Attainability in Repeated Games with Vector Payoffs*

Dario Bauso[†], Ehud Lehrer[‡], and Eilon Solan[§]

July 30, 2012

Abstract

Motivated by problems from control theory we introduce the concept of attainable sets of payoffs in two-player repeated games with vector payoffs. A set of payoff vectors is *attainable* by Player 1 if he can ensure that there is a finite horizon T such that after time T the distance between the set and the cumulative payoff is arbitrarily small, regardless of the strategy Player 2 implements. This paper focuses on the case where the attainable set consists of a single payoff vector, called an attainable vector. We study properties of the set of attainable vectors, and characterize when a specific vector is attainable and when every vector is attainable.

Keywords: Attainability, repeated games with vector payoffs, dynamic games, approachability.

JEL classification: C73, C72

*This research was supported in part by the Google Inter-university center for Electronic Markets and Auctions. Lehrer acknowledges the support of the Israel Science Foundation, Grant #762/045. Solan acknowledges the support of the Israel Science Foundation, Grant #212/09.

[†]Dipartimento di Ingegneria Chimica, Gestionale, Informatica, Meccanica, Università di Palermo, viale delle Scienze 90128, Palermo, Italy. e-mail: dario.bauso@unipa.it.

[‡]The School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel and INSEAD, Boulevard de Constance, 77305 Fontainebleau, France. e-mail: lehrer@post.tau.ac.il.

[§]The School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel. e-mail: eilons@post.tau.ac.il.

1 Introduction

There are various dynamic situations in which the stage payoff is multi-dimensional and the goal of the decision maker is to drive the total vector payoff as close as possible to a given target set. One such example is dynamic network models, which include a variety of different logistic applications such as production, distribution and transportation networks. In the literature on dynamic network flow control [2, 3, 8, 9, 17], the supplier tries to meet a multi-dimensional demand. His goal is to ensure that the difference between the *total* demand and the *total* supply converges with time to a desirable target. One can model such a situation as a two-player repeated game, where Player 1 is the decision maker, and Player 2 represents the adversarial market that controls the demand. For instance, in the distribution network scenario, the supplier has a desirable multi-dimensional inventory level that he would like to maintain in his warehouse, despite erratic behavior of the demand side. Having to deal with an adversarial opponent requires the supplier to cope with the worst case possible. This motivates our main objective: to find conditions that characterize when a specific target vector x can be attained against any possible demand pattern exhibited by the market.

A second example is the Capital Adequacy Ratio. The third Basel Accord states that (a) the bank's Common Equity Tier 1 must be at least 4.5% of its risk-weighted assets at all times, (b) the bank's Tier 1 Capital must be at least 6.0% of its risk-weighted assets at all times, and (c) the total capital, that is, Tier 1 Capital plus Tier 2 Capital, must be at least 8.0% of the bank's risk weighted assets at all times. To accommodate this example in our setup, consider the following 3-dimensional vector. The first coordinate stands for the per-period difference between the bank's Tier 1 Capital and 6.0% of its risk-weighted assets; the second coordinate stands for the per-period difference between the bank's Tier 1 Capital and 6.0% of its risk-weighted assets; and the third stands for the per-period difference between the total capital and 8.0% of the bank's risk weighted assets. According to the Capital Adequacy Ratio the coordinates of this vector should be nonnegative. Here, Player 1 represents the bank's managers who control its assets, and Player 2 represents market behavior, which is unpredictable and thought of as adversarial. The goal of Player 1 is to design a strategy that would drive the 3-dimensional total payoff to the target set, which is the nonnegative orthant. In practice, to ensure that they fulfill the requirements of the Basel Accord, banks try to hold a capital buffer on top of the regulatory minimum, and they periodically adjust their assets to be at the top of the buffer [15, 27].

To model such situations we study two-player repeated games with vector payoffs in continuous time. We say that a set A in the payoff space is *attainable* by Player 1 if there is a time T such that for every level of proximity, $\varepsilon > 0$, Player 1 has a strategy guaranteeing that against every possible strategy of Player 2, the distance between A and the cumulative payoff up to any time t greater than T is smaller than ε. When a set A is attainable, Player 1 can plan his actions, based on historical inventory and market data, to ensure that the inventory level would converge to A. In the case where A consists of one point, the supplier may guarantee that his inventory level will be as close as he wishes to an ideal level.

In the game that we discuss, players are allowed to use a special type of behavior strategies. These strategies are characterized by an increasing sequence of positive real numbers, that divide the time span $[0, \infty)$ into subinterval. The play of the player in each interval depends on the play of the other player *before* the interval starts, but is independent of the play of the other player during that time interval. This is equivalent to saying that before the game starts, a player sets an alarm clock to ring in certain times, and whenever the clock rings, the player looks at the historical play path up to that point and determines how to play until the next time the clock rings. In Section 3.1 we discuss the interpretation of this type of strategies.

The definition of attainability is close in spirit to the concept of approachable sets [6], which refers to the average stage payoff rather than the cumulative payoff. While a set A is attainable by Player 1 if he can ensure that the cumulative payoff converges to it, A is approachable by him if he can ensure that the average payoff converges to the set.

In case a point, say x, is attainable by Player 1, it implies that the long run inventory level is stable around x. On the other hand, if x is approachable, it merely guarantees that the average payoff converges to x. This may happen even when the actual payoff does not converge to x, and, in fact, even when any fixed running average does not converge to x. This observation suggests that although the notions of attainability and approachability look similar, and indeed are related, the flavor of the results and their proofs are completely different. For example, in the setup of supply networks, approachability theory can control the average stock, yet it cannot control cases where the warehouse becomes full or empty, in which the supplier might suffer a substantial loss. Attainability theory, on the other hand, controls the amount of stock in the warehouse, and does not concern itself with the average stock.

When a vector x is attainable by a player, the player can ensure that the total payoff converges to that vector. In particular, he can ensure that the total payoff gets close to x, and remains around x forever. This implies that once some vector x is attainable, so is the vector $\vec{0}$. The first main result, Theorem 1, characterizes when $\vec{0}$ is attainable. The result implies that, with regard to the vector $\vec{0}$, attainability and approachability coincide: the vector $\vec{0}$ is attainable by a player if and only if it is approachable by him in the game in discrete time.

The second main result, Theorem 2, characterizes when a given vector x is attainable. It turns out that a vector x is attainable by a player if and only if (i) the vector $\vec{0}$ is attainable by him, and (ii) the vector $\vec{0}$ is attainable by him in a game whose payoffs are translated by δx for some $\delta > 0$. Condition (i) is shared by all x: in order for any x to be attainable, the vector $\vec{0}$ must be attainable. Condition (ii), on the other hand, is point specific, and it uses the attainability of $\vec{0}$ in a modified version of the original game.

The last result, Theorem 3, characterizes the cases in which the decision maker has full freedom in setting his target level, in the sense that any vector x is attainable.

Related control and approachability literature. We highlight two main streams of related literature, one from the control area and the second from the approachability area. These two bodies of literature have in common the interest towards robustness.

Connections between robust control and noncooperative game theory has a long history (see, e.g., [1]). Robust control is the area of control theory that looks for strategies that "control" the state of a dynamical system, for instance, drive it to a given set, despite the effects of disturbances (see the seminal paper [5]). Among the foundations of robust control we find two main notions that can be related to attainability and are surveyed in [7]. The first notion, *robust global attractiveness*, refers to the property of a set to "attract" the state of the system under a proper control strategy, independently of the effects of the disturbance. The second notion, *robustly controlled invariance*, describes the property of a set to bound the state trajectory under a proper control strategy, independently of the effects of the disturbance. Both notions are widely exploited in a variety of works that contribute to the use of robust control in dynamic network flow models [2, 3, 4, 8, 9, 17].

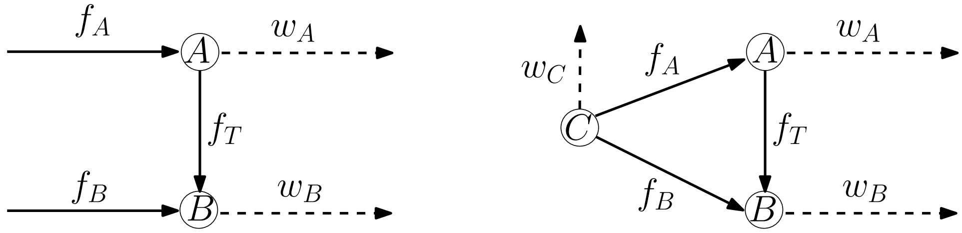
With regard to the second stream of literature, the definition of attainability is close in spirit to the concept of approachability (see [6, 16, 19, 21, 22, 23, 24, 25, 26]). The two concepts differ in that approachability aims at controlling the average payoff whereas attainability aims at controlling the cumulative payoff. Approachability theory has been applied in several areas of game theory, such as allocation processes in coalitional games [18], regret minimization [14, 20], and adaptive learning [10, 11, 12, 13], just to name a few. Approachability theory has been originally formulated for discrete-time repeated games with finite-dimensional payoffs. Since then it has been extended to infinite-dimensional payoffs [19] and to continuous-time repeated games, showing common elements with Lyapunov theory [14].

In between the two aforementioned streams, there is a literature on decision problems related to dynamic multi-inventory in continuous time (see for instance, the continuous-time control strategy in [8]). The control literature up to this point refers to one-person (the controller) decision problems facing uncertainty. As far as we know, this paper is the first taking a strategic approach to the problem.

Organization. This paper is organized as follows. In Section 2 we provide motivating examples. In Section 3 we introduce the model and main definitions. In Section 4 we introduce and discuss the results. Section 5 is devoted to the discussion of a few aspects related to the definition of attainability and to the type of continuous strategies that we are using. Proofs are relegated to Section 6.

2 A motivating example

This section details one motivation of our study, namely, distribution networks. Consider a distributer of a certain product who has two warehouses A and B in different regions. Every month the distributer can order products from factories to each of the warehouses, and he can transport products between the two warehouses, while vendors order products from the warehouses. This situation is described graphically in Figure 1(a).



(a) Three distribution flows f_A, f_T, f_B and two vendors requests w_A, w_B.

(b) Factory manager can sell directly to vendors: node C represents factory.

Figure 1: Distribution network with warehouses A and B.

In Figure 1(a), f_A and f_B are the number of products that are sent from factories to the two warehouses A and B respectively, f_T is the number of products that are transported from warehouse A to warehouse B, and w_A and w_B are the number of products sent from the two warehouses to vendors. Negative flows are interpreted as flows in the opposite direction; e.g., if vendors return products to warehouse A (resp. to warehouse B), then w_A (resp. w_B) is negative; if products are transported from warehouse B to warehouse A, then

f_T is negative. We analyze this situation in continuous time. The change of stock in the two warehouses is given by the 2-dimensional vector

$$u(a_1^t, a_2^t) = \underbrace{\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix}}_{F} \underbrace{\begin{pmatrix} f_A^t \\ f_T^t \\ f_B^t \end{pmatrix}}_{a_1^t} - \underbrace{\begin{pmatrix} w_A^t \\ w_B^t \end{pmatrix}}_{a_2^t},$$

where $a_1^t = (f_A^t, f_B^t, f_T^t)$ is the decision variable of the distributer at time t, and $a_2^t = (w_A^t, w_B^t)$ is the uncontrolled market demand at time t.

Suppose that the number of products that can be ordered by vendors at each time instance is bounded by 2, and the number of products that can be returned by vendors to each warehouse at every time instance is 3. In other words, w_A^t and w_B^t are in $[-3, 2]$. Suppose also that the amount of products that the distributer can order from or return to the factories and transport between the two warehouses is bounded by 5.

This situation can be described by a two-person game as follows. Player 1 represents the distributer and has 8 actions, each of which is given by a 3-dimensional vector

$$(5,5,5), (5,5,-5), (5,-5,5), (5,-5,-5), (-5,5,5), (-5,5,-5), (-5,-5,5), (-5,-5,-5).$$

These 8 actions are the extreme points of set of decisions the distributer can take, so that the set of mixed actions corresponds to the set of all decisions the distributer can take at any given time. Player 2 represents the market demand or nature and has 4 actions, each of which is given by a 2-dimensional vector

$$(-3,-3), (-3,2), (2,-3), (2,2).$$

These 4 actions are the extreme points of the set of possible demands from the two warehouses, so that the set of mixed actions of Player 2 corresponds to the set all possible market demands in each given time. The payoffs correspond to the change of stock in the two warehouses, and are given by the following table:

	(-3,-3)	(-3,2)	(2,-3)	(2,2)
(5,5,5)	(3,13)	(3,8)	(-2,13)	(-2,8)
(5,5,-5)	(3,3)	(3,-2)	(-2,3)	(-2,-2)
(5,-5,5)	(13,3)	(13,-2)	(8,3)	(8,-2)
(5,-5,-5)	(13,-7)	(13,-12)	(8,-7)	(8,-12)
(-5,5,5)	(-7,13)	(-7,8)	(-12,-13)	(-12,8)
(-5,5,-5)	(-7,3)	(-7,-2)	(-12,3)	(-12,-2)
(-5,-5,5)	(3,3)	(3,-2)	(-2,3)	(-2,-2)
(-5,-5,-5)	(3,-7)	(3,-12)	(-2,-7)	(-2,-12)

Figure 2: The strategic-form game corresponding to the situation.

At every time instance the two players choose their actions. Each market behavior translates into a mixed action of Player 2, and each behavior of the distributer corresponds to a mixed action of Player 1. The (2-dimensional) total payoff up to time t is the number of products that are stored in each of the two warehouses at time t. The goal of the distributer is to ensure that the total number of products in each warehouse does not exceed its capacity, that is, that the total payoff does not exceed a certain (2-dimensional) bound.

Figure 1(b) describes the case where the factory manager can sell directly to vendors, bypassing the distribution to warehouses. This situation can be represented by adding an additional node C modeling the factory, and an edge that represents direct sales from the factory to vendors. The stock is now a 3-dimensional vector, as we have to take into account the inventory available at the factory, and consequently the change in the stock modifies as shown below:

$$u(a_1^t, a_2^t) = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & -1 \end{pmatrix} \begin{pmatrix} f_A^t \\ f_T^t \\ f_B^t \end{pmatrix} - \begin{pmatrix} w_A^t \\ w_B^t \\ w_C^t \end{pmatrix}.$$

A recurrent question in the network flow control literature [3, 2, 8, 9, 17] is finding conditions that ensure the existence of a control strategy that drives the excess supply vector to a desired target level in $\mathbb{R}^m$ regardless of the unpredictable realization of the demand. The equivalence between the excess supply and the cumulative payoff in the dynamic game motivates our study. The rest of the paper is devoted to the analysis of conditions under which Player 1 has a strategy ensuring the attainability of a given point, regardless of the behavior of Player 2.

Situations where the goal is to control the total payoff occur also in production and transportation networks. Production networks describe production processes and activities necessary to turn raw materials into intermediate products and eventually into final products. The nodes of the networks represent raw materials and intermediate/final products. The buffer at each single node i models the amount of material or product of type i currently stored or produced at that node, and hyper-arcs describe the materials or products consumed (tail nodes) and produced (head nodes) in each activity or process. Transportation networks model the flow of commodities, information, or traffic; nodes of the networks represent hubs and the buffers at the nodes describe the quantity of commodities present in the hubs. The edges describe transportation routes.

3 Attainability

In the first part of this section we introduce the mathematical model of repeated game in continuous time and elaborate on the type of strategies used by the players. In the remaining part we provide a formal definition of attainability.

3.1 The model

We study a two-player repeated game with vector payoffs in continuous time Γ. The set of players is $N = \{1, 2\}$, and the finite set of actions of each player i is A_i. The instantaneous payoff is given by a function $u : A_1 \times A_2 \to \mathbb{R}^m$, where m is a natural number. We assume w.l.o.g. that payoffs are bounded by 1, so that $u : A_1 \times A_2 \to [-1, 1]^m$. We extend u to the set of mixed-action pairs, $\Delta(A_1) \times \Delta(A_2)$, in a bilinear fashion. The one-shot vector-payoff game (A_1, A_2, u) is denoted by G and we will say that the game in continuous time Γ is *based on* G. If $i \in \{1, 2\}$, then $-i$ denotes the player who is not i.

The game Γ is played over the time interval $[0, \infty)$. We assume that the players use nonanticipating behavior strategies with delay, which we define now. Roughly, a nonanticipating behavior strategy with delay divides time into blocks. The behavior of a player in a given block depends on the behavior of the other player up to the beginning of that block. In other words, the way a player plays during a given block of time does not affect the way the other player plays during that block. Still, it may affect the other player's play in subsequent blocks.

Denote by $\mathcal{C}_i$ the set of all *controls* of player i, that is, the set of all measurable functions from the time space, $[0, \infty)$, to player i's mixed actions.

$$\mathcal{C}_i := \{a_i : [0, \infty) \to \Delta(A_i),\ a_i \text{ is measurable}\}.$$

Definition 1 *A function $\sigma_i : \mathcal{C}_{-i} \to \mathcal{C}_i$ is a* behavior strategy with delay *(or simply a* strategy*) for player i, if there exists an increasing sequence of real numbers $(\tau_i^k)_{k\in\mathbb{N}}$ such that for every $a_{-i}, a'_{-i} \in \mathcal{C}_{-i}$ and every $k \in \mathbb{N}$,*

$$a_{-i}(t) = a'_{-i}(t) \quad \forall t \in [0, \tau_i^k) \quad \Longrightarrow (\sigma_i(a_{-i}))(t) = (\sigma_i(a'_{-i}))(t) \quad \forall t \in [0, \tau_i^{k+1}),$$

where $\tau_i^0 = 0$.

Remark 1 *A strategy as we defined here is called a* nonanticipating strategy with delay *in the literature of differential games. An equivalent formulation, that may look more transparent to game theorists, is as follows. A strategy of player i is a vector $(\tau_i^k, \sigma_i^k)_{k\in\mathbb{N}}$ where $(\tau_i^k)_{k\in\mathbb{N}}$ is an increasing sequence of real numbers, and for each $k \in \mathbb{N}$, σ_i^k is a function that maps play paths (of both players) on the interval $[0, \tau_i^{k-1})$ to plays of player i in the interval $[\tau_i^{k-1}, \tau_i^k)$.*

When defining a strategy and when referring to a strategy we will usually take the equivalent formulation given in this remark.

In the sequel we will refer to the real numbers $(\tau_i^k)_{k\in\mathbb{N}}$ in Definition 1 as *the updating times related to σ_i.*

Remark 2 *Continuous time is usually used as a convenient model for discrete time, when the gap between two stages is small. This is the case here as well. Suppose that time is discrete, and the time difference between any two decision moments is extremely small. Suppose that observation of the actions of the other player is time consuming and possibly costly, so that players cannot observe each other at every decision point. Thus, the players are in fact playing a game in discrete time, in which they can randomize at every decision point, but they observe the actions of the other player only rarely, relative to the frequency in which they take actions. By improving the observation technology a player can observe the actions of the other player more frequently, but this frequency will always be significantly slower than the frequency in which actions are taken.*

Every pair of strategies $\sigma = (\sigma_1, \sigma_2)$ uniquely determines a play path $(a^t(\sigma))_{t\in\mathbb{R}_+}$. The payoff (vector) up to time T associated with the pair of strategies σ is given by

$$\gamma^T(\sigma) = \int_0^T u(a^t(\sigma))\mathrm{d}t \in \mathbb{R}^m. \tag{1}$$

When we wish to emphasize that the payoff is in the game based on G we will write γ_G^T rather than γ^T. Since payoffs are bounded by 1, the integral in (1), which is the cumulative payoff up to time T, is well-defined.

3.2 Attainability: the definition

The subject matter of this paper is the concept of attainable sets: a set of vectors is attainable by a player if he can guarantee that the distance between the set and the total payoff converges to 0, regardless of the strategy implemented by the other player. We provide here one definition for this concept. Two alternative definitions are given and discussed in Section 5.

Definition 2 *The set $Y \subseteq \mathbb{R}^m$ is* attainable *by Player 1 if there is $T > 0$ such that for every $\varepsilon > 0$ there is a strategy σ_1 of Player 1 such that*[1]

$$d(\gamma^t(\sigma_1, \sigma_2), Y) \leq \varepsilon, \quad \forall t \geq T, \forall \sigma_2.$$

A set Y is attainable if there is a finite horizon T such that Player 1 can ensure, against any possible strategy of Player 2, that the distance between Y and the cumulative payoff up to any time $t \geq T$ is at most ε. Note that the time T is uniform across all levels of precision. That is, in order for Y to be attainable by Player 1, Player 1 must be able to guarantee that the distance between Y and the cumulative payoff at any time longer than T is at most ε. However, different ε's might require different strategies employed by Player 1. It might therefore happen that although Y is attainable by Player 1, the cumulative payoff would never touch Y itself. We say that the strategy σ_1 in Definition 2 *attains the set Y up to ε.*

The definition of an attainable set looks close in spirit to that of approachable set in games played over discrete set of times [6]. There is, however, a significant difference between the two concepts. A set is approachable if the average payoff converges to it, while a set is attainable if the cumulative payoff converges to the set. In other words, approachability refers to the convergence of the *average* payoff, while attainability to the convergence of the *cumulative* payoff. Indeed, the results characterizing approachable sets and attainable sets are significantly different, both in contents and in spirit.

When the set attainable by Player 1, Y, contains a single vector x, we say that the vector x is attainable by Player 1. Denote by W the set of attainable vectors.

[1]The distance and the norm referred to throughout the paper is $\|.\|_2$. For instance, $d(x, A) = \min_{y \in Y} \|x - y\|_2$.

4 Results

This section presents the three main results of our study. The first result, Theorem 1, focuses on the conditions under which the vector $\vec{0}$ is attainable. Attainability of $\vec{0}$ turns out to be crucial if we wish to inspect attainability of a single vector $x \neq \vec{0}$, and this constitutes our second result (see Theorem 2). The third result builds upon the previous two results, and provides a stronger condition that ensures that any vector $x \in \mathbb{R}^m$ is attainable.

We start with a simple observation regarding the notion of attainability, which holds due to the continuous time setup. Its proof is differed to the last section.

Proposition 1 *The set W is a convex cone.*

The next theorem characterizes when the vector $\vec{0}$ is attainable. To state this result we need the following notation. Let $\lambda \in \mathbb{R}^m$. Denote[2] by $\langle \lambda, G \rangle$ the zero-sum one-shot game whose set of players and their action sets are as in the game G, and the payoff that Player 2 pays to Player 1 is $\langle \lambda, u(a_1, a_2) \rangle$ for every $(a_1, a_2) \in A_1 \times A_2$. As a zero-sum one-shot game, the game $\langle \lambda, G \rangle$ has a value, denoted v_λ.

For every mixed action $p \in \Delta(A_1)$ denote

$$D_1(p) = \{u(p, q) \colon q \in \Delta(A_2)\}.$$

$D_1(p)$ is the set of all payoffs that might be realized when Player 1 plays the mixed action p. If $v_\lambda \geq 0$ (resp. $v_\lambda > 0$), then there is a mixed action $p \in \Delta(A_1)$ such that $D_1(p)$ is a subset of the closed half space $\{x \in \mathbb{R}^m \colon \langle \lambda, x \rangle \geq 0\}$ (resp. half space $\{x \in \mathbb{R}^m \colon \langle \lambda, x \rangle > 0\}$).

Theorem 1 *The following conditions are equivalent.*

B1 *The vector $\vec{0} \in \mathbb{R}^m$ is attainable by Player 1;*

B2 $v_\lambda \geq 0$ *for every $\lambda \in \mathbb{R}^m$.*

Corollary 2 in [6] implies that the vector $\vec{0}$ is approachable in the game in discrete time with payoff function u if and only if condition **B2** holds. We thus deduce the following corollary to Theorem 1.

Corollary 1 *The vector $\vec{0} \in \mathbb{R}^m$ is attainable by Player 1 in Γ if and only if it is approachable by Player 1 in the repeated game in discrete time with payoff function u.*

[2] The inner product is defined by $\langle x, y \rangle := \sum_{i=1}^m x_i y_i$ for every $x, y \in \mathbb{R}^m$.

The following result characterizes when a given vector x is attainable. We will need the following notation. For every $y \in \mathbb{R}^m$ denote by $(G - y)$ the two-player one-shot game that is identical to G except for its payoff function. The payoff function of $(G - y)$ is $(u - y)$, where $(u - y)(a_1, a_2) = u(a_1, a_2) - y$ for every $a_1 \in A_1$ and $a_2 \in A_2$.

Theorem 2 *Let $\vec{0} \neq x \in \mathbb{R}^m$. The vector x is attainable by Player 1 if and only if*

B1 *The vector $\vec{0} \in \mathbb{R}^m$ is attainable by Player 1*

and either one of the following conditions holds:

B3 *There is $\delta_0 > 0$ such that for every $q \in \Delta(A_2)$ there is $p \in \Delta(A_1)$ and $\delta > \delta_0$ satisfying $u(p, q) = \delta x$.*

B4 *There is $\delta > 0$ such that the vector $\vec{0} \in \mathbb{R}^m$ is attainable by Player 1 in the game based on $(G - \delta x)$.*

Theorem 2 implies that whenever any vector x is attainable, so is the vector $\vec{0}$. Since attainability is concerned with the cumulative payoff, once a target level is (almost) reached, this level should be maintained in the long run. This means that once a neighborhood of a target level x is reached, from that point on the level $\vec{0}$ ought to be attained. This is the reason why $\vec{0}$ is attainable when any vector x is, and why $\vec{0}$ plays a major role in the theory of attainability. However, condition **B1** alone is not sufficient for the attainability of other vectors.

Condition **B3** states that for every $q \in \Delta(A_2)$ there is a strategy of Player 1 such that the payoff $u(p, q)$ is x multiplied by a scalar δ, which is bounded away from 0. Condition **B4** states that $\vec{0}$ is attainable in some game whose payoff is a translation of the original payoff function by δx.

The following theorem deals with the case where all the vectors are attainable.

Theorem 3 *The following statements are equivalent:*

C1 *Every vector $x \in \mathbb{R}^m$ is attainable by Player 1;*

C2 $v_\lambda > 0$ *for every* $\lambda \in \mathbb{R}^m \setminus \vec{0}$.

Remark 3 *If condition **C2** is satisfied, then for every open half space H of $\mathbb{R}^m$ there is a mixed action $p \in \Delta(A_1)$ such that $D_1(p) \subseteq H$. Standard continuity and compactness arguments imply that in this case there is $\delta_1 > 0$ such that for every half space H there is $p \in \Delta(A_1)$ satisfying $d(D_1(p), H) \geq \delta_1$. Stated differently, there is $\delta_2 > 0$ such that for every vector λ whose ℓ_1-norm is 1, $\langle \lambda, u(p, q) \rangle > \delta_2$ for every $q \in \Delta(A_2)$.*

Note the difference between condition **B2** of Theorem 1 and condition **C2** of Theorem 3. In the former, the value of the scalar-payoff game with payoffs $\langle \lambda, u(p,q) \rangle$ is nonnegative for every direction $\lambda \in \mathbb{R}^m \setminus \vec{0}$, while in the latter it is strictly positive. The former guarantees attainability of the vector $\vec{0}$, while the latter guarantees that every vector is asymptotically attainable.

5 Discussion

5.1 Continuous time versus discrete time.

The characterization presented in Theorem 1 depends crucially on the continuous time setting. The following example shows that this result is invalid when time is discrete.

Example 1 Consider a game in *discrete* time where payoffs are one-dimensional and each player has two actions. Payoffs are given by the following matrix:

	L	R
U	$-2-1$	$-2+1$
B	$2-1$	$2+1$

$=$

	L	R
U	-3	-1
B	1	3

Figure 3: The payoff function in Example 1.

The payoffs in this game are the sum of two numbers, one is determined by Player 1 (-2 if he plays U, 2 if he plays B), and the other by Player 2 (-1 if she plays L, 1 if she plays R).

Condition **B2** is satisfied, and therefore 0 is attainable by Player 1. The following strategy guarantees that the distance between 0 and the cumulative payoff up to time t is[3] at most $9 \cdot 2\eta$, for any $t > 2$; the details of the proof can be found in the proof of Theorem 1. Divide the time line into countably many blocks, where the length of the k-th block is $\frac{\eta}{k}$. In the k-th block Player 1 plays U if the cumulative payoff at the beginning of the block is positive, and he plays B otherwise.

We show that 0 is not attainable by Player 1 in the game in discrete time. When time is discrete, a behavior strategy of a player is a function that assigns a mixed action to each past history. For every $\ell \in \mathbb{N}$ let p^ℓ be the mixed action played by Player 1 at stage ℓ. The mixed action p^ℓ depends on past play. Let σ_2 be the strategy that at each stage ℓ plays L if $p^\ell(U) \geq \frac{1}{2}$, and R otherwise. The stage payoff is then at least 2 whenever Player 2 plays

[3]The extra 9 appears because in this example the payoffs are not bounded by 1, and $3^2 = 9$.

R, and at most -2 whenever Player 2 plays L. In particular, if the total payoff up to stage ℓ is in the interval $[-\frac{1}{2}, \frac{1}{2}]$, then the payoff up to stage $\ell+1$ lies outside this interval. Thus, the cumulative payoff does not converge to 0. ■

Example 1 suggests that the characterization of the set of attainable vectors in games in discrete time is more challenging that the characterization in continuous time.

5.2 Alternative definitions of attainability

We introduce two alternative definitions of the concept of attainability, termed asymptotic attainability and weak asymptotic attainability. We then explore some relations between the definition of attainability discussed above and the two additional definitions.

For every set $Y \subseteq \mathbb{R}^m$ we denote by $B(Y, \varepsilon)$ the set of all points whose distance from at least one point in Y is less than ε, that is,

$$B(Y, \varepsilon) := \{x \in \mathbb{R}^m \colon d(x, Y) < \varepsilon\}.$$

When Y is a single point x, we write $B(x, \varepsilon)$ instead of $B(\{x\}, \varepsilon)$.

Definition 3 *(i) The set $Y \subseteq \mathbb{R}^m$ is* asymptotically attainable *by Player 1 if there is a strategy σ_1 for Player 1 such that for every strategy σ_2 of Player 2,*

$$\lim_{T \to \infty} d(\gamma^T(\sigma_1, \sigma_2), Y) = 0. \tag{2}$$

(ii) The set Y is weakly asymptotically attainable *by Player 1, if the set $B(Y, \varepsilon)$ is asymptotically attainable by Player 1 for every $\varepsilon > 0$.*

Asymptotic attainability requires that a set is asymptotically reached by the cumulative payoff without putting any bound on the time it takes to reach the set. Attainability, on the other hand, requires that a set is approximately reached in a bounded time, independent of the degree of approximation. This property of the notion of attainability seems to us most appealing for applications, which is why we adopted it throughout the paper.

Weak asymptotic attainability relaxes both time boundedness and the level of the approximation precision. A set Y is weakly asymptotically attainable if any neighborhood $B(Y, \varepsilon)$ of Y can be asymptotically attained, without having a universal bound on the time at which this neighborhood is reached.

Any attainable set is also weakly asymptotically attainable and any asymptotically attainable set is weakly asymptotically attainable as well. Analogously to Proposition 1, the set of asymptotically attainable vectors and the set of weakly asymptotically attainable vectors are convex cones. The definition implies that the set of weakly attainable vectors is also closed.

Using Theorem 1 we now show that attainability of a vector does not imply its asymptotic attainability. This implies in particular that these two concepts are not identical.

Example 2 We provide an example where the vector $\vec{0}$ is attainable but not asymptotically attainable. Consider the following game where payoffs are 2-dimensional, each player has 2 actions, and the payoffs are scalar and given by:

	L	R
U	1	0
B	0	-1

Figure 4: The payoff function in Example 2.

In this game $v_\lambda = 0$ for every $\lambda \in \mathbb{R}$. Thus, for every $\lambda \in \mathbb{R}^2$ one has $v_\lambda \geq 0$, and therefore Theorem 1 implies that the vector $\vec{0}$ is attainable by Player 1. We argue that $\vec{0}$ is not asymptotically attainable by Player 1. Assume that Player 1 implements a strategy σ_1. In an initial time interval the strategy σ_1 plays one of the rows with a positive probability. Consider the strategy σ_2 of Player 2 that plays constantly a column that generates a non-zero payoff in that initial interval. For instance, if σ_1 plays the action U with positive probability in the initial time interval, then σ_2 plays the action B always. The initial period produces a non-zero payoff and this payoff is not diminishing to zero because Player 2 keeps playing the same column forever. This example shows that $\vec{0}$ is attainable but not asymptotically attainable by Player 1.

We point out that the argument mentioned above shows in fact that $\vec{0}$ is not attainable in the corresponding game in discrete time as well.

The following example shows that a weakly asymptotically attainable vector need not be attainable.

Example 3 Consider a two-player game where payoffs are 2-dimensional, Player 1 has 3 actions, Player 2 has 2 actions, and the payoff function is given by the left-hand side matrix in Figure 5.

	L	R
U	$(1,1)$	$(0,1)$
M	$(0,0)$	$(1,1)$
B	$(0,0)$	$(0,0)$

The game G

	L	R
U	$(1-\delta,1-\delta)$	$(-\delta,1-\delta)$
M	$(-\delta,-\delta)$	$(1-\delta,1-\delta)$
B	$(-\delta,-\delta)$	$(-\delta,-\delta)$

The game $G-\delta(1,1)$

Figure 5: The payoff functions of the games G and $G-\delta(1,1)$ in Example 3.

The vector $(0,0)$ is attainable by Player 1, using the strategy that always plays B. The vector $x := (1,1)$ is weakly asymptotically attainable according to Definition 3. Indeed, given $\varepsilon > 0$ consider the strategy σ_1^ε, with updating times $(\tau_1^k)_{k\in\mathbb{N}}$ defined by $\tau_1^k = k\varepsilon$ for $k \in \mathbb{N}$, that is defined as follows.

- If the total payoff up to time τ_1^k is not in the set $B((1,1),\varepsilon)$, during the time interval $[\tau_1^k, \tau_1^{k+1})$ play the mixed action $[\varepsilon(U),(1-\varepsilon)(M)]$.
- If the total payoff up to time τ_1^k is in the set $B((1,1),\varepsilon)$, during the time interval $[\tau_1^k, \tau_1^{k+1})$ play the action B.

For every $t \geq \frac{1}{\varepsilon}$ one has $d(\gamma^t(\sigma_1^\varepsilon,\sigma_2),(1,1)) < \varepsilon$, so that the vector x is indeed weakly asymptotically attainable by Player 1.

The vector x, however, is not attainable by Player 1 (according to Definition 2). To show that x is not attainable by Player 1 we use Theorem 2 and show that condition **B3** does not hold for x. Indeed, fix $\delta_0 > 0$ and set $q := [(1-\frac{\delta_0}{2})(L), \frac{\delta_0}{2}(R)]$. Let $p \in \Delta(A_1)$ be arbitrary. If $u(p,q) = \delta x = (\delta,\delta)$ for $\delta > 0$ then necessarily $p_U = 0$. One can verify that $u(p,q)$ cannot be equal to δx for $\delta > \delta_0$, and therefore condition **B3** does not hold for x.

Remark 4 *The proof of Theorem 3 shows that every vector $x \in \mathbb{R}^m$ is attainable by Player 1 if and only if every vector $x \in \mathbb{R}^m$ is asymptotically attainable by Player 1. Example 2 shows that attainability does not imply asymptotic attainability. We are unable to tell whether or not asymptotic attainability implies attainability.*

5.3 Alternative strategies in continuous time.

The strategies we use here are nonanticipating strategies with delay. In these strategies the times $(\tau_i^k)_{k\in\mathbb{N}}$ at which a player observes past play are independent of the play of the other player. One could consider a broader class of strategies in which $(\tau_i^k)_{k\in\mathbb{N}}$ are stopping times.

In other words, τ_i^{k+1} is a time that depends on (that is, it is measurable with respect to) the information available to player i at time τ_i^k, for each $k \in \mathbb{N}$. In this type of strategies the updating times $(\tau_i^k)_{k\in\mathbb{N}}$, are not pre-determined real numbers, as in Definition 1. Our results remain valid even if Player 2 is allowed to use a strategy from this broader class of strategies.

5.4 Remaining open problems

The problem setting and the results illustrated above give rise to a number of additional questions.

1. In the applications that we provided, both in control theory and in banking, the target sets are often not singletons. All the results above refer to attainability of singletons and not to attainability, asymptotic attainability or weak asymptotic attainability of sets. Characterizing when a set of payoffs is attainable (according to these three definitions) remains open.

2. Planning in continuous time provides more flexibility to the controller. Nevertheless, it would be interesting to find the extent to which restricting updating times to a discrete set affects the controller's ability to attain target sets.

6 Proofs

6.1 Proof of Proposition 1

The following two lemmas prove the proposition.

Lemma 1 *The set W is a cone.*

Lemma 2 *The set W is convex.*

Proof of Lemma 1. Suppose that $x \in W$, and fix $\beta > 0$. We will show that $\frac{1}{\beta}x \in W$ as well.

For every strategy σ_i of player i, let σ_i^β be the strategy σ_i accelerated by a factor β. That is, $(\sigma_i^\beta(a_{-i}))(t) := (\sigma_i(\widehat{a}_{-i}))(\beta t)$, where $\widehat{a}_{-i}(t) = a_{-i}(\beta t)$.

By the definition of attainability, there is $T > 0$ such that for every $\varepsilon > 0$ there is a strategy σ_1^ε of Player 1 such that

$$d(x, \gamma^t(\sigma_1, \sigma_2)) \leq \varepsilon, \quad \forall t \geq T, \forall \sigma_2.$$

For every strategy σ_2 of Player 2 one has

$$\begin{aligned}\gamma^t(\sigma_1^\beta, \sigma_2) &= \int_0^t u(a^s(\sigma_1^\beta, \sigma_2))\mathrm{d}s \\ &= \frac{1}{\beta}\int_0^{\beta t} u(a^s(\sigma_1, \sigma_2^{1/\beta}))\mathrm{d}s \\ &= (1/\beta)\gamma^{\beta t}(\sigma_1, \sigma_2^{1/\beta}).\end{aligned} \tag{3}$$

We deduce that

$$d(\frac{x}{\beta}, \gamma^T(\sigma_1^\beta, \sigma_2)) \le \frac{\varepsilon}{\beta}, \quad \forall t \ge \beta T, \forall \sigma_2,$$

and therefore $\frac{1}{\beta}x$ is attainable by Player 1, as desired. ■

Proof of Lemma 2. Let $x, y \in W$ and let $0 < \beta < 1$. We show that $\beta x + (1-\beta)y \in W$. From Lemma 1, βx and $(1-\beta)y$ are attainable by Player 1.

By the definition of attainability there are $T_x, T_y > 0$ such that for every $\varepsilon > 0$ there are strategies σ_1^x and σ_1^y of Player 1 such that

$$d(\gamma^t(\sigma_1^x, \sigma_2), \beta x) < \varepsilon, \quad \forall t \ge T_x, \forall \sigma_2, \tag{4}$$

$$d(\gamma^t(\sigma_1^y, \sigma_2), \beta y) < \varepsilon, \quad \forall t \ge T_y, \forall \sigma_2. \tag{5}$$

Partition the time line $[0, \infty)$ into two sets, $\mathbb{T}_1$ and $\mathbb{T}_2$ as follows:

- $\mathbb{T}_1 = \cup_{\ell=0}^\infty [\ell, \ell + \beta)$;
- $\mathbb{T}_2 = \cup_{\ell=0}^\infty [\ell + \beta, \ell + 1)$.

Thus, each integer interval $[\ell, \ell + 1)$ is partitioned into two sets: one set with Lebesgue measure β is included in $\mathbb{T}_1$, and the other set with Lebesgue measure $1 - \beta$ is included in $\mathbb{T}_2$. We construct a strategy σ_1 of Player 1 that uses σ_1^x on $\mathbb{T}_1$ and σ_1^y on $\mathbb{T}_2$, and show that this strategy attains $\beta x + (1-\beta)y$ up to ε. Formally, for each $j \in \{1, 2\}$ define $\varphi_j(t)$ to be the Lebesgue measure of the set $[0, t) \cap \mathbb{T}_j$. Given a control $a_{-i} \in \mathcal{C}_{-i}$ define two auxiliary controls $\widehat{a}_{-i}^1$ and $\widehat{a}_{-i}^2$ by

$$\widehat{a}_{-i}^j(t) := a_{-i}(\varphi_j^{-1}(t)) \text{ for every } t \ge 0, j \in \{1, 2\}.$$

The control $\widehat{a}_{-i}^j$ consists of that part of a_{-i} that corresponds to times in $\mathbb{T}_j$. The strategy σ_1 is defined by

$$\sigma_1(a_{-i})(t) := \begin{cases} (\sigma_1^x(\widehat{a}_{-i}^1))(t), & t \in \mathbb{T}_1, \\ (\sigma_1^y(\widehat{a}_{-i}^2))(t), & t \in \mathbb{T}_2. \end{cases}$$

Fix a strategy σ_2 of Player 2.

$$\begin{aligned}\gamma^T(\sigma_1,\sigma_2) &= \int_0^T u(a^t(\sigma_1,\sigma_2))\mathrm{d}t \\ &= \int_{[0,T)\cap\mathbb{T}_1} u(a^{\varphi_1^{-1}(t)}(\sigma_x,\sigma_{2,x}))\mathrm{d}t + \int_{[0,T)\cap\mathbb{T}_2} u(a^{\varphi_2^{-1}(t)}(\sigma_y,\sigma_{2,y}))\mathrm{d}t,\end{aligned}$$

where $\sigma_{2,x}$ and $\sigma_{2,y}$ are the strategies of Player 2 induced on $\mathbb{T}_1$ and $\mathbb{T}_2$. The right-hand side converges to $\beta x + (1-\beta)y$ up to ε, which completes the proof. ■

6.2 Proof of Theorem 1

To see that condition **B1** implies condition **B2**, assume to the contrary that **B1** holds but condition **B2** does not. Then, there is λ such that $v_\lambda < 0$. That is, there is $q \in \Delta(A_2)$ and $\delta > 0$ such that $\langle \lambda, u(p,q)\rangle < \delta < 0$ for every $p \in \Delta(A_1)$. Denote by σ_2 the stationary strategy of Player 2 that plays constantly the mixed action q. This strategy guarantees that for every strategy σ_1 of Player 1,

$$\langle \lambda, \gamma^T(\sigma)\rangle = \int_0^T \langle \lambda, u_i((a^t(\sigma))\rangle \mathrm{d}t < T\delta.$$

In particular, the distance between $\int_0^T u_i(a^t(\sigma))\mathrm{d}t$ and $\vec{0}$ does not go to 0, and therefore the vector $\vec{0}$ is not attainable by Player 1, which contradicts condition **B1**.

Suppose now that condition **B2** is satisfied. The argument used in the proof of Lemma 1 of accelerating the time shows that, given $T > 0$, if for every ε there is a strategy σ_1 of Player 1 such that $d(\gamma^t(\sigma_1,\sigma_2),\vec{0}) < \varepsilon$ for every $t \geq T$ then for every ε there is a strategy $\widehat{\sigma}_1$ of Player 1 such that $d(\gamma^t(\widehat{\sigma}_1,\sigma_2),\vec{0}) < \varepsilon$ for every $t \geq \frac{T}{2}$. It follows that to prove that condition **B1** is satisfied it is sufficient to prove that condition **B2** implies that for every $\varepsilon > 0$ there is $T > 0$ and a strategy σ_1 of Player 1 such that $d(\gamma^t(\sigma_1,\sigma_2),\vec{0}) < \varepsilon$ for every $t \geq T$.

Consider then the following strategy σ_1 of Player 1 that depends on a parameter $\eta > 0$. The updating times of the strategy σ_1, $(\tau_1^k)_{k\in\mathbb{N}}$, are defined by

$$\tau_1^k := \sum_{\ell=1}^{k} \frac{\eta}{\ell}. \tag{6}$$

Denote the payoff up to time τ_1^k by S_k. For $\tau_1^k \leq t < \tau_1^{k+1}$ we set $\sigma_1(t)$ to be an optimal strategy of Player 1 in the game $\langle -S_k, G\rangle$. That is, σ_1 is constant in the interval $[\tau_1^k,\tau_1^{k+1})$; in this interval $\sigma_1(t)$ is equal to a mixed action that guarantees that the payoff and S_k lie on different sides of the hyperplane perpendicular to S_k. This means that if σ_2 is the strategy played by Player 2, then $\langle S_k, u(a^t(\sigma))\rangle \leq 0$, for every $t \in [\tau_1^k,\tau_1^{k+1})$.

By definition, $S_k = \gamma^{\tau_1^k}(\sigma_1, \sigma_2)$. Thus,

$$S_k = S_{k-1} + \int_{\tau_1^{k-1}}^{\tau_1^k} u(\sigma(t))\mathrm{d}t.$$

Consequently,

$$\begin{aligned}
\|S_k\|^2 &= \|S_{k-1}\|^2 + 2\left\langle S_k, \int_{\tau_1^{k-1}}^{\tau_1^k} u(\sigma_1(t),\sigma_2(t))\mathrm{d}t \right\rangle + \left\| \int_{\tau_1^{k-1}}^{\tau_1^k} u(\sigma_1(t),\sigma_2(t))\mathrm{d}t \right\|^2 && (7)\\
&= \|S_{k-1}\|^2 + 2\int_{\tau_1^{k-1}}^{\tau_1^k} \langle S_k, u(\sigma_1(t),\sigma_2(t))\rangle \mathrm{d}t + \left\| \int_{\tau_1^{k-1}}^{\tau_1^k} u(\sigma_1(t),\sigma_2(t))\mathrm{d}t \right\|^2 && (8)\\
&\leq \|S_{k-1}\|^2 + \left(\int_{\tau_1^{k-1}}^{\tau_1^k} \|u(\sigma_1(t),\sigma_2(t))\|\mathrm{d}t \right)^2 && (9)\\
&\leq \|S_{k-1}\|^2 + \left(\frac{\eta}{k}\right)^2. && (10)
\end{aligned}$$

Continuing this way, one obtains, $\|S_k\|^2 \leq \eta^2 \sum_{\ell=1}^{k} \frac{1}{\ell^2} < 4\eta^2$. Thus, the distance between S_k and $\vec{0}$ does not exceed 2η for any k.

Note that when $k \geq e^{\frac{1}{\eta}}$, then $\tau_1^k := \sum_{\ell=1}^{k} \frac{\eta}{\ell} \geq \eta \log(k) \geq 1$. It means that the cumulative payoff at any time τ_1^k that exceeds 1 is within 2η from $\vec{0}$. Finally, since the length of the time segments $[\tau_1^k, \tau_1^{k+1})$ is $\frac{\eta}{k}$, the cumulative payoff up to $t > 1+\eta$ is within $2\eta + \frac{\eta}{e^{\frac{1}{\eta}}}$ from $\vec{0}$. Since η is arbitrary, it follows that the vector $\vec{0}$ is indeed attainable, with[4] $T = 2$, by Player 1. Condition **B1** is therefore established. ■

6.3 Proof of Theorem 2

Part 1: If the vector x is attainable by Player 1 then $\vec{0}$ is attainable by Player 1.

Assume to the contrary that the vector x is attainable by Player 1 but $\vec{0}$ is not attainable by Player 1. Let $T_0 > 0$ be arbitrary. Since the vector $\vec{0}$ is not attainable by Player 1, there is $\varepsilon_0 > 0$ such that for every strategy σ_1 of Player 1 there is a strategy σ_2 of Player 2 and $t \geq T_0$ such that $d(\gamma^t(\sigma_1,\sigma_2), \vec{0}) > \varepsilon_0$.

Let σ_1 be a strategy of Player 1 that attains x. In particular, there is $T > 0$ and a strategy σ_1 of Player 1 such that

$$d(\gamma^t(\sigma_1,\sigma_2), x) \leq \frac{\varepsilon_0}{2}, \quad \forall \sigma_2, \forall t \geq T. \tag{11}$$

[4]Recall that the definition of attainability (Definition 2) requires the existence of a uniform time T where the cumulative payoff is within distance ε from $\vec{0}$. Here we set $T = 2$ because it is greater than $1+\eta$ and it depends neither on η nor on the strategy employed by Player 1.

Let $\widehat{\sigma}_1$ be the strategy σ_1 from time T and on. Formally, let k be the minimal integer such that $\tau_1^k > T$. The updating times of $\widehat{\sigma}_1$ are $(\tau_1^\ell)_{\ell=1}^\infty$, and $\widehat{\sigma}_1(a_2) = \sigma_1(a_2^* \circ a_2)$ for every $a_2 \in \mathcal{C}_2$, where a_2^* is the play of Player 2 according to σ_2 up to time T (when Player 1 plays σ_1), and $a_2^* \circ a_2$ is the concatenation of a_2^* and a_2 into a long history that starts with a_2^* and continues with a_2.

Let $\widehat{\sigma}_2$ be a strategy of Player 2 and $t \geq T_0$ be a positive real such that $d(\gamma^t(\widehat{\sigma}_1, \widehat{\sigma}_2), \vec{0}) > \varepsilon_0$. Consider now the strategy of Player 2 that plays arbitrarily up to time T, and from time T is follows $\widehat{\sigma}_2$. Then $d(\gamma^t(\sigma_1, \widehat{\sigma}_2), \vec{0}) > \frac{\varepsilon_0}{2}$, which contradicts (11).

Part 2: If the vector x is attainable by Player 1, then condition **B3** is satisfied.

Suppose to the contrary that condition **B3** is not satisfied. That is, for every $\delta_0 > 0$ there is $q \in \Delta(A_2)$ such that for every $p \in \Delta(A_1)$ one has $u(p,q) \neq \delta x$ for every $\delta > \delta_0$. We divide the argument into two cases.

Case A: There is $q \in \Delta(A_2)$ such that for every $p \in \Delta(A_1)$, and every $\delta > 0$, $u(p,q) \neq \delta x$. We show that by playing constantly q (a strategy that we denote by q^*) Player 2 can prevent Player 1 from attaining x. Let σ_1 be a strategy of Player 1. Denote by p_t the average mixed action played by Player 1 up to time t; that is, $p^t = \frac{1}{t}\int_0^t \sigma_1(s)\mathrm{d}s$. We obtain, $\gamma^t(\sigma_1, q^*) = tu(p_t, q)$. Thus, $\gamma^t(\sigma_1, q^*)$ is in the cone generated by $R_1(q) := \{u(p,q);\ p \in \Delta(A_1)\}$. This cone is closed and by assumption, does not contain x. Thus, there is a positive distance between x and this cone, implying that $\gamma^t(\sigma_1, q^*)$ cannot get arbitrarily close to x. This contradicts the fact that x is attainable.

Case B: For every $q \in \Delta(A_2)$ there is $p \in \Delta(A_1)$ such that $u(p,q) = \delta x$, but the δ's are *not* bounded away from zero. Thus, for every $\delta > 0$, there is $q_\delta \in \Delta(A_2)$ such that $\delta \geq \max\{\delta';\ \exists p \text{ s.t. } u(p,q) = \delta' x\}$. We show that for every $\delta > 0$, if Player 2 plays q_δ all the time (a strategy that we denote by q_δ^*), then there is $\varepsilon > 0$ such that for every σ_1, $\|\gamma^T(\sigma_1, q_\delta^*) - x\| < \varepsilon$ implies $T > 1/(4\delta)$.

Fix $\delta > 0$. Denote

$$\delta_0 := \max\{\delta' \colon \exists p \text{ s.t. } u(p, q_\delta) = \delta' x\} < \delta.$$

In particular, $\delta_0 x \in R_1(q_\delta)$, and $\delta' x \notin R_1(q_\delta)$ for every $\delta' > \delta_0$. Let E be the convex hull of $R_1(q_\delta)$ and $\vec{0}$. That is, $E = \operatorname{conv}\left(R_1(q_\delta) \cup \{\vec{0}\}\right)$. The set E is convex, compact and it does not contain $\delta' x$ for every $\delta' > \delta_0$. In particular, $2\delta_0 x \notin E$. Thus, there is an open ball $F = B(2\delta_0 x, \eta)$ which is disjoint of E. By the hyperplane separation theorem there is a non-zero vector $\alpha \in \mathbb{R}^m$ (without loss of generality we may assume that $\|\alpha\| = 1$) such that $\langle e, \alpha\rangle \leq \langle f, \alpha\rangle$ for every $e \in E$ and $f \in F$. It implies that $0 = \langle \vec{0}, \alpha\rangle \leq \langle f, \alpha\rangle$ for every $f \in F$.

We claim that $0 < \langle x, \alpha \rangle$. Indeed, if $0 = \langle x, \alpha \rangle$, then every $f \in F$ can be expressed as $f = 2\delta_0 x + v$, where $v = v(f) \in B(\vec{0}, \eta)$. In particular, $0 \leq \langle f, \alpha \rangle = \langle v, \alpha \rangle$. It follows that $\langle v, \alpha \rangle = 0$ for every $v \in B(\vec{0}, \eta)$, which implies that $\alpha = 0$, contradicting the fact that $\|\alpha\| = 1$.

Suppose that $e \in R_1(q_\delta)$ and $T \cdot e \in B(x, \varepsilon)$, with $\varepsilon = \langle x, \alpha \rangle / 2$. Then, $T \cdot e = x + z$, where $\|z\| \leq \varepsilon$. Thus, $\langle T \cdot e, \alpha \rangle = \langle x + z, \alpha \rangle$. Since $e \in E$ and $2\delta_0 x \in F$,

$$\langle e, \alpha \rangle \leq \langle 2\delta_0 x, \alpha \rangle \leq \langle 2\delta x, \alpha \rangle.$$

Hence,

$$T = \frac{\langle x + z, \alpha \rangle}{\langle e, \alpha \rangle} \geq \frac{\langle x, \alpha \rangle + \langle z, \alpha \rangle}{2\langle \delta x, \alpha \rangle} \geq \frac{\langle x, \alpha \rangle - \varepsilon}{2\langle \delta x, \alpha \rangle} = \frac{1}{4\delta}. \tag{12}$$

Recall that q^*_δ denotes the strategy of Player 2 that plays q_δ all the time. To show that condition **B1** holds, that is, that x is not attainable, we need to show that for every T there is $\varepsilon > 0$ such that for every strategy σ_1 of Player 1 there is a strategy σ_2 of Player 2 and $t \leq T$ satisfying $d(\gamma^t(\sigma_1, \sigma_2), x) > \varepsilon$. Fix a strategy σ_1 of Player 1, and suppose that the cumulative payoff up to T is within ε from x. In other words, $\|\gamma^T(\sigma_1, q^*_\delta) - x\| \leq \varepsilon$. Let $p^T := \frac{1}{T}\int_0^T \sigma_1(s)\mathrm{d}s$ be the average mixed action played by σ_1 until time T. Thus, $Tu(p_T, q_\delta) = x + z$, where $\|z\| \leq \varepsilon$. Letting $e = u(p_T, q_\delta)$, we obtain by Eq. (12) that $T > \frac{1}{4\delta}$. In words, the time it takes to reach $B(x, \varepsilon)$ is at least $\frac{1}{4\delta}$. This shows that there is no uniform bound on the time at which the total payoff gets close to x. Thus, x is not attainable, which contradicts the assumption.

<u>Part 3:</u> If conditions **B1** and **B3** hold, then condition **B4** holds.

Let $\langle \lambda, (G - \delta x) \rangle$ be an auxiliary two-player zero-sum one-shot game where the sets of actions of the two players are A_1 and A_2, and the payoff function is $r(a_1, a_2) = \langle \lambda, u(a_1, a_2) - \delta x \rangle$ for every $(a_1, a_2) \in A_1 \times A_2$.

We prove that there is $\delta > 0$ such that for every vector $\lambda \in \mathbb{R}^m$ one has $\mathrm{val}(\langle \lambda, (G - \delta x) \rangle) > 0$. Theorem 1 would then imply that the vector $\vec{0}$ is attainable in the game $\langle \lambda, (G - \delta x) \rangle$, so that condition **B4** would hold.

To this end we show that there is $\delta > 0$ such that for every $\lambda \in \mathbb{R}^m$ there is $p \in \Delta(A_1)$ such that for every $q \in \Delta(A_2)$, $\langle \lambda, u(p, q) - \delta x \rangle \geq 0$, or equivalently, $\langle \lambda, u(p, q) \rangle \geq \langle \lambda, \delta x \rangle$.

Fix $\lambda \in \mathbb{R}^m$. Assume first that $\langle \lambda, x \rangle < 0$. By condition **B1** we have that $v_\lambda \geq 0$, and therefore there is $p \in \Delta(A_1)$ such that for every $q \in \Delta(A_2)$, $\langle \lambda, u(p, q) \rangle \geq 0$. Hence, $\langle \lambda, u(p, q) \rangle \geq \delta \langle \lambda, x \rangle = \langle \lambda, \delta x \rangle$ for every $\delta > 0$, as desired. We can therefore assume that $\langle \lambda, x \rangle \geq 0$. From condition **B3** we know that for every $q \in \Delta(A_2)$ there is $p \in \Delta(A_1)$ such that $u(p, q) = \delta x$ with δ bounded away form 0. Therefore, there is δ_0 such that for every $q \in \Delta(A_2)$ there is $p \in \Delta(A_1)$, such that $\langle \lambda, u(p,q) \rangle = \langle \lambda, \delta_q x \rangle \geq \langle \lambda, \delta_0 x \rangle$. By the minmax theorem, there is $p \in \Delta(A_1)$ such that for every $q \in \Delta(A_2)$, $\langle \lambda, u(p,q) \rangle \geq \langle \lambda, \delta_0 x \rangle$, as desired.

Part 4: If conditions **B1** and **B4** hold then the vector x is attainable by Player 1.

Let $\eta > 0$ and let $\varepsilon > 0$ satisfy $\frac{\varepsilon}{\delta T} < \eta$. By condition **B4** the vector $\vec{0}$ is attainable by Player 1 in the game $(G - \delta x)$, so that there is $T > 0$ such that for every $\varepsilon > 0$ there is a strategy σ_1 of Player 1 satisfying

$$d(\gamma^t_{(G-\delta x)}(\sigma_1, \sigma_2), \vec{0}) \leq \varepsilon, \quad \forall \sigma_2, \forall t \geq T. \tag{13}$$

By Eq. (3), for every $\beta > 0$ one has

$$d(\frac{1}{\beta}\gamma^{\beta t}_{(G-\delta x)}(\sigma_1^{1/\beta}, \sigma_2), \vec{0}) \leq \varepsilon, \quad \forall \sigma_2, \forall t \geq T. \tag{14}$$

It follows that

$$d(\gamma^{\beta t}_{(G-\delta x)}(\sigma_1^{1/\beta}, \sigma_2), \vec{0}) \leq \beta\varepsilon, \quad \forall \sigma_2, \forall t \geq T, \tag{15}$$

so that

$$d(\gamma^{t}_{(G-\delta x)}(\sigma_1^{1/\beta}, \sigma_2), \vec{0}) \leq \beta\varepsilon, \quad \forall \sigma_2, \forall t \geq \beta T. \tag{16}$$

Setting $\beta = \frac{1}{\delta T}$ we deduce that

$$d(\gamma^{t}_{(G-\delta x)}(\sigma_1^{\delta T}, \sigma_2), \vec{0}) \leq \frac{\varepsilon}{\delta T}, \quad \forall \sigma_2, \forall t \geq \frac{1}{\delta}. \tag{17}$$

Setting $t = \frac{1}{\delta}$ in (17) it follows in particular that

$$d(\gamma^{t}_{G}(\sigma_1^{\delta T}, \sigma_2), x) \leq \frac{\varepsilon}{\delta T} < \eta, \quad \forall \sigma_2. \tag{18}$$

The strategy $\sigma_1^{\delta T}$ ensures that the payoff at time $\frac{1}{\delta}$ is close to x. It follows that the strategy of Player 1 that follows $\sigma_1^{\delta T}$ up to time $\frac{1}{\delta}$, and follows a strategy that attains $\vec{0}$ in the game G thereafter, is a strategy that attains x in G. ■

6.4 Proof of Theorem 3

We first prove that condition **C2** implies condition **C1**. If condition **C2** holds then condition **B2** holds as well. By Theorem 1 the vector $\vec{0}$ is attainable by Player 1. By Theorem 2, to prove that condition **C1** holds, it is sufficient to prove that condition **B4** holds for every $x \in \mathbb{R}^m$.

Assume then to the contrary that condition **B4** does not hold for some $x \in \mathbb{R}^m$. Then, for every $\delta > 0$ the vector $\vec{0}$ is not attainable by Player 1 in the game $(G - \delta x)$. By Theorem

1 this implies that there is $\lambda_\delta \in \mathbb{R}^m$ such that $v_{\lambda_\delta}(G - \delta x) < 0$. Since this inequality does not hold for $v_{\vec{0}}(G - \delta x) = 0$ for every $\delta > 0$, we can assume w.l.o.g. that $\|\lambda_\delta\| = 1$ for every $\delta > 0$. Since $v_{\lambda_\delta}(G - \delta x) < 0$, there exists $q_\delta \in \Delta(A_2)$ such that

$$\langle u(p, q_\delta) - \delta x, \lambda_\delta \rangle < 0, \quad \forall p \in \Delta(A_1).$$

Take a subsequence $(\delta_k)_{k\in\mathbb{N}}$ that converges to 0 such that the sequence $(q_{\delta_k})_{k\in\mathbb{N}}$ converges to a limit $q_* \in \Delta(A_2)$ and the sequence $(\lambda_{\delta_k})_{k\in\mathbb{N}}$ converges to a limit $\lambda_* \in \mathbb{R}^m$. Note that $\|\lambda_*\| = 1$, so that $\lambda_* \neq \vec{0}$. Then

$$\langle u(p, q_*), \lambda_* \rangle \leq 0, \quad \forall p \in \Delta(A_1).$$

Ii implies that $v_{\lambda_*} \leq 0$, contradicting **C2**.

It remains to show that condition **C1** implies **C2**. Assume that the payoff function u does not satisfy condition **C2**. Then, there is $\lambda \in \mathbb{R}^m$ such that $v_\lambda \leq 0$. It follows that there is $q \in \Delta(A_2)$ such that $D_2(q) := \{u(p, q) \colon p \in \Delta(A_1)\} \subseteq H := \{x \colon \langle \lambda, x \rangle \leq 0\}$. Therefore, if Player 2 employs the strategy that always plays q, the total payoff is in H. Thus, every payoff vector in the complement of H is not attainable by Player 1. ∎

References

[1] T. Başar and P. Bernhard. *H-infinity Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach.* Birkhäuser, 1991.

[2] D. Bauso, F. Blanchini, and R. Pesenti. Robust control policies for multi-inventory systems with average flow constraints. *Automatica*, 42(8):1255–1266, 2006.

[3] D. Bauso, F. Blanchini, and R. Pesenti. Optimization of long-run average-flow cost in networks with time-varying unknown demand. *IEEE Trans. on Automatic Control*, 55(1):20–31, 2010.

[4] D. Bauso, L. Giarrè, and R. Pesenti. Robust control of uncertain multi-inventory systems via linear matrix inequality. *International Journal of Control*, 83(8):1723–1740, 2010.

[5] D. P. Bertsekas and I. B. Rhodes. On the minimax reachability of target sets and target tubes. *Automatica*, 7:233–241, 1971.

[6] D. Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific J. Math.*, 6(1):1–8, 1956.

[7] F. Blanchini. Set invariance in control a survey. *Automatica*, 35(11):1747–1768, 1999.

[8] F. Blanchini, S. Miani, and W. Ukovich. Control of production-distribution systems with unknown inputs and system failures. *IEEE Trans. on Automatic Control*, 45(6):1072–1081, 2000.

[9] F. Blanchini, F. Rinaldi, and W. Ukovich. Least inventory control of multi-storage systems with non-stochastic unknown input. *IEEE Trans. on Robotics and Automation*, 13(5):633–645, 1997.

[10] N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning and Games.* Cambridge University Press, 2006.

[11] D. Foster and R. Vohra. Regret in the on-line decision problem. *Games and Economic Behavior*, 29:7–35, 1999.

[12] S. Hart. Adaptive heuristics. *Econometrica*, 73:1401–1430, 2005.

[13] S. Hart and A. Mas-Colell. A general class of adaptive strategies. *Journal of Economic Theory*, 98:2654, 2001.

[14] S. Hart and A. Mas-Colell. Regret-based continuous-time dynamics. *Games and Economic Behavior*, 45:375–394, 2003.

[15] F. Heid, D. Porath, and S. Stoltz. Does capital regulation matter for bank behavior evidence for german savings banks. *Deutsche Bundesbank, Discussion Paper 03/04*, 2004.

[16] T.-F. Hou. Approachability in a two-person game. *The Annals of Mathematical Statistics*, 42:735–744, 1971.

[17] E. Khmelnitsky and M. Tzur. Parallelism of continuous- and discrete-time production planning problems. *IIE Transactions*, 36:611–628, 2004.

[18] E. Lehrer. Allocation processes in cooperative games. *International Journal of game Theory*, 31:341351, 2002.

[19] E. Lehrer. Approachability in infinite dimensional spaces. *International Journal of game Theory*, 31(2):253–268, 2002.

[20] E. Lehrer. A wide range no-regret theorem. *Games and Economic Behavior*, 42, 2003.

[21] E. Lehrer and E. Solan. Approachability with bounded memory. *Games and Economic Behavior*, 66:995–1004, 2009.

[22] E. Lehrer, E. Solan, and D. Bauso. Repeated games over networks with vector payoffs: the notion of attainability. In *Proceedings of the NetGCoop 2011*, Paris, France, October 2011.

[23] E. Lehrer and S. Sorin. Minmax via differential inclusion. *Convex Analysis*, 14(2):271–273, 2007.

[24] E. Roxin. The axiomatic approach in differential games. *J. Optim. Theory Appl.*, 3:153–163, 1969.

[25] A. S. Soulaimani, M. Quincampoix, and S. Sorin. Approchability theory, discriminating domain and differential games. *SIAM Journal of Control and Optimization*, 48(4):2461–2479, 2009.

[26] X. Spinat. A necessary and sufficient condition for approachability. *Mathematics of Operations Research*, 27:31–44, 2002.

[27] P. van Roy. The impact of the 1988 basel accord on banks capital ratios and credit risk-taking: an international study. *EFMA 2004 Basel Meetings available at SSRN: http://ssrn.com/abstract=497264*, 2005.