

On equilibrium strategies in noncooperative dynamic games

Citation for published version (APA):

Groenewegen, L. P. J., & Wessels, J. (1978). *On equilibrium strategies in noncooperative dynamic games*. (Memorandum COSOR; Vol. 7826). Technische Hogeschool Eindhoven.

Document status and date:

Published: 01/01/1978

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Department of Mathematics

PROBABILITY THEORY, STATISTICS, OPERATIONS RESEARCH AND SYSTEMS THEORY GROUP

Memorandum COSOR 78-26

On equilibrium strategies in noncooperative
dynamic games

by

Luuk P.J. Groenewegen^{*)} and Jaap Wessels

^{*)} Rijkswaterstaat, Data Processing Division, Rijswijk (Z.H.).

Eindhoven, December 1978

The Netherlands

On equilibrium strategies in noncooperative dynamic games

Luuk P.J. Groenewegen and Jaap Wessels

Rijkswaterstaat, Data Processing Division, Rijswijk (Z.H.), the Netherlands.

Department of Mathematics, Eindhoven University of Technology, Eindhoven, the Netherlands

SUMMARY. In this paper a characterization is given for equilibrium strategies in noncooperative dynamic games. These dynamic games are formulated in a very general way without any topological conditions. For a Nash-equilibrium concept, it is shown that equilibrium strategies are conserving and equalizing. Moreover, it is shown that a set of strategies with these properties satisfies the equilibrium conditions.

With these characterization earlier characterizations for one-person decision processes, gambling houses and dynamic games have been generalized. Especially, this paper shows that such a characterization is basic for a very general class of dynamic games and does not depend on special structure. Of course, in dynamic games with more structure a more refined formulation of the characterization is possible.

1. INTRODUCTION

In the analysis of any type of decision processes (with one or more decision makers) one may distinguish three essentially different kinds of activities:

1. The construction of a decent mathematical model based on the decision structure and the propulsion mechanism, usually both being formulated only conditionally in local time.

In discrete time, this activity is usually not very difficult, since it can make use of the well-known Ionescu Tulcea-construction for handling random elements. In continuous time this activity presents essential difficulties, but there are techniques for handling these difficulties. These are based for instance on the so-called Girsanov measure transformations.

2. The proof of the existence of "good" strategies of a nice type.

"Good" can mean here: optimal or nearly optimal, Nash-equilibrium or nearly Nash-equilibrium, Pareto-optimal etc. A nice type of strategies can mean: memoryless or even stationary, pure, monotone etc.

3. The search for necessary (and preferably sufficient) conditions for "good" strategies.

Such conditions can have the form of a set of optimality equations (resulting from Bellman's optimality principle), a maximum principle, a set of Hamilton-

Jacobi equations etc. As is well-known all these types of conditions are strongly related.

In the literature, the activities of the second and third type are often intensively interwoven, since results on conditions are often obtained in order to use them for obtaining existence results. In this paper however, we will concentrate on the search for necessary and sufficient conditions for good strategies. Namely, it appears that the formulation of such conditions is possible in a very general way without using the specific structure of the actual problem. Of course, specification of the generally formulated conditions for actual problems can give extra insight. However, it is equally important to see the general principles which produce the results.

Bellman's *optimality principle* [1, chapter 3, section 3], which runs as follows: *an optimal policy has the property that, whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision*, is a good starting point for an analysis of a very general decision process. Although it is rather vague in its initial formulation, the optimality principle has proved to constitute a very stimulating guide for the analysis. In the theory of one-person decision processes this has led to characterizations for optimal strategies in different situations to begin with Bellman's *optimality equations* in [1]. Dubins and Savage [4] and Sudderth [17] gave precise and elegant formulations for the situation of gambling houses. They were the first who made it explicit that a condition in the form of optimality equations is only sufficient for optimality of a strategy if there is some strong form of fading in the process. If not, a supplementary condition should be added. Essentially, the optimality equations-type-of-condition says that the strategy is such that the player does not give any potential reward away. Therefore, this type of condition has been called the *conservingness* property. If the process or reward does not fade away in time, one should add a condition which requires a strategy to cash its potential rewards. This condition has been called the *equalizingness* property. Together these two conditions have appeared to be necessary and sufficient for optimality in several types of one-person decision processes. As mentioned already, this has been proved for gambling houses. Later, Hordijk [9] gave a similar formulation for a certain class of Markov decision processes. More recently, this has been generalized to discrete time decision processes with a much more general propulsion mechanism and reward structure by Kertz and Nachman in [11]. In another line of development, typical control theoretical structures (with continuous time) have been treated by several authors. Relatively general formulations for the fading case have been given by Striebel in [16] and by Boel and Varaiya in [2]. These lines of thought have been combined and further generalized by Groenewegen in his monograph [7]. However, the most important fea-

ture of Groenewegen's approach is that it is very straightforward and relying on intuitively clear notions. In doing so, it is proved on one hand that conservingness and equalizingness are very central and essential notions, on the other hand it is proved that equalizing and conserving provide necessary and sufficient conditions for optimality of a strategy in many types of situations nearly without any specific side conditions e.g. of a topological nature.

So far for one-person decision processes.

For multi-person decision processes or dynamic games an analogous development can be traced. The two basic papers for this topic are Kuhn [12] and Shapley [15]. Kuhn formulates the optimality principle for multi-stage games and Shapley formulates the optimality equations for discounted (i.e. strongly fading) stochastic games. The most striking feature of these papers is, that they have been written before their one-person counterparts. For later developments the one-person theory took the lead. For stochastic games with some strong fading mechanism the conservingness property has been proved to be necessary and sufficient for optimality by many authors (for an overview of such conditions, we refer to Parthasarathy and Stern [13] and to van der Wal and Wessels [18]). Analogous results may be found in the literature about differential games (both deterministic and stochastic), see e.g. the book by Isaacs [10] and the paper by Elliott [5]. The first direct attempts for the establishment of a characterization of optimality in dynamic games can be found in Groenewegen and Wessels [8], Groenewegen [6], and Couwenbergh [3]. However, in his monograph [7], Groenewegen seems to be presenting the most elegant and intuitively appealing approach for the characterization problem in noncooperative dynamic games. Like for one-person decision processes, this approach does not require any extra conditions (e.g. of a topological nature) and it also allows for a very general propulsion mechanism and reward structure. The analysis in this paper will be based essentially on the approach of that monograph.

In section 2 the set-up of our general dynamic (noncooperative) game will be given. Section 3 gives some examples. Section 4 contains our main results and in section 5 some ramifications are indicated.

2. THE SET-UP

In this section we will present a (nonconstructive) set-up for a rather general class of noncooperative stochastic dynamic games for an arbitrary number of players.

L = set of players;

X = state space; X is endowed with a σ -field \mathcal{X} ;

T = time space; for simplicity we will take T as $\{0,1,2,\dots\}$ or $[0,\infty)$;

$A^{(\ell)}$ = action space for player ℓ ($\ell \in L$); $A^{(\ell)}$ is endowed with a σ -field $\mathcal{A}^{(\ell)}$; ν = starting distribution, so ν is a probability measure on (X, \mathcal{X}) .

The idea is that a starting state in X is determined by random selection from X with probability distribution ν . At any time $t \in T$, the state of the system (an element of X) is observed by the players and they all may choose an action from their action space. These actions have some influence on the behaviour of the system. In order to define this behaviour, we have first to introduce strategies for the players. Since these strategies may depend on the history of the process, we start by introducing such histories.

$A = \prod_{\ell \in L} A^{(\ell)}$, A is Cartesian product of the individual action spaces; A is endowed with the appropriate product- σ -field; an element of A denotes a compound action of all the players;

$H^{(t)} = [\prod_{\tau < t} (X \times A)] \times X$, is the space of state-action paths ending with the state of the system at time t ; $H^{(t)}$ is endowed with the appropriate product- σ -field $\mathcal{H}^{(t)}$. Similarly, $H = \prod_{\tau \in T} (X \times A)$, endowed with the appropriate product- σ -field \mathcal{H} . H is the set of all allowed realizations with respect to state and compound actions.

$U^{(\ell)}$ = space of allowed strategies (or controls) for player ℓ ; an arbitrary element $u^{(\ell)}$ of $U^{(\ell)}$ gives for any $t \in T$ and any history $h_t \in H^{(t)}$ a probability distribution $u^{(\ell)}(t, h_t, \cdot)$ on $A^{(\ell)}$, it is required that $u^{(\ell)}(t, \cdot, \cdot)$ is a transition probability from $H^{(t)}$ to $A^{(\ell)}$; moreover, it is required that $U^{(\ell)}$ is closed with respect to tail exchanges of individual controls, this means: if $u_1^{(\ell)}, u_2^{(\ell)} \in U^{(\ell)}$, $t \in T$, $B \in \mathcal{H}^{(t)}$, then $u^{(\ell)} \in U^{(\ell)}$ with $u^{(\ell)}(\tau, h_\tau, \cdot) = u_2^{(\ell)}(\tau, h_\tau, \cdot)$ for $\tau \geq t$, h_t (the restriction of h_τ until time t) $\in B$ and $u^{(\ell)}(\tau, h_\tau, \cdot) = u_1^{(\ell)}(\tau, h_\tau, \cdot)$ elsewhere.

$U = \prod_{\ell \in L} U^{(\ell)}$, $u \in U$ is a compound strategy.

Now we are able to formulate the main assumption, viz.

for any (starting) state $x \in X$ and any compound strategy $u \in U$ a probability measure $\mathbb{P}_{x,u}$ on (H, \mathcal{H}) is given; moreover the probabilities are measurable as a function of x .

By this assumption we circumvent the obligation to construct a probabilistic structure from more elementary data. With these probabilities we can easily construct the probability measures for the decision process for the given starting distribution ν by

$$\mathbb{P}_u(H') := \int_X \mathbb{P}_{x,u}(H') \nu(dx) \quad \text{for any } H' \in \mathcal{H}.$$

In fact we need a slight extension of the assumption, namely we need the probabilities for the remainder of the process for any given path until time t . Conditioning of \mathbb{P}_u only gives them almost surely on $H^{(t)}$; especially in the case of continuous time the exception set might grow out of hands. Therefore we prefer to assume the existence of all these conditioned probabilities (note that these assumptions cover exactly the activity that is described in the introduction as model construction): For any $h_t \in H^{(t)}$ ($t \in T$), and any compound strategy $u \in U$ there exists a probability measure $\mathbb{P}_{h_t, u}$ on (H, \mathcal{H}) such that

$$\equiv \mathbb{P}_{h_t, u} \text{ is concentrated on } \{h_t\} \times A \times \prod_{\substack{\tau \in T \\ \tau > t}} (X \times A)$$

$$\equiv \mathbb{P}_{h_t, u}(H') \text{ is } H^{(t)}\text{-measurable as a function of } h_t \text{ for any } H' \in \mathcal{H}$$

$$\equiv \mathbb{P}_{h_t, u}(H^{(t)}, A', \prod_{\tau > t} (X \times A)) = u(t, h_t, A') \text{ for all } t \in T, u \in U, A' \in \mathcal{A}$$

$$\equiv (\text{nonanticipativity}) \mathbb{P}_{h_t, u_1}(H') = \mathbb{P}_{h_t, u_2}(H') ,$$

if $H' = B \times \prod_{\tau > s} (X \times A)$ with B a measurable subset of $\prod_{\tau \leq s} (X \times A)$ for some $s > t$ and $u_1(\sigma, \cdot, \cdot) = u_2(\sigma, \cdot, \cdot)$ for all σ with $t \leq \sigma < s$.

\equiv (conditioning properties)

$$\text{a) } \int_H f(h) \mathbb{P}_{h_t, u}(dh) = \int_H \int_H f(h) \mathbb{P}_{h_s, u}(dh) \mathbb{P}_{h_t, u}(dh'')$$

for any $t \leq s$ and any nonnegative \mathcal{H} -measurable function f .

$$\text{b) } \int_H f(h)g(h) \mathbb{P}_{h_t, u}(dh) = f(h'_t) \int_H g(h) \mathbb{P}_{h_t, u}(dh)$$

for any $t \in T$ and any nonnegative \mathcal{H} -measurable f and g with f only depending on h_t .

Now the process-part of the dynamic game has been defined appropriately. Only the criterion is still to be defined. This will also be done in a very general way:

$r^{(\ell)}$ is a real valued measurable function on H for any $\ell \in L$ and denotes the reward of player ℓ as a function of the realization of the game.

$r^{(\ell)}$ is supposed to be quasi-integrable with respect to \mathbb{P}_u for all $u \in U$.

For given $h_t \in H^{(t)}$, $u \in U$ the expected reward for player ℓ is defined by

$$v_t^{(\ell)}(h_t, u) := \begin{cases} \mathbb{E}_{h_t, u} r^{(\ell)}, & \text{if } r^{(\ell)} \text{ is quasi-integrable} \\ -\infty, & \text{otherwise.} \end{cases}$$

Now we are able to introduce our equilibrium concept:

$u_* \in U$ is a *compound equilibrium strategy* iff

$$v_t^{(\ell)}(h_t, u_*) \geq v_t^{(\ell)}(h_t, u_*/u^{(\ell)}) \quad \mathbb{P}_{u_*}\text{-a.s.} \quad \text{for all } \ell, u^{(\ell)},$$

where $u_*/u^{(\ell)}$ denotes the compound strategy u_* with $u_*^{(\ell)}$ replaced by $u^{(\ell)}$.

3. SOME EXAMPLES

a) *Markov or stochastic games* (compare references [13,15,18]). $T = \{0,1,\dots\}$,
 $X = \{1,2,\dots\}$, $L = \{1,2,\dots\}$.

The probability measures $\mathbb{P}_{v,u}$ are generated by the transition probabilities $p(i,j;a_1,a_2,\dots)$, denoting the probability for finding state j at time $t+1$, if at time t the system is in state i and the players choose actions a_1, a_2, \dots respectively.

In this type of problem the utility is usually based on the local income function $r_\ell(i;a_1,a_2,\dots)$ denoting the actual reward for player ℓ at time t if the system is in state i and the players choose actions a_1, a_2, \dots respectively. Standard forms for the utility then become

$$r^{(\ell)}(h) := \sum_{t=0}^{\infty} r_\ell(x_t; a^{(t)}) \quad \text{or} \quad \sum_{t=0}^{\infty} \beta^t r_\ell(x_t; a^{(t)})$$

and

$$r^{(\ell)}(h) := \liminf_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=0}^{t-1} r_\ell(x_\tau; a^{(\tau)}),$$

where $h = (x_0, a^{(0)}, x_1, a^{(1)}, \dots)$, $x_t \in X$, $a^{(t)} \in A$.

In the literature many variants of such models can be found.

b) *Differential (and difference) games* (references [5,10]). $T = [0, \tau]$ or $[0, \infty)$,
 $X = \mathbb{R}^m$.

The propulsion mechanism of the process is

$$\dot{x}(t) = f(x(t), a),$$

which generates the path of the process and its probability distribution if the instantaneous compound action \bar{a} is chosen according to some (mixed) strategy. The utility function for a given state realization over time $x(t)$ and a given compound action a as a function of time $a(t)$ can be

$$r^{(\ell)}(x(\cdot), a(\cdot)) = \int_0^\tau f_\ell(x(t), a(t)) dt.$$

With the same utility function the propulsion mechanism may also be stochastic e.g.

$$dx(t) = f(x(t), a) dt + A(x(t)) dB(t),$$

where $B(t)$ is Brownian motion.

In the literature many variants of such models are studied.

4. THE BASIC CONCEPTS

Here we will introduce the basic concepts of our characterization. These concepts are formulated in terms of the value functions:

$$(4.1) \quad \psi_t^{(\ell)}(h_t, u_0) := \sup_{u^{(\ell)}} v_t^{(\ell)}(h_t, u_0 / u^{(\ell)}) .$$

The optimality concept can now be rewritten as:

u is an equilibrium strategy iff

$$\psi_t^{(\ell)}(h_t, u) = v_t^{(\ell)}(h_t, u) \quad \text{for all } \ell, t \quad \mathbb{P}_u\text{-a.s.}$$

Now we obtain for an arbitrary strategy u for $\tau \geq t$ and all ℓ

$$v_t^{(\ell)}(h_t, u) = \mathbb{E}_{h_t, u} r^{(\ell)} = \mathbb{E}_{h_t, u} \mathbb{E}_{h_\tau, u} r^{(\ell)} = \mathbb{E}_{h_t, u} v_\tau^{(\ell)}(h_\tau, u) \quad \mathbb{P}_u\text{-a.s.}$$

Using this relation, we obtain for an equilibrium strategy that the functions $\psi_t^{(\ell)}$ have the martingale property:

Lemma. If u is an equilibrium strategy, then for all $\ell, \tau \geq t$

$$(4.2) \quad \psi_t^{(\ell)}(h_t, u) = \mathbb{E}_{h_t, u} \psi_\tau^{(\ell)}(h_\tau, u) \quad \mathbb{P}_u\text{-a.s.}$$

The question now arises whether any strategy which satisfies (4.2) is an equilibrium strategy. An extremely simple example shows that this is not true.

Counterexample. (one-person game)



This example has 2 states. State 2 is absorbing and brings no rewards.

In state 1 the only player of the game has two options: staying another period ($T = \{0, 1, 2, \dots\}$) without reward or jumping to state 2 with reward 1.

Apparently, the strategy "stay in 1" satisfies (4.2) and is not optimal.

So (4.2) is a necessary condition for equilibriumness but not a sufficient one.

(4.2) only requires from a strategy that the players don't lose their prospective rewards. However, it does not guarantee that the players really cash their prospective rewards (see the example above). For this reason we call a strategy that satisfies (4.2) a *conserving* strategy.

For finding an additional condition which might ensure optimality if it is combined with conservingness, we turn to the definition of equilibrium strategy.

Suppose u is an equilibrium strategy, then we have by definition

$$\mathbb{E}_u [\psi_t^{(\ell)}(h_t, u) - v_t^{(\ell)}(h_t, u)] = 0 .$$

So we obtain the following trivial statement. If u is optimal, then

$$(4.3) \quad \lim_{t \rightarrow \infty} \mathbb{E}_u [\psi_t^{(\ell)}(h_t, u) - v_t^{(\ell)}(h_t, u)] = 0 \quad \text{for all } \ell .$$

This statement says that in the end the prospective reward is really cashed. A strategy which satisfies (4.3) is said to be *equalizing*. So, we have two properties for equilibrium strategies, namely conservingness saying that prospective reward should be maintained - and equalizingness - saying that prospective reward should be cashed in the end. Now we can hope that these two conditions are (also) sufficient for a strategy to be an equilibrium strategy.

Theorem. A necessary and sufficient condition for a strategy to be an equilibrium strategy is that it is conserving (4.2) and equalizing (4.3).

Proof. The necessity has already been proved. So suppose that u satisfies (4.2) and (4.3). For any t and τ ($\tau \geq t$) we have from (4.2):

$$(4.4) \quad \mathbb{E}_u \psi_t^{(\ell)}(h_t, u) = \mathbb{E}_u \mathbb{E}_{h_t, u} \psi_\tau^{(\ell)}(h_\tau, u) = \mathbb{E}_u \psi_\tau^{(\ell)}(h_\tau, u) .$$

With (4.3) this implies (remember (4.4) holds for any $\tau \geq t$)

$$\mathbb{E}_u \psi_\tau^{(\ell)}(h_\tau, u) = \lim_{\tau \rightarrow \infty} \mathbb{E}_u v_\tau^{(\ell)}(h_\tau, u) = \mathbb{E}_u r^{(\ell)} = \mathbb{E}_u v_t^{(\ell)}(h_t, u) .$$

Since

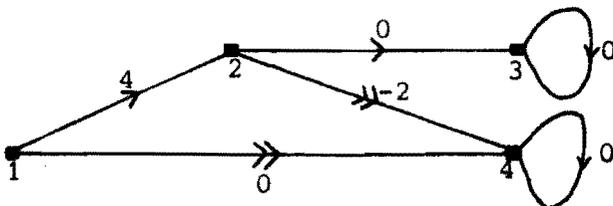
$$v_t^{(\ell)}(h_t, u) \leq \psi_t^{(\ell)}(h_t, u)$$

and since both functions have equal expectations, we may conclude that they are equal P_u -a.s.

5. SOME RAMIFICATIONS

A compound equilibrium strategy is not necessarily a sensible strategy. To illustrate one weakness of the concept we give an example of a deterministic 2-person, 0-sum game.

Example.



States 3 and 4 are absorbing states without reward. In state 1 the second player can choose between going to 2 (which costs him 4) and going to 4 (which costs him nothing). In state 2 the first player may choose between going to state 3 (without reward) and going to 4 (which costs him 2). Now the strategy for player 1 "go to 3 if the state is 2 at $t = 0$, otherwise go to 4" is part of an equilibrium strategy when it is combined with "go to 4" for the second player. However, in this way the first player takes unnecessary risks. Namely, if the second player would play stupid, the first player might win 4 units and now only wins 2.

From the example we see that the equilibrium concept might be improved. In fact some improvements have been suggested in the literature (see e.g. Selten [14]).

Below we present 3 types of equilibrium concepts, which are largely based on concepts from the literature. The third implies the second, the second implies the first and the first implies the concept from section 4. The notations are largely self-evident, but will be explained after the definitions.

A strategy u_* is *semi-subgame perfect* iff for all t, ℓ, u

$$\psi_t^{(\ell)}(h_t, u_*) = v_t^{(\ell)}(h_t, u_* /_t u^{(\ell)}) \quad \mathbb{P}_{u_* /_t u^{(\ell)}}\text{-a.s.}$$

A strategy u_* is *tail optimal* iff for all t, k, ℓ, u

$$\psi_t^{(\ell)}(h_t, u_* /_t u^{(k)}) = v_t^{(\ell)}(h_t, u_* /_t u^{(k)}) \quad \mathbb{P}_{u_* /_t u^{(k)}}\text{-a.s.}$$

A strategy u_* is *subgame perfect* iff for all t, ℓ, u

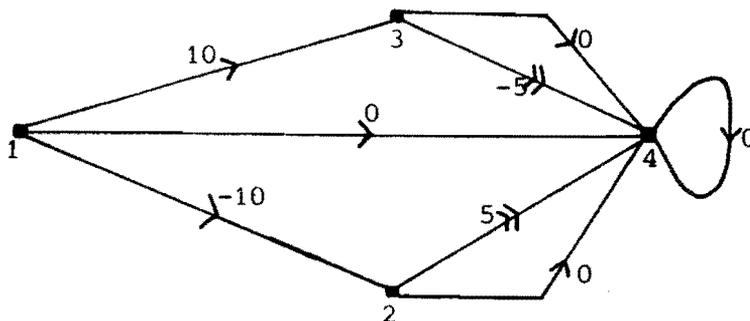
$$\psi_t^{(\ell)}(h_t, {}_t u u_*) = v_t^{(\ell)}(h_t, {}_t u u_*) \quad \mathbb{P}_u\text{-a.s.}$$

Here $u_* /_t u^{(\ell)}$ means the strategy u_* with u_* before time t replaced by $u^{(\ell)}$.

${}_t u u_*$ means the strategy which combines the strategies u (before time t) and u_* from time t on.

The difference between these equilibrium concepts can best be seen from examples as the following of a deterministic 2-person 0-sum game in discrete time:

Example.



State 4 is absorbing without any further reward. In state 3 the first player can choose between rewards 0 and -5 (without influence for the other player). In state 2 the second player can choose between losses 0 and 5. In state 1 both players have two actions. The reward is 0 if both choose the same action and the first earns 10 for the combination (1,2) and loses 10 for the combination (2,1).

Consider the strategy for player 1 which chooses always action 1 in state 1 and in state 3 uses action 1 if the game starts in 3 and action 2 otherwise.

For player 2 we consider the analogous strategy.

This pair of strategies is semi-subgame perfect but not tail optimal.

Completely analogous to the situation in section 4 for the standard equilibrium concept, one can define conservingness and equalizingness related to these stronger equilibrium concepts. Equally similar one proves that the appropriate conservingness and equalizingness are necessary and sufficient for a compound strategy to be an equilibrium strategy in the related sense (for details see [7, Ch. 6].

Another extension of our theory may be found by putting somewhat more structure on the dynamic game. An important example of such a structure is recursiveness, which requires basically that the process (allowed action set and propulsion mechanism) from time t on do not depend on the history before t and it requires that the future rewards except for additive and multiplicative factors only depend on the future of the process. In such a structure it is possible to reformulate the characterization of optimality in local quantities instead of the global quantities of section 4. For one-person decision processes such a reformulation can be found in [7, Ch. 3,4]. For dynamic games it will be worked out in a forthcoming paper. In such a reformulation the characterization is more akin to the usual optimality conditions.

REFERENCES

- [1] R. Bellman, Dynamic programming. Princeton, Princeton University Press, 1957.
- [2] R. Boel, P. Varaiya, Optimal control of jump processes. SIAM J. Control Optimization 15 (1977), 92-119.
- [3] H.A.M. Couwenbergh, Characterization of strong (Nash) equilibrium points in Markov games. Memorandum COSOR 77-09 (April 1977) Eindhoven University of Technology (Dept. of Math.).
- [4] L.E. Dubins, L.J. Savage, How to gamble if you must. New York, McGraw-Hill, 1965.

- [5] R.J. Elliott, The existence of optimal strategies and saddle in stochastic differential games. p. 123-135 in P. Hagedorn, H.W. Knobloch, G.J. Olsder (eds.), Differential games and Applications. Springer (Lecture Notes in Control and Information Sciences no. 3), Berlin 1977.
- [6] L.P.J. Groenewegen, Markov games; properties of and conditions for optimal strategies. Memorandum COSOR 76-24 (November 1976) Eindhoven University of Technology (Dept. of Math.).
- [7] L.P.J. Groenewegen, Characterization of optimal strategies in dynamic games. MC-tracts no. 90, Mathematical Centre, Amsterdam 1979 (to appear).
- [8] L.P.J. Groenewegen, J. Wessels, On the relation between optimality and saddle-conservation in Markov games. pp. 183-211 in Dynamische Optimierung, Bonn, Math. Institut der Universität Bonn (Bonner Mathematische Schriften 98) 1977.
- [9] A. Hordijk, Dynamic programming and Markov potential theory. MC-tract no. 51, Mathematical Centre, Amsterdam 1974.
- [10] R. Isaacs, Differential games. J. Wiley, New York 1975 (2nd ed.).
- [11] R.P. Kertz, D.C. Nachman, Persistently optimal plans for non-stationary dynamic programming: the topology of weak convergence case. Annals of probability (to appear).
- [12] H.W. Kuhn, Extensive games and the problem of information. Annals of Mathematics Study 28 (1953), p. 193-216.
- [13] T. Parthasarathy, M. Stern, Markov games - a survey. Technical Report of University of Illinois at Chicago Circle, Chicago 1976.
- [14] R. Selten, Reexamination of the perfectness concept for equilibrium points in extensive games. Internat. J. Game Th. 4 (1975), 25-55.
- [15] L.S. Shapley, Stochastic games. Proc. Nat. Acad. Sci. USA 39 (1953), 1095-1100.
- [16] C. Striebel, Optimal control of discrete time stochastic systems. Springer (Lecture Notes in Econ. and Math. Systems no. 110), Berlin 1975.
- [17] W.D. Sudderth, On the Dubins and Savage characterization of optimal strategies. Ann. Math. Statist. 43 (1972), 498-507.
- [18] J. van der Wal, J. Wessels, Successive approximation methods for Markov games. p. 39-55 in H.C. Tijms, J. Wessels (eds.), Markov decision theory. MC-tract 93, Mathematical Centre, Amsterdam 1977.