question one could ask is how much is lost if the demands are lowered from evolutionary to neutral stability. After all, the fact that there is no ESS does by itself not make neutral stability a more stable concept. It therefore seems worth trying to find out exactly how stable or unstable those NSS'es in the repeated prisoners dilemma are. As the only difference between the definitions of an ESS and an NSS is that the latter allows for invasions by neutral mutants, the question then becomes how much harm these neutral mutants can do. If all that happens is that they drift in and out of the population, not being selected against nor being selected for, then the equilibrium could be considered to be stable, especially if the mutants do not change the behaviour on the equilibrium path and actual behaviour in the whole population therefore remains unchanged. But if drift can take the population to a state in which further mutations can arise that actually do have a selective advantage, then the first mutant opens the door for the invasion of a second one. As paths out of equilibrium typically have such a stepping stone structure, we would like to consider whether all strategies can fall victim to a succession of mutants or if there are also NSS'es that are immune to these two-stage invasions. In order to be able to make this distinction, a refinement of the concept of an NSS is suggested. This refinement allows for inconsequential neutral mutants but excludes mutants that open doors for other mutants.

The definition is less concise than for instance the definition of an ESS or an NSS. It will also take some effort to indicate why this definition captures the difference between harmless and harmful mutants. Its most important feature however is that it excludes the following. Let $P$ be a (symmetric) Nash equilibrium. Suppose furthermore that there is an $\alpha \in (0,1)$ and that there are strategies $Q$ and $R$ for which the following holds:

$$\Pi(P,P) = \Pi(Q,P)$$
$$\Pi(P,Q) = \Pi(Q,Q)$$
$$\Pi(R, \alpha P + (1-\alpha) Q) > \Pi(P, \alpha P + (1-\alpha) Q)$$

[1] In a finite automata setting with complexity costs and lexicographic preferences Binmore & Samuelson (1992, 1997) relax the requirement of an ESS to a MESS (see also Swinkels & Samuelson (2003) for a perfectly accurate and meaningful characterization of the different definitions). Their results are also in favour of efficiency.

The first two equations make $Q$ a neutral mutant, as they imply that $\Pi(P, \alpha P + (1-\alpha) Q) = \Pi(Q, \alpha P + (1-\alpha) Q)$ for all $\alpha \in [0,1]$. The third formalises what harm that could do. The neutral mutant $Q$ can by random drift attain a share $\alpha$ in the population for which a second mutant $R$ outperforms both $P$ and $Q$. As this is the way out of equilibrium that the theorems in Section 3 focus on, it is possible for readers to go there directly and still understand how equilibria in repeated games can always be upset.[2] It is however also worthwhile to give a positive (rather than a negative) definition that can properly be placed in between ESS and NSS, that are also positive definitions. Therefore we define robustness against indirect invasions and below we will also indicate why a strategy that is robust against indirect invasions, although not necessarily as stable as an ESS, is fundamentally more stable than an NSS that is not robust against indirect invasions.

**Definition 5** *A strategy $P$ is robust against indirect invasions if*

*1)* $\Pi(P,P) \geq \Pi(Q,P)$ *for all* $Q$ *and*

*2) if* $\Pi(P,P) = \Pi(Q,P)$ *then* $\Pi(P,Q) \geq \Pi(Q,Q)$

*3) if* $\Pi(P,P) = \Pi(Q,P)$ *and* $\Pi(P,Q) = \Pi(Q,Q)$ *then* $\Pi(Q,Q) \geq \Pi(R,Q)$ *for all* $R$ *and*

*4) if* $\Pi(P,P) = \Pi(Q,P)$ *and* $\Pi(P,Q) = \Pi(Q,Q) = \Pi(R,Q)$ *then* $\Pi(Q,R) \geq \Pi(R,R)$

It is not too hard to see that this definition indeed excludes the path out of equilibrium that is described above. Conditions *1)* and *3)* imply that no strategy $R$ can ever do better against any mix of $P$ and a neutral mutant $Q$ than $P$ and $Q$ themselves. Taking $R$ for $Q$ in *1)* we know that $\Pi(P,P) \geq \Pi(R,P)$. Together with *3)* this directly gives that $\Pi(R, \alpha P + (1-\alpha) Q) = \alpha \Pi(R,P) + (1-\alpha) \Pi(R,Q) \leq \alpha \Pi(P,P) + (1-\alpha) \Pi(P,Q) = \Pi(P, \alpha P + (1-\alpha) Q)$ for all $\alpha \in [0,1]$. We also know that the first two equalities in condition *3)* imply that $\Pi(P, \alpha P + (1-\alpha) Q) = \Pi(Q, \alpha P + (1-\alpha) Q)$ for all $\alpha \in [0,1]$ and therefore both $P$ and $Q$ do at least as good as $R$ against any mix of the first two.

While condition *2)* rules out Nash equilibria that can be invaded by best responses that outperform $P$ against themselves, condition *4)* does the same for indirect invasions. Showing this is perhaps slightly tedious, but other than that also straightforward and insightful. If there is an $\alpha \in [0,1]$ for which $\Pi(R, \alpha P + (1-\alpha) Q) = \Pi(P, \alpha P + (1-\alpha) Q) = \Pi(Q, \alpha P + (1-\alpha) Q)$, then conditions *2)* and *4)* imply that the mix $\alpha P + (1-\alpha) Q$ performs at least as good against $R$ as $R$ does against itself. In order to see why, it is worth noting that equality of $\Pi(R, \alpha P + (1-\alpha) Q)$ and $\Pi(P, \alpha P + (1-\alpha) Q)$ then either holds

[2] Please note that two mutants feature also in Boyd & Lorberbaum (1987), Lorberbaum (1994) and Lorberbaum, Bohning, Shastri & Sine (2002). The difference is that they assume that some mutations, although having a selective disadvantage, will always be present in the population in some proportion. We however look at *neutral* (first) mutants, that is, mutants that do not have a selective disadvantage. Their share in the population can therefore in- or decrease by random drift.

for all $\alpha \in [0,1]$, or only for $\alpha = 0$, or only for $\alpha = 1$. We will deal with these cases separately.

If all three strategies perform equally well against both $P$ and $Q$ - that is: $\Pi(P,P) = \Pi(Q,P) = \Pi(R,P)$ and $\Pi(P,Q) = \Pi(Q,Q) = \Pi(R,Q)$ - then $\Pi(R, \alpha P + (1-\alpha) Q) = \Pi(P, \alpha P + (1-\alpha) Q)$ holds for all $\alpha \in [0,1]$. In this case, we know from taking $R$ for $Q$ in *2)* that $\Pi(P,R) \geq \Pi(R,R)$. From *4)* we also know that $\Pi(Q,R) \geq \Pi(R,R)$. Thus $\Pi(\alpha P + (1-\alpha) Q, R) = \alpha \Pi(P,R) + (1-\alpha) \Pi(Q,R) \geq \alpha \Pi(R,R) + (1-\alpha) \Pi(R,R) = \Pi(R,R)$.

If all three strategies perform equally well against $Q$, but not against $P$, that is: $\Pi(P,P) = \Pi(Q,P) > \Pi(R,P)$ and $\Pi(P,Q) = \Pi(Q,Q) = \Pi(R,Q)$, then $\Pi(R, \alpha P + (1-\alpha) Q) = \Pi(P, \alpha P + (1-\alpha) Q)$ holds for $\alpha = 0$ only. But then the 'mix' consists purely of $Q$, and from *4)* we know that $\Pi(Q,R) \geq \Pi(R,R)$.

If all three strategies perform equally well against $P$, but not against $Q$, that is: $\Pi(P,P) = \Pi(Q,P) = \Pi(R,P)$ and $\Pi(P,Q) = \Pi(Q,Q) > \Pi(R,Q)$, then $\Pi(R, \alpha P + (1-\alpha) Q) = \Pi(P, \alpha P + (1-\alpha) Q)$ only holds for $\alpha = 1$. Here the 'mix' consists purely of $P$, and from *2)* we know that $\Pi(P,R) \geq \Pi(R,R)$.

In Section 5 we will return to this definition, explore the scope for further nesting, and see how it relates to other concepts such as for instance the ES set. Here we only observe that there are some inclusions that follow directly from the definitions. If $\Delta^{ESS}$ is the set of ESS'es, $\Delta^{NSS}$ is the set of NSS'es, $\Delta^{NE}$ is the set of Nash equilibria and $\Delta^{RII}$ is the set of equilibria that are robust against indirect invasions, then

$$\Delta^{ESS} \subset \Delta^{RII} \subset \Delta^{NSS} \subset \Delta^{NE}$$

For non-trivial repeated games we know that $\Delta^{ESS}$ does not contain finite mixtures (Theorem 3) and that $\Delta^{NSS}$ is a very rich set (Bendor & Swistak, 1995, 1997, 1998). Below we will show however that for sufficiently low probability of breakdown also $\Delta^{RII}$ does not contain finite mixtures, which means that all finite NSS'es are vulnerable to indirect invasions. If a strategy is a neutrally stable, but not robust against indirect invasions, we will sometimes also refer to it as indirectly invadable.

# 3 Stepping stones in either direction

This section contains two theorems that state conditions under which strategies are not robust against indirect invasions. The first one shows that any positive level of cooperation can be undermined by a succession of two mutations. The second states that if there are possible gains from (increased) cooperation and the probability of continuation is sufficiently high, then a stepping stone route into more cooperation exists. Together they imply that no equilibrium in interesting repeated games with low enough probability of breakdown is robust against indirect invasions, and mostly there are ways out of equilibrium in the direction of in- as well as in the direction of decreasing cooperation. Both theorems come in a pure strategy version for expositional clarity and a mixed version for generality.

We start with a few formal definitions. Consider a symmetric one-shot 2-player game $g$ characterized by a set of players $I = \{1, 2\}$, an action space $A$, equal for both players, and a payoff function $\pi : A \times A \to \mathbb{R}^2$. Using a discount factor $\delta$, interpreted as a continuation probability, this one-shot game is turned into a repeated one, which will be called $\Gamma(\delta)$. A history at time $t$ is a list of the actions played up to and including time $t-1$, where an empty pair of brackets is used to denote the history 'no history'. If $a_{t,i}$ is the action played by player $i$ at time $t$, then these histories are:

$$h_1 = ()$$
$$h_t = ((a_{1,1}, a_{1,2}), ..., (a_{t-1,1}, a_{t-1,2})), \qquad t = 2, 3, ...$$

Sometimes we will also write $(h_t, (a_{t,1}, a_{t,2}))$ for a history $h_{t+1}$. The set of possible histories at time $t$ is:

$$H_1 = \{h_1\}$$
$$H_t = \prod_{i=1}^{t-1} (A \times A) \qquad t = 2, 3, ...$$

and the set of all possible histories is:

$$H = \bigcup_{t=1}^{\infty} H_t.$$

It will furthermore be useful to have a way of writing down a history with the roles of the players reversed. Given a history $h_t$ as they are defined above, its mirror image $h_t^{\leftarrow}$ is found by simply renumbering the players:

$$h_1^{\leftarrow} = ()$$
$$h_t^{\leftarrow} = ((a_{1,2}, a_{1,1}), ..., (a_{t-1,2}, a_{t-1,1})), \qquad t = 2, 3, ...$$

The reason why histories with roles reversed are needed, is that we assume that both players label themselves as player 1 and the other as player 2 and therefore face mirrored histories as they go along.

A strategy is a function that maps histories to the action space: $S : H \to A$. For two strategies, say $S$ and $T$, the course of actions is determined by recursion; all actions at all stages are determined by the initiation

$$h_1^{S,T} = ()$$

and the recursion step

$$
\begin{aligned}
a_t^{S,T} &= \left(S\left(h_t^{S,T}\right), T\left(h_t^{S,T\leftarrow}\right)\right) \\
h_{t+1}^{S,T} &= \left(h_t, a_t^{S,T}\right), \qquad\qquad t = 1, 2, ...
\end{aligned}
$$

The discounted normalised payoffs to (a player that uses) strategy $S$ against strategy $T$ is given by:

$$
\Pi(S,T) = (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\pi_1\left(a_t^{S,T}\right)
$$

With these definitions, we can prove the first theorem. Note that strategies here are pure, and that we write that $S$ is an equilibrium strategy, which is short for $(S,S)$ being a symmetric equilibrium of the game $\Gamma(\delta)$.

**Theorem 6** *Let $S$ be a strategy in the game $\Gamma(\delta)$ and let there be a time $\tau$ at which $a_\tau^{S,S}$ is not an equilibrium of the stage game. Then $S$ is not robust against indirect invasions.*

**Proof.** Assume that condition *1)* of Definition 5 is satisfied and $S$ is an equilibrium. Let $T$ be the strategy that equals $S$ for all histories, except for those that are elements of the set $\widehat{H} = \{h_t \mid t > \tau,\ a_{\tau,2} = \arg\max_{a\in A}\pi_2\left(S\left(h_\tau^{S,S}\right), a\right)\}$. These histories only occur off the equilibrium path, since it is assumed that players playing $S$ against each other do not play an equilibrium of the stage game at time $\tau$. For those histories $h_t \in \widehat{H}$ we take $T(h_t) = S\left(h_t^{S,S}\right)$. Obviously, the paths of $T$ against $S$, $T$ against $T$, $S$ against $S$ and $S$ against $T$ are all the same; $h_t^{T,S} = h_t^{T,T} = h_t^{S,S} = h_t^{S,T}\ \forall\, t$. Consequently the corresponding payoffs are also equal; $\Pi(T,S) = \Pi(S,S) = \Pi(T,T) = \Pi(S,T)$.

Now let $U$ be the strategy that equals $S$, except for $h_\tau^{S,S}$, for which we take $U\left(h_\tau^{S,S}\right) = \arg\max_{a\in A}\pi_1\left(a, S\left(h_\tau^{S,S}\right)\right)$ and except for histories that are elements of the set $\widetilde{H} = \{h_t \mid t > \tau,\ a_{\tau,1} = \arg\max_{a\in A}\pi_1\left(a, S\left(h_\tau^{S,S}\right)\right)\}$, for which we take $U(h_t) = S\left(h_t^{S,S}\right), h_t \in \widetilde{H}$.

It is obvious that $\Pi(U,S) \leq \Pi(S,S)$, for $S$ is an equilibrium, and it is also clear that $\Pi(U,T) > \Pi(T,T) = \Pi(S,T)$, because $U$ improves itself at time $\tau$ without being punished by $T$. As a result of this, requirement *3)* of Definition 5 is not satisfied.

Note that $\Pi(U,T) > \Pi(T,T) = \Pi(S,S) \geq \Pi(U,S)$, and therefore that $T \neq S$. In other words, if $T = S$, then $U$ does strictly better against $S$ than $S$ itself and that contradicts $S$ being an equilibrium. ■

What this theorem indicates is that as soon as there are equilibrium actions that must be upheld by the threat of punishment, then there can be mutants that do not punish, and subsequently there can be other mutants that takes advantage of the first mutant not punishing. One thing worth noting is that the proof constructs only one way out of equilibrium. While this particular stepping stone path changes behaviour for histories that are elements of rather moderate sets $\widehat{H}$ and $\widetilde{H}$, other ways out of equilibrium may come with changes on larger, and sometimes also more natural sets of histories, as we will see later in Section 4. But what the theorem shows is that if there is cooperation in equilibrium, at least the existence of an indirect way out is guaranteed.

While the reference point in Theorem 6 is the equilibrium of the one-shot game, we will now focus on departures from what in non-trivial games is the other extreme: the maximally feasible symmetric payoffs. Therefore we define $\pi_{\max} = \max_{a \in A} \pi_1 (a, a)$ and $a_{\max} = \arg\max_{a \in A} \pi_1 (a, a)$. Note that $a_{\max}$ is an action, while $a_\tau^{S,S} = \left(S\left(h_\tau^{S,S}\right), S\left(h_\tau^{S,S}\right)\right)$ is an action profile. The following theorem states that if there is a point in the course of play of an equilibrium strategy at which unilaterally initiating cooperation could be offset by future gains from (increased) cooperation, then the strategy is not robust against indirect invasions.

**Theorem 7** *Let $S$ be a strategy in the game $\Gamma(\delta)$ and let there be a time $\tau$, for which the following holds:*

1. $\pi_1\left(a_\tau^{S,S}\right) - \pi_1\left(a_{\max}, S\left(h_\tau^{S,S}\right)\right) < \sum_{t=\tau+1}^{\infty} \delta^{t-\tau}\left(\pi_{\max} - \pi_1\left(a_t^{S,S}\right)\right)$.
2. $a_{\max} \neq S\left(h_\tau^{S,S}\right)$

*Then $S$ is not robust against indirect invasions.*

**Proof.** Assume that condition *1)* of Definition 5 is satisfied and $S$ is an equilibrium. Let $T$ be the strategy that equals $S$ for all histories, except for those that are elements of the set $\widehat{H} = \{h_t \mid t > \tau,\ a_{u,2} = a_{\max}, u \geq \tau\}$. These histories only occur off the equilibrium path, as it is assumed that $a_{\max} \neq S\left(h_\tau^{S,S}\right)$. For those histories $h_t \in \widehat{H}$ we take $T(h_t) = a_{\max}$. Obviously, the paths of $T$ against $S$, $T$ against $T$, $S$ against $S$ and $S$ against $T$ are all the same; $h_t^{T,S} = h_t^{T,T} = h_t^{S,S} = h_t^{S,T}\ \forall\ t$. Consequently the corresponding payoffs are also equal; $\Pi(T,S) = \Pi(S,S) = \Pi(T,T) = \Pi(S,T)$.

Now let $U$ be the strategy that equals $S$, except for the history $h_\tau^{S,S}$, for which we choose $U\left(h_\tau^{S,S}\right) = a_{\max}$ and except for the histories that are elements of the set $\widetilde{H} = \{h_t \mid t > \tau,\ a_{\tau,1} = a_{\max} \text{ and } a_{u,2} = a_{\max}, u > \tau\}$, for which we take $U(h_t) = a_{\max}, h_t \in \widetilde{H}$.

It is obvious that $\Pi(U,S) \leq \Pi(S,S)$, for $S$ is an equilibrium, and it is also clear that $\Pi(U,T) > \Pi(T,T) = \Pi(S,T)$, because that follows directly from the first requirement of the theorem. As a result of this, requirement *3)* of Definition 5 is not satisfied.

As in the proof of Theorem 6, $S$ being an equilibrium implies that $T \neq S$. ■

The requirements in this theorem are slightly less simple to check for than those in Theorem 6, but when translated to prisoners dilemma's, it turns out to imply something that is relatively easy to handle. Before doing so, however, it good to realize that discounted, normalised payoffs that belong to a combination of two strategies can vary with $\delta$ and that they do so in different ways. If we look at symmetric equilibria, then it might be that two different strategies that both play against themselves have the same discounted, normalised payoff for a given $\delta$, while a higher $\delta$ increases them for one and decreases them for the other strategy.

The next theorem states that for repeated prisoners dilemma's, all symmetric equilibria with payoffs less than $\pi_1(C,C) - (1-\delta)[\pi_1(C,C) - \pi_1(C,D)]$ are not robust against indirect invasions. If we take more or less standard values, that is $\pi_1(D,C) = 4$, $\pi_1(C,C) = 3$, $\pi_1(D,D) = 1$, $\pi_1(C,D) = 0$, then this amounts to $3\delta$; all strategies $S$ with with payoffs $\Pi(S,S)$ less then $3\delta$ are indirectly invadable. There may be many other equilibria that are also not robust against indirect invasions, but Theorem 8 shows that at least all strategies with relatively low payoffs satisfy the criteria for Theorem 7. It also means that the closer $\delta$ gets to 1, the more strategies are shown to be vulnerable to indirect invasions with increasing cooperation, and for any strategy $S$ with payoff $\Pi(S,S) < \pi_1(C,C)$ there is a $\overline{\delta} \in (0,1)$ such that $S$ indirectly invadable for all $\delta \in (\overline{\delta}, 1)$. Together with Theorem 6 that implies that for sufficiently high $\delta$ no symmetric equilibrium strategy is robust against indirect invasions.

**Theorem 8** *In a repeated prisoner's dilemma, all strategies $S$ with $\Pi(S,S) < \pi_1(C,C) - (1-\delta)[\pi_1(C,C) - \pi_1(C,D)]$ are not robust against indirect invasions.*

**Proof.** First realize that $S$ is $\pi_1(C,C) - \Pi(S,S)$ short from full, symmetric efficiency. Then choose as time $\tau$ in Theorem 7 the first period that $S$ plays defect. The second requirement of the same theorem is then automatically fulfilled.

The following can then be derived

$$\Pi(S,S) < \pi_1(C,C) - (1-\delta)(\pi_1(C,C) - \pi_1(C,D)) \Rightarrow$$

$$\Pi(S,S) < \pi_1(C,C) - (1-\delta)\delta^\tau(\pi_1(C,C) - \pi_1(C,D)) \Leftrightarrow$$

$$\pi_1(D,D) - \pi_1(C,D) < \tfrac{1}{(1-\delta)\delta^\tau}[\pi_1(C,C) - \Pi(S,S)] - [\pi_1(C,C) - \pi_1(D,D)] \Leftrightarrow$$

$$\pi_1\left(a_\tau^{S,S}\right) - \pi_1\left(a_{\max}, S\left(h_{\tau-1}^{S,S}\right)\right) < \sum_{t=\tau}^{\infty} \delta^{t-\tau}\left(\pi_{\max} - \pi_1\left(a_t^{S,S}\right)\right) - \left(\pi_{\max} - \pi_1\left(a_\tau^{S,S}\right)\right)$$

This satisfies the first requirement of Theorem 7. ■

Again, the proof of Theorem 7 only gives one stepping stone route out of equilibrium, but there may be lots of ways in which successive mutants can throw an equilibrium off balance with an increasing level of cooperation.

## 3.1 Mixed strategies

In evolutionary as well as in standard game theory, equilibrium concepts usually allow for mixed strategies. While the standard setting of symmetric 2-person bi-matrix games (see Weibull, 1995) naturally comes with definitions in terms of mixed strategies, the literature on repeated games is much more focussed on pure equilibria (with exceptions such as for instance Binmore & Samuelson, 1992, and Samuelson & Swinkels, 2003). It seems however no less natural to include mixed strategies here too, especially since the paths out of equilibrium at least at first lead away from pure strategies (or homogeneous populations) and into mixtures of strategies. While Theorem 3 shows that there is also no mixed ESS, Theorems 6 and 7 do not yet exclude the possibility that there is a mixture of strategies that is robust against indirect invasions. In this subsection we therefore give the equivalents of those theorems for finite mixtures. Here we will directly focus on repeated prisoners dilemma's rather than repeated games in general. This will keep notation simpler, it hopefully helps the intuition and still captures the essentials. Also, $\Pi(D, D)$ will be used to denote $(1-\delta)\sum_{t=0}^{\infty}\delta^{t-1}\pi_1(D, D) = \pi_1(D, D)$, which is the normalised discounted payoff of *AllD* against *AllD*.

**Theorem 9**

*Let $P$ be a finite mixture of strategies in $\Gamma(\delta)$.*

*If $\Pi(P, P) > \Pi(D, D)$ then $P$ is not robust against indirect invasions.*

**Proof.** See Appendix B ■

As with the pure strategy version, the proof in the appendix just constructs one particular way out of equilibrium, while there may be many other stepping stone paths, some of which can be considered to be more likely than others. But the theorem shows that indirect invasions are always possible for equilibria with cooperation.

In order to formulate the mixed strategy counterpart for increasing cooperation, it will be helpful to define the following. Let $P_1, ..., P_n$ be the composing pure strategies of $P$ and let $p_1, ..., p_n$, with $\sum_{i=1}^{n} p_i = 1$, be the probabilities with which they are played in $P$. For any defection that occurs along a path of interaction between any two strategies $P_i$ and $P_j$ from $P$ we can discount the possible gains in the future and compare it to the current period loss of switching from $D$ to $C$ as an initiation of cooperation. Therefore we first define $E_i(j)_t = \left\{P_l \mid h_t^{P_i,P_l} = h_t^{P_i,P_j}\right\}$, which makes it the set of strategies against which the history of $P_i$ at time $t$ is the same as against $P_j$. Since we assume that $P$ is a finite mixture, we know that $\lim_{t\to\infty} E_i(j)_t = E_i(j)$, where $E_i(j)$ is defined (see also the proof of Theorem 9) as $E_i(j) = \left\{P_l \mid a_t^{P_i,P_l} = a_t^{P_i,P_j}\ \forall\, t\right\}$. For any combination of strategies $(P_i, P_j)$ and any time $t$ we can compute $\delta_{ij,t}$ as follows:

$$\delta_{ij,t} = \begin{cases} \delta \mid \sum\limits_{P_l \in E_i(j)_t} p_l \left(\pi_1\left(a_t^{P_i,P_l}\right) - \pi_1\left(C, a_{t,2}^{P_i,P_l}\right)\right) = \sum\limits_{P_l \in E_i(j)_t} p_l \sum\limits_{u=t+1}^{\infty} \delta^{u-t}\left(\pi_1(C,C) - \pi_1\left(a_u^{P_i,P_l}\right)\right) & \\ & \text{if } a_{t,1}^{P_i,P_j} = D \text{ and the equation has a solution } \delta \in (0,1) \\ 1 & \text{otherwise} \end{cases}$$

This definition greatly simplifies the formulation of the next theorem.

**Theorem 10**

*Let $P$ be a finite mixture of strategies in $\Gamma(\delta)$.*

*If $\min_{i,j,t} \delta_{ij,t} < \delta \leq 1$ then $P$ is not robust against indirect invasions.*

**Proof.** See Appendix B.1 ■

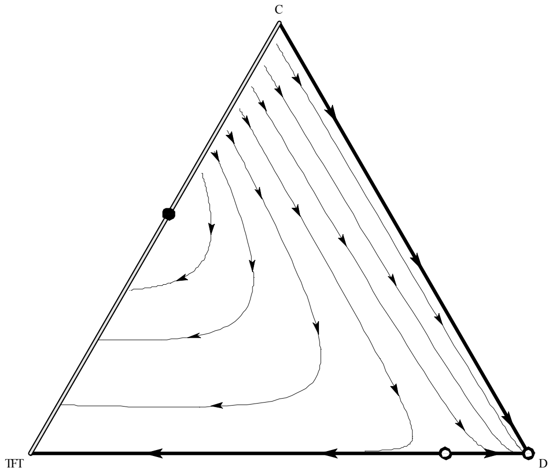
# 4 Out of equilibrium

In Section 3 it was already mentioned that the proofs only provide one stepping stone path out of equilibrium. For most equilibria however there are many ways out, some of which can be considered to be more natural than the ones provided. In this section we will consider two examples, with simple strategies that have received some attention in the literature. We will also look at what happens after the indirect invasion.

For the prisoners dilemma we use the same payoffs as in the previous section: $\pi_1(D,C) = 4$, $\pi_1(C,C) = 3$, $\pi_1(D,D) = 1$, $\pi_1(C,D) = 0$. For the probability of continuation we choose $\delta = 0.75$. The first strategy we look at is *Tit-for-tat*, where *AllC* is the first and *AllD* is the second mutant. The payoffs of these strategies against themselves and each other are easily computed (see Appendix C) and summarized in the following payoff matrix.

| | $TFT$ | $C$ | $D$ |
|---|---|---|---|
| $TFT$ | 3 | 3 | 0.75 |
| $C$ | 3 | 3 | 0 |
| $D$ | 1.75 | 4 | 1 |

All points on the boundary face between $TFT$ and $C$ are obviously rest points of the dynamics. We can compute the payoff of $D$ against a mixture of the first two - $\Pi(D, \alpha TFT + (1-\alpha)C)$ - and compare it to the payoff of $TFT$ and $C$ against the same mixture. They are equal for $\alpha = \frac{4}{9}$, so for $\alpha > \frac{4}{9}$ the population will directly be pushed back onto the boundary face after an invasion of $D$. Then there is a range of $\alpha$'s smaller than $\frac{4}{9}$, but larger than $\frac{1}{6}$, for which the trajectory after invasions of $D$ first moves away from the boundary face, and then takes a detour to return to the boundary face, but then at the other side of $\alpha = \frac{4}{9}$. Then for the last part of the boundary face, with $0 \leq \alpha < \frac{1}{6}$, invasions of $D$ result in convergence towards the vertex where all play $D$.




If the strategy space would be restricted to those three strategies only, $D$ would be the only ESS. Although on the larger part of the boundary face between $TFT$ and $C$, invasions of $D$ are (eventually) driven out - reducing the share of the first mutant in the process - there is an $\overline{\alpha} = \frac{1}{6}$ such that if the share of $TFT$ falls below it, mutants playing $D$ will not just successfully invade, but also take over.

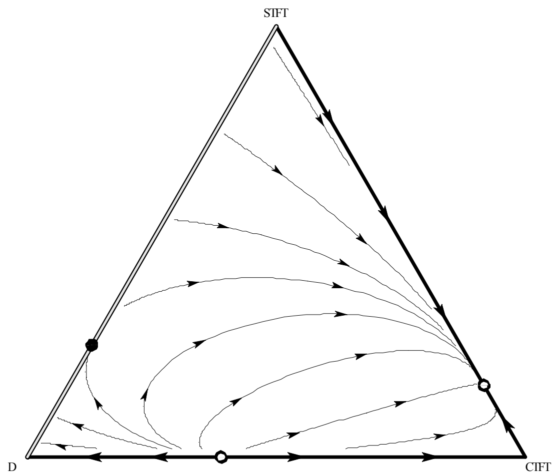
The second example is *AllD*, with *Suspicious Tit-for-tat* (a.k.a. tat-for-tit) as the neutral first, and *Cooperate-Tit-for-tat* as the second mutant. $STFT$ differs from tit-for-tat in that it starts with playing $D$. Afterwards, it also copies the previous move of the other player. $CTFT$ starts with playing $C$ in periods 1 and 2, and then copies the previous move of the other player.[3] The payoffs of these strategies against themselves and each other are again computed in Appendix C and summarized in a payoff matrix.

| | $D$ | $STFT$ | $CTFT$ |
|---|---|---|---|
| $D$ | 1 | 1 | $2\frac{5}{16}$ |
| $STFT$ | 1 | 1 | $3\frac{1}{4}$ |
| $CTFT$ | $\frac{9}{16}$ | $2\frac{1}{4}$ | 3 |

All points on the boundary face between $D$ and $STFT$ are rest points. We can compute the payoff of $CTFT$ against a mixture of the first two - $\Pi\left(CTFT, \alpha D + (1-\alpha)\, STFT\right)$ -

[3] Results would not change if we took $TF2T$ (tit-for-two-tats, see Bendor & Swistak, 1995) to replace $CTFT$.

and compare it to the payoff of $TFT$ and $C$ against the same mixture. They are equal for $\alpha = \frac{20}{27}$, so for $\alpha > \frac{20}{27}$ the population will directly be pushed back to the boundary face after an invasion of $CTFT$. For $\alpha$'s smaller than $\frac{20}{27}$ invasions of $CTFT$ result in convergence towards the mixed equilibrium where $\frac{1}{6}$ plays $STFT$ and $\frac{7}{6}$ plays $CTFT$.




If the strategy space would be restricted to those three strategies only, the mixed equilibrium $\left(0, \frac{1}{6}, \frac{7}{6}\right)$ would be the only ESS, and for the larger part of the boundary face between $D$ and $STFT$, invasions of $CTFT$ lead to this equilibrium.

Although they are first of all examples of natural stepping stone paths out of an NSS, the two phase plots are also representative for the indirect invasions that are constructed in the proofs of Theorem 9 and 10, or their pure strategy equivalents 6 and 7. This can be seen if we consider the typical payoff matrix that represents an indirect invasion in a strategy $P$ that is an NSS.

$$\begin{array}{cccc} & P & Q & R \\ P & a & a & d \\ Q & a & a & e \\ R & b & c & f \end{array}$$

The block of $a$'s reflects that $P$ and its neutral mutant $Q$ earn the same payoff when they play themselves and each other; $\Pi(P,P) = \Pi(Q,P) = \Pi(P,Q) = \Pi(Q,Q)$. In both proofs the sequence of mutations given violates condition *3)* of the definition of robustness against indirect invasions. This means that $c > a$ and hence also that $b < a$. (The strictness of the second inequality follows from $P$ being NSS and not just a Nash equilibrium; if $c > a$ and $b = a$, then $P$ would not be an NSS). The shape of the dynamics in the three-dimensional simplex therefore depends on how $d, e$ and $f$ are ordered.

Inspection of the proof of Theorem 9 or 6 shows that the two mutants are constructed such that $f > d > e$. (It is also true there that $a > f$, but the replicator dynamics are invariant with respect to adding or substracting constant to or from the rows, so that is not informative about the behaviour of the dynamics). Because $c > a$ and $f > e$, it follows that on the whole boundary face between $Q$ and $R$ the dynamics go in the direction of $R$. Because $b < a$ and $f > d$ the boundary face between $P$ and $R$ has an unstable rest point, with dynamics on either side going towards the pure extremes. Furthermore strategy $P$ weakly dominates $Q$, because $d > e$.

In the proof of Theorem 10 and 7 the mutants imply payoffs for which $e > f > d$. (Here $d > a$, but again that does not affect the dynamics). Because $c > a$ and $e > f$, the boundary face between $Q$ and $R$ has a mixed rest point, and on this boundary face the dynamics point towards it. Because $b < a$ and $f > d$ the boundary face between $P$ and $R$ again has an unstable rest point, with dynamics on either sides going towards the pure extremes. Furthermore $e > d$, implying that now $Q$ weakly dominates $P$.

## 5 Indirect invasions and ES sets

The definition of robustness against indirect invasions is designed to single out equilibria that are not susceptible to (two-stage) indirect invasions. It seems natural to also consider whether or not we should include three-stage or higher order indirect invasions in this or perhaps in another definition. While conditions *1)* and *2)* of Definition 5 imply that no mutant has a selective advantage over a strategy $P$ that is robust against indirect invasions, *3)* and *4)* imply that no mutant has a selective advantage over any mixture of $P$ and any neutral mutant $Q$. One could therefore wonder whether there is scope for further nesting, as it is not excluded that there are mutants $R$ that themselves are neutral for all mixtures of $P$ and a neutral mutant $Q$. A further mutant $S$ might then have a selective advantage over a mixture of $P, Q$ and $R$.

As the construction of indirect invasions in the proofs of Theorem 9 and 10 already indicates, almost all of the higher order indirect invasions can be rewritten as two-stage indirect invasions. This will be shown formally in Theorem 11. Situations to which that theorem does not apply are those where neutral mutations first have to completely replace the incumbent strategy in order to allow for other neutral mutants. Although perhaps not the most likely possibility, such paths are not excluded by the definition of robustness against indirect invasions. We will illustrate that with a simple game and show that the concept of robustness against indirect invasions nonetheless fits in a sequence of definitions with increasing stability properties.

The scenario with higher order invasions mentioned above is that a second mutant $R$ is neutral for a mixture of $P$ and a neutral mutant $Q$. If $P$ is robust against indirect invasions, and $R$ is neutral for $\alpha P + (1-\alpha) Q$ with $\alpha \in (0,1)$, then it must be also be neutral for all $\alpha$. Condition (*) in the following theorem captures exactly that, and drift can then take the population to any convex combination of $P, Q$ and $R$. The case where $R$ is only neutral for $\alpha P + (1-\alpha) Q$ if $\alpha = 0$ is a different one, and will be discussed right after this one.

**Theorem 11**

*If a strategy $P$ is robust against indirect invasions and*

$$\begin{aligned} &\Pi(P,P) = \Pi(Q,P) = \Pi(R,P), \\ &\Pi(P,Q) = \Pi(Q,Q) = \Pi(R,Q) \text{ and } \qquad (*) \\ &\Pi(P,R) = \Pi(Q,R) = \Pi(R,R). \end{aligned}$$

*then for all $\alpha, \beta$, $0 \le \alpha \le 1$, $0 \le \beta \le 1$, $0 \le \alpha + \beta \le 1$ and all strategies $S$*

*5)* $\Pi(P, \alpha P + \beta Q + (1-\alpha-\beta) R) \ge \Pi(S, \alpha P + \beta Q + (1-\alpha-\beta) R)$ *and*

*6) if* $\Pi(P, \alpha P + \beta Q + (1-\alpha-\beta) R) = \Pi(S, \alpha P + \beta Q + (1-\alpha-\beta) R)$ *then* $\Pi(\alpha P + \beta Q + (1-\alpha-\beta) R, S) \ge \Pi(S,S)$

**Proof.** See Appendix B.2 ■

The following example illustrates what is not excluded by the definition of robustness against indirect invasions. Consider the payoff matrix below, where the letters now indicate pure strategies:

$$\begin{array}{ccccc} & S & T & U & V \\ S & 1 & 1 & 1 & 1 \\ T & 1 & 1 & 1 & 1 \\ U & 0 & 1 & 1 & 1 \\ V & 0 & 0 & 1 & 2 \end{array} \tag{1}$$

In this game there is a number of different equilibria. First $V$ is an ESS. Then $U$ is a Nash equilibrium, but not an NSS. Furthermore $T$ is an NSS, but not robust against indirect invasions. Finally $S$ is robust against indirect invasions, but not ESS.

Looking at the replicator dynamics for this game, we observe that there is a path along which drift and a sequence of neutral mutations can take a population from $S$ to $V$. If first the neutral mutant $T$ arises, and drift drives $S$ extinct, and second the mutant $U$ that is neutral for $T$ arises, and drift drives $T$ extinct, then $V$ can successfully invade and take over the population. This possibility is not excluded by robustness against indirect invasions. It should however be noted that $U$ is not neutral for $S$. Hence a reintroduction of $S$ anywhere between $T$ and $U$ throws the population back onto the boundary face between $S$ and $T$, because $U$ has a selective disadvantage anywhere but at the boundary face between $T$ and $U$ itself. It is therefore reasonable to think that $S$ is more stable than $T$ and $T$ is more stable than $U$ (while $V$ off course is the most stable). It is also worth noting that if we would make $U$ neutral for both $S$ and $T$ by changing $\Pi(U, S)$ from 0 to 1, then $S$ is no longer robust against indirect invasions.

Generalisations of this game are also possible, if we define a game with strategies $S_1$ to $S_n$ as follows:

$$\Pi(S_i, S_j) = \begin{cases} 1 & \text{if } i \leq j+1 \text{ and } i \neq n \\ 0 & \text{if } i > j+1 \\ 2 & i = j = n \end{cases}$$

If we take $n \geq 4$, then $S_n$ is an ESS, $S_{n-1}$ is a Nash equilibrium, but not an NSS, $S_{n-2}$ is an NSS, but not robust against indirect invasions, and $S_1, ..., S_{n-3}$ are all robust against indirect invasions, but not ESS. Still it is clear that the higher $n$, the more stable $S_1$ and its neutral mutant $S_2$ - or the boundary face between $S_1$ and $S_2$ - is.

Robustness against indirect invasions puts restrictions on how (third) mutants perform against a set of strategies, namely its neutral mutants. It can therefore, for finite strategy spaces, be useful to relate that to an ES set, which is a setwise generalisation of the ESS concept (Thomas, 1985, and see also Weibull, 1995, p51 & p105 for the related concept of an ES* set). An example from Weibull (1995, p106-108) shows that a set of equilibria that is asymptotically stable need not be an ES set. All the equilibria in that example are also robust against indirect invasions, and the remainder of the set of equilibria are all neutral mutants, so the same game can also be used to show that a strategy that is robust against indirect invasions, together with its neutral mutants, need not be an ES set. Here we give another game for which the same holds.

$$\begin{array}{cccc} & S & T & U \\ S & 1 & 1 & 1 \\ T & 1 & 1 & 0 \\ U & 1 & 0 & 1 \end{array} \tag{2}$$

The set of $S$ and all of its neutral mutants - the faces between $S$ and $T$ and between $S$ and $U$ - is asymptotically stable, but not an ES set. Please note that if $Q$ is a neutral mutant of $P$, then so is $\alpha Q + (1-\alpha) P$ for all $\alpha \in (0, 1)$. The example however illustrates that if $Q$ and $R$ are neutral mutants of $P$, then $\alpha Q + (1-\alpha) R$ for $\alpha \in (0, 1)$ need not be one too.

**Theorem 12** *Suppose $\mathcal{S}$ is a finite set. If $X$ is an ES set, and $P \in X$, then $P$ is robust against indirect invasions*

**Proof.** From the observation in Weibull (1995, p51) that $X$ must be a subset of $\Delta^{NSS}$ it follows that requirements *1)* and *2)* are satisfied. Furthermore, if $Q$ is a neutral mutant of $P$ - that is: $\Pi(P,P) = \Pi(Q,P)$ and $\Pi(P,Q) = \Pi(Q,Q)$ - then all strategies $Q_\alpha = \alpha Q + (1-\alpha) P$ for $\alpha \in [0,1]$ must also be elements of $X$; if not, then by closedness of ES sets, there is a $Q_\alpha \in X$ for which every neighbourhood $U$ of $Q_\alpha$ contains a strategy $Q_\beta \notin X$, for which $\Pi(Q_\alpha, Q_\beta) = \Pi(Q_\beta, Q_\beta)$, which contradicts that $X$ is an ES set. By the same observation from Weibull (1995) these must also be NSS, implying that requirements *3)* and *4)* are also satisfied. ■

For repeated games however an ES set does not seem to be a particularly useful concept. One reason is that it is not obvious what topology should be used on the set of strategies $\mathcal{S}$ (see Spreij & Van Veelen, ...., and references therein), but more important is that Theorems 9 and 10 show that robustness against indirect invasions already cannot be satisfied.

# 6 Conclusion and discussion

Theorems 9 and 10 show that with sufficiently large continuation probability $\delta$, there is no strategy in the repeated prisoners dilemma that is robust against indirect invasions. In other words: every equilibrium can be upset, either by a mutant, if the strategy is not neutrally stable, or by a succession of mutants if the strategy is NSS. The richness of the strategy space therefore excludes that there is an equilibrium refinement, or a static stability concept, that by only looking at the game itself can predict what happens in a population with random matching, mutation and selection.

One thing that can be learned from this, is that what we expect to evolve must depend on further assumptions. Besides - obviously - the true size of $\delta$, the most important of those will have to concern mutation probabilities. The proofs of the results show that there are stepping stone paths out of equilibrium with in- and with decreasing levels of cooperation. Whether we can expect cooperation to in- or decrease, however, depends on how many more of these paths there are, and, more importantly, on the probabilities with which the different mutations occur. Also the starting point might matter, although it seems that a natural starting point for evolution is the strategy to always defect.

Another possibility to get stability results is to restrict the strategy space. Here it is worthwhile noticing that a restriction of the strategy space to, say, a subset $\mathcal{T}$ of $\mathcal{S}$ is in fact a special case of a combination of a starting point (somewhere within $\mathcal{T}$) and an assumption concerning mutation probabilities (they are zero for all mutations from elements of $\mathcal{T}$ to elements of $\mathcal{S} \backslash \mathcal{T}$).

# References


[1] Abreu, D. (1988). "On the theory of infinitely repeated games with discounting," *Econometrica* **56**, 383-396.

[2] Aumann, R., and L. Shapley (1976). "Long term competition: a game theoretic analysis," mimeo, Hebrew University.

[3] Axelrod, R. (1984). *The evolution of cooperation.* New York: Basic Books.

[4] Bendor, J. and P. Swistak (1995). "Types of evolutionary stability and the problem of cooperation". *Proceedings of the National Academy of Sciences* **92**, 3596-360.

[5] Bendor, J. and P. Swistak (1997). "The evolutionary stability of cooperation," *American Political Science Review* **91(2)**, 290-307.

[6] Bendor, J. and P. Swistak (1998). "Evolutionary equilibria: characterization theorems and their implications," *Theory and Decision* **45 (2)**, 99-159.

[7] Binmore, K.G., and L. Samuelson (1992). "Evolutionary stability in repeated games played by finite automata," *Journal of Economic Theory* **57**, 278-305.

[8] Binmore, K.G., and L. Samuelson (1997). "Muddling through: noisy equilibrium selection," *Journal of Economic Theory* **74**, 235-265.

[9] Boyd, R. and J. P. Lorberbaum (1987). "No pure strategy is stable in the repeated prisoner's dilemma game, *Nature* **327**, 58–59.

[10] van Damme, E. (1989). "Renegiotion-proof equilibria in repeated prisoner's dilemma," *Journal of Economic Theory* **47**, 206-217.

[11] Farrell, J. and E. Maskin (1989). "Renegotiation proof equilibria: reply," *Journal of Economic Theory* **49**, 376–378.

[12] Farrell, J. and R. Ware (1989). "Evolutionary stability in the repeated prisoner's dilemma," *Theoretical Popolation Biology* **36**, 161-166.

[13] Friedman, J. (1971). "A noncooperative equilibrium for supergames," *Review of Economic Studies* **38**, 1-12.

[14] Fudenberg, D. and E. Maskin (1986). "The folk theorem in repeated games with discounting or with incomplete information," *Econometrica* **54**, 533-554.

[15] Fudenberg, D., and E. Maskin (1990). "Evolution and cooperation in noisy repeated games," *American Economic Review* **80 (2)**, 274-279.

[16] Lorberbaum, J.P (1994). "No strategy is evolutionarily stable in the repeated prisoner's dilemma," *Journal of Theoretical Bilology* **168**, 117–130.

[17] Lorberbaum, J.P., D.E. Bohning, A. Shastri and L.E Sine (2002). "Are there really no evolutionarily stable strategies in the iterated prisoner's dilemma?" *Journal of Theoretical Bilology* **214(2)**, 155-69.

[18] Maynard Smith, J. and G.R. Price (1973). "The logic of animal conflict," *Nature* **246**, 15-18.

[19] Maynard Smith, J (1982). *Evolution and the theory of games.* Cambridge: Cambridge University Press.

[20] Rubinstein, A. (1979) "Equilibrium in supergames with the overtaking criterion," *Journal of Economic Theory* **21**, 1-9.

[21] Rubinstein, A. (1980) "Strong perfect equilibrium in supergames," *International Journal of Game Theory* **9**, 1-12.

[22] Samuelson, L. and J.M. Swinkels (2003). "Evolutionary stability and lexicographic preferences," *Games and Economic Behavior, Elsevier* **44 (2)**, 332-342

[23] Selten, R. and P. Hammerstein (1984). "Gaps in Harley's argument on evolutionarily stable learning rules and in the logic of "tit for tat"," *Behavioral and Brain Sciences* **7**, 115-116.

[24] Thomas, B. (1985). "On evolutionarily stable sets," *Journal of Mathemathical Biology* **22**, 105-115

[25] Weibull, J.W. (1995). *Evolutionary Game Theory.* Cambridge MA: MIT Press.

# A A metric for $\mathcal{S}$

Let $f : H \times \mathcal{S} \times \mathcal{S} \to \{0, 1\}$ be defined by

$$f(h_t, S, T) = \begin{cases} 0 & \text{if } S(h_t) = T(h_t) \\ 1 & \text{if } S(h_t) \neq T(h_t) \end{cases}$$

We assume that the action space $A$ is finite, and that it has $k$ elements, $a_1, ..., a_k$. The number of possible histories in $H_t$ therefore is $k^{2t-2}$.

Define the distance between $S$ and $T$, both $S, T \in \mathcal{S}$, as follows:

$$d(S, T) = \sum_{t=1}^{\infty} \rho^t \sum_{h_t \in H_t} |f(h_t, S, T)|$$

with $\rho = \frac{\delta}{k^2}$ and $\delta \in (0, 1)$.

If we take for $\mathcal{S}_t \subset \mathcal{S}$ the set of strategies in $\mathcal{S}$ that all play $a_1$ for all histories $h_u$ with $u > t$, then it is a finite set; it has $k^{\left(\sum_{v=1}^{t} k^{2t-2}\right)} = k^{\left(\frac{k^{2t}-1}{k^2-1}\right)}$ elements. The set $\bigcup_{t=1}^{\infty} \mathcal{S}_t$ is therefore countable, but it is easy to see that it is dense in $\mathcal{S}$.

# B Proofs of theorems

**Proof of Theorem 9** Assume that condition *1)* of Definition 5 is satisfied and $P$ is an equilibrium. Let $P_1, ..., P_n$ be the composing pure strategies of $P$ and let $p_1, ..., p_n, \sum_{i=1}^{n} p_i = 1$, be the probabilities with which they are played in $P$. If $\Pi(P,P) > \Pi(D,D)$, then obviously not all combinations of $P_i$ and $P_j$, with $1 \leq i, j \leq n$, can always play $D$ when they interact. So there must be at least one $i$ and one $j$, with $1 \leq i, j \leq n$, and a time $\tau$ for which $a_\tau^{P_i,P_j} \neq (D,D)$. First it is clear that there cannot only be a finite number of times that $C$ is played in the mixture. Suppose that were true, and there is a time $\tau$ and an $i$ and a $j$, with $1 \leq i, j \leq n$, for which $a_\tau^{P_i,P_j} \neq (D,D)$ and $a_t^{P_i,P_j} = (D,D)$ for all $i$ and $j$ and $t > \tau$, then the mixture is not an equilibrium; without restricting generality we can assume that $P_i\left(h_\tau^{P_i,P_j}\right) = C$ and then a strategy that equals $P_i$ for all histories at times $t < \tau$ and plays $D$ for all histories at times $t \geq \tau$ earns a higher payoff than $P_i$ and therefore also higher than all other composing strategies. Hence $C$ must be played infinitely many times in the mixed population. Since there is only a finite number of combinations $(P_i, P_j)$, it also follows that there is at least one in which $P_i$ plays $C$ an infinite number of times.

Let $(P_i, P_j)$ be a combination of strategies in which $P_i$ plays $C$ infinitely often. Let $E(i,j)$ be the set of combinations of strategies $(P_k, P_l)$ for which $a_t^{P_k,P_l} = a_t^{P_i,P_j}\ \forall\ t$, that is, strategies $P_k$ and $P_l$ that follow the same path as when $P_i$ interacts with $P_j$. Given that $P$ is a finite mixture, there is a finite time $\tau'$ which is sufficiently large to determine whether or not $(P_k, P_l) \in E(i,j)$, that is, there is a $\tau'$ such that if $a_t^{P_k,P_l} = a_t^{P_i,P_j}\ \forall\ t \leq \tau'$ then $a_t^{P_k,P_l} = a_t^{P_i,P_j}\ \forall\ t$. Let $\tau'' > \tau'$ be the first time $t$ after $\tau'$ at which $a_{t,1}^{P_i,P_j} = C$.

Let $E_i(j)$ be the set of strategies $P_l$ such that $(P_i, P_l) \in E(i,j)$. For all $P_l \in E_i(j)$ one can define $Q_l$ as the pure strategy that equals $P_l$ for all histories, except for those in the set $\widehat{H} = \left\{h_t \mid t > \tau'', a_u = a_u^{P_j,P_i} \text{ for } u \leq \tau' \text{ and } a_{\tau'',2} = D\right\}$. These histories only occur off all equilibrium paths, since the history up to and including $\tau'$ implies that this history does not occur along an equilibrium path outside $E(i,j)$, as experienced by $j$, while the remainder implies that it does not occur along equilibrium paths in $E(i,j)$. For the histories $h_t \in \widehat{H}$ we take $Q_l(h_t) = a_t^{P_j,P_i} = P_l\left(h_t^{P_j,P_i}\right)$. Obviously, the path of $Q_l$ against $P_m$ is the same as the path of $P_l$ against $P_m$ for all $m$, $1 \leq m \leq n$ and all $P_l \in E_i(j)$. Define $Q$ as the strategie that plays $Q_l$ with probability $p_l$ for $P_l \in E_i(j)$ and $P_l$ with probability $p_l$ for all $P_l \notin E_i(j)$. For this strategy we have that $\Pi(Q,P) = \Pi(P,P) = \Pi(Q,Q) = \Pi(P,Q)$.

Let $R$ be the strategy that equals $P_i$, except for $h_{\tau''}^{P_i,P_j}$, for which we take $R\left(h_{\tau''}^{P_i,P_j}\right) = D$ and except for histories that are elements of the set $\widetilde{H} = \left\{h_t \mid t > \tau'', a_u = a_u^{P_i,P_j} \text{ for } u \leq \tau' \text{ and } a_{\tau'',1} = D\right\}$, for which we take $R(h_t) = P_i\left(h_t^{P_i,P_j}\right), h_t \in \widetilde{H}$.

Because $P$ is an equilibrium, it must be that $\Pi(R,P) \leq \Pi(P,P)$. It is also clear that $\Pi(R,Q) > \Pi(Q,Q) = \Pi(P,Q)$, because $R$ improves itself against strategies $Q_l \in E_i(j)$ at time $\tau''$ without being punished and remains unchanged against strategies that are not in $E_i(j)$. As a result of this, requirement *3)* of definition 5 is not satisfied.

Note that if $Q_l = P_l \ \forall \ P_l \in E_i(j)$, that would contradict $P$ being an equilibrium, because if $P = Q$ then $\Pi(R,Q) > \Pi(Q,Q)$ would contradict that $\Pi(R,P) \leq \Pi(P,P)$.

## B.1

**Proof of Theorem 10** Assume that condition *1)* of Definition 5 is satisfied and $P$ is an equilibrium. Take $i, j$ and $\tau$ such that $\delta_{ij,\tau} = \min_{k,l,t} \delta_{kl,t}$. For all $P_l \in E_i(j)_\tau$ one can define $Q_l$ as the pure strategy that equals $P_l$ for all histories, except for those that are elements of the set $\widehat{H} = \{h_t \mid t > \tau, \ a_{u,2} = C, u \geq \tau\}$. These histories only occur off all equilibrium paths; the assumption implies that $\delta_{ij,\tau} < 1$ and hence it is not possible that $a_{\tau,2}^{P_l,P_i} = C$, for that would make $\pi_1\left(a_\tau^{P_i,P_l}\right) - \pi_1\left(C, a_\tau^{P_i,P_l}\right) = 0 \ \forall \ P_l \in E_i(j)_\tau$. For those histories $h_t \in \widehat{H}$ we take $Q_l(h_t) = C$. Obviously, the path of $Q_l$ against $P_m$ is the same as the path of $P_l$ against $P_m$ for all $m$, $1 \leq m \leq n$ and all $P_l \in E_i(j)_\tau$. Define $Q$ as the strategie that plays $Q_l$ with probability $p_l$ for $P_l \in E_i(j)_\tau$ and $P_l$ with probability $p_l$ for all $P_l \notin E_i(j)$. Consequently the corresponding payoffs are also equal; $\Pi(Q,P) = \Pi(P,P) = \Pi(Q,Q) = \Pi(P,Q)$.

Now let $R$ be the strategy that equals $P_i$, except for the history $h_\tau^{P_i,P_j}$, for which we choose $R\left(h_\tau^{P_i,P_j}\right) = C$ and except for the histories that are elemens of the set $\widetilde{H} = \left\{h_t \mid t > \tau, a_u = a_u^{P_i,P_j} \text{ for } u < \tau, \ a_{\tau,1} = C \text{ and } a_{u,2} = C, u > \tau\right\}$, for which we also take $R(h_t) = C, h_t \in \widetilde{H}$.

Because $P$ is an equilibrium, it must be that $\Pi(R,P) \leq \Pi(P,P)$. It is also clear that $\Pi(R,Q) > \Pi(Q,Q) = \Pi(P,Q)$, because $R$ improves itself against strategies $Q_l \in E_i(j)$ at time $\tau''$ without being punished and remains unchanged against strategies that are not in $E_i(j)_\tau$. As a result of this, requirement *3)* of definition 5 is not satisfied.

Note again that if $Q_l = P_l \ \forall \ P_l \in E_i(j)_\tau$, that would contradict $P$ being an equilibrium, because if $P = Q$ then $\Pi(R,Q) > \Pi(Q,Q)$ would contradict that $\Pi(R,P) \leq \Pi(P,P)$

## B.2

**Proof of Theorem 11** The robustness against indirect invasions of $P$ will be exploited a number of times in this proof, by making specific choices for the $Q$ and $R$ in Definition 5. In order to avoid confusion, we will number these choices with a subscript as they follow each other in the proof.

First observe that if $\alpha = 1$, then *5)* and *6)* in this theorem are exactly point *1)* and *2)* of Definition 5, if we take $Q_1 = S$. Therefore we only need to prove the theorem for $\alpha' < 1$.

To prove *5)*, define $Q_2 = \frac{\beta}{1-\alpha}Q + \frac{1-\alpha-\beta}{1-\alpha}R$ and $R_2 = S$. From condition (*) it follows that $\Pi(P,P) = \Pi(Q_2,P)$ and $\Pi(P,Q_2) = \Pi(Q_2,Q_2)$ $\forall\ \alpha,\beta$. Since $P$ is robust against indirect invasions, we know from condition *3)* in Definition 5 that then $\Pi(P,Q_2) = \Pi(Q_2,Q_2) \geq \Pi(R_2,Q_2)$ for all $\alpha,\beta$. Taking $Q_3 = S$ in condition *1)* of the same definition tells that $\Pi(P,P) \geq \Pi(S,P)$. Hence $\Pi(P,\alpha P + (1-\alpha)Q_2) \geq \Pi(S,\alpha P + (1-\alpha)Q_2)$ for all $\alpha,\beta$, which is *5)*.

Please realize that condition (*) together with *5)* can also be written as:

$\Pi(P,P) = \Pi(Q,P) = \Pi(R,P) \geq \Pi(S,P)$,
$\Pi(P,Q) = \Pi(Q,Q) = \Pi(R,Q) \geq \Pi(S,Q)$ and
$\Pi(P,R) = \Pi(Q,R) = \Pi(R,R) \geq \Pi(S,R)$.

To prove *6)*, we begin with $\Pi(P,P) \geq \Pi(S,P)$, which above was shown to hold. First consider the possibility that $\Pi(P,P) > \Pi(S,P)$. Then the condition in *6)* can only hold for $\alpha = 0$, and then it reads $\Pi(P,Q_2) = \Pi(R_2,Q_2)$. This makes the conditions in requirement *4)* of Definition 5 hold, so it implies that $\Pi(Q_2,R_2) \geq \Pi(R_2,R_2)$, which prooves *6)* for $\Pi(P,P) > \Pi(S,P)$.

Then we consider the possibility that $\Pi(P,P) = \Pi(S,P)$. By taking $Q_3 = S$ again, this by *2)* implies that $\Pi(P,S) \geq \Pi(S,S)$. Because we have already shown that *5)* holds, we also know that $\Pi(P,Q) \geq \Pi(S,Q)$. First consider the possibility that $\Pi(P,P) = \Pi(S,P)$ and $\Pi(P,Q) > \Pi(S,Q)$. Here the condition in *6)* can only hold for $\beta = 0$, and then it reads $\Pi(P,\alpha P + (1-\alpha)R) = \Pi(S,\alpha P + (1-\alpha)R)$ which by $\Pi(P,P) = \Pi(S,P)$ is equivalent to $\Pi(P,R) = \Pi(S,R)$. Taking $Q_4 = R$ and $R_4 = S$ in condition *4)* of Definition 5, we must conclude that $\Pi(R,S) \geq \Pi(S,S)$. But then $\Pi(\alpha P + (1-\alpha)R,S) \geq \Pi(S,S)$.

Then consider the possibility that $\Pi(P,P) = \Pi(S,P)$ and $\Pi(P,Q) = \Pi(S,Q)$. Because we have already shown that *5)* holds, we also know that $\Pi(P,R) \geq \Pi(S,R)$. First consider the possibility that $\Pi(P,P) = \Pi(S,P)$ and $\Pi(P,Q) = \Pi(S,Q)$ and $\Pi(P,R) > \Pi(S,R)$. Here the condition in *6)* can only hold for $\alpha + \beta = 1$, and then it reads $\Pi(P,\alpha P + (1-\alpha)Q) = \Pi(S,\alpha P + (1-\alpha)Q)$, which by $\Pi(P,P) = \Pi(S,P)$ is equivalent to $\Pi(P,Q) = \Pi(S,Q)$. Taking $Q_5 = Q$ and $R_5 = S$ in condition *4)* of Definition 5, we must conclude that $\Pi(Q,S) \geq \Pi(S,S)$. But then $\Pi(\alpha P + (1-\alpha)Q,S) \geq \Pi(S,S)$.

Finally consider the possibility that $\Pi(P,P) = \Pi(S,P)$ and $\Pi(P,Q) = \Pi(S,Q)$ and $\Pi(P,R) = \Pi(S,R)$. If we use $Q_2$ and $R_2$ again, but now in *4)*, we conclude that $\Pi(Q_2,R_2) \geq \Pi(R_2,R_2)$. Because $\Pi(P,S) \geq \Pi(S,S)$, as was shown above, this prooves *6)*.

## C Payoffs of strategies for examples Section 4

Payoffs of strategies against themselves and each other in the first example, with $\delta = 0.75$.

$$\Pi(C,C) = \Pi(C,TFT) = \Pi(TFT,C) = \Pi(TFT,TFT) = (1-\delta)\sum_{i=0}^{\infty} 3\delta^i = 3$$
$$\Pi(TFT,D) = (1-\delta)\left(0 + \sum_{i=1}^{\infty} 1\delta^i\right) = \delta = 0.75$$
$$\Pi(C,D) = (1-\delta)\sum_{i=0}^{\infty} 0\delta^i = 0$$
$$\Pi(D,TFT) = (1-\delta)\left(4 + \sum_{i=1}^{\infty} 1\delta^i\right) = 4(1-\delta) + \delta = 1.75$$
$$\Pi(D,C) = (1-\delta)\sum_{i=0}^{\infty} 4\delta^i = 4$$
$$\Pi(D,D) = (1-\delta)\sum_{i=0}^{\infty} 1\delta^i = 1$$

$$\Pi(D, \alpha TFT + (1-\alpha)C) = \Pi(TFT, \alpha TFT + (1-\alpha)C)$$
$$(4-3\delta)\alpha + 4(1-\alpha) = 3 \Leftrightarrow \alpha = \tfrac{1}{3\delta} = \tfrac{4}{9}$$

$$\Pi(TFT, \alpha TFT + (1-\alpha)D) = \Pi(D, \alpha TFT + (1-\alpha)D)$$
$$3\alpha + (1-\alpha)\delta = \alpha(4-3\delta) + (1-\alpha) \Leftrightarrow \alpha = \tfrac{1-\delta}{2\delta} = \tfrac{1}{6}$$

Payoffs of strategies against themselves and each other in the second example, with $\delta = 0.75$.

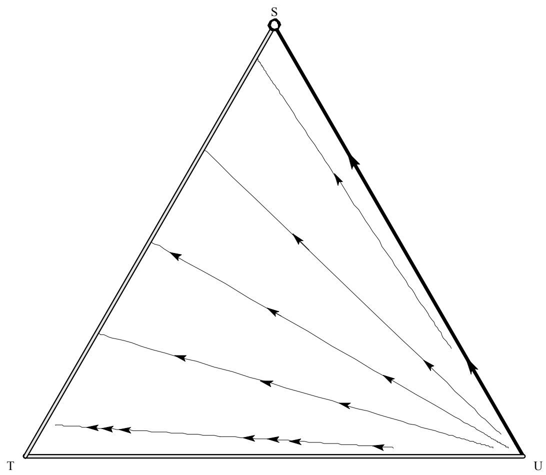
$$\Pi(D,D) = \Pi(D,STFT) = \Pi(STFT,D) = \Pi(STFT,STFT) = 1$$
$$\Pi(D,CTFT) = (1-\delta)\left(4 + 4\delta + \sum_{i=2}^{\infty} 1\delta^i\right) = 4\left(1-\delta^2\right) + \delta^2 = \tfrac{37}{16}$$
$$\Pi(STFT,CTFT) = (1-\delta)\left(4 + \sum_{i=1}^{\infty} 3\delta^i\right) = 4(1-\delta) + 3\delta = \tfrac{13}{4}$$
$$\Pi(CTFT,D) = (1-\delta)\left(0 + 0\delta + \sum_{i=2}^{\infty} 1\delta^i\right) = \delta^2 = \tfrac{9}{16}$$
$$\Pi(CTFT,STFT) = (1-\delta)\left(0 + \sum_{i=1}^{\infty} 3\delta^i\right) = 3\delta = \tfrac{9}{4}$$
$$\Pi(CTFT,CTFT) = \Pi(C,C) = 3$$

$$\Pi(D, \alpha D + (1-\alpha)CTFT) = \Pi(CTFT, \alpha D + (1-\alpha)CTFT)$$
$$\alpha + (1-\alpha)\left(4-3\delta^2\right) = \alpha\delta^2 + 3(1-\alpha) \Leftrightarrow \alpha = \tfrac{1}{2\delta^2}\left(3\delta^2 - 1\right) = \tfrac{11}{18}$$

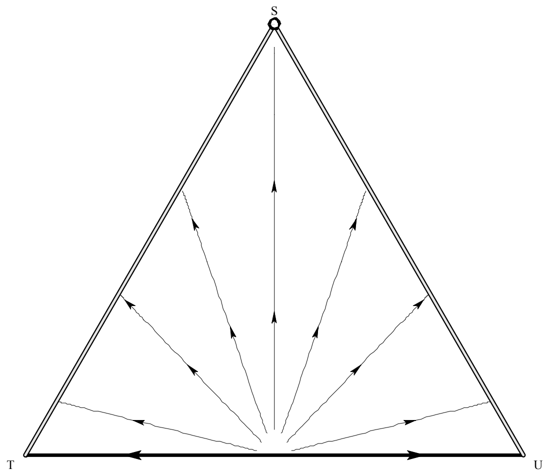
$$\Pi(D, \alpha D + (1-\alpha)STFT) = \Pi(CTFT, \alpha D + (1-\alpha)STFT)$$
$$\alpha + (1-\alpha) = \alpha\delta^2 + 3(1-\alpha)\delta \Leftrightarrow \alpha = \tfrac{3\delta-1}{3\delta-\delta^2} = \tfrac{20}{27}$$

$$\Pi(STFT, \alpha STFT + (1-\alpha)CTFT) = \Pi(CTFT, \alpha STFT + (1-\alpha)CTFT)$$
$$\alpha + (1-\alpha)(4-\delta) = 3\delta\alpha + 3(1-\alpha) \Leftrightarrow \alpha = \tfrac{1-\delta}{2\delta} = \tfrac{1}{6}$$

# D Phase plots for examples Section 5




Example 1




Example 2

# E Acknowledgements

I would like to thank Gerard van der Laan, Arthur Schram, Peter Spreij and Aljaz Ule. Phase plots were made with Dynamo Version 1.4.1 by William H. Sandholm and Emin Dokumaci. Financial Support by the Netherlands' Organization for Scientific Research (NWO) is gratefully acknowledged.