Decision Support

Subgame-perfection in free transition games

J. Flesch [a], J. Kuipers [b], G. Schoenmakers [b],*, K. Vrieze [b]

[a] Department of Quantitative Economics, Maastricht University, P.O. Box 616, 6200MD Maastricht, The Netherlands
[b] Department of Knowledge Engineering, Maastricht University, P.O. Box 616, 6200MD Maastricht, The Netherlands

ARTICLE INFO

ABSTRACT

We prove the existence of a subgame-perfect ε-equilibrium, for every $\varepsilon > 0$, in a class of multi-player games with perfect information, which we call free transition games. The novelty is that a non-trivial class of perfect information games is solved for subgame-perfection, with multiple non-terminating actions, in which the payoff structure is generally not (upper or lower) semi-continuous. Due to the lack of semi-continuity, there is no general rule of comparison between the payoffs that a player can obtain by deviating a large but finite number of times or, respectively, infinitely many times. We introduce new techniques to overcome this difficulty.

Our construction relies on an iterative scheme which is independent of ε and terminates in polynomial time with the following output: for all possible histories h, a pure action a_h^1 or in some cases two pure actions a_h^2 and b_h^2 for the active player at h. The subgame-perfect ε-equilibrium then prescribes for every history h that the active player plays a_h^1 with probability 1 or respectively plays a_h^2 with probability $1 - \delta(\varepsilon)$ and b_h^2 with probability $\delta(\varepsilon)$. Here, $\delta(\varepsilon)$ is arbitrary as long as it is positive and small compared to ε, so the strategies can be made "almost" pure.

© 2013 Published by Elsevier B.V.

1. Introduction

In many practical situations, decision makers have to make a number of choices sequentially, with perfect knowledge of all past decisions. To model and to analyse such situations, the mathematical framework of non-cooperative games with perfect information has served as a useful tool. In those games, a widely accepted solution concept is the one of Nash equilibrium, which is a collection of strategies, one for each player, such that no player can profit by individually deviating from his strategy. In other words, in a Nash equilibrium, every player plays a strategy that is a best response to the strategies of his opponents. Yet, if a player or some players make choices, either on purpose or by mistake, that differ from the prescriptions of their strategies, the continuation strategies may no longer enjoy the above mentioned equilibrium property in the remaining subgame. Hence, in certain cases the strategies may prescribe non-equilibrium behavior for the players, who consequently may become reluctant to follow them. Realizing this shortcoming, Selten [15] and [16] proposed the notion of a subgame-perfect (Nash) equilibrium, as a selection criterion amongst the Nash equilibria of the game. A subgame-perfect equilibrium requires that, in every subgame, the continuation strategies must form a Nash equilibrium. Due to the lack of existence of such exact solutions, we often need to allow for a positive error term ε. While,

in fairly general perfect information games, the existence of a Nash ε-equilibrium, for every $\varepsilon > 0$, has long been known due to Mertens and Neyman, cf. Mertens [10], and later to Thuijsman and Raghavan [20], Flesch et al. [5] have recently shown that not all perfect information games admit a subgame-perfect ε-equilibrium. It is still an open question which classes of perfect information games do.

Given the large number and the diversity of applications of subgame-perfect equilibrium, also in perfect information games, this existence question is relevant for both theory and practice, and has received notable attention in recent years. Areas where subgame-perfect equilibria play an important role include bargaining (cf., e.g. Osborne and Rubinstein [12], Muthoo [11], Kim and Jeon [6], Weg and Zwick [23]), coalition formation (cf. Bloch [1], Vartiainen [21], Vidal-Puga [22]), entry deterrence (cf. Selten [17]), matching (cf. Ramsey [14]) and political science (cf. Bueno de Mesquita and Lalman [2]).

In this paper we examine a natural class of perfect information games, which we call free transition games, and prove that they admit a subgame-perfect ε-equilibrium for all $\varepsilon > 0$. Moreover, the corresponding strategies are "almost" pure: either one action is prescribed with probability 1, or one action with probability $1 - \delta$ and another one with probability δ, where $\delta > 0$ is arbitrarily small.

A free transition game is played as follows: at every period during play, one of the players is called active. The active player can either terminate the game, or he can announce any other player,

* Corresponding author.
  E-mail address: gm.schoenmakers@maastrichtuniversity.nl (G. Schoenmakers).

who then becomes the active player. If a player decides to terminate the game, then depending on the identity of this player, every player receives a payoff. If no player ever terminates the game, then every player receives payoff zero.

One could think of a group of friends, who want to spend the evening together. Initially, one of them, say person j, is supposed to organize the evening activities. He can either organize the evening himself (termination at j), or ask someone else, say person k, to replace him as the organizer. In the latter case, person k has the following options: organize the evening (termination at k), reject person j's request (person k nominates j), or ask a third person to replace him as the organizer. In the last two cases, the procedure continues as before. Since the evening activities may depend on the organizer, the payoff of each person may depend on the identity of the organizer. If no one ever decides to organize the evening, the friends cannot meet and they all receive payoff zero.

We remark that such a game can be viewed as a special type of multi-player stochastic game with the following properties: (1) it is a game with perfect information (i.e., in every state, only one of the players is to choose an action), (2) the payoffs are recursive (i.e., the only payoffs are upon termination), and (3) all the transitions are deterministic.

In this paper we prove that every free transition game admits a subgame-perfect ε-equilibrium for every $\varepsilon > 0$. In view of an instructive example in Solan and Vieille [19], cf. Example 1 below, this result cannot be strengthened to pure strategies or to $\varepsilon = 0$. Our construction uses an iterative method that results in a non-empty collection of plays that can be a part of a subgame-perfect ε-equilibrium, for every $\varepsilon > 0$. The main difficulty arises by negative termination payoffs, which may dissuade certain players from terminating the game.

For free transition games with only non-negative termination payoffs, the one-deviation principle applies, cf. Kuipers et al. [7]: if with respect to a strategy profile, no player can improve his payoff in any subgame by deviating exactly once, then this strategy profile is a subgame-perfect 0-equilibrium. Indeed, this property implies that a finite number of deviations cannot be profitable either, and because the termination payoffs are non-negative, no player can profit from deviating infinitely many times (that can only happen if play never terminates). By making use of this one-deviation principle, Kuipers et al. [7] showed the existence of a pure subgame-perfect 0-equilibrium.

Within the more general framework of perfect-information games, there has been significant development in recent years on the existence of subgame-perfect ε-equilibria. Solan [18], Mashiah-Yaakovi [8] and [9] proved the existence of (approximate) subgame-perfect ε-equilibria in different subclasses in which the active player has only one non-terminating action. When multiple non-terminating actions are allowed, the game becomes strategically more complex. Iterative methods have been proven useful in a number of subclasses. Kuipers et al. [7] and Flesch et al. [3] applied such methods, but in both cases it is crucial to assume that the termination payoffs are non-negative. These results have been generalized by Flesch et al. [4], who proved the existence of subgame-perfect ε-equilibria, provided that the transitions are deterministic and that the payoff functions of the players, defined as a function of the whole play, are bounded and lower-semicontinuous. The assumption of lower-semicontinuity guarantees that deviating infinitely many times is never better, up to a small error-term, than deviating a large but finite number of times, and hence the one-deviation principle still applies. Subsequently, Purves and Sudderth [13] proved the same result, by using a different approach, when the payoffs are upper semicontinuous. Due to upper semicontinuity, the opposite is true here: deviating a large but finite number of times is never better, up to a small error-term, than deviating infinitely many times.

Our model of free transition games, with possibly negative termination payoffs, differs from the above mentioned classes with multiple non-terminating actions, because the payoff functions are generally not (upper or lower) semi-continuous. Consequently, there is no general rule of comparison between the payoffs that a player can obtain by deviating finitely many times and the payoffs that he can obtain by deviating infinitely many times, and this issue requires a delicate treatment.

2. Game model and strategic concepts

A *free transition game* is given by a triple $G_j = (N, r, j)$, where $N = \{1, \ldots, n\}$ is a set of players with $n \geqslant 2$, $j \in N$, and $r : N \to \mathbb{R}^n$ is a function that associates a payoff vector $r(t) \in \mathbb{R}^n$ with each player $t \in N$. The game is played at periods in $\mathbb{N}$ as follows. At every period, one of the players is called active, with j being the active player at period 1. The active player, say t, always has the option to terminate the game. In this case, the players receive payoffs according to $r(t)$. The active player can also announce any player other than himself, who then becomes the active player at the next period, and so on. If no player ever terminates the game, each player's payoff is zero.

For convenience, we assume that (i) the active player can terminate the game by announcing himself as the next active player and (ii) if he does so, then play goes on, but this player has to announce himself at all further periods. So, play is always infinite and results in a sequence $(t_m)_{m \in \mathbb{N}}$ of active players such that if $t_M = t_{M+1}$ for some $M \in \mathbb{N}$ then $t_m = t_M$ for all $m \geqslant M$. If a player $s \in N$ announces himself, we say that termination occurs at s.

The *history* at some period is the sequence that consists of the active players at all previous periods and the currently active player, in chronological order. Note that every history in the game G_j starts with j. For G_j, we denote by H_j the set of all possible histories, and by H_{ji} the set of all possible histories in which i is the current active player. A strategy σ_i for player $i \in N$ is a map that assigns to every history $h \in H_{ji}$ a probability distribution $\sigma_i(h)$ on N such that if termination has occurred in h at i then $\sigma_i(h)$ places probability 1 on i. The interpretation is that σ_i prescribes for player i to choose the next active player according to $\sigma_i(h)$, after every $h \in H_{ji}$. A vector of strategies $\sigma = (\sigma_i)_{i \in N}$, one for each player, is called a strategy profile. For a strategy profile σ, we denote by $\phi_i(\sigma)$ the expected payoff for player $i \in N$ if every player $s \in N$ plays according to σ_s.

A strategy profile σ is called a (Nash) ε-equilibrium in G_j, for some $\varepsilon \geqslant 0$, if no player can gain more than ε by individually deviating from his strategy, i.e. for every player i and every strategy $\widetilde{\sigma_i}$ of player i we have

$$\phi_i(\sigma_1, \ldots, \sigma_{i-1}, \widetilde{\sigma_i}, \sigma_{i+1}, \ldots, \sigma_n) \leqslant \phi_i(\sigma) + \varepsilon.$$

Consider a history $h = (t_m)_{m=1}^L \in H_{js}$. We denote by H_{si}^h the set of all histories $h' = (t'_m)_{m=1}^{L'} \in H_{si}$ such that $(h, h') := (t_1, \ldots, t_{L-1}, t'_1, \ldots, t'_{L'})$ is a history in H_{ji}. Note that H_{si}^h is empty exactly when termination occurred in h at s and $i \neq s$. For a strategy σ_i, the continuation strategy σ_i^h for h is defined as the map that assigns to every history $h' \in H_{si}^h$ the probability distribution $\sigma_i(h, h')$. Intuitively, the recommendation by σ_i^h after h' is just as if h had happened before h'.

A strategy profile σ is a *subgame-perfect ε-equilibrium* in G_j, for some $\varepsilon \geqslant 0$, if for every history $h \in H_j$, with some final player s, the continuation strategy profile $\sigma^h = (\sigma_i^h)_{i \in N}$ is an ε-equilibrium in G_s, i.e. for every player i and every strategy $\widetilde{\sigma_i}$ of player i we have

$$\phi_i(\sigma_1^h, \ldots, \sigma_{i-1}^h, \widetilde{\sigma_i}^h, \sigma_{i+1}^h, \ldots, \sigma_n^h) \leqslant \phi_i(\sigma^h) + \varepsilon.$$

A *deterministic game-plan*, or simply *plan*, is an infinite sequence $g = (t_m)_{m \in \mathbb{N}}$ with $t_m \in N$ for all $m \in \mathbb{N}$ and with the property

$$[\exists M \in \mathbb{N} : t_M = t_{M+1}] \Rightarrow [\forall m \geqslant M : t_m = t_M].$$

If $t_M = t_{M+1} = s$ then g corresponds to a plan or play where player s eventually terminates the game. We say that plan g *terminates* at s. Otherwise, i.e., if $t_m \neq t_{m+1}$ for all $m \in \mathbb{N}$, we say that g is *non-terminating*, which corresponds to a plan or play, where no player ever terminates the game. We write $t \in g$ if $m \in \mathbb{N}$ exists with $t_m = t$. If $t \in N$ is such that $t \neq t_1$, then the sequence $(t, t_1, t_2, \ldots)$ is also a plan, and we will denote this plan as (t,g).

For every plan g, let $\phi(g)$ denote the corresponding payoff vector in $\mathbb{R}^n$. Formally, $\phi(g) = r(t)$ if g terminates at t and $\phi(g) = (0, \ldots, 0)$ if g does not terminate.

Define Ω as the set of real vectors $\alpha \in \mathbb{R}^n$ with $\alpha_t \geq r_t(t)$ for every $t \in N$. Now take $\alpha \in \Omega$. We say that a plan g is α-*viable* if $\phi_t(g) \geq \alpha_t$ for all $t \in g$. This means that, whenever play is according to g and player t is the active player, then t will receive a payoff of at least α_t. For $t \in N$, let $V_t(\alpha)$ denote the set of plans starting at t that are α-viable. Say that player $t \in N$ is α-*positive* if $\alpha_t > 0$ and that t is α-*anchored* if $\alpha_t = r_t(t)$. The α-anchored players are the players who do not mind to terminate the game with respect to α. Say that player t is α-*safe at* $u \in N \backslash \{t\}$ if $\phi_t(g) \geq \alpha_t$ for all $g \in V_u(\alpha)$. This means that if player t moves to u then he is guaranteed a payoff of at least α_t, provided that an α-viable plan is executed from u. Notice that if t is α-*safe at* $u \in N \backslash \{t\}$ and $t \in g$ for some $g \in V_u(\alpha)$, then player t will again be α-*safe at* u. Say that player t is α-*vulnerable* if $\alpha_t \leq 0$ and if at least one of the following holds:

V1. a plan $g \in V_u(\alpha)$ exists for some $u \in N \backslash \{t\}$ such that g terminates at t and there is a player $x \in g, x \neq t$, with $r_t(x) \leq \alpha_t$;
V2. an α-anchored player $x \in N \backslash \{t\}$ exists with $r_t(x) \leq \alpha_t$.

The motivation behind this definition is the following: Consider the situation in which player t is supposed to terminate the game. If player t is α-vulnerable, then indeed he can be persuaded to terminate. First suppose that player t is α-vulnerable due to condition V1, with corresponding players u, x and plan g, and that he moves to player u instead of terminating the game. Then, the players can execute the plan g with the modification that player x only follows g with probability close to 1 and terminates with a small but positive probability. If x follows g then t is again asked to terminate (implying that t is not better off), whereas if x terminates then t gets punished for not terminating. This would be particularly useful if we could find a plan as in condition V1 for all players $u \neq t$, because then it would not matter which player is chosen by player t when deciding not to terminate. This is not going to be the case completely, but it is certainly true for all players $u' \neq t$ who do not mind termination at player t. They can simply move[1] to player u and then play can continue at u just as above. As for all other players, who do not want termination at t, we show – at least for a special value of α, which we call α^* – that they can start up a plan that also punishes player t without t ever becoming active.

Now suppose that player t is α-vulnerable due to condition V2, and he decides to move to another player w, instead of terminating the game. Our construction will provide a plan $\tilde{g}$ – once again, for α^* – that starts at w such that player t is not better off by plan $\tilde{g}$. Now, we let the players execute $\tilde{g}$ with probability close to 1 and the following plan with a small but positive probability: player w moves[1] to player x and subsequently x terminates the game. Notice that it is essential that player t does not become the active player according to the latter plan, when w moves to x, so he cannot deviate again, and that x is α-anchored, otherwise he would refuse to terminate the game.

The following lemma establishes two fairly intuitive properties of the concepts we just introduced.

Lemma 1. *Let* $\alpha \in \Omega$.

1. *Let* $t, u, s \in N$ *be three distinct players. If* t *is* α-*safe at* u *and* u *is* α-*safe at* s, *then* t *is* α-*safe at* s.
2. *If player* t *is not* α-*vulnerable and* $\alpha_t \leq 0$, *then* t *is* α-*safe at every* $s \in N \backslash \{t\}$.

Proof. First we prove part 1. Assume that t is α-safe at u and that u is α-safe at s. To prove that t is α-safe at s, choose an arbitrary plan $g \in V_s(\alpha)$. Since u is α-safe at s, we have $\phi_u(g) \geq \alpha_u$, hence $(u,g) \in V_u(\alpha)$. Then, since t is α-safe at u, we have $\phi_t(u,g) \geq \alpha_t$, hence $\phi_t(g) \geq \alpha_t$. Since g was chosen arbitrarily in $V_s(\alpha)$, it follows that t is α-safe at s.

Now we prove part 2. Assume that t is not α-vulnerable and $\alpha_t \leq 0$. Let $s \in N \backslash \{t\}$. If $V_s(\alpha)$ is empty, then we are done. Otherwise, choose a plan $g \in V_s(\alpha)$. If g terminates at t, then the α-viability of g ensures that $\phi_t(g) \geq \alpha_t$. If g terminates at $s' \neq t$, then $r_t(s') > \alpha_t$, since condition V2 does not hold. Hence, $\phi_t(g) \geq \alpha_t$. If g does not terminate, then $\phi_t(g) \geq \alpha_t$ follows from $\alpha_t \leq 0$. This shows that t is α-safe at s. $\square$

We say that the plan $g = (t_m)_{m \in \mathbb{N}}$ is α-*safe* if.

S1. g is α-viable;
S2. for all $m \in \mathbb{N}$ such that $t_m \neq t_{m+1}$, it holds that t_m is α-safe at t_{m+1};
S3. if g terminates at s, then s is α-positive or s is α-vulnerable.

The reason why s is required to be α-vulnerable if g terminates at s and $\alpha_s \leq 0$ is that we want that s can indeed be persuaded to terminate. Recall that this is exactly what vulnerability is defined for. All other conditions are obviously useful.

We say that $\alpha \in \Omega$ is *stable* if for every player $t \in N$, an α-safe plan starting at t exists. The following lemma demonstrates that stable vectors exist.

Lemma 2. *The vector* $\rho = (r_t(t))_{t \in N}$ *is stable.*

Proof. Trivially, we have $\rho \in \Omega$. Let $t \in N$. Notice that the plan $(t, t, \ldots)$, in which player t immediately terminates, is ρ-viable. So, if t is ρ-positive or if t is ρ-vulnerable, we choose the plan $(t, t, \ldots)$. Otherwise, $\rho_t \leq 0$ and t is not ρ-vulnerable. Take an arbitrary $u \in N \backslash \{t\}$. By Lemma 1.2, t is ρ-safe at u. Therefore, if u is ρ-positive or if u is ρ-vulnerable $(t, u, u, \ldots)$ is the required plan. Otherwise, $\rho_u \leq 0$ and u is not ρ-vulnerable. By Lemma 1.2 once more, u is ρ-safe at t, and therefore $(t, u, t, u, \ldots)$ is the required plan. $\square$

3. The main result

In this section, we first provide a method for updating a stable vector $\alpha \in \Omega$, such that the updated vector, say $\delta(\alpha)$, is also a stable vector in Ω, and satisfies $\delta(\alpha) \geq \alpha$ coordinatewisely. This will imply the existence of a stable vector $\alpha^* \in \Omega$ with the property $\delta(\alpha^*) = \alpha^*$. This property ensures that a deviation, say by player t, from a prescribed α^*-viable plan g can be answered by prescribing a new α^*-viable plan g' such that $\phi_t(g') \leq \alpha_t^*$. If the prescribed plans during play do not terminate or terminate at α^*-positive players, this actually constitutes a subgame-perfect 0-equilibrium. However, if termination at a player s with $\alpha_s^* < 0$ is prescribed, it may be necessary to have a secondary threat plan available that is played with small probability ε to enforce termination at s. Now, the properties of α^* will be shown to be such that s is α^*-vulnerable, and hence an appropriate threat plan is indeed available.

---

[1] Here we specifically make use of the free transition structure of the game.

For $t, u \in N$ with $u \neq t$, and $\alpha \in \Omega$, we define

$$\beta_t^+(u, \alpha) = \min\{\phi_t(v) | v \in V_u(\alpha)\},$$
$$\beta_t^-(u, \alpha) = \min\{\phi_t(v) | v \in V_u(\alpha) \text{ and } v \text{ does not terminate at } t\}.$$

Intuitively, $\beta_t^+(u, \alpha)$ is the minimal payoff to player t if he moves to player u provided that from player u an α-viable plan is executed. Furthermore, $\beta_t^-(u, \alpha)$ is similar with the restriction that the plan does not terminate at t. Since only α-positive and α-vulnerable players can be persuaded to terminate, we define

$$\beta_t(u, \alpha) = \begin{cases} \beta_t^-(u, \alpha) & \text{if } \alpha_t \leqslant 0 \text{ and } t \text{ is not } \alpha - \text{vulnerable,} \\ \beta_t^+(u, \alpha) & \text{otherwise} \end{cases}$$

$$\beta_t(t, \alpha) = r_t(t).$$

For $t \in N$ and $\alpha \in \Omega$, we define

$$\delta_t(\alpha) = \max\{\beta_t(u, \alpha) | u \in N\},$$

and

$$B_t(\alpha) = \{u \in N | \beta_t(u, \alpha) = \delta_t(\alpha)\}.$$

So, $B_t(\alpha)$ is the set of "best" moves for player t, i.e. those moves that guarantee the highest payoff given by $\beta_t(\cdot, \alpha)$.

We use the convention $\min \emptyset = \infty$, so that $\beta_t(u, \alpha)$ and $\delta_t(\alpha)$ are well defined. Moreover, by the definition of $\delta_t(\alpha)$, the set $B_t(\alpha)$ is always non-empty.

Lemma 3. *Let $\alpha \in \Omega$ be stable and let $t \in N$. Then $\alpha_t \leqslant \delta_t(\alpha) < \infty$.*

Proof. First we show that $\alpha_t \leqslant \delta_t(\alpha)$. Since α is stable, an α-safe plan starting at t exists. If $(t, t, \ldots)$ is an α-safe plan, then $\alpha_t = r_t(t) = r_t(t) = \beta_t(t, \alpha) \leqslant \delta_t(\alpha)$. Otherwise, $u \in N\setminus\{t\}$ exists such that t is α-safe at u, which implies

$$\alpha_t \leqslant \min\{\phi_t(v) | v \in V_u(\alpha)\} \leqslant \beta_t(u, \alpha) \leqslant \delta_t(\alpha).$$

Now we prove that $\delta_t(\alpha) < \infty$. Choose $u \in B_t(\alpha)$. If $u = t$, then $\delta_t(\alpha) = \beta_t(t, \alpha) = r_t(t) < \infty$. So assume further that $u \neq t$. Let v denote an α-safe plan starting at u. If $\alpha_t > 0$ or if t is α-vulnerable, then $\delta_t(\alpha) = \beta_t^+(u, \alpha) \leqslant \phi_t(v) < \infty$. Otherwise, if $\alpha_t \leqslant 0$ and t is not α-vulnerable, then by definition of an α-safe plan, v does not terminate at t. Thus $\delta_t(\alpha) = \beta_t^-(u, \alpha) \leqslant \phi_t(v) < \infty$. □

Notice that if $\delta_t(\alpha) < \infty$ holds for some $\alpha \in \Omega$ and $t \in N$, then $\beta_t(u, \alpha) < \infty$ for every $u \in N$. Hence, Lemma 3 has the following corollary.

Corollary 4. *Let $\alpha \in \Omega$ be stable. Then, $V_u(\alpha)$ is non-empty for every $u \in N$. Moreover, if for $t \in N$ it holds that $\alpha_t \leqslant 0$ and t is not α-vulnerable, then for every $u \in N$ with $u \neq t$ there exists a plan $v \in V_u(\alpha)$ that does not terminate at t.*

Lemma 5. *Let $\alpha \in \Omega$ be stable, let $t \in N$, and let δ be the vector that results from α when coordinate α_t is replaced with $\delta_t(\alpha)$. Then*

1. *$V_u(\delta) \subseteq V_u(\alpha)$ for all $u \in N$.*
2. *If $s, u \in N$ with $s \neq t$, and if s is α-safe at u, then s is δ-safe at u.*

Proof. By Lemma 3, we have $\delta \geqslant \alpha$. Then, if a plan is δ-viable, it is obviously also α-viable. This proves part 1.

Let $s, u \in N$ with $s \neq t$, and assume that s is α-safe at u. Choose an arbitrary plan $g \in V_u(\delta)$. Then $g \in V_u(\alpha)$ by part 1, and since s is α-safe at u, it follows that $\phi_s(g) \geqslant \alpha_s$. Since $s \neq t$, we have $\alpha_s = \delta_s$, hence $\phi_s(g) \geqslant \delta_s$. This proves part 2. □

Lemma 6. *Let $\alpha \in \Omega$ be stable, let $t \in N$, and let δ be the vector that results from α when coordinate α_t is replaced with $\delta_t(\alpha)$. Then δ is stable.*

Proof. We may assume $\delta_t > \alpha_t$, as otherwise the lemma is trivially true. To prove the lemma, we need to demonstrate, for every $u \in N$, the existence of a δ-safe plan starting at u. In part A of the proof we show how a δ-safe plan starting at t can be constructed, and in part B we give a construction for $u \neq t$.

A: Choose $u \in B_t(\alpha)$. Notice that the assumption $\delta_t > \alpha_t$ implies $t \notin B_t(\alpha)$, hence $u \neq t$. We first show that t is δ-safe at u. To see this, let g be a δ-viable plan starting at u (if there is no such plan, then t is δ-safe at u by definition). If $t \in g$, then $\phi_t(g) \geqslant \delta_t$ follows by the δ-viability of g. If $t \notin g$, then obviously g does not terminate at t. Since g is also α-viable, it follows that $\phi_t(g) \geqslant \beta_t^-(u, \alpha) \geqslant \beta_t(u, \alpha) = \delta_t$. So indeed, t is δ-safe at u.

Since α is stable, there exists an α-safe plan g starting at u. We now argue that g does not terminate at t. Suppose by way of contradiction that g terminates at t. Then, since g is an α-safe plan, $\alpha_t > 0$ or t is α-vulnerable. Hence, $\delta_t = \beta_t^+(u, \alpha) \leqslant \phi_t(g) = r_t(t)$, which contradicts our assumption that $\delta_t > \alpha_t$. Thus, g does not terminate at t as claimed. We distinguish two cases.

A1: g terminates at some $s \neq t$. In this case, we prove that $g' := (t, u, s, s, \ldots)$ is a δ-safe plan.

Proof of condition S1: We have $\phi_t(g') = \phi_t(g) \geqslant \beta_t(u, \alpha) = \delta_t$, and by the α-viability of g, we also have $\phi_u(g') = \phi_u(g) \geqslant \alpha_u = \delta_u$ and similarly $\phi_s(g') \geqslant \delta_s$. Hence, S1 holds indeed.

Proof of condition S2: We know already that t is δ-safe at u. Thus, if $u = s$ then S2 holds. If $u \neq s$, then by Lemmas 1.1 and 5.2, u is δ-safe at s, and S2 holds in this case as well.

Proof of condition S3: As g is α-safe, we have that s is α-anchored, and moreover either α-positive or α-vulnerable. If s is α-positive, then s is δ-positive too, and S3 holds. So, assume that s is α-vulnerable. First suppose that s is α-vulnerable due to V1, so that there exists a plan $v \in V_y(\alpha)$ starting at some $y \in N\setminus\{s\}$ such that v terminates at s and there is an $x \in v$, $x \neq s$ with $r_s(x) \leqslant \alpha_s$. The plan $(x, s, s, \ldots)$ when $x \neq t$ and respectively the plan g' when $x = t$ show that s is δ-vulnerable (due to V1 in both cases). Finally, assume that s is α-vulnerable due to V2, so that an α-anchored player $x \in N\setminus\{s\}$ exists with $r_s(x) \leqslant \alpha_s$. If $x \neq t$, then s is δ-vulnerable (due to V2), and if $x = t$, then the plan g' shows that s is δ-vulnerable (due to V1).

A2: g does not terminate. In this case, $g' := (t, g)$ is δ-safe when $t \notin g$, and $g'' = (t, u, t, u, \ldots)$ is δ-safe when $t \in g$. Indeed, condition S1 holds for either case because $\delta_t = \beta_t(u, \alpha) \leqslant \phi_t(g) = 0$, and because g is α-viable. Condition S2 is valid due to t being δ-safe at u, and additionally to Lemma 5.2 for g' and to Lemmas 5.2 and 1.1 for g''. Condition S3 needs no verification.

B: We construct a δ-safe plan starting at any $u \neq t$ as follows. If u is δ-safe at t, then choose a δ-safe plan $\bar{v}$ starting at t and $(u, \bar{v})$ is the required δ-safe plan starting at u. So assume from here that u is not δ-safe at t. Let v denote an α-safe plan starting at u. We claim that v is the required δ-safe plan. Observe that $t \notin v$: otherwise u is α-safe at t by Lemma 1.1, hence u is δ-safe at t by Lemma 5.2, contradicting our assumption that u is not δ-safe at t. Therefore $\alpha_x = \delta_x$ for all $x \in v$, which shows that v is δ-viable, so condition S1 holds for v and δ. Also, v consists entirely of δ-safe moves by Lemma 5.2, so condition S2 also holds for v and δ. Condition S3 for v and δ is satisfied when v does not terminate, or when v terminates at an α-positive player. So assume that v terminates at s such that $\alpha_s \leqslant 0$. Since v is α-safe, s must be α-vulnerable. Notice that s is not δ-safe at t: if $s = u$ then s is not δ-safe at t due to our assumption that u is not δ-safe at t, whereas if $s \neq u$ then s is not δ-safe at t because u is δ-safe at s and u is not δ-safe at t, and because of the transitivity of safe moves (cf. Lemma 1.1). Hence,

$g \in V_t(\delta)$ exists with $\phi_s(g) < \delta_s = \alpha_s \leqslant 0$. Notice that g must terminate, since $\phi_s(g) \neq 0$, and g does not terminate at s, since g is δ-viable and $\phi_s(g) < \delta_s$. This implies that a δ-anchored player $x \in N \backslash \{s\}$ exists with $r_s(x) < \delta_s$. It follows that s is δ-vulnerable, and condition S3 for v and δ is satisfied. So indeed, v is the required plan. $\square$

Lemma 7. *A stable vector $\alpha^* \in \Omega$ exists such that $\delta_t(\alpha^*) = \alpha_t^*$ for all $t \in N$. Moreover, for all $t \in N$, we have*

$$\alpha_t^* < 0 \Rightarrow t \text{ is } \alpha^* - \text{vulnerable}.$$

Proof. For any $\alpha \in \Omega$, we denote by $N(\alpha)$ the set of players $t \in N$ for which $\delta_t(\alpha) > \alpha_t$. Let $\alpha^0 = (r_t(t))_{t \in N}$. By Lemma 2, the vector α^0 is stable. If $N(\alpha^0) = \emptyset$, then define $\alpha^* = \alpha^0$. Otherwise, choose an arbitrary $t^0 \in N(\alpha^0)$, and let α^1 be the vector that results from α^0 when coordinate $\alpha_{t^0}^0$ is replaced with $\delta_{t^0}(\alpha^0)$. In view of Lemmas 3 and 6, $\alpha^1 \in \Omega$ and α^1 is stable too, and it is in one coordinate strictly larger than α^0, while in all other coordinates they are equal. If $N(\alpha^1) = \emptyset$, then define $\alpha^* = \alpha^1$. Otherwise, we can increase α^1 in one coordinate, just as above, and so on. Since the set of payoffs is finite, Lemma 3 guarantees that, in a finite number of steps, we obtain a stable vector $\alpha^* \in \Omega$ such that $\delta_t(\alpha^*) = \alpha_t^*$ for all $t \in N$.

Assume that $\alpha_t^* < 0$ for some $t \in N$. We need to show that t is α^*-vulnerable. Suppose by way of contradiction that t is not α^*-vulnerable. Take an arbitrary player $u \neq t$, and let $v \in V_u(\alpha^*)$ be such that v does not terminate at t and $\phi_t(v) = \beta_t^-(u, \alpha^*)$. Note that such a plan v exists by the definition of $\beta_t^-(u, \alpha^*)$ and by Corollary 4. Since $\beta_t^-(u, \alpha^*) = \beta_t(u, \alpha^*) \leqslant \delta_t(\alpha^*) = \alpha_t^* < 0$, the plan v terminates, and by assumption, v terminates at some $s \neq t$. Hence, s is α^*-anchored and $r_t(s) = \phi_t(v) \leqslant \alpha_t^*$. In view of condition V2 of vulnerability, t is α^*-vulnerable, which is a contradiction. $\square$

Theorem 8. *A subgame-perfect ε-equilibrium exists, for every $\varepsilon > 0$.*

Proof. Fix an $\varepsilon > 0$ and take a stable vector α^* as in Lemma 7. Take an $R > \varepsilon$ large enough so that all payoffs in the game belong to $[-R, R]$.

Choosing plans: The most important part of the subgame-perfect ε-equilibrium that we will construct is the specification of how the players should react when some player t deviates from a prescribed plan by nominating some player $u \neq t$. We will distinguish two types of deviations, depending on t and u.

When a deviation of type 1 occurs, the players will react by switching to a plan g_{tu}^1 starting at u.

For a deviation of type 2, however, we define a primary plan g_{tu}^2 as well as a secondary plan h_{tu}^2. These two plans will coincide up to a certain player y_{tu}. The players are expected to follow these plans until y_{tu} becomes active, who is then asked to play the action corresponding to g_{tu}^2 with probability $1 - \frac{\varepsilon}{2R}$ and to play the action corresponding to h_{tu}^2 with probability $\frac{\varepsilon}{2R}$. Subsequently, the players should execute g_{tu}^2 in the former case, and h_{tu}^2 in the latter case.

So, consider a pair (t, u) where $t, u \in N$ and $u \neq t$. Choose an α^*-viable plan g starting at u such that $\phi_t(g) \leqslant \alpha_t^*$. Note that such a plan exists because $\delta_t(\alpha^*) = \alpha_t^*$. If it is possible, choose g such that $t \notin g$.

Case 1: We say that (t, u) is of type 1 if $\alpha_t^* \geqslant 0$ or $t \notin g$. In this case, we define $g_{tu}^1 = g$.

Case 2: We say that (t, u) is of type 2 if $\alpha_t^* < 0$ and $t \in g$. Note that in this case g terminates, since $\phi_t(g) \leqslant \alpha_t^* < 0$. Then in fact g terminates at t, as otherwise a choice of g with

$t \notin g$ would have been possible, due to free transitions. Moreover, by Lemma 7, player t is α^*-vulnerable. There are two subcases.

Case 2.1: Condition V1 is satisfied for α^* and t, i.e., a plan $g' \in V_s(-\alpha^*)$ exists starting at some $s \in N \backslash \{t\}$ such that g' terminates at t and there is an $x \in g'$, $x \neq t$, with $r_t(x) \leqslant \alpha_t^*$. Note that, if $s \neq u$, the plan (u, g') is also α^*-viable, since $\phi_u((u, g')) = r_u(t) = \phi_u(g) \geqslant \alpha_u^*$. We may therefore assume that $g' \in V_u(\alpha^*)$. In this subcase, we define $g_{tu}^2 = g'$, we define h_{tu}^2 as the plan that follows g' up to x and subsequently terminates at x, and we define $y_{tu} = x$. (Possibly, $x = u$.)

Case 2.2: Condition V1 is not satisfied for α^* and t, but condition V2 is, i.e., an α^*-anchored player $x \in N \backslash \{t\}$ exists such that $r_t(x) \leqslant \alpha_t^*$. In this subcase, we define $g_{tu}^2 = g$, $h_{tu}^2 = (u, x, x, \ldots)$, and $y_{tu} = u$. (Notice that x cannot be the second player of g, since condition V1 is not satisfied. So indeed, g_{tu}^2 and h_{tu}^2 coincide up to y_{tu}.)

Properties of these plans: The properties below follow directly from the definitions. For every $t, u \in N$ with $u \neq t$, we have:

(A) *Viability*: When (t, u) is of type 1, g_{tu}^1 is α^*-viable. When (t, u) is of type 2, g_{tu}^2 is α^*-viable and the part of h_{tu}^2 beyond y_{tu} is α^*-viable, i.e., if y_{tu} does not terminate, then $\phi_s(h_{tu}^2) \geqslant \alpha_s^*$ for every player s who appears in h_{tu}^2 after y_{tu}.

(B) *Retaliation of t^**: When (t, u) is of type 1, $\phi_t(g_{tu}^1) \leqslant \alpha_t^*$. When (t, u) is of type 2, $\phi_t(g_{tu}^2) \leqslant \alpha_t^*$ and $\phi_t(h_{tu}^2) \leqslant \alpha_t^*$.

(C) *Player t cannot deviate infinitely often*: This is true, since for a deviation of type 1, player t does not appear in g_{tu}^1, and for a deviation of type 2, player t does not appear in the secondary plan h_{tu}^2.

(D) *The executor of a lottery is not t (the deviator) and the executer of the lottery places high probability on the plan that he prefers:* When (t, u) is of type 2, $y_{tu} \neq t$ and $\phi_{y_{tu}}(g_{tu}^2) \geqslant \phi_{y_{tu}}(h_{tu}^2)$.

The profile σ: Now we define a strategy profile σ, and afterwards we show that σ is a subgame-perfect ε-equilibrium. Suppose that the initial player is s, and choose an arbitrary α^*-viable plan v starting at s. Let σ prescribe to play as follows: Follow v as long as all players choose the action prescribed by v. If some player t deviates from v by terminating, then play is strategically over. If t deviates from v by nominating another player u, then determine the type of (t, u) and the corresponding plans, and play according to them as described above, until some player t' deviates by choosing an action u' that he should have used with probability zero. If t' deviates in such a way, and by doing so he terminates, i.e. $u' = t'$, then play is strategically over. If $u' \neq t'$, then just as above, determine the type of (t', u') and the corresponding plans, and play according to them, until some player t'' deviates by choosing an action u'' that he should have used with probability zero, and so on.

We remark that, if (t', u') is of type 2, then player $y_{t'u'}$ can change the probabilities with which he chooses between two plans, without retaliation of his opponents. (Property (D) above makes sure that he cannot substantially improve his payoff by such changes.)

The profile σ is a subgame-perfect ε-equilibrium: Consider an arbitrary subgame $G(h)$, after some history h, and a player t. Notice that, according to property (A), player t's payoff without deviation would be at least

$$\left(1 - \frac{\varepsilon}{2R}\right) \cdot \alpha_t^* + \frac{\varepsilon}{2R} \cdot (-R) \geqslant \alpha_t^* - \varepsilon. \tag{1}$$

We distinguish three cases for a single deviation by player t:

Case (i): t deviates in $G(h)$ by terminating. In this case, t's payoff is $r_t(t) \leqslant \alpha_t^*$, which is an improvement of at most ε.

Case (ii): t is the player who is supposed to execute a lottery between two plans, and t deviates by taking another probability distribution on the two actions that he should use. In this case, t's payoff can increase by at most ε due to property (D).

Case (iii): t deviates by playing a non-terminating action that he should not use at all, and by doing so he nominates player u. In this case, due to property (B), the new plans give player t at most α_t^*, and hence this deviation improves t's payoff by at most ε.

Now we examine what happens if t plans to deviate once more. For this purpose, suppose that, after t's first deviation, t appears again, and no termination has occurred yet. Then, t is not assigned to execute a lottery. Indeed, if t's first deviation belongs to case (ii) then no new lottery is planned yet, whereas if t's first deviation belongs to case (iii) then t is not assigned a new lottery due to property (D) above. Hence, t's second deviation would fall into a case similar to case (i) or case (iii) above, and t's payoff then would be at most α_t^*. In conclusion, in view of (1), player t's improvement is still at most ε.

Now it follows by induction, that if t decides to deviate only a finite number of times, then his payoff can only increase by ε.

It only remains to verify what happens if player t deviates infinitely often. This of course means that play does not terminate and player t receives payoff 0. If $\alpha_t^* < 0$, property (C) ensures that a plan will eventually be executed that does not contain player t. So, $\alpha_t^* \geqslant 0$ must hold. But then, by (1), player t can only gain at most $0 - (\alpha_t^* - \varepsilon) \leqslant \varepsilon$. This completes the proof. $\square$

A polynomial time algorithm: It is obvious from the iterative method that α^* can be calculated in polynomial time. Subsequently, the plans g_{tu}^1, g_{tu}^2 and h_{tu}^2 in the proof of Theorem 8 can also be determined in polynomial time. Hence, we obtain a polynomial time algorithm whose output directly yields a subgame-perfect ε-equilibrium for every $\varepsilon > 0$. Note that the algorithm itself is independent of ε.

4. Examples

In this section, we illustrate the construction with some examples. We start with an instructive example by Solan and Vieille, cf. Example 3 in [19].

Example 1. Consider the free transition game G which has two players, with respective payoff-vectors

$$r(1) = (-1, 2), \quad r(2) = (-2, 1).$$

The main idea of the game is that player 2 would like termination, and preferably at player 1, who is not easily inclined to terminate due to his negative payoff. Solan and Vieille provided the following subgame-perfect ε-equilibrium σ, for $\varepsilon \in (0,1)$: player 1 always terminates with probability 1 and player 2 always terminates with probability ε. Notice that player 1 is forced into termination, as trying to never terminate would eventually result in termination by player 2, which would give player 1 only payoff -2. Player 2, on the other hand, is only playing an ε-best response in σ, as he could improve his payoff by ε by not terminating at all. Solan and Vieille pointed out that this game does not admit a subgame-perfect 0-equilibrium.

Let us follow the iterative method given in Section 3.

Step 0: Initially, $\alpha^0 = (-1,1)$. Note that only the plans that eventually terminate at player 1, and $(2,2,\ldots)$ are α^0-viable.
First consider player 1. For termination we have $\beta_1(1,\alpha^0) = -1$, whereas moving to player 2 gives $\beta_1(2,\alpha^0) = -2$ due to the terminating plan $(2,2,\ldots)$. In conclusion for player 1, $\delta_1(\alpha^0) = -1 = \alpha_1^0$.
Now consider player 2. We have $\beta_2(1,\alpha^0) = 2$ and $\beta_2(2,\alpha^0) = 1$, hence $\delta_2(\alpha^0) = 2$. So, we update player 2 in step 1 and obtain $\alpha^1 = (-1,2)$.

Step 1: We continue with $\alpha^1 = (-1,2)$. The only α^1-viable plans, starting at either player, are the plans that eventually terminate at player 1. It is easy to verify that $\delta_1(\alpha^1) = -1 = \alpha_1^1$ and $\delta_2(\alpha^1) = 2 = \alpha_2^1$. This means that the iterative method terminates at step 1, with $\alpha^* = \alpha^1 = (-1,2)$. By following the construction in the proof of Theorem 8, we obtain exactly the subgame-perfect ε-equilibrium that Solan and Vieille found in this game.

Example 2. Consider the free transition game G which has two players, with respective payoff-vectors

$$r(1) = (-1,-2), \quad r(2) = (-2,-1).$$

There are two obvious subgame-perfect 0-equilibria. One is when each player always terminates with probability 1. Another, and a more natural one, is when no player ever terminates, inducing payoff 0 to both players.

Let us follow the iterative method given in Section 3.

Step 0: Initially, $\alpha^0 = (-1,-1)$. Note that only the following plans are α^0-viable: the plans $(1,1,\ldots)$, $(2,2,\ldots)$, and the two non-terminating plans $(1,2,1,2,\ldots)$ and $(2,1,2,1,\ldots)$.

First consider player 1. For termination we have $\beta_1(1,\alpha^0) = -1$, whereas moving to player 2 gives $\beta_1(2,\alpha^0) = -2$ due to the terminating plan $(2,2,\ldots)$. Hence, $\delta_1(\alpha^0) = -1 = \alpha_1^0$. Due to symmetry, we also obtain for player 2 that $\delta_2(\alpha^0) = -1 = \alpha_2^0$.

This means that the iterative method terminates at step 0, with $\alpha^* = \alpha^0 = (-1,-1)$. By following the construction in the proof of Theorem 8, we obtain both subgame-perfect 0-equilibria that we mentioned above (depending on the initially chosen α^*-viable plan).

5. Extensions

1. **Subgame-perfect 0-equilibria and subgame-perfect ε-equilibria in pure strategies:** Example 1, which is a game by Solan and Vieille, demonstrated that our main result cannot be extended in these directions.
2. **Stationary strategies:** Within the class of stationary strategies, subgame-perfect ε-equilibria fail to exist in certain free transition games. Consider the following example. There are 3 players, with payoff-vectors

$$r(1) = (2,1,4), \quad r(2) = (4,2,1), \quad r(3) = (1,4,2).$$

It is shown in Kuipers et al. [7] that this game has no stationary subgame-perfect 0-equilibrium, and similar arguments can be used to prove the non-existence of stationary subgame-perfect ε-equilibrium for small $\varepsilon > 0$.

3. **Games without free transitions:** It is an interesting question if the main result can be extended to similar games without the assumption of free transitions. Once the players are allowed to have more than one state and have restrictions in choosing the next state, the problem becomes more difficult. Because our

construction heavily relies on the transition structure of the game, it is evident that more advanced combinatorial techniques are needed to tackle such games.

6. Technical remarks

6.1. The distinction between $\beta_t^+(u, \alpha)$ and $\beta_t^-(u, \alpha)$

Notice that, for any $t, u \in N$ satisfying $t \neq u$ and for any $\alpha \in \Omega$, we have $\beta_t^+(u, \alpha) \leqslant \beta_t^-(u, \alpha)$. Moreover, if t is not α-anchored, i.e., $\alpha_t > -r_t(t)$, then no α-viable plan can terminate at t and hence $\beta_t^+(u, \alpha) = \beta_t^-(u, \alpha)$. Thus, regarding the iterative procedure in Section 3, the distinction between $\beta_t^+(u, \alpha)$ and $\beta_t^-(u, \alpha)$ is only relevant as long as t is an α-anchored player.

6.2. Monotonicity of $\delta_t(\cdot)$

The operator δ_t has the useful property that $\delta_t(\alpha) \leqslant \delta_t(\hat{\alpha})$ whenever $\alpha, \hat{\alpha} \in \Omega$ satisfy $\alpha \leqslant \hat{\alpha}$ in a coordinatewise sense. We will now prove this. Suppose that $\alpha \leqslant \hat{\alpha}$ for some $\alpha, \hat{\alpha} \in \Omega$. It suffices to prove that $\beta_t(u, \alpha) \leqslant \beta_t(u, \hat{\alpha})$ for every $u \in N \setminus \{t\}$. Take an arbitrary $u \in N \setminus \{t\}$. Recall that, as mentioned in the first concluding remark, $\beta_t^+(u, \alpha) \leqslant \beta_t^-(u, \alpha)$, and similarly for $\hat{\alpha}$. We distinguish three cases:

Case 1: $\beta_t(u, \alpha) = \beta_t^+(u, \alpha)$. In this case,

$$\beta_t(u, \alpha) = \beta_t^+(u, \alpha) \leqslant \beta_t^+(u, \hat{\alpha}) \leqslant \beta_t(u, \hat{\alpha}).$$

Case 2: $\beta_t(u, \hat{\alpha}) = \beta_t^-(u, \hat{\alpha})$. In this case,

$$\beta_t(u, \alpha) \leqslant \beta_t^-(u, \alpha) \leqslant \beta_t^-(u, \hat{\alpha}) = \beta_t(u, \hat{\alpha}).$$

Case 3:

$$\beta_t^+(u, \alpha) < \beta_t(u, \alpha) = \beta_t^-(u, \alpha) \tag{2}$$
$$\beta_t^+(u, \hat{\alpha}) = \beta_t(u, \hat{\alpha}) < \beta_t^-(u, \hat{\alpha}). \tag{3}$$

We prove that this case is impossible. In view of the first concluding remark, (2) and (3) imply $r_t(t) = \alpha_t = \hat{\alpha}_t$. Thus, by the definition of $\beta_t(u, \alpha)$ and $\beta_t(u, \hat{\alpha})$, we may conclude that t is not α-vulnerable and that t is $\hat{\alpha}$-vulnerable. Due to the latter, either V1 or V2 holds with respect to $\hat{\alpha}$.

Case 3.1 : V1 holds with respect to $\hat{\alpha}$. In this case, as $V_u(\alpha) \supseteq V_u(\hat{\alpha})$ and $\alpha_t = \hat{\alpha}_t$, condition V1 also holds with respect to α. This is a contradiction, as t is not α-vulnerable.

Case 3.2 : V2 holds with respect to $\hat{\alpha}$. In this case, as $\alpha \leqslant \hat{\alpha}$ and $\alpha_t = \hat{\alpha}_t$, condition V2 also holds with respect to α. This is a contradiction, as t is not α-vulnerable.

6.3. The uniqueness of α^*

For a free transition game G, let $\Delta(G)$ denote the set of stable vectors $\alpha \in \Omega$ such that $\delta_t(\alpha) = \alpha_t$ for all $t \in N$. In Lemma 7 we proved, for every G, that $\Delta(G)$ is non-empty, which was the crucial step in the construction of a subgame-perfect ε-equilibrium for G. We remark that $\Delta(G)$ is not necessarily a singleton, as the following example demonstrates. Consider the free transition game $\widehat{G}$ which has only two players, with respective payoff-vectors

$$r(1) = (1, 2), \quad r(2) = (2, 1).$$

We argue that $(1, 1), (1, 2) \in \Delta(\widehat{G})$. First, $\alpha = (1, 1)$ and $\alpha' = (1, 2)$ are both stable, because the plan $(1, 1, \ldots)$ starting at player 1 is both α-safe and α'-safe, whereas the plan $(2, 1, 1, \ldots)$ starting at player 2 is both α-safe and α'-safe. Since it is straightforward that $\delta_t(\alpha) = \alpha_t$ and $\delta_t(\alpha') = \alpha'_t$ for all $t = 1, 2$, we have $(1, 1), (1, 2) \in \Delta(\widehat{G})$ indeed. In fact, due to symmetry, $(2, 1) \in \Delta(\widehat{G})$ holds as well.

Take an arbitrary free transition game G. All inequalities between two vectors below are meant coordinatewisely, just as the terms non-decreasing sequence and minimal element. Recall that the iterative procedure in the proof of Lemma 7 produces a non-decreasing sequence of stable vectors, starting with $r = (r_t(t))_{t \in N}$ and terminating with $\alpha^* \in \Delta(G)$. We will now show that $\alpha^* \leqslant \hat{\alpha}$ for every $\hat{\alpha} \in \Delta(G)$, which means that α^* is the minimal element of $\Delta(G)$. Let $\hat{\alpha} \in \Delta(G)$. Notice that if α is stable and $\alpha \leqslant \hat{\alpha}$, then $\delta_t(\alpha) \leqslant \delta_t(\hat{\alpha}) = \hat{\alpha}_t$ holds because of the second concluding remark. This means that the iterative procedure only generates stable vectors that are at most $\hat{\alpha}$. In particular, $\alpha^* \leqslant \hat{\alpha}$ indeed. Thus, α^* is the minimal element of $\Delta(G)$ as claimed.

As a consequence, the iterative procedure in Section 3 terminates at the same α^*, irrespectively of the order in which the players are updated. The only important point is that at every step a player is updated whose coordinate strictly increases by this update.

References

[1] F. Bloch, Sequential formation of coalitions in games with externalities and fixed payoff division, Games and Economic Behavior 14 (1996) 90–123.
[2] B. Bueno de Mesquita, D. Lalman, Domestic opposition and foreign war, The American Political Science Review 84 (1990) 747–765.
[3] J. Flesch, J. Kuipers, G. Schoenmakers, K. Vrieze, Subgame perfection in positive recursive games with perfect information, Mathematics of Operations Research 35 (2010) 193–207.
[4] J. Flesch, J. Kuipers, A. Mashiah-Yaakovi, G. Schoenmakers, E. Solan, K. Vrieze, Perfect-information games with lower-semi-continuous payoffs, Mathematics of Operations Research 35 (2010) 742–755.
[5] J. Flesch, J. Kuipers, A. Mashiah-Yaakovi, G. Schoenmakers, E. Shmaya, E. Solan, K. Vrieze, A perfect-information game without subgame-perfect ε-equilibria, preprint (2011).
[6] T. Kim, Y. Jeon, Stationary perfect equilibria of an n-person noncooperative bargaining game and cooperative solution concepts, European Journal of Operational Research 194 (2009) 922–932.
[7] J. Kuipers, J. Flesch, G. Schoenmakers, K. Vrieze, Pure subgame-perfect equilibria in free transition games, European Journal of Operational Research 199 (2008) 442–447.
[8] A. Mashiah-Yaakovi, Periodic stopping games, International Journal of Game Theory 38 (2009) 169–181.
[9] A. Mashiah-Yaakovi, Subgame-perfect equilibrium in stopping games with perfect information, preprint (2009).
[10] J.-F. Mertens, Repeated games, in: Proceedings of the International Congress of Mathematicians, American Mathematical Society, Providence, RI, 1987, pp. 1528–1577.
[11] A. Muthoo, On the strategic role of outside options in bilateral bargaining, Operations Research 43 (1995) 292–297.
[12] M.J. Osborne, A. Rubinstein, Bargaining and Markets, Academic Press, 1990.
[13] R.A. Purves, W.D. Sudderth, Perfect information games with upper-semicontinuous payoffs, preprint (2010).
[14] D. Ramsey, Partnership formation based on multiple traits, European Journal of Operational Research 216 (2012) 624–637.
[15] R. Selten, Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit, Zeitschrift für die gesamte Staatswissenschaft 121 (1965) 301–324, 667–689.
[16] R. Selten, A simple model of imperfect competition, where four are few and six are many, International Journal of Game Theory 2 (1973) 141–201.
[17] R. Selten, The chain store paradox, Theory and Decision 9 (1978) 127–159.
[18] E. Solan, Subgame-perfection in quitting games with perfect information, Mathematics of Operations Research 30 (2005) 51–72.
[19] E. Solan, N. Vieille, Deterministic multi-player Dynkin games, Journal of Mathematical Economics 39 (2003) 911–929.
[20] F. Thuijsman, T.E.S. Raghavan, Perfect information stochastic games and related classes, International Journal of Game Theory 26 (1997) 403–408.
[21] H. Vartiainen, Dynamic coalitional equilibrium, Journal of Economic Theory 146 (2011) 672–698.
[22] J. Vidal-Puga, Forming coalitions and the Shapley NTU value, European Journal of Operational Research 190 (2008) 659–671.
[23] E. Weg, R. Zwick, Toward the settlement of the fairness issues in ultimatum games, Journal of Economic Behavior and Organization 24 (1994) 19–34.