Printed at the Mathematical Centre, 413 Kruislaan, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

MATHEMATICAL CENTRE TRACTS 90

CHARACTERIZATION OF OPTIMAL STRATEGIES IN DYNAMIC GAMES

L.P.J. GROENEWEGEN

MATHEMATISCH CENTRUM 1981

1980 Mathematics subject classification: 90D15, 93E20

ISBN 90 6196 156 4

CONTENTS

ACKOWLEDGEMENTS

This monograph is a revised version of my doctoral thesis, written at the Department of Mathematics of the Eindhoven University of Technology. The research leading to this thesis and monograph has been supervised by prof. dr. J. Wessels. I am very grateful to him for the opportunity he gave me and for our regular stimulating discussions.
I wish to thank also professors dr. S.T.M. Ackermans, dr.ir. M.L.J. Hautus, dr. F.H. Simons and dr. F.W. Steutel for their valuable suggestions.
Moreover, it turned out to be very pleasant as well as most inspiring to do this research within the group around Jaap Wessels with members as Kees van Hee, Jo van Nunen, Kees Verhoeven, Jan van der Wal and Jacob Wijngaard.
Furthermore, I thank the Mathematical Centre for the opportunity to publish this monograph in their series Mathematical Centre Tracts and all those at the Mathematical Centre who have contributed to its technical realization.
Finally, I want to mention Maud Kronenburg for the outstanding way she converted my miniatures into readable text, and Carla Vermin-Anderson and Rita Kamlade-Spruit for doing the corrections perfectly in style.

CHAPTER 1

INTRODUCTION

It is well known, that for rather general Markov decision processes with additive reward functions, strategies are optimal if and only if they are conserving and equalizing (references will be given presently). A strategy is conserving, if no irrecoverable loss can be expected at any step. A strategy is equalizing if for each large time instant almost all profit, that might be obtained from that time on, is indeed obtained. Partial results of the above type are also known in continuous-time stochastic control.

In this monograph the characterization of optimal strategies is derived for a fairly general decision process. By imposing more structure on the reward function and on the process, we can also give more structure to the concepts of conservingness and equalizingness. Without difficulty we can generalize the derivation of the characterization to decision processes with more than one decision maker or player. At first we restrict ourselves to a characterization of Nash optimality. Afterwards, the generalization to processes with several players leads to the characterization of stronger types of optimality.

The remaining part of this introductory chapter is built up as follows. We start by sketching the structure of the decision process. The relation of our work to that of others is described thereafter. Further we introduce some notation. Finally, the contents of this monograph are summarized chapter by chapter.

The decision process we study, can be sketched as follows (for the sake of simplicity this sketch is restricted to the discrete-time case). At successive time instants t from a time space T, a system is observed to be in states x_t from a state space X. This observation is made by all n players of the system (the number n is not necessarily finite). Then each player chooses an action from his own action space, and *thereafter* he observes which action is chosen by the other players. These choices cause the system to move into a next state, which is observed by all players. The transition mechanism is determined by a probability distribution, defined on the state space, and may depend on the history up to the time of the transition. The action chosen by a player has to be admissible, and the admissibility of an action may depend on all preceding

observations. However, the choice of an action at a certain time by a given player is not allowed to depend on the choices of the other players *at that time*. In other words, the process is "noncooperative".
A strategy is a rule, which determines where and when what action must be chosen by each player. Thus every strategy determines a measure on the space of possible paths (these paths are sequences of the following form: state, action, state, action, etc.). By means of a utility function each path has a certain value, hence each strategy has a value, namely the expected utility value. A strategy is called optimal if the expected utility value is maximal in a certain sense: we will restrict ourselves to optimality concepts of the Nash type. Precisely this type of optimality will be characterized by the properties conservingness and equalizingness mentioned before.
Intuitively the idea of characterizing optimality by these two properties is so selfevident, that one cannot expect it to be new. And indeed, this type of characterization can already be found in the work of Dubins and Savage (1965), Sudderth (1972) and Hordijk (1974) and more recently in a paper of Kertz and Nachman (1977). Also the discussion at the end of Blackwell (1970) contains some remarks about this characterization. (The concept of thriftiness arising in some of these papers, means that also a special action - the stopping action - is conserving.) However, in the literature mentioned the proofs of the characterization make essential use of the specific structure of the process or of the utility function.
In Groenewegen (1975) a different proof is given, based on the principle of optimality from Bellman (1957). This technique has led to generalizations for the case of a two-person zero-sum game (see Groenewegen and Wessels (1977) and Groenewegen (1976)). Meanwhile Groenewegen and van Hee (1977) found another proof of this characterization, using a martingale approach. Rieder (1976) also uses martingale theory in establishing a characterization. Some of his results are closely related to those in Groenewegen and van Hee (1977).
Above we mentioned Bellman's principle of optimality. Apparently there is some confusion about the exact meaning of this concept. In Bellman (1957) chapter 3 section 3 it is formulated 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". This is essentially the same as the assertion of lemma 1 in Groenewegen (1975), and the result in Gavish

and Schweitzer (1976) which they call a principle of optimality. These three authors formulated the result without a reference to Bellman. Although Bellman states that he uses the principle of optimality in the derivation of the dynamic programming equation, he is not extremely careful in giving this derivation. This is probably the reason that the dynamic programming equation is also called Bellman's optimality principle by some authors. Since the dynamic programming equation and conservingness are the same thing, this explains why in control theory one uses the term optimality principle for the conserving property (see e.g. Striebel (1975), Boel and Varaiya (1977)).

In control theory much has been written on the relation between Pontryagin's maximum principle, Hamilton Jacobi equations and the optimality principle or conservingness, so there is no need for us to discuss it here. A good reference for this topic is Berkovitz (1974), chapter 5 section 2.

In the sequel we will use the following notation to classify the decision processes we are interested in: a/b/c/n, with $a \in \{C,D\}$, denoting that the time space is discrete $(a = D)$ or continuous $(a = C)$, with $b \in \{F,D,G\}$, denoting that the state space is finite $(b = F)$, denumerable $(b = D)$ or general $(b = G)$, with $c \in \{F,D,G\}$, denoting that for each player the action space is finite $(c = F)$, denumerable $(c = D)$ or general $(c = G)$, and with n a cardinal number, denoting the number of players. For instance, a C/D/F/2 process is a continuous-time, two-person decision process on a denumerable state space with a finite action space for each player.

The discrete as well as the continuous time spaces are supposed to have the usual ordering ($\leq$). Both have a lowest element and are unbounded to the right. It is not difficult to see how processes which actually do not continue after a terminal time τ, can be fitted into our model with an infinite time space. This can be done by defining the transition mechanism in such a way that after time τ the process stays with probability 1 in the state it has reached at time τ, whatever actions are chosen.

The contents of this monograph are as follows. After this introductory chapter 1, two chapters are devoted to the D/G/G/1 process: in chapter 2 we discuss its general model and derive the characterization of optimality for a general utility. In chapter 3 it is established that for so-called tail vanishing utilities the characterization has a "nice" form. The concept of a tail vanishing utility is stronger than the concept of recursiveness, introduced in Furukawa and Iwamoto (1973) and also treated here in chapter 3. Chapter 4 gives the analogous results for the C/G/G/1 process. The D (or C)/G/G/2 process with a zero-sum utility is studied in chapter 5. Several optimality concepts are discussed, and characterized in terms of conservingness and equalizingness. The analogous results for the D (or C)/G/G/n process are given in chapter 6.

CHAPTER 2

THE D/G/G/1 PROCESS WITH A GENERAL UTILITY

As has already been said in chapter 1, the D/G/G/1 process is a discrete-time (the D) decision process on a general state space (the first G) with a general action space (the second G), controlled by 1 player (the 1). The D/G/G/1 process is formulated in the first section of this chapter, and some conventions, notations and definitions are given there too. In section 2.2 we give a characterization of ν-optimality by means of ν-conservingness and ν-equalizingness. Since this characterization is given in the general situation, where the utility has no special structural properties (as e.g. additivity), this characterization is fairly global. So, at least in this chapter, our concepts of conservingness and equalizingness look a bit different from those introduced for gambling houses (Dubins and Savage (1965), Sudderth (1972)) and for Markov decision processes (Hordijk (1974), Groenewegen (1975), Rieder (1976)).

2.1. The D/G/G/1 PROCESS

To begin with we present a definition of the D/G/G/1 process that is closely related to the set-up given in Hinderer (1970). The general D/G/G/1 (decision) process is defined as a tuple $(T,(X,\mathcal{X}),(A,\mathcal{A}),(L_t \mid t \in T),(p_t \mid t \in T),r)$ together with a set of requirements.

- $T = \{t_0,t_0+1,\ldots\}$ is the *time space* (usually we will take $T = \mathbb{N} = \{0,1,\ldots\}$);
- X is the *state space*, endowed with a σ-field $\mathcal{X}$;
- A is the *action space*, endowed with a σ-field $\mathcal{A}$;
- for each $t \in T$ the symbol L_t denotes a subset of $\overset{t}{\underset{k=0}{\times}} (X \times A)$, $\overset{t}{\underset{k=0}{\times}}$ and $\times$ denoting a Cartesian product. If $(x_0,a_0,\ldots,x_t,a_t) \in L_t$, then a_t is called an *admissible action in* $(x_0,a_0,\ldots,x_t)$;
- $(p_t \mid t \in T)$ is the family of *transition functions*;
- r is the *utility function*.

For the description of the components and the behaviour of the process, we also introduce

- the *sample space* (i.e. the set of all sample paths or histories)

$(H,\mathcal{H}) := (\overset{\infty}{\underset{k=0}{\times}} (X \times A), \overset{\infty}{\underset{k=0}{\otimes}} (\mathcal{X} \otimes \mathcal{A}))$, where $\mathcal{X} \otimes \mathcal{A}$ denotes the product σ-field of $\mathcal{X}$ and $\mathcal{A}$;

- for each $t \in T$ the space $(K_t, \mathcal{K}_t) := (\overset{t}{\underset{k=0}{\times}} (X \times A), \overset{t}{\underset{k=0}{\otimes}} (\mathcal{X} \otimes \mathcal{A}))$;
- for each $t \in T$ the space of histories up to time t

  $(H_t, \mathcal{H}_t) := (K_{t-1} \times X, \mathcal{K}_{t-1} \otimes \mathcal{X})$.

Note that the empty product disappears from the above expressions. The sets $L_t \subset K_t$, $t \in T$ satisfy the following requirements:

(i) $L_t \in \mathcal{K}_t$;

(ii) for any $h_t \in H_t$ the set $L_{t\,h_t}$, the *projection of the* h_t*-section of* L_t (i.e. $L_{t\,h_t} = \{a \in A \mid (h_t, a) \in L_t\}$), is nonempty. The set $L_{t\,h_t}$ is called the *set of admissible actions in* h_t.

Note that $L_{t\,h_t}$ is an $\mathcal{A}$-measurable set (see e.g. Neveu (1965), th. III, 1.2).

Now we give the requirements for the transition mechanism.
The transitions made by the process from one coordinate of the sample space to the next are determined in part by the transition functions in the family $(p_t \mid t \in T)$. Any element p_t of this family is a transition probability from $(K_t, \mathcal{K}_t)$ into $(X, \mathcal{X})$, i.e. $p_t((x_0, a_0, \ldots, x_t, a_t), .)$ is a probability measure on $(X, \mathcal{X})$ for each $(x_0, a_0, \ldots, x_t, a_t) \in K_t$, and $p_t(., B)$ is a measurable function on $(K_t, \mathcal{K}_t)$ for each $B \in \mathcal{X}$.

For the other part, the transition mechanism of the process is determined by a *strategy* $\pi = (\pi_0, \pi_1, \ldots)$. This is a sequence of functions π_t, $t \in T$, such that π_t is a transition probability from $(H_t, \mathcal{H}_t)$ into $(A, \mathcal{A})$, with the condition that for each $h_t = (x_0, a_0, \ldots, a_{t-1}, x_t) \in H_t$ the probability measure $\pi_t(h_t, .)$ is concentrated on the set $L_{t\,h_t}$ of admissible actions in h_t. The $\mathcal{A}$-measurability of $L_{t\,h_t}$ has been noted before.

The set of all strategies is denoted by Π.

In the sequel we shall use the following convention: let $f: H \to \mathbb{R}$ be measurable with respect to the σ-field on H induced by $\mathcal{H}_t$, then we write $f(k_t)$ instead of $f(h)$.

Now, the Ionescu Tulcea theorem can be applied (see Neveu (1965) th.V.1.1 and its corollaries) to construct a probability measure for the process on the sample space. Since $(H,\mathcal{H})$ is a product space of measurable spaces, and since for each choice of a strategy π all the relevant transition probabilities are determined, it may be concluded that for every $x_0 \in X$ there exists a probability measure $\mathbb{P}_{x_0,\pi}$ on $(H,\mathcal{H})$, with the following properties.

Let $f: H \to \mathbb{R}$ be nonnegative. If f is measurable with respect to the σ-field on H, induced by $\mathcal{K}_k$, then it holds that

$$\int_H f(h)\ \mathbb{P}_{x_0,\pi}(dh) =$$

$$= \int_A\int_X\int_A \cdots \int_X\int_A f(x_0,a_0,\ldots,a_k)\ \pi_k((x_0,a_0,\ldots,x_k),\ da_k)\,.\ \ldots$$

$$\ldots\ .\ p_0((x_0,a_0),dx_1)\ \pi_0(x_0,da_0)\,;$$

and if f is measurable w.r.t. the σ-field on H, induced by $\mathcal{H}_k$, the

$$\int_H f(h)\ \mathbb{P}_{x_0,\pi}(dh) =$$

$$= \int_A\int_X \cdots \int_A\int_X f(x_0,a_0,\ldots,x_k)p_{k-1}((x_0,\ldots,a_{k-1}),dx_k)\,.\ \ldots$$

$$\ldots\ .\ p_0((x_0,a_0),\ dx_1)\ \pi_0(x_0,da_0)\,.$$

This $\mathbb{P}_{x_0,\pi}$ is the uniquely determined probability measure for the process which starts in x_0, with a transition mechanism prescribed by π and the family $(p_t \mid t \in T)$.

Let ν be a probability measure on $(X,\mathcal{X})$, called the *starting distribution*. In the Ionescu Tulcea theorem it is also asserted, that there exists a probability measure $\mathbb{P}_{\nu,\pi}$ on $(H,\mathcal{H})$, defined by

$$\mathbb{P}_{\nu,\pi}(H') = \int_X \mathbb{P}_{x,\pi}(H')\ \nu(dx) \text{ for all } H' \in \mathcal{H}.$$

It may even be concluded that for every $h_t \in H_t$ and $k_t \in K_t$ there exist probability measures $\mathbb{P}_{h_t,\pi}$ and $\mathbb{P}_{k_t,\pi}$ respectively, which are a version of the conditional probability measures for the process, given h_t and k_t, respectively. They satisfy the following conditions:

$$\mathbb{P}_{h_t,\pi}(H') = \int_A \mathbb{P}_{k_t,\pi}(H')\ \pi_t(h_t,da_t)$$

and $$\mathbb{P}_{k_t,\pi}(H') = \int_X \mathbb{P}_{h_{t+1},\pi}(H')\ p_t(k_t,\ dx_{t+1}) \quad \text{for all } H' \in \mathcal{H}.$$

REMARK. In the sequel we will use the probability measures $\mathbb{P}_{x_0,\pi}$ as well as $\mathbb{P}_{\nu,\pi}$. It would be convenient if every $\mathbb{P}_{x_0,\pi}$ could be considered as a special case of $\mathbb{P}_{\nu,\pi}$. Unfortunately, we cannot in general construct a starting distribution ν concentrated on the set $\{x_0\}$, since it is not necessary that $\{x_0\} \in \mathcal{X}$ for all $x_0 \in X$.
However, there always exists a ν such that $\mathbb{P}_{\nu,\pi} = \mathbb{P}_{x_0,\pi}$.
In fact, if we define for all $B \in \mathcal{X}$

$$\nu(B) = \mu_{x_0}(B) := \begin{cases} 1 & \text{if } x_0 \in B \\ 0 & \text{otherwise} \end{cases}$$

then ν is a probability measure on $(X,\mathcal{X})$ with

$$\mathbb{P}_{\nu,\pi}(D) = \int_X \mathbb{P}_{x,\pi}(D)\ \nu(dx) = \mathbb{P}_{x_0,\pi}(D) \qquad \text{for all } D \in \mathcal{H}.$$

Hence the $\mathbb{P}_{x_0,\pi}$ case is contained in the $\mathbb{P}_{\nu,\pi}$ case.

Since any strategy π selects with probability one only admissible actions, the following theorem is intuitively clear.

2.1.1. THEOREM. For every $x_0 \in X$ and $\pi \in \Pi$ we have

$$\mathbb{P}_{x_0,\pi}\left(\bigcap_{k=0}^{\infty} (L_k \times X \times A \times X \times A \times \ldots)\right) = 1.$$

PROOF. It is sufficient to prove, that for all $k \in T$

$$\mathbb{P}_{x_0,\pi}(L_k \times X \times A \times X \times A \times \ldots) = 1$$

(note that $L_k \times X \times A \times \ldots \in H$, since $L_k \in K_k$).

We reason as follows:

$$\mathbb{P}_{x_0,\pi}(L_k \times X \times A \times \ldots) = \int_A\int_X \ldots \int_X\int_A 1_{\{L_k\times X\times A\times\ldots\}}(h) \cdot$$

$$\cdot\ \pi_k((x_0,a_0,\ldots,x_k),da_k)\ p_{k-1}((x_0,\ldots,a_{k-1}),\ dx_k)\cdot \ldots \cdot$$

$$\cdot\ \pi_0(x_0,\ da_0) =$$

$$= \int_A\int_X \ldots \int_X 1\ p_{k-1}((x_0,\ldots,a_{k-1}),\ dx_k)\ \cdot \ldots \cdot\ \pi_0(x_0,da_0) = 1. \qquad \Box$$

Note that the sets L_t, determining the admissible actions, play no essential role in the description of the model for the D/G/G/1 process. However, the sets L_t restrict the set of possible strategies. This set-up is not unusual in papers on Markov decision processes, see Hinderer (1970), Blackwell (1965).

An important property of the set of strategies Π, which follows directly from our definition of a strategy, is the following. Let $\pi,\pi' \in \Pi$. Then a new strategy $\pi'' \in \Pi$ is specified by

$\pi''_k = \pi_k$ for $0 \le k < t$, $\pi''_k = \pi_k$ on $B \times A \times \mathop{\times}\limits_{\ell=t+1}^{k} (X \times A)$ and $\pi''_k = \pi'_k$ on

$B^c \times A \times \mathop{\times}\limits_{\ell=t+1}^{k} (X \times A)$ for $k \ge t$. This π'' is a strategy indeed, since each

π''_k is a transition probability from H_k to A, for each history h_k concentrated on

the corresponding set of admissible actions.
To be able to handle "heads" and "tails" of strategies appropriately, we introduce the following notations.
Let $\pi = (\pi_0, \pi_1, \ldots) \in \Pi$. Then ${}_t\pi = (\pi_0, \pi_1, \ldots, \pi_{t-1})$ is called the *head of* π *until time t*. Furthermore, for any $h_t = (x_0, a_0, \ldots, a_{t-1}, x_t) \in H_t$ we define $\pi(t;h_t) = (\pi'_0, \pi'_1, \ldots)$ with

$$\pi'_\tau(h'_\tau, \cdot) = \pi_{t+\tau}((x_0, a_0, \ldots, a_{t-1}, x'_0, a'_0, x'_1, \ldots, x'_\tau), \cdot)$$

for all $h'_\tau = (x'_0, a'_0, \ldots, x'_\tau) \in H_\tau$ and $\tau \in T$.

Thus, $\pi(t;h_t)$ is a strategy for the process which starts at time t, and $\pi(t;h_t)$ causes this process to behave stochastically the same as the original process from time t on, if the "history" h_t without x_t has occurred, and if the transitions depend on the strategy π. We call this strategy $\pi(t;h_t)$ the *tail of* π *given a history* h_t *before time t*. Note that in the case of a Markov process and a Markov strategy π, the occurrence of h_t in the tail $\pi(t;h_t)$ is not essential.
For each $t \in T$ we define F_t as the σ-field in H, generated by sets of type $H'_t \times A \times X \times A \times X \times \ldots$, with $H'_t \in H_t$. Moreover we introduce the following random variables: H denoting the whole history, H_t denoting the history up to time t, X_t denoting the state at time t, and A_t denoting the action at time t. More formally we say that H, H_t, X_t and A_t with $t \in T$ are measurable functions from H into H, such that H is the identical function, H_t is the projection from H into H_t, X_t is the projection on the 2t+1-th coordinate of H and A_t is the projection on the 2(t+1)-th coordinate. Instead of $r \circ H(h)$ we will use the notation $r(H)$. It should be noted that the σ-field F_t is precisely the σ-field generated by H_t.

$E_{x_0,\pi}$, $E_{\nu,\pi}$ etc. are the expectation operators with respect to the probability measures $\mathbb{P}_{x_0,\pi}$, $\mathbb{P}_{\nu,\pi}$ etc.

Now we are in a position to give another notation for tails of strategies. Consistently with the notation $\pi(t;h_t)$, we use $\pi(t;H_t)$ to denote $\pi(t;h_t)$ if $H_t = h_t$. The symbol $\pi(t;H_t)$ is called the *tail of* π *from time t on*. It is also possible to concatenate heads and tails of strategies, so each $\pi \in \Pi$ can be written as ${}_t\pi\pi(t;H_t)$.

We have given a description of all stochastic processes involved, and we have introduced some notations. Now we have reached the point where the decision part comes in.
The player of the D/G/G/1 process chooses a strategy, and in this way he "controls" the process. In order to attach a certain value to each strategy, we have the following requirements for the utility function. The utility function r is supposed to be a real valued measurable function on $(H,\mathcal{H})$. Moreover r is supposed to be quasi integrable with respect to each $\mathbb{P}_{\nu,\pi}$ with $\pi \in \Pi$ and ν a fixed starting distribution, i.e. either

$E_{\nu,\pi}\, r^+(H) < \infty$ or $E_{\nu,\pi}\, r^-(H) < \infty$, with $r^+(h) = \max(0,r(h))$ and $r^-(h) = \max(0,-r(h))$.

REMARK. ν is supposed to be fixed throughout this monograph.

Now the value of a strategy can be defined. For each $t \in T$ the *value of strategy* π*, given the history* $h_t = (x_0,a_0,\ldots,a_{t-1},x_t)$ *up to time* t, is a function $v_t : H_t \times \Pi \to \overline{\mathbb{R}}$ with $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty,\infty\}$, such that

$$v_t\,(h_t,\pi) = \begin{cases} E_{h_t,\pi}\, r(H) & \text{if this integral exists,} \\ -\infty & \text{otherwise.} \end{cases}$$

Consistently with this definition we will choose from now on, for all $t \in T$ and the starting distribution ν, the function $v_t(H_t,\pi)$ as our fixed representative for $E_{\nu,\pi}^{\mathcal{F}_t}\, r\,(H)$,

where $E_{\nu,\pi}^{\mathcal{F}_t}$ denotes the conditional expectation of $E_{\nu,\pi}$ w.r.t. $\mathcal{F}_t$.
(By the above assumptions about r the right-hand side is $\mathbb{P}_{\nu,\pi}$-almost everywhere defined.) This is called the *value of strategy* π*, given* H_t. The *value of the game, given* h_t, henceforth called: *the value given* h_t or the *value function*, is a function

$$w_t : H_t \to \overline{\mathbb{R}} \quad \text{with}$$

$$w_t\,(h_t) = \sup_{\pi\in\Pi} v_t\,(h_t,\pi),$$

and therefore the *value, given* H_t for a fixed starting distribution ν is

$$w_t(H_t)=\sup_{\pi\in\Pi} v_t(H_t,\pi) = \sup_{\pi\in\Pi} E^{F_t}_{\nu,\pi}\, r(H).$$

The higher the value of a strategy, the more the player prefers this strategy. Thus, we arrive at the concept of ν-optimality.

2.1.2. DEFINITION. A strategy $\pi^* \in \Pi$ is called *ν-optimal*, iff

$$v_t(H_t,\pi^*) = w_t(H_t) \quad \mathbb{P}_{\nu,\pi^*}\text{ - a.s.} \quad \text{for all } t \in T$$

REMARK. The value function w_t is not necessarily measurable. At the end of the next section we give some references where this problem is discussed.

We conclude this section with a few remarks on the model. Since the transition probabilities p_t depend on the action at time t, and on the history up to time t instead of only the state at time t, the model describes a class of decision processes, which is much more general than the class of Markov decision processes. The strategies we allow may also depend on the history, and may select randomized actions, the class of strategies under study is the class of randomized behavioural strategies. This class is fairly general, since by a result of Aumann (1964) so-called mixed strategies may be replaced by behavioural strategies, if for instance the history up to time t is known at every time t. The precise definitions of mixed and randomized behavioural strategies can also be found in Aumann (1964).
The utility functions we allow, are of the same generality as those in Kreps (1977). In the next chapter the more restrictive recursive utilities, as introduced in Furukawa and Iwamoto (1973), will arise quite naturally.

2.2. CHARACTERIZATION OF ν-OPTIMAL STRATEGIES

In this section it will be shown that the class of ν-optimal strategies coincides with the class of strategies that are both ν-conserving and ν-equalizing. As said before in chapter 1, conservingness means, that at every step prescribed by the strategy, you loose nothing, and equalizingness means, that in the long run the value of the strategy comes arbitrarily close to the value one can hope for from then on.

First it will be shown, that for each strategy π the value function w_t is a supermartingale, if some measurability conditions are satisfied. This generalizes a result in Groenewegen and van Hee (1977), where this property is proved for a special class of utility functions and for Markov strategies in the context of a D/D/D/1 Markov decision process. Recall that in general the value function is not measurable.

2.2.1. THEOREM.

Let ν be a starting distribution and π a strategy, such that for all $t \in T$ the value w_t is $\mathbb{P}_{\nu,\pi}$-almost equal to a measurable function. If the following condition is satisfied: for all probability measures μ on H_t, $t \in T$, all $\varepsilon > 0$ and all $m \in \mathbb{R}$ there exist strategies π',π'' such that

$$v_t(H_t,\pi') > w_t(H_t) - \varepsilon \qquad \mu - \text{a.s. on } \{h_t \in H_t \mid w_t(h_t) < \infty\},$$

and

$$v_t(H_t,\pi'') > m \qquad \mu - \text{a.s. on } \{h_t \in H_t \mid w_t(h_t) = \infty\},$$

then the value function is a supermartingale, i.e.

$$w_t(H_t) \geq E_{\nu,\pi}^{F_t}\, w_{t+1}(H_{t+1}) \qquad \mathbb{P}_{\nu,\pi} - \text{a.s.}$$

REMARK. For X and A complete separable metric spaces the result is well known, see Strauch (1966), Hinderer (1970), Shreve (1977). The condition in the theorem is satisfied, if there exists a μ-almost everywhere measurable selection from tails of strategies.
This is the point, where so-called selection theorems play a role, see the survey on this topic by Wagner (1977). Since theorem 2.2.1. is not really used in the sequel, we will not discuss a possible derivation of the conditions in the theorem from other conditions.

PROOF. Without loss of generality we may restrict ourselves to the case that w_t is finite.
Suppose there exist $\varepsilon > 0$, $t \in T$, $\pi \in \Pi$ and a starting distribution ν such that

$$E^{\mathcal{F}_t}_{\nu,\pi}\, w_{t+1}(H_{t+1}) > w_t(H_t) + \varepsilon \qquad \text{on } F_t \in \mathcal{F}_t \text{ with } \mathbb{P}_{\nu,\pi}(F_t) > 0.$$

By the condition in the theorem with μ the marginal probability corresponding to $\mathbb{P}_{\nu,\pi}$ on the (2t+1)-th coordinate, there exists a $\pi' \in \Pi$ such that

$$v_{t+1}(H_{t+1},\pi') > w_{t+1}(H_{t+1}) - \frac{1}{2}\varepsilon \qquad \mathbb{P}_{\nu,\pi} \text{ - a.s.}$$

Since $v_{t+1}(h_{t+1},\pi')$ does not depend on ${}_{t+1}\pi'$, we may conclude that

$$v_{t+1}(H_{t+1},\pi'') > w_{t+1}(H_{t+1}) - \frac{1}{2}\varepsilon \qquad \mathbb{P}_{\nu,\pi} \text{ - a.s.}$$

with ${}_{t+1}\pi'' = {}_{t+1}\pi$ and $\pi''(t+1;\,.) = \pi(t+1;.)$. Hence

$$v_t(H_t,\pi'') = E^{\mathcal{F}_t}_{\nu,\pi}\, v_{t+1}(H_{t+1},\pi'') > E^{\mathcal{F}_t}_{\nu,\pi}\, w_{t+1}(H_{t+1}) - \frac{1}{2}\varepsilon > w_t(H_t) + \frac{1}{2}\varepsilon$$

$\mathbb{P}_{\nu,\pi}$ - a.s. on F_t.

which is contradicted by the definition of w_t. □

This result in fact means, that the best the player can hope for at any given time, is not less than what he can hope for after the next step taken. In this light it seems plausible, that for a ν-optimal $\pi^* \in \Pi$ it is necessary that $(w_t(H_t) \mid t \in T)$ should be a martingale with respect to $\mathbb{P}_{\nu,\pi^*}$. This result is contained indeed in theorem 2.2.4., where we call this martingale property the ν-conservingness of π^*.

Before formulating and proving a characterization theorem for optimal strategies, we need the concepts of ν-conserving and ν-equalizing strategies.

2.2.2. DEFINITION. A strategy $\pi^* \in \Pi$ is called ν-*conserving* iff for all $t \in T$

$$w_t(H_t) = E^{\mathcal{F}_t}_{\nu,\pi^*}\, w_{t+1}(H_{t+1}) \qquad \mathbb{P}_{\nu,\pi^*} \text{ - a.s.,}$$

i.e. $(w_t(H_t) \mid t \in T)$ is a martingale with respect to $\mathbb{P}_{\nu,\pi^*}$.

(In this definition it is supposed that the right-hand side of the equation is well defined.)

The concept of conservingness used by Kreps (1977) is stronger, as his concept of optimality is stronger. His optimality concept is in fact the analogue of subgame perfectness, introduced in Selten (1965). We will come back to this in chapter 5 section 2.

2.2.3. DEFINITION. A strategy $\pi^* \in \Pi$ is called ν-*equalizing* iff

$$\lim_{t\to\infty} E_{\nu,\pi^*} [w_t(H_t) - v_t(H_t,\pi^*)] = 0.$$

(The left-hand side of the equation is supposed to be well defined.)

REMARK. Since

$$\lim_{t\to\infty} E_{\nu,\pi^*} v_t(H_t,\pi^*) = \lim_{t\to\infty} E_{\nu,\pi^*} E_{\nu,\pi^*}^{F_t} r(H) = E_{\nu,\pi^*} r(H),$$

ν-equalizingness of π^* can also be defined by

$$\lim_{t\to\infty} E_{\nu,\pi^*} w_t(H_t) = \lim_{t\to\infty} E_{\nu,\pi^*} v_t(H_t,\pi^*).$$

2.2.4. THEOREM. A necessary and sufficient condition for the ν-optimality of $\pi^* \in \Pi$ is that π^* is ν-conserving and ν-equalizing.

PROOF. Suppose π^* is ν-optimal. First we prove the ν-conservingness, using the definition of ν-optimality in the first and in the last equality.

$$w_t(H_t) = v_t(H_t,\pi^*) = E_{\nu,\pi^*}^{F_t} r(H) = E_{\nu,\pi^*}^{F_t} E_{\nu,\pi^*}^{F_{t+1}} r(H) =$$

$$= E_{\nu,\pi^*}^{F_t} v_{t+1}(H_{t+1},\pi^*) = E_{\nu,\pi^*}^{F_t} w_{t+1}(H_{t+1}) \quad \mathbb{P}_{\nu,\pi^*} \text{ - a.s.}$$

To show that π^* is ν-equalizing, we use the definition of ν-optimality. We have

$$E_{\nu,\pi^*} [w_t(H_t) - v_t(H_t,\pi^*)] = E_{\nu,\pi^*} [w_t(H_t) - w_t(H_t)] = 0,$$

for all $t \in T$. Taking the limit for $t \to \infty$, we obtain that ν-optimal strategies are both ν-conserving and ν-equalizing.
Now suppose π^* is ν-conserving and ν-equalizing, then

$$w_t(H_t) = E^{\mathcal{F}_t}_{\nu,\pi^*}\, w_{t+1}(H_{t+1}) = E^{\mathcal{F}_t}_{\nu,\pi^*}\, E^{\mathcal{F}_{t+1}}_{\nu,\pi^*}\, w_{t+2}(H_{t+2}) =$$

$$= E^{\mathcal{F}_t}_{\nu,\pi^*}\, w_{t+2}(H_{t+2}) = \ldots = E^{\mathcal{F}_t}_{\nu,\pi^*}\, w_\tau(H_\tau) \qquad \mathbb{P}_{\nu,\pi^*}\text{-a.s.}$$

for all $t,\tau \in T$ with $\tau \geq t$, since π^* is ν-conserving. Hence, using the ν-equalizingness

$$E_{\nu,\pi^*} w_t(H_t) = \lim_{\tau\to\infty} E_{\nu,\pi^*}\, w_\tau(H_\tau) = \lim_{\tau\to\infty} E_{\nu,\pi^*}\, v_\tau(H_\tau,\pi^*) =$$

$$= E_{\nu,\pi^*}\, r(H) = E_{\nu,\pi^*}\, v_t(H_t,\pi^*).$$

And since $v_t(H_t,\pi^*) \leq w_t(H_t)$ $\mathbb{P}_{\nu,\pi^*}$-a.s. it follows that

$$v_t(H_t,\pi^*) = w_t(H_t) \quad \mathbb{P}_{\nu,\pi^*}\text{-a.s.}$$ □

Let us make a few remarks about this last theorem and its proof. The part of the proof where it is shown that ν-optimality implies ν-conservingness can also be found in Kreps (1977).

The concepts of conserving and equalizing strategies can be found already in Dubins and Savage (1965), where they have been introduced and used in a characterization of optimal strategies in gambling situations. In Hordijk (1974) this characterization is given for the convergent dynamic programming case. His proof depends rather heavily on the special type of utility he considers, the so-called charge structure. In Groenewegen (1975) and Groenewegen and van Hee (1977) two different proofs of this characterization can be found in practically the same situation as in Hordijk, and these proofs can both be extended to the case of a more general utility and more players (for a two-person zero-sum Markov game this is partly done in Groenewegen and Wessels (1977) and Groenewegen (1976)). The proof in Groenewegen (1975) gives insight in the result itself, the proof in Groenewegen and van Hee (1977), however, is more concise. In the next chapter we come to speak about these proofs in more detail.
It should be noted that the ν-equalizingness of strategy π^* implies

the existence of the L^1-limit of $[w_t(H_t) - v_t(H_t,\pi^*)]$ for $t \to \infty$ with respect to the measure $\mathbb{P}_{\nu,\pi^*}$, since $w_t(H_t) - v_t(H_t,\pi^*)$ is nonnegative $\mathbb{P}_{\nu,\pi^*}$-a.s. Moreover we emphasize the fact that the problem of the value function w not being measurable, as extensively discussed in Blackwell, Freedman and Orkin (1974) and more recently in Shreve (1977), does not play any role at all here. This is so because we deal merely with a characterization of optimality. If there exists an optimal strategy π^*, then the value function equals the value of π^*, which, of course, is measurable indeed. When proving the other part of the characterization, the (quasi) integrability of the value function is implicit in the definitions of conservingness and equalizingness.

As a final remark within this chapter we want to make more clear why in the definition of ν-optimality the equality $v_t(\cdot,\pi^*) = w_t(.)$ is supposed to hold for all $t \in T$ and not only for $t = t_0$. It may be observed that there are cases where the equality for all $t \in T$ follows from the equality for $t = t_0$, so one may ask whether this is true in general. However, this depends on some topological requirements, which allow the use of some selection theorem to derive the equality for all $t \in T$. These requirements can be formulated in different ways, each of them corresponding to its own selection theorem. In order to avoid the choice of whatever set of topological requirements, we prefer to incorporate this property in the definition of ν-optimality itself.

CHAPTER 3

THE D/G/G/1 PROCESS WITH A RECURSIVE UTILITY

In this chapter we study the D/G/G/1 process with a recursive utility. It will be seen that the recursiveness enables us to reformulate the ν-conservingness and the ν-equalizingness. Thus we obtain a new form of the characterization of ν-optimality, which is more similar to the formulation given in e.g. Dubins and Savage (1965) and Hordijk (1974). In addition we will give two more proofs of this characterization, not depending on theorem 2.2.4.

The first of these two proofs makes rather explicit use of Bellman's optimality principle for t-recursive utilities, expressed in corollary 3.1.5. We quite agree with Gavish and Schweitzer (1976), who say that in various cases it is precisely this optimality principle, which is behind the proofs. This certainly applies to the characterization given here, since it actually was the use of the optimality principle, which motivated our study. (See Groenewegen (1975), Groenewegen and Wessels (1977), Groenewegen (1976).)

The second of the two extra proofs for the characterization of ν-optimality is the generalization of the concise proof in Groenewegen and van Hee (1977).

We begin this chapter with a section on t-recursive utilities, in which we have gathered some results for later use.

3.1. t-RECURSIVE UTILITIES

The first aim of this section is to give a sufficient condition for the following property to hold: the tail (from time t on) of an optimal strategy is itself optimal in those states the system can be in at time t. As it is formulated here, this is precisely the optimality principle as used in Bellman (1957), Gavish and Schweitzer (1976) and Groenewegen (1975). It turns out that t-recursiveness of the utility function suffices for the optimality principle to hold at a fixed time t.

A second result for t-recursive utilities, derived in this section, and needed for the validity of the optimality principle, also plays a role in the sequel. This result guarantees the possibility of splitting up $w_t(H_t)$, the value given H_t, into two parts: the first part only depends on the history up to time t, and the second part is just the value of a new decision process, which is "the tail from time t on" of the original decision process.

Let us denote the original D/G/G/1 process by Σ, so

$$\Sigma := (T,(X,\mathcal{X}),(A,\mathcal{A}),(L_\tau|\tau\in T),(p_\tau|\tau \in T),r).$$

Let $h_t = (x_0,a_0,\ldots,a_{t-1},x_t) \in H_t$. We introduce the "tail" of Σ from time t on given a history h_t, called the *t-delayed process given h_t* and denoted by $\Sigma^{[h_t]}$, as follows

$$\Sigma^{[h_t]} := (T^{[t]},(X,\mathcal{X}),(A,\mathcal{A}),(L_\tau^{[h_t]}|\tau\in T^{[t]}),$$
$$(p_\tau|\tau \in T^{[t]}),\ r^{[h_t]}).$$

Here

$$T^{[t]} := \{t_0+t,\ t_0+t+1,\ \ldots\},$$

$L_\tau^{[h_t]}$ is the $(x_0,a_0,\ldots,x_{t-1},a_{t-1})$-section of L_τ,

$r^{[h_t]}$ is a function $r^{[h_t]} : H \to \mathbb{R}$ with

$$r^{[h_t]}(x_t',a_t',x_{t+1}',a_{t+1}',\ldots) = r(x_0,a_0,\ldots,x_{t-1},a_{t-1},x_t',a_t')$$

(When in the sequel some symbol with the superscript $[h_t]$ is used, we mean the analogue for the process $\Sigma^{[h_t]}$ of what that symbol means in the process Σ. Sometimes $[h_t]$ may be replaced by $[t]$.) Note that $L_\tau^{[h_t]}$ is again measurable and also that each $(x_t',a_t,\ldots,a_{\tau-1},x)$ - section of $L_\tau^{[h_t]}$ is nonempty. We denote by $\Pi^{[h_t]}$ the set of strategies for the process $\Sigma^{[h_t]}$. This $\Pi^{[h_t]}$ is defined by

$$\Pi^{[h_t]} := \{\pi(t;h_t)|\pi \in \Pi\}.$$

As already concluded directly after the definition of the tail of a strategy, such a $\pi(t;h_t)$ is a strategy itself for the process which starts at time t, so the set $\Pi^{[h_t]}$ is well defined as a set of strategies.

The measure corresponding to a strategy $\pi(t,h_t) \in \Pi^{[h_t]}$ is defined by $\mathbb{P}^{[h_t]}_{x_t',\pi(t;h_t)}$.

3.1.1. DEFINITION. Let $t \in T$ be fixed. The D/G/G/1 process is called *t-separable* iff for all histories $h_t \in H_t$ and all strategies $\pi \in \Pi$ there exists a strategy $\pi^* \in \Pi$, such that for all $h_t'' \in H_t$ and all $x_t' \in X$

$$\mathbb{P}^{[h_t]}_{x_t',\pi(t;h_t)} = \mathbb{P}^{[h_t'']}_{x_t',\pi^*(t;h_t'')}$$

It is not difficult to see that a sufficient condition for t-separability is: the transition probabilities p_τ, $\tau \geq t$ as well as the admissibility of actions at times $\tau, \tau \geq t$ do not depend on $x_0,a_0,\ldots,x_{t-1},a_{t-1}$. Namely, choosing $h_t = (x_0,a_0,\ldots,x_{t-1},a_{t-1},x_t) \in H_t$ and $\pi \in \Pi$, we define π^* such that for each $h_\tau'' = (x_0'',a_0'',\ldots,x_\tau'') \in H_\tau$

$$\pi^*_\tau(h_\tau'',.) = \begin{cases} \pi_\tau((x_0,a_0,\ldots,x_{t-1},a_{t-1},x_t'',a_t'',\ldots,x_\tau''),.) & \text{if } \tau \geq t \\ \pi_\tau(h_\tau'',.) & \text{otherwise} \end{cases}$$

This is possible since for $\tau \geq t$ the admissibility of actions does not depend on $x_0,a_0,\ldots,a_{t-1}$. Since also the transition probabilities p_τ, $\tau \geq t$ do not depend on $x_0,\ldots,a_{t-1}$, it follows from the definition of π^* that $\pi^*(t;h_t'') = \pi(t;h_t)$, so

$$\mathbb{P}^{[h_t]}_{x_t',\pi(t;h_t)} = \mathbb{P}^{[h_t'']}_{x_t',\pi(t;h_t)} \quad \text{for all } h_t'' \in H_t \text{ and all } x_t' \in X$$

In the following definition we use the transformation $\zeta: H \to H$ with $\zeta(h) = \zeta(x_0,a_0,x_1,a_1,\ldots) = (x_1,a_1,x_2,a_2,\ldots)$ for all $h \in H$. We also use ζ on finite sequences: $\zeta: H_t \to H_{t-1}$ with $\zeta(h_t) = (x_1,a_1,\ldots,x_t)$.

3.1.2. DEFINITION. Let the D/G/G/1 process be t-separable for some $t \in T$. The utility r is called *t-recursive* iff

$$r(h) = \theta_t(h_t) + \chi_t(h_t) \cdot \rho(\zeta^t(h)),$$

where $\theta_t: H_t \to \mathbb{R}$, $\chi_t: H_t \to \mathbb{R}^+$ (the nonnegative real halfline) are measurable and integrable, and $\rho: H \to \mathbb{R}$ is measurable and quasi integrable, with respect to every $\mathbb{P}_{\nu,\pi}$ (or restriction of $\mathbb{P}_{\nu,\pi}$ to H_t) with ν our fixed chosen starting distribution. (Integrability of a measurable function f means $E_{\nu,\pi} |f| < \infty$.)

In other words, t -recursiveness means, that the utility function can be split up into a part which depends on the history up to time t, and a part which depends on the sample path beginning at time t.
Note that, though the admissibility of actions at a time $\tau, \tau \geq t$, does not depend on the history before time t, a certain action a may be admissible in state j at time 0, and in-admissible in j at time t, since the admissibility of action a still may depend on the time t itself.
Examples of t-recursive utilities can be easily given. The examples we give, are also examples of recursive utilities, which will be introduced in the beginning of the next section. The first example is the total reward or additive utility: corresponding to the action chosen at time t, $t \in T$, the player immediately receives a one step reward, which depends on the action chosen, on the state at time t and on the state at time t + 1 (cf. Blackwell (1970), Strauch (1966)). Then θ_t is the sum of the one step rewards up to time t - 1, χ_t is 1, and ρ is the sum of the one step rewards from time t on. In the case of a discounted (discount factor $\alpha > 0$) additive reward the function θ_t is the sum of the discounted one step rewards up to time t - 1, χ_t is α^t, and ρ is the sum of the discounted one step rewards from time t on. Another interesting example is the average reward: θ_t is 0, χ_t is 1, and $\rho(\zeta^t(h))$ is $r(h)$.

3.1.3. LEMMA. Let Σ be a t-separable process. For a given $x_t' \in X$ the set of measures $\mathbb{P}^{[h_t]}_{x_t',\pi(t;h_t)}, \pi \in \Pi$ does not depend on h_t.

PROOF. Actually this lemma follows directly from the definition of t-separability, since choosing $x_t' \in X$ and $h_t, h_t'' \in H_t$ there exists for each measure $\mathbb{P}^{[h_t]}_{x_t',\pi(t;h_t)}, \pi \in \Pi$ a measure $\mathbb{P}^{[h_t'']}_{x_t',\pi^*(t;h_t'')}, \pi^* \in \Pi$ such that both are equal. □

Denote the value of a strategy $\pi \in \Pi^{[h_t]}$ by $v_\tau^{[h_t]}(\cdot,\pi)$, $\tau \in T^{[h_t]}$, and the value function of the process $\Sigma^{[h_t]}$ by $w_\tau^{[h_t]}$, $\tau \in T^{[h_t]}$. The next lemma shows how the functions v_t and w_t can be separated into different parts.

3.1.4. LEMMA. If r is a t-recursive utility, then for all $h_t \in H_t$

$$v_t(h_t,\pi) = \theta_t(h_t) + \chi_t(h_t)\, v_t^{[h_t]}(x_t,\pi(t;h_t)),$$

and $$w_t(h_t) = \theta_t(h_t) + \chi_t(h_t)\, w_t^{[h_t]}(x_t).$$

PROOF. Choose $h_t \in H_t$, then we have

$$v_t(h_t,\pi) = E_{h_t,\pi}\, r(H) = E_{h_t,\pi}\, [\theta_t(H_t) + \chi_t(H_t)\rho(\zeta^t(H))] =$$

$$= \theta_t(h_t) + \chi_t(h_t) \cdot E_{h_t,\pi}\, \rho(\zeta^t(H)) =$$

$$= \theta_t(h_t) + \chi_t(h_t) \cdot E_{x_t,\pi(t;h_t)}\, \rho(H) = \theta_t(h_t) + \chi_t(h_t) v_t^{[h_t]}(x_t,\pi(t;h_t)).$$

From lemma 3.1.3. it follows that

$$w_t(h_t) = \theta_t(h_t) + \chi_t(h_t) \sup_{\pi\in\Pi} v_t^{[h_t]}(x_t,\pi(t;h_t)) =$$

$$= \theta_t(h_t) + \chi_t(h_t)\, w_t^{[h_t]}(x_t).$$ □

Since $v_\tau^{[h_t]}(\cdot,\pi(t;h_t))$ only depends on h_t via $\pi(t;h_t)$, we shall write $v_\tau^{[t]}(\cdot,\cdot)$ instead of $v_\tau^{[h_t]}$. Likewise we shall use $w_\tau^{[t]}(\cdot)$ instead of $w_\tau^{[h_t]}$, since by lemma 3.1.3. $w_\tau^{[h_t]}$ does not depend on h_t at all.

Now we come to the formulation of the optimality principle. For its formulation we need a new notation. Define the probability measure μ_x on $(X,\mathcal{X})$ as $\mu_x(B) = 1$ if $x \in B$ and $\mu_x(B) = 0$ otherwise, for any $B \in \mathcal{X}$ (cf. our remark in section 2.1, where we say that the measures $\mathbb{P}_{x_0,\pi}$ are contained in the measures $\mathbb{P}_{\nu,\pi}$).

3.1.5. COROLLARY. Let r be a t-recursive utility. If $\pi^* \in \Pi$ is ν-optimal, then for $\mathbb{P}_{\nu,\pi^*}$-almost all (a.a.) $h_t \in H_t$ with $\chi_t(h_t) > 0$, the strategy $\pi^*(t;h_t) \in \Pi^{[ht]}$ is μ_{x_t} - optimal for the process $\Sigma^{[ht]}$.

PROOF. By definition 2.1.2. we have for $\mathbb{P}_{\nu,\pi^*}$-a.a $h_t \in H_t$

$$v_t(h_t,\pi^*) = w_t(h_t).$$

By lemma 3.1.4. this means

$$\theta_t(h_t) + \chi_t(h_t)\, v_t^{[t]}(x_t,\pi^*(t;h_t)) = \theta_t(h_t) + \chi_t(h_t) w_t^{[t]}(x_t).$$

Hence for $\mathbb{P}_{\nu,\pi^*}$- a.a. $h_t \in H_t$ with $\chi_t(h_t) > 0$ we have

$$v_t^{[t]}(x_t,\pi^*(t;h_t)) = w_t^{[t]}(x_t),$$

which establishes the result. □

3.1.6. DEFINITION. If the ν-optimality of $\pi^* \in \Pi$ implies that $\pi^*(t;h_t)$ is μ_{x_t}-optimal in the process $\Sigma^{[h_t]}$ for $\mathbb{P}_{\nu,\pi^*}$- a.a. $h_t \in H_t$, then we say that the *optimality principle holds* for the process Σ.

Note that if the utility r is t-recursive for all $t \in T$, then in fact it is a direct consequence of the definition of ν-optimality that the optimality principle holds.

3.2. CHARACTERIZATION OF ν-OPTIMALITY IF THE UTILITY IS RECURSIVE

After the preliminary results of the preceding section, the recursiveness of the utility is now introduced. Recursiveness together with another condition, called the ν-vanishing tail, turns out to be sufficient to give formulations of ν-conservingness and ν-equalizingness, that are analogous to the formulations given in Hordijk (1974).

This more special form of the characterization of ν-optimality will be derived in this section in three different ways. The first proof uses the characterization established in theorem 2.2.4 and lemma 3.1.4. The second proof uses the optimality principle, and the third proof uses a suitably chosen function which gives rise to a martingale. Therefore this proof is referred to as the martingale approach, although martingale properties are not really used there.

We start by introducing the concept of recursiveness.

3.2.1. DEFINITION. Let the D/G/G/1 process Σ be t-separable for all $t \in T$. The utility r is called *recursive* iff for all $t \in T$ and all $\tau \in T^{[t]}$ there exist functions $\theta_\tau^{[t]}$, $\chi_\tau^{[t]}$ and $r^{[t]}$, such that

$$r^{[0]} = r,$$

$$r^{[t]}(h) = \theta_\tau^{[t]}(h_\tau) + \chi_\tau^{[t]}(h_\tau)\; r^{[\tau]}(\zeta^{\tau-t}(h))$$

for each sequence $(x_t,a_t,x_{t+1},a_{t+1},\dots) =: h =: (h_\tau,a_\tau,x_{\tau+1},\dots) \in \mathop{X}_{k=t}^{\infty} (X\times A)$

with $\theta_\tau^{[t]} : [\mathop{X}_{k=t}^{\tau-1} (X \times A)] \times X \to \mathbb{R}$,

$$\chi_\tau^{[t]} : [\mathop{X}_{k=t}^{\tau-1} (X \times A)] \times X \to \mathbb{R}^+,$$

$$r^{[t]} : \mathop{X}_{k=t}^{\infty} (X \times A) \to \mathbb{R},$$

both $\theta_\tau^{[t]}$ and $\chi_\tau^{[t]}$ measurable and integrable, and $r^{[t]}$ measurable and quasi integrable (with respect to the σ-fields generated by products of X and A and with respect to the probability measures induced by the measures $\mathbb{P}_{\nu,\pi}, \pi \in \Pi$). To ensure the uniqueness of the decomposition of $r^{[t]}$ we define $\chi_\tau^{[t]}(x_t,a_t,\dots,a_{\tau-1},x_\tau) = 0$ iff $r^{[\tau]}(h') =$ constant for all $h' = (x'_\tau,a'_\tau,x'_{\tau+1},a'_{\tau+1},\dots)$ with $x'_\tau = x_\tau$.
(As before, the empty product disappears from the above expressions.)

So a recursive utility is t-recursive for each $t \in T$, with $\rho = r^{[t]}$. Recursiveness implies that the $\theta_\tau^{[t]}$'s and the $\chi_\tau^{[t]}$'s are related in a certain sense.

3.2.2. LEMMA. Let r be a recursive utility. Then for each $\tau \in T$ and $h_\tau = (x_0,a_0,\dots,x_\tau) \in H_\tau$ we have

(i) $\quad \theta_\tau^{[0]}(h_\tau) = \sum_{k=1}^{\tau} [\prod_{\ell=1}^{k-1} \chi_\ell^{[\ell-1]}(x_{\ell-1},a_{\ell-1},x_\ell)]\; \theta_k^{[k-1]}(x_{k-1},a_{k-1},x_k),$

(ii) $\chi_\tau^{[0]}(h_\tau) = \prod_{k=1}^{\tau} \chi_k^{[k-1]}(x_{k-1},a_{k-1},x_k)$.

PROOF. We will prove this by induction. Note that $h_1 = (x_0,a_0,x_1)$, so that both assertions are obviously true for $\tau = 1$.

Suppose (i) and (ii) are true for $\tau = \sigma$. Then choosing $h = (x_0,a_0,\ldots) \in H$, and defining

$$\alpha(h_{\sigma+1}) = \alpha(x_0,a_0,\ldots,x_{\sigma+1}) =$$

$$= \sum_{k=1}^{\sigma+1} [\prod_{\ell=1}^{k-1} \chi_\ell^{\ell-1} (x_{\ell-1},a_{\ell-1},x_\ell)] \cdot \theta_k^{[k-1]} (x_{k-1},a_{k-1},x_k)$$

and $\beta(h_{\sigma+1}) = \beta(x_0,a_0,\ldots,x_{\sigma+1}) = \prod_{k=1}^{\sigma+1} \chi_k^{[k-1]}(x_{k-1},a_{k-1},x_k)$,

we may write

$$r^{[0]}(h) = \theta_\sigma^{[0]}(h_\sigma) + \chi_\sigma^{[0]}(h_\sigma)\ r^{[\sigma]}(\zeta^\sigma(h)) =$$

$$= \theta_\sigma^{[0]}(h_\sigma) + \chi_\sigma^{[0]}(h_\sigma)\ [\theta_{\sigma+1}^{[\sigma]}(x_\tau,a_\tau,x_{\tau+1}) + \chi_{\sigma+1}^{[\sigma]}(x_\tau,a_\tau,x_{\tau+1}) \cdot$$

$$\cdot\ r^{[\sigma+1]}(\zeta^{\sigma+1}(h)].$$

Using the induction hypothesis we get

$$r^{[0]}(h) = \alpha(h_{\sigma+1}) + \beta(h_{\sigma+1})\ r^{[\sigma+1]}(\zeta^{\sigma+1}(h)).$$

On the other hand

$$r^{[0]}(h) = \theta_{\sigma+1}^{[0]}(h_{\sigma+1}) + \chi_{\sigma+1}^{[0]}(h_{\sigma+1}) r^{[\sigma+1]}(\zeta^{\sigma+1}(h)).$$

Since the $\theta_\tau^{[0]}$'s and $\chi_\tau^{[0]}$'s are uniquely determined (note that $\chi_{\sigma+1}^{[0]}(h_{\sigma+1}) = 0 \Leftrightarrow \beta(h_{\sigma+1}) = 0$), the proof is completed. □

REMARK. It is easy to see, that we have obtained the following recursion relations for $\tau \geq 1$

$$\theta_{\tau+1}^{[0]}(h_{\tau+1}) = \theta_\tau^{[0]}(h_\tau) + \chi_\tau^{[0]}(h_\tau)\ \theta_{\tau+1}^{[\tau]}(x_\tau,a_\tau,x_{\tau+1}),$$

$$\chi_{\tau+1}^{[0]}(h_{\tau+1}) = \chi_\tau^{[0]}(h_\tau)\ \chi_{\tau+1}^{[\tau]}(x_\tau,a_\tau,x_{\tau+1}).$$

For a recursive utility, we can give an equivalent formula for ν-conservingness.

3.2.3. THEOREM. If r is a recursive utility, then the condition

$$3.2.3.1.\quad w_t^{[t]}(X_t) = E_{\nu,\pi}^{F_t} [\theta_{t+1}^{[t]}(X_t,A_t,X_{t+1}) + \chi_{t+1}^{[t]}(X_t,A_t,X_{t+1})\ w_{t+1}^{[t+1]}(X_{t+1})]$$

$\mathbb{P}_{\nu,\pi}$- a.s. for all $t \in T$, is a necessary and sufficient condition for the strategy $\pi \in \Pi$ to be ν-conserving.

Especially in the situation of an additive utility (see the examples given after definition 3.1.2), the interpretation of this theorem is intuitively obvious. Since in that case $\chi_{t+1}^{[t]} = 1$, the theorem says that a utility is conserving iff the value function equals the expected one-step reward plus the value in the next state.

PROOF. The following four assertions are equivalent, and the arguments leading to the equivalence of assertion j and j+1 are given directly after the (j+1)-th assertion. The first statement is the definition of ν-conservingness (definition 2.2.2).

(i) $\quad w_t(H_t) = E_{\nu,\pi}^{F_t}\ w_{t+1}(H_{t+1})\quad \mathbb{P}_{\nu,\pi}$ - a.s. for all $t \in T$.

(ii) $\quad \theta_t^{[0]}(H_t) + \chi_t^{[0]}(H_t)\ w_t^{[t]}(x_t) =$

$$= E_{\nu,\pi}^{F_t} [\theta_{t+1}^{[0]}(H_{t+1}) + \chi_{t+1}^{[0]}(H_{t+1})\ w_{t+1}^{[t+1]}(X_{t+1})]$$

$\mathbb{P}_{\nu,\pi}$-a.s. for all $t \in T$. Use lemma 3.1.4, and the fact that $\zeta^t(H_t) = X_t$.

(iii) $\theta_t^{[0]}(H_t) + \chi_t^{[0]}(H_t)\ w_t^{[t]}(X_t) = E_{\nu,\pi}^{F_t}\ \theta_t^{[0]}(H_t) + E_{\nu,\pi}^{F_t}\ \chi_t^{[0]}(H_t)\ .$

$$\cdot\ [\theta_{t+1}^{[t]}(X_t,A_t,X_{t+1}) + \chi_{t+1}^{[t]}(X_t,A_t,X_{t+1}) w_{t+1}^{[t+1]}(X_{t+1})]$$

$\mathbb{P}_{\nu,\pi}$-a.s. for all $t \in T$. Use the formulae of lemma 3.2.2.

(iv) $w_t^{[t]}(X_t) = E_{\nu,\pi}^{F_t} [\theta_{t+1}^{[t]}(X_t,A_t,X_{t+1}) + \chi_{t+1}^{[t]}(X_t,A_t,X_{t+1})\, w_{t+1}^{[t+1]}(X_{t+1})]$

$\mathbb{P}_{\nu,\pi}$-a.s. for all $t \in T$. Use the F_t-measurability of H_t.
Note that the last step is valid both for $\chi_t^{[0]} \neq 0$ and $\chi_t^{[0]} = 0$. □

REMARK. Completely analogously it can be proved, that formula 3.2.3.1 with " > " instead of " = " is equivalent to

$$w_t(H_t) > E_{\nu,\pi}^{F_t} w_{t+1}(H_{t+1}) \quad \mathbb{P}_{\nu,\pi} \text{ - a.s. for all } t \in T,$$

provided that r is recursive.

To make a reformulation of the ν-equalizingness possible, we assume the utility to be ν-tail vanishing.

3.2.4. DEFINITION. The utility r is called ν-*tail vanishing* (or is said to have a ν-*vanishing tail*) iff it is recursive and for all $\pi \in \Pi$

$$\lim_{t\to\infty} E_{\nu,\pi}\, \chi_t^{[0]}(H_t)\, v_t^{[t]}(X_t,\pi(t;H_t)) = 0.$$

REMARK. The function $v_t^{[t]}(X_t,\pi(t;H_t)$ is measurable, since $\chi_t^{[0]}(H_t)\, v_t^{[t]}(X_t,\pi(t;H_t)) = v_t(H_t,\pi) - \theta_t^{[0]}(H_t)$. The property of definition 3.2.4 implies, that $\lim_{t\to\infty} E_{\nu,\pi}\, \theta_t^{[0]}(H_t) = E_{\nu,\pi}\, v_0(H_0,\pi)$. This equality holds e.g. in the case of an additive utility, if $v_0(H_0,\pi) = E_{\nu,\pi}^{F_0} \sum_{k=0}^{\infty} \theta_{k+1}^{[k]}(X_k,A_k,X_{k+1})$ (i.e. the value of π equals the expected sum of one-step rewards) and $E_{\nu,\pi}^{F_0} \sum_{k=0}^{\infty} |\, \theta_{k+1}^{[k]}(X_k,A_k,X_{k+1}) \,| < \infty$.

Actually, this situation is described in Hordijk (1974), and he calls this property of the utility function the *charge structure*.

3.2.5. THEOREM. If r is a ν-tail vanishing utility, and $\pi \in \Pi$, then the following two assertions are equivalent.

(i) $\lim_{t\to\infty} E_{\nu,\pi} [w_t(H_t) - v_t(H_t,\pi)] \geqq 0,$

(ii) $\lim_{t\to\infty} E_{\nu,\pi}\, \chi_t^{[0]}(H_t)\, w_t^{[t]}(X_t) \gtreqqless 0.$

These formulae should both be read with equality, or both with strict inequality. Note that (ii) is equivalent to $\underline{\lim}_{t\to\infty} E_{\nu,\pi}\, w_t(H_t) \gtreqqless \underline{\lim}_{t\to\infty} E_{\nu,\pi}\, \theta_t^{[0]}(H_t)$.

PROOF. Choose $\pi \in \Pi$. The following three assertions are equivalent.

(i) $\lim_{t\to\infty} E_{\nu,\pi}\, [w_t(H_t) - v_t(H_t,\pi)] \gtreqqless 0.$

(ii) $\lim_{t\to\infty} E_{\nu,\pi}\, [\theta_t^{[0]}(H_t) + \chi_t^{[0]}(H_t)\, w_t^{[t]}(X_t) - \theta_t^{[0]}(H_t) +$

$- \chi_t^{[0]}(H_t)\, v_t^{[t]}(X_t,\, \pi(t;H_t))] \gtreqqless 0.$

Use lemma 3.1.4.

(iii) $\lim_{t\to\infty} E_{\nu,\pi}\, \chi_t^{[0]}(H_t)\, w_t^{[t]}(X_t) \gtreqqless 0,$

because r is ν-tail vanishing □

It is worth noting that theorem 3.2.5, read with equality signs, gives an equivalent criterion for ν-equalizingness. Therefore a combination of theorem 2.2.4 with theorems 3.2.3 and 3.2.5 leads to a new chracterization of ν-optimality.

3.2.6. COROLLARY. Let r be a ν-tail vanishing utility. Then a necessary and sufficient condition for the ν-optimality of $\pi^* \in \Pi$ is the validity of both

3.2.6.1. $w_t^{[t]}(X_t) = E_{\nu,\pi^*}^{F_t}\, [\theta_{t+1}^{[t]}(X_t,A_t,X_{t+1}) +$

$+ \chi_{t+1}^{[t]}(X_t,A_t,X_{t+1})\, w_{t+1}^{[t+1]}(X_{t+1})] \quad \mathbb{P}_{\nu,\pi^*}\text{-a.s.}$

for all $t \in T$, and

3.2.6.2. $\lim_{t\to\infty} E_{\nu,\pi^*}\, \chi_t^{[0]}(H_t)\, w_t^{[t]}(X_t) = 0$

We also mention here that by the nonnegativity of the expression in part (i) of theorem 3.2.5, we get the nonnegativity of the expression in part (ii).

3.2.7. COROLLARY. If r is a ν-tail vanishing utility, then

3.2.7.1. $\lim_{t\to\infty} E_{\nu,\pi}\, X_t^{[0]}(H_t)\cdot w_t^{[t]}(X_t) \geq 0$ for all $\pi \in \Pi$

if this expression is well defined.

3.2.8. DEFINITION. Let r be a recursive utility. The property of formula 3.2.7.1 is called the property *anne* of the value function (anne is the abbreviation of asymptotically nonnegative expectation).

So the property anne for the value function, introduced in Hordijk (1974) (definition 3.7 and theorem 3.9), holds far more generally than only in the situation, where the utility has a so-called charge structure (see Hordijk (1974) definition 2.12, and our remark after definition 3.2.4). By now we have seen a first proof of the result stated in corollary 3.2.6. A second proof of the same result utilizes the optimality principle in an essential way, and so it throws a somewhat different light on the situation. Actually, this way of attacking the characterization problem was the instigation to this monograph.
The optimality principle was used for the first time in this manner in Groenewegen (1975), to prove a result of Hordijk (1974). Afterwards this method turned out to be succesful in deriving a similar characterization for special kinds of optimality in two-person zero-sum Markov games (Groenewegen and Wessels (1977), Groenewegen (1976)), and in Markov games with countably many players (Couwenbergh (1977)). These results in game theory can be found in chapter 5 and 6 of this monograph.

3.2.9. SECOND PROOF OF COROLLARY 3.2.6. Suppose $\pi^* \in \Pi$ is ν-optimal. We first establish 3.2.6.1. Choose $t \in T$. Then

$$w_t^{[t]}(X_t) = v_t^{[t]}(X_t, \pi^*(t;h_t)) \text{ for } \mathbb{P}_{\nu,\pi^*}\text{ - a.s. } h_t \in H_t,$$

by the optimality principle. The right-hand side of this relation equals

$$E^{F_t}_{\nu,\pi^*}\, r^{[t]}(\zeta^t(H)) =$$

$$= E^{F_t}_{\nu,\pi^*}\, [\theta^{[t]}_{t+1}(X_t,A_t,X_{t+1}) + \chi^{[t]}_{t+1}(X_t,A_t,X_{t+1})\; r^{[t+1]}(\zeta^{t+1}(H))] =$$

$$= E^{F_t}_{\nu,\pi^*}\, [\theta^{[t]}_{t+1}(X_t,A_t,X_{t+1}) + E^{F_{t+1}}_{\nu,\pi^*}\, [\chi^{[t]}_{t+1}(X_t,A_t,X_{t+1})\; r^{[t+1]}(\zeta^{t+1}(H))]]=$$

$$= E^{F_t}_{\nu,\pi^*}\, [\theta^{[t]}_{t+1}(X_t,A_t,X_{t+1}) + \chi^{[t]}_{t+1}(X_t,A_t,X_{t+1})\; E^{F_{t+1}}_{\nu,\pi^*}\, r^{[t+1]}(\zeta^{t+1}(H))] =$$

$$= E^{F_t}_{\nu,\pi^*}\, [\theta^{[t]}_{t+1}(X_t,A_t,X_{t+1}) + \chi^{[t]}_{t+1}(X_t,A_t,X_{t+1}) \cdot$$

$$\cdot v^{[t+1]}_{t+1}(X_{t+1},\pi^*\cdot(t+1;H_{t+1}))] =$$

$$= E^{F_t}_{\nu,\pi^*}\, [\theta^{[t]}_{t+1}(X_t,A_t,X_{t+1}) + \chi^{[t]}_{t+1}(X_t,A_t,X_{t+1})\; w^{[t+1]}_{t+1}(X_{t+1})]\ ,$$

where we have used the optimality principle again in the last step.

Hence $w^{[t]}_t(X_t) = E^{F_t}_{\nu,\pi^*}\, [\theta^{[t]}_{t+1}(X_t,A_t,X_{t+1}) + \chi^{[t]}_{t+1}(X_t,A_t,X_{t+1})\; w^{[t+1]}_{t+1}(X_{t+1})]$

$\mathbb{P}_{\nu,\pi^*}$ - a.s.

Next, we establish formula 3.2.6.2, using the optimality principle.

$$\lim_{t\to\infty} E_{\nu,\pi^*}\, \chi^{[0]}_t(H_t)\; w^{[t]}_t(X_t) =$$

$$= \lim_{t\to\infty} E_{\nu,\pi^*}\, \chi^{[0]}_t(H_t)\; v^{[t]}_t(X_t,\pi^*(t;H_t)) = 0$$

by the ν-tail vanishing property. This completes the proof of the necessity of formulae 3.2.6.1 and 3.2.6.2.

Now suppose formulae 3.2.6.1 and 3.2.6.2 are valid. Then using 3.2.6.1

iteratively,

$$w_t^{[t]}(X_t) = E_{\nu,\pi^*}^{F_t} [\theta_{t+1}^{[t]}(X_t,A_t,X_{t+1}) + \chi_{t+1}^{[t]}(X_t,A_t,X_{t+1})\, w_{t+1}^{[t+1]}(X_{t+1})] =$$

$$= E_{\nu,\pi^*}^{F_t} [\theta_{t+1}^{[t]}(X_t,A_t,X_{t+1}) + \chi_{t+1}^{[t]}(X_t,A_t,X_{t+1})\, E_{\nu,\pi^*}^{F_{t+1}}[\theta_{t+2}^{[t+1]}(X_{t+1},A_{t+1},X_{t+2}) +$$

$$+ \chi_{t+2}^{[t+1]}(X_{t+1},A_{t+1},X_{t+2})\, w_{t+2}^{[t+2]}(X_{t+2})]] .$$

By Lemma 3.2.2. it follows that

$$w_t^{[t]}(X_t) = E_{\nu,\pi^*}^{t} [\theta_{t+2}^{[t]}(X_t,A_t,\ldots,X_{t+2}) + \chi_{t+2}^{[t]}(X_t,A_t,\ldots,X_{t+2})\, w_{t+2}^{[t+2]}(X_{t+2})] =$$

$$= E_{\nu,\pi^*}^{F_t} [\theta_\tau^{[t]}(X_t,A_t,\ldots X_\tau) + \chi_\tau^{[t]}(X_t,A_t,\ldots,X_\tau)\, w_\tau^{[\tau]}(X_\tau)].$$

Using 3.2.6.2. and the ν-tail vanishingness, we have

$$E_{\nu,\pi^*}\, w_t^{[t]}(X_t) = \lim_{\tau\to\infty} E_{x,\pi^*} [\theta_\tau^{[t]}(X_t,A_t,\ldots,X_\tau) + \chi^{[t]}(X_t,A_t,\ldots X_\tau).$$

$$.\, v_\tau^{[\tau]}(X_\tau,\pi^*(\tau;H_\tau))] = E_{\nu,\pi^*}\, v_t^{[t]}(X_t,\pi^*(t;H_t)).$$

By lemma 3.1.4. this implies

$$E_{\nu,\pi^*}\, w_t(H_t) = E_{\nu,\pi^*}\, v_t(H_t,\pi^*).$$

Since $w_t(H_t) \geq v_t(H_t,\pi^*)$ $\mathbb{P}_{\nu,\pi^*}$ - a.s., it follows that $w_t(H_t) = v_t(H_t,\pi^*)$

$\mathbb{P}_{x,\pi^*}$ - a.s.; hence π^* is ν-optimal. □

A third approach to corollary 3.2.6 is given in a proof of Groenewegen and van Hee (1977). In this proof a martingale is used, that is introduced in Mandl (1974) in connection with the average cost criterion for the optimal control of a Markov chain. We will use an analogous martingale here, without exploiting its martingale properties. This martingale can be described as follows: at each instant of time it is the one-step loss you incur by choosing some action, minus the expected one-step loss you incur by that action. We will refer to this third approach to corollary 3.2.6 as the *martingale approach*.

3.2.10. THIRD PROOF OF COROLLARY 3.2.6. This proof is only valid under the assumption of theorem 2.2.1. The expected one-step loss Λ, incurred by choosing strategy π, in state x_t at time t given history $h_t = (x_0, a_0, \dots, a_t)$, is defined for $\mathbb{P}_{\nu,\pi}$-a.a. histories h_t by

$$\Lambda(x_t, \pi(t;h_t)) =$$

$$= E_{h_t,\pi}\,[\theta^{[t]}_{t+1}\,(H^{[t]}_{t+1}) + \chi^{[t]}_{t+1}\,(H^{[t]}_{t+1})\; w^{[t+1]}_{t+1}(X_{t+1})] - w^{[t]}_{t}(x_t).$$

Here $H^{[t]}_t$ denotes $X_\tau, A_\tau, X_{\tau+1}, \dots, A_{t-1}, X_t$.
Then the function Λ is nonnegative, since the first term of the right-hand side equals $E_{h_t,\pi}\; w^{[t]}_{t+1}(X_t, A_t, X_{t+1})$, and since by theorem 2.2.1 the value function is a supermartingale. Let $\tau \in T$ be arbitrarily chosen. For each proces $\Sigma^{[h_\tau]}$ we introduce the one-step loss minus the expected one-step loss at time $t \geq \tau$ by means of a quantity Y^τ_t. This Y^τ_t is a real valued measurable function on $(H^{[\tau]}_{t+1}, \mathcal{H}^{[\tau]}_{t+1})$, defined for each $\tau, t \in T$ with $t \geq \tau$

$$Y^\tau_t = \chi^{[\tau]}_t(H_t)\;\theta^{[t]}_{t+1}(X_t, A_t, X_{t+1}) + \chi^{[\tau]}_{t+1}(H_{t+1})\; w^{[t+1]}_{t+1}(X_{t+1}) +$$

$$- \chi^{[\tau]}_t(H_t)\; w^{[t]}_t(X_t) - \chi^{[\tau]}_t(H_t)\;\Lambda(X_t, \pi(t;H_t)).$$

Note that Y^τ_t only depends on h_τ through τ.
Then for all $x_\tau \in X$, $\pi \in \Pi$ and $\tau, t \in T$ with $t \geq \tau$

$$E^{\mathcal{F}_t}_{x_\tau,\pi}\; Y^{[\tau]}_t = E^{\mathcal{F}_t}_{x_\tau,\pi}\;\{\, \chi^{[\tau]}_t(H_t)\;\theta^{[t]}_{t+1}(X_t, A_t, X_{t+1}) + \chi^{[\tau]}_{t+1}(H_{t+1})\;\cdot$$

$$\cdot\; w^{[t+1]}_{t+1}(X_{t+1}) - \chi^{[\tau]}_t(H_t)\; w^{[t]}_t(X_t) - \chi^{[\tau]}_t(H_t)\; E^{\mathcal{F}_t}_{x,\pi}\,[\theta^{[t]}_{t+1}(X_t, A_t, X_{t+1}) +$$

$$+ \chi^{[t]}_{t+1}(X_t, A_t, X_{t+1})\; w^{[t+1]}_{t+1}(X_{t+1}) - w^{[t]}_t(X_t)\,]\}.$$

As $\chi_t^{[\tau]}(H_t)$ is F_t-measurable,

$$E_{x_\tau,\pi}^{F_t} Y_t^\tau = E_{x_\tau,\pi}^{F_t} \{ \chi_t^{[\tau]}(H_t)\ \theta_{t+1}^{[t]}(X_t,A_t,X_{t+1}) + \chi_{t+1}^{[\tau]}(H_{t+1})\ w_{t+1}^{[t+1]}(X_{t+1}) +$$

$$- \chi_t^{[\tau]}(H_t)\ w_t^{[t]}(X_t) - \chi_t^{[\tau]}(H_t)\ \theta_{t+1}^{[t]}(X_t,A_t,X_{t+1}) +$$

$$- \chi_{t+1}^{[\tau]}(H_{t+1})\ w_{t+1}^{[t+1]}(X_{t+1}) + \chi_t^{[\tau]}(H_t)\ w_t^{[t]}(X_t) \} = 0.$$

Hence

$$0 = E_{\nu,\pi}^{F_\tau} \sum_{k=\tau}^{t} Y_k^\tau =$$

$$= E_{\nu,\pi}^{F_\tau} \{\theta_{t+1}^{[\tau]}(H_{t+1}) + \sum_{k=\tau+1}^{t+1} \chi_k^{[\tau]}(H_k)\ w_k^{[k]}(X_k) - \sum_{k=\tau}^{t} \chi_k^{[\tau]}(H_k)\ w_k^{[k]}(X_k) +$$

$$- \sum_{k=\tau}^{t} \chi_k^{[\tau]}(H_k)\ \Lambda(X_k,\pi(k;H_k))\} =$$

$$= E_{\nu,\pi}^{F_\tau} \{\theta_{t+1}^{[\tau]}(H_{t+1}) + \chi_{t+1}^{[\tau]}(H_{t+1})\ w_{t+1}^{[t+1]}(X_{t+1}) - w_\tau^{[\tau]}(H_\tau)$$

$$- \sum_{k=\tau}^{t} \chi_k^{[\tau]}(H_k)\ \Lambda(X_k,\pi(k;H_k))\}.$$

We let $t \to \infty$, and conclude on account of the ν-vanishing tail of r, putting the second and the fourth term together, that

3.2.10.1. $$\lim_{t\to\infty} E_{\nu,\pi}\ \chi_{t+1}^{[\tau]}(H_{t+1})\ w_{t+1}^{[t+1]}(X_{t+1}) - E_{\nu,\pi} \sum_{k=\tau}^{\infty} \chi_k^{[\tau]}(H_k)\ \cdot$$

$$\cdot \Lambda(X_k,\pi(k;H_k)) = E_{\nu,\pi}\ w_\tau^{[\tau]}(H_\tau) - \lim_{t\to\infty} E_{\nu,\pi}\ \theta_t^{[\tau]}(H_t) =$$

$$= E_{\nu,\pi}\ [w_\tau^{[\tau]}(H_\tau) - v_\tau^{[\tau]}(H_\tau,\pi)].$$

The first term in the top line of formula 3.2.10.1 is nonnegative by the property anne, corollary 3.2.7. The second term on the same line is non-positive by the first remark of this proof. The integrand in the bottom line (of 3.2.10.1) is nonnegative by definition. Hence π is ν-optimal, or equivalently formulated,

$$w_\tau(H_\tau) = v_\tau(H_\tau,\pi) \quad \mathbb{P}_{\nu,\pi} - \text{a.s. for all } \tau \in T.$$

iff the two terms on the top line of formula 3.2.10.1 vanish. By the definition of recursiveness, the last term on the same line vanishes iff for all $k \in T$

$$\Lambda(X_k,\pi(k;H_k)) = 0 \quad \mathbb{P}_{\nu,\pi} - \text{a.s.}$$

This in turn is nothing else than formula 3.2.6.1, while the vanishing of the first term on the top line of formula 3.2.10.1 is nothing else than formula 3.2.6.2. □

3.3. REMARKS AND EXAMPLES

We conclude this chapter with some remarks and examples.

A: Kreps (1977) uses a stronger optimality concept than ν-optimality. He calls a strategy $\pi^* \in \Pi$ optimal iff for all $t \in T$, $x_0 \in X$ and $\pi \in \Pi$

$$w_t(h_t) = v_t(h_t,\pi^*(t;h_t)) \quad \text{for } \mathbb{P}_{x_0,\pi} - \text{a.a. } h_t.$$

This means that every tail of π^* is optimal even for those histories that are possible at the beginning of the tail only by choosing nonoptimal actions prior to t. This kind of optimality is equivalent to the subgame perfectness from Selten (1975). This stronger optimality concept can also be characterized by means of conservingness and equalizingness. We do not give this characterization now, but we return to it in chapters 5 and 6.

B: The average reward as it is often used in Markov decision processes, is a recursive utility, since in that case we can choose $\theta_{t+1}^{[t]} = 0$ and $\chi_{t+1}^{[t]} = 1$ for all $t \in T$. However, it should be noted that in general the ν-tail vanishing condition is not fulfilled. Hence in the average-reward case ν-optimality is characterized by 3.2.6.1 and ν-equalizingness (definition 2.2.3 . The optimality principle (corollary 3.1.5) remains valid.

C: If the utility is recursive and all strategies are ν-equalizing (which happens for instance if the reward structure is additive, if there is a discount factor β, $0 \le \beta < 1$ and if r is bounded; cf. van Nunen (1976) and van Hee, Hordijk and van der Wal (1977)), then formula 3.2.6.1 is necessary and sufficient for ν-optimality. For a fixed $t \in T$ this formula depends only on π^* by π^*_t.

Let $\pi^* \in \Pi$ be ν-conserving. Suppose for a moment that $w_0^{[0]} < \infty$. Then it is intuitively clear that almost all actions a_t, selected by π^*_t in state x_t for a given history h_t, should have the property

3.3.C.1. $$w_t^{[t]}(x_t) = \int_X [\theta_{t+1}^{[t]}(x_t,a_t,x_{t+1}) + \chi_{t+1}^{[t]}(x_t,a_t,x_{t+1})\, w_{t+1}^{[t+1]}(x_{t+1})] \cdot \cdot\, p_t((x_t,a_t),dx_{t+1}).$$

The condition $w_0^{[0]} < \infty$, is satisfied, whenever there exists a ν-conserving strategy for the situation C, as can be seen by the following reasoning . If π is ν-conserving, then $E_{\nu,\pi}\, w_0^{[0]}(X_0) = E_{\nu,\pi}\, \theta_t^{[0]}(H_t) + E_{\nu,\pi}\, \chi_t^{[0]}(H_t) \cdot \cdot w_t^{[t]}(X_t)$. Since ν is also ν-equalizing and $\theta_t^{[0]}$ is $\mathbb{P}_{\nu,\pi}$-integrable, it follows that $w_0^{[0]}$ is $\mathbb{P}_{\nu,\pi}$ -integrable, so $w_0^{[0]}(x_0) < \infty$ for ν-a.a. $x_0 \in X$.

Let us call an action, satisfying 3.3.C.1., a conserving action, and let us suppose that $\{a\} \in A$ for all $a \in A$. If there exists a ν-conserving strategy π, we can construct a strategy π^* which selects always the same conserving action in $\mathbb{P}_{\nu,\pi}$-almost all states x, that can be reached with strategy π and starting distribution ν. The strategy π^* is ν-conserving since it prescribes conserving actions only. It is Markov since it depends only on the last state of the history. It is stationary since the choice of the action does not depend on the time. And it is nonrandomized since only one action is chosen. (See 3.3.F for a counterexample against this result, if the condition of the recursiveness is somewhat weakened.)

Hence we may conclude, that for Markov decision processes with a recursive utility and only ν-equalizing strategies for a given ν, the ν-optimality of a strategy implies the existence of a nonrandomized stationary Markov strategy which is also ν-optimal. Actually the above idea to derive the existence of stationary optimal strategies given the existence of an arbitrary optimal strategy, is quite commonly used (see e.g. Blackwell (1965) theorem 6).

D: The essential negative case (EN). Suppose r is a recursive and ν-tail vanishing utility. Define

$$m_t(X_t) = \sup_{\pi\in\Pi} E_{\nu,\pi}^{F_t} \sum_{k=t}^{\infty} [\prod_{\ell=t}^{k-1} \chi_{\ell+1}^{[\ell]}(X_\ell, A_\ell, X_{\ell+1})][\theta_{k+1}^{[k]}(X_k, A_k, X_{k+1})]^+ .$$

Suppose furthermore

(i) $\lim_{t\to\infty} E_{\nu,\pi}\, m_t(X_t) = 0$ for all $\pi \in \Pi$.

The condition (i) is a weakened version of the condition C^+ in Hinderer (1971), and also of the condition C in Schäl (1975). Clearly it is satisfied if

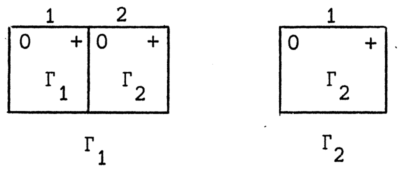
(ii) $\sum_{k=0}^{\infty} \| \prod_{\ell=0}^{k-1} \chi_{\ell+1}^{[\ell]}(x_\ell, a_\ell, x_{\ell+1})\ \theta_{k+1}^{k}(x_k, a_k, x_{k+1}) \| < \infty,$

with $\|\ \|$ the usual supremum norm. The case where the "additive analogue" of condition (ii) holds, can be found in Hinderer (1970), where it is called *the essential negative case (EN)*. We will use this term for the analogous situation, covered by condition (i) and the ν-tail vanishing property. Evidently, each strategy $\pi \in \Pi$ is ν-equalizing in the EN case, since (i) holds, and on the other hand the property anne holds (corollary 3.2.7). So ν-conservingness is necessary and sufficient for ν-optimality. This is also established for a more special model in Striebel (1975). And by remark C it follows for an EN Markov decision process, that if there exists an optimal strategy, then there exists a nonrandomized stationary Markov strategy which is also ν-optimal. This generalizes a result of Strauch (1966) for the case that the utility has the ν-tail vanishing property.

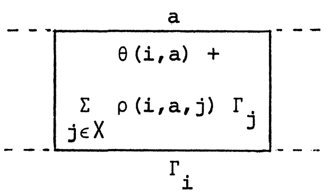
E: The condition in D that r is ν-tail vanishing is essential, as is shown by the following example.

3.3.1. THEOREM. COUNTEREXAMPLE. If r is not ν-tail vanishing but only recursive, then condition (i) in D does not imply that all strategies are ν-equalizing.

PROOF. We introduce the following D/F/F/1 process



Here the notation

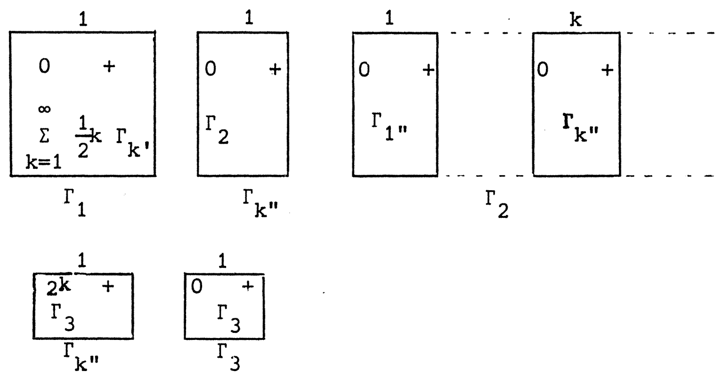


means the following. In state i a "game" Γ_i is played, i.e. if in state i the player chooses action a, then the system moves with probability p(i,a,j) to state j, and a one step reward $\theta(i,a)$ is earned, not depending on the time t and the state j, in other words $\theta_{t+1}^{[t]}(i,a,j) = \theta(i,a)$ for all $t \in T$ and $j \in X$. (Note that we have used a superfluously complicated notation. This notation will be needed later for a more complicated case.) The utility function is defined as

$$r(h) = \begin{cases} -1 & \text{if } h = (1,1,1,1,\dots,1,1,\dots), \\ 0 & \text{otherwise.} \end{cases}$$

In fact, this utility is the usual average gain in a Markov decision process, where the one-step gain in state 1 with action 1 equals -1, and all the other gains equal 0. It is easily verified that the strategy which prescribed always action 1 in state 1 is not equalizing. □

F: We here give an example, showing that the result of remark 3.3.C. does not hold if the assumptions are only weakened in such a way, that $\theta_{\tau}^{[t]}$ is allowed to be quasi integrable. It turns out that $w_0^{[0]}$ may be infinite, in which case it may be impossible to construct an optimal nonrandomized stationary Markovian strategy from a given optimal strategy.
We introduce the following D/D/D/1 process



$r(h) = \sum_{k=0}^{\infty} \theta(x_k, a_k)$, ν is concentrated on $\{1\}$. Let π be such, that on time 2 in state 2 action k_0 is selected with probability 1 at time 1, if the system was in state k_0'. Then $v_0^{[0]}(1,\pi) = \infty$, hence $w_0^{[0]}(1) = \infty$. Each strategy is ν-tail vanishing (except that $\theta_\tau^{[t]}$ is quasi integrable instead of integrable) and ν-equalizing. However, the value in state 1 of an arbitrary nonrandomized Markov strategy is finite. Nevertheless, there exists a randomized Markov strategy π^* which is ν-optimal. (Choose π^* such that on time 2 in state 2 action k is selected with probability $\frac{1}{2^k}$.)

CHAPTER 4
THE C/G/G/1 PROCESS

In continuous-time stochastic control processes most attention is paid to the case with a finite time horizon. In that situation one can expect, that all strategies are equalizing, and that optimality may be characterized by some kind of conservingness only. Actually, this type of characterization is known already for certain classes of control processes: for processes that can be written as the sum of a "nice" deterministic process and a Wiener process (Davis and Varaiya (1973)), and for jump processes (Boel and Varaiya (1977)). As in Richel (1970) this type of conservingness is called the "principle of optimality", or "Bellman's principle of optimality" (cf. our remarks in the introductory chapter).
Even in the case where Boel and Varaiya (1977) admit an infinite time horizon, the equalizing property holds for all strategies, since the situation considered there is the continuous-time analogue of the essential negative (EN) case with additive rewards (see remark D in chapter 3 section 3).
Again we shall characterize optimality in this chapter. In section 1 we set up the underlying model. The infinite-time horizon case will be treated in section 2. In section 3 we study recursive utilities, and we give some examples.

4.1. THE DESCRIPTION OF THE C/G/G/1 PROCESS

In the introduction to the D/G/G/1 process we started with a description of the transition mechanism of the process. The transition structure enabled us to use Ionescu Tulcea's theorem to construct probability measures on the sample space. However, in the continous-time case there is no result like Ionescu Tulcea's theorem. Therefore we prefer to start with measures on the (measurable) sample space. This approach turns out to be quite adequate in deriving the characterization of optimality as given in section 2 of this chapter.
Thus we avoid the type of problems, that are treated in Doshi (1976) for continuous-time Markov decision processes by techniques from Dynkin (1965). In stochastic control, problems of this kind are treated with measure

transformations: Beneš (1971) introduced the so-called Girsanov measure transformations (see Girsanov (1960)) to handle Brownian noise (see also Davis and Varaiya (1973)), and Boel, Varaiya and Wong (1975) introduced a similar technique to treat jump processes.
The general C/G/G/1 process is defined here as a tuple

$$\Sigma := (T,\ (X,\mathcal{X}),\ (A,\mathcal{A}),\ U,\ \{\mathbb{P}_{x_0,u} \mid x_0 \in X,\ u \in U\},r) \text{ together}$$

with a set of requirements.

- T, the *time space*, is some subset of $\mathbb{R}$, containing a lowest element, say t_0;
- X is the *state space*, endowed with a σ-field $\mathcal{X}$;
- A is the *action space*, endowed with a σ-field $\mathcal{A}$;
- U is the set of *controls* (in the continous-time case it is usual to talk about the set of controls U, whereas the discrete-time analogue is called the set of strategies Π);
- $\{\mathbb{P}_{x_0,u} \mid x_0 \in X,\ u \in U\}$ is a set of probability measures on the *sample space* $(H,\mathcal{H}) = (\underset{t\in T}{\times} (X \times A)_t,\ \underset{t\in T}{\otimes} (\mathcal{X} \otimes \mathcal{A})_t)$;
- r is the *utility function*.

The set of requirements will be specified further down.
Let us first remark that by the notation $\mathcal{H} = \underset{t\in T}{\otimes}(\mathcal{X} \otimes \mathcal{A})_t$ we denote the σ-field, generated by sets of type $\underset{t\in T}{\times}(X_t \times A_t)$ with $X_t \in \mathcal{X}$ and $A_t \in \mathcal{A}$ and $X_t \times A_t$ unequal to $X \times A$ for only finitely many $t \in T$. These generating sets are called *cylinders of a finite base*. For each $t \in T$ we define a σ-field $\mathcal{F}_t$, which is generated by a special subset of these cylinders, namely by sets of type $\underset{\tau\in T}{\times}(X_\tau \times A_\tau)$ with $X_\tau \in \mathcal{X}_\tau, A_\tau \in \mathcal{A}_\tau$, $X_\tau \times A_\tau$ unequal to $X \times A$ for finitely many τ, and $X_\tau = X$ if $\tau > t$ and $A_\tau = A$ if $\tau \geq t$. Hence for each sequence $t_0 < t_1 < t_2 < \ldots$ in T it follows that $\mathcal{F}_{t_0} \subset \mathcal{F}_{t_1} \subset \ldots \subset \mathcal{H}$.

Define $H_t = [\underset{\substack{\tau\in T\\ \tau<t}}{\times} (X \times A)] \times X$, the truncation of H at time t. The set H_t is called the set of histories up to time t. The symbol $\mathcal{H}_t$ denotes the σ-field induced by $\mathcal{H}$ on H_t. If $h \in H$, then h_t is the restriction of h to H_t.

We have the following requirements for the controls $u \in U$. The set of controls U is a set of functions $u: T \times H \times A \to [0,1]$, such that $u(t,.,.)$ is a transition probability from $(H_t, \mathcal{H}_t)$ into $(A, \mathcal{A})$. So u is *nonanticipative*, i.e. $u(t,h',\cdot) = u(t,h'',\cdot)$ for all $t \in T$ and all $h', h'' \in H$ with $h'_t = h''_t$. This enables us to write $u(t,h_t,B)$ instead of $u(t,h,B)$ for any $t \in T$ and $B \in \mathcal{A}$.

Moreover, we assume that U is closed under exchange of tails, i.e. if $u',u'' \in U$, $t \in T$ and $B \in \mathcal{H}_t$, then there exists a $u \in U$ such that for all $D \in \mathcal{A}$, $\tau \in T$ and $h_\tau \in H_\tau$

$$u(\tau,h_\tau,D) = \begin{cases} u''(\tau,h_\tau,D) & \text{if } \tau \geq t \text{ and } h_t \in B, \\ \\ u'(\tau,h_\tau,D) & \text{otherwise,} \end{cases}$$

with h_t the restriction of h_τ $(\tau \geq t)$ to H_t.

Next we formulate the requirements for the probability measures $\mathbb{P}_{x,u}$. Let ν be a fixed probability measure on $(X, \mathcal{X})$, the so-called *starting distribution*. We assume that $\mathbb{P}_{x,u}$ is measurable in x, and we define

$$\mathbb{P}_{\nu,u} := \int_X \mathbb{P}_{x,u}\, \nu(dx).$$

We also assume the existence of a probability measure $\mathbb{P}_{h_t,u}$ on H for each $h_t \in H_t$ and $u \in U$, such that $\mathbb{P}_{h_t,u}$ is an $\mathcal{F}_t$-measurable function of h_t, and moreover that $\mathbb{P}_{h_t,u}(h_t \times A' \times \underset{\tau \in T}{\underset{\tau>t}{\times}} (X \times A)) = u(t,h_t,A')$ for all $t \in T$, $u \in U$, $A' \in \mathcal{A}$.

We suppose, that $\mathbb{P}_{h_t,u}$ depends nonanticipatively on u, i.e. for all $\tau \in T$ and all $B \in \mathcal{F}_\tau$ we have $\mathbb{P}_{h_t,u'}(B) = \mathbb{P}_{h_t,u''}(B)$ whenever $u',u'' \in U$ satisfy $u'(\sigma,\cdot,\cdot) = u''(\sigma,\cdot,\cdot)$ for $t \leq \sigma < \tau$.

It follows that the expectation operator, corresponding to $\mathbb{P}_{\nu,u}$ and written as $E_{\nu,u}$, has a similar nonanticipativity property, viz. $E_{\nu,u'}y = E_{\nu,u''}y$ if y is $\mathcal{F}_t$-measurable and $u'(\tau,\cdot,\cdot) = u''(\tau,\cdot,\cdot)$ for $\tau \leq t$. A similar result holds for $E_{h_t,u}$, the expectation operator corresponding to $\mathbb{P}_{h_t,u}$.

The correspondence between the measures $\mathbb{P}_{h_t,u}$, $t \in T$ is given by

(i) $$\int_H f(h)\ \mathbb{P}_{h'_s,u}(dh) = \int_H\int_H f(h)\ \mathbb{P}_{h''_t,u}(dh)\ \mathbb{P}_{h'_s,u}(dh'')$$

for any $s,t \in T$ with $s \le t$, and nonnegative H-measurable function f,

(ii) $$\int_H f(h)\cdot g(h)\ \mathbb{P}_{h_t,u}(dh) = f(h_t)\int_H g(h)\ \mathbb{P}_{h_t,u}(dh)$$

for any $t \in T$, nonnegative H-measurable function g and nonnegative F_t-measurable function f (i.e. $f(h') = f(h'')$ if $h'_t = h''_t$, and we write $f(h'_t)$ instead of $f(h')$).

As before, the random variable H denotes the whole history, and the random variable H_t denotes the history up to time t, including X_t the state at time t, and excluding A_t the action at time t.

Formula (i) and (ii) mean, that the fundamental properties of the conditional expectation $E^{F_t}_{\nu,u} f(H)$ hold even everywhere on H.

Finally, we give the requirements for the utility function. The utility function $r: H \to \mathbb{R}$ is assumed to be measurable and quasi integrable w.r.t. any $\mathbb{P}_{\nu,u}$, $u \in U$ arbitrarily chosen and ν fixed.

All foregoing requirements (for the controls, the probability measures and the utility function) together with the tuple Σ define a C/G/G/1 process. We shall refer to this process as the process Σ.

Corresponding to the C/G/G/1 process Σ, there exists for each $t \in T$ and $h_t \in H_t$ a C/G/G/1 process $\Sigma^{[h_t]}$, defined by

$$\Sigma^{[h_t]} = (T^{[h_t]}, (X,X), (A,A), U^{[h_t]},$$
$$\{\mathbb{P}^{[h_t]}_{x'_t,u^{[h_t]}} \mid x'_t \in X, u^{[h_t]} \in U^{[h_t]}\}, r^{[h_t]}).$$

$\Sigma^{[h_t]}$ is called the *t-delayed process (given the history before time t).*

$T^{[h_t]} := \{\tau \in T \mid \tau \geq t\}$.

$U^{[h_t]} := \{u^{[h_t]} : T^{[h_t]} \times H^{[h_t]} \times A \to [0,1] \mid$ there exists a $u \in U$ such that $u^{[h_t]}(\tau, h', B) = u(\tau, h_t h', B)$, on $T^{[h_t]} \times H^{[h_t]} \times A\}$,

with $h_t h'$ the concentenation of h_t and h' in the sense that x_t, the last component of h_t, has disappeared (cf. the introduction of $\pi(t;h_t)$ in section (2.2), and with $H^{[h_t]} := \mathop{X}\limits_{\tau \in T^{[h_t]}} (X \times A)$, the sample space of the t-delayed process $\Sigma^{[h_t]}$. Instead of $u^{[h_t]}$ we write $u_t(h_t)$, and this symbol is called *the tail of u from time t on, given h_t*. For all $\tau \in T^{[h_t]}$ we define $H_\tau^{[h_t]} := (\mathop{X}\limits_{\substack{\sigma \in T^{[h_t]} \\ \sigma < \tau}} (X \times A)) \times X$, the set of histories up to time τ of the process $\Sigma^{[h_t]}$. Let $H^{[h_t]}, F_\tau^{[h_t]}, H_\tau^{[h_t]}$ be the σ-fields on $H^{[h_t]}, H^{[h_t]}, H_\tau^{[h_t]}$, which are the restriction of H, F_τ, H_τ respectively. It follows directly, that $u_t(h_t)(\tau, \cdot, \cdot)$ is a transition probability, that $u_t(h_t)$ is nonanticipative, and that $U^{[h_t]}$ is closed under exchange of tails, since u has these properties.

For all $\sigma \in T^{[h_t]}, B \in F_\tau^{[h_t]}$ we define

$$\mathbb{P}^{[h_t]}_{x_0', u_t(h_t)}(B) = \mathbb{P}_{h_t x_0', u}(\mathop{X}\limits_{\substack{s \in T \\ s < t}} (X \times A) \times B),$$

$$\mathbb{P}^{[h_t]}_{h_\sigma', u_t(h_t)}(B) = \mathbb{P}_{h_t h_\sigma', u}(\mathop{X}\limits_{\substack{s \in T \\ s < t}} (X \times A) \times B).$$

Note that the probability measures on the right-hand side of each of both equations do not depend on the part of the control before time t, since these measures are nonanticipative.

The nonanticipativity of $\mathbb{P}^{[h_t]}_{h'_\sigma, u_t(h_t)}$ follows also from the nonanticipativity of $\mathbb{P}_{h_t h'_\sigma, u}$. To verify the properties (i) and (ii) for $\mathbb{P}^{[h_t]}_{h'_\sigma, u_t(h_t)}$ we need the following notation. For a function $f\colon H \to \mathbb{R}$, define $f^{[h_t]}\colon H^{[h_t]} \to \mathbb{R}$ by $f^{[h_t]}(h) = f(h_t h)$. Then, using (i) in the second step, we have for any $h'_\sigma \in H^{[h_t]}_\sigma$ and all $h'''_\tau \in H^{[h_t]}_\tau$

$$\int_{H^{[h_t]}} f^{[h_t]}(h")\ \mathbb{P}^{[h_t]}_{h'_\sigma, u_t(h_t)}(dh") = \int_H f(h")\ \mathbb{P}_{h_t h'_\sigma, u}(dh") =$$

$$= \int_H \int_H f(h")\ \mathbb{P}_{\tilde{h}_\tau, u}(dh")\ \mathbb{P}_{h_t h'_\sigma, u}(d\tilde{h}) =$$

$$= \int_H \int_H f(h")\ \mathbb{P}_{h_t h'''_\tau, u_\tau}(dh")\ \mathbb{P}_{h_t h'_\sigma, u}(dh''') =$$

$$= \int_{H^{[h_t]}} \int_{H^{[h_t]}} f^{[h_t]}(h")\ \mathbb{P}^{[h_t]}_{h'''_\tau, u_t(h_t)}\, d(h")\ \mathbb{P}^{[h_t]}_{h'_\sigma, u}(dh''').$$

This establishes property (i) for $\mathbb{P}^{[h_t]}_{h'_\sigma, u_t(h_t)}$. Property (ii) can be proved analogously.

Finally we define $r^{[h_t]}$ by $r^{[h_t]}\colon H^{[h_t]} \to \mathbb{R}$, such that $r^{[h_t]}(h) = r(h_t h)$. The measurability of $r^{[h_t]}$ follows from the measurability of r. The utility $r^{[h_t]}$ is not necessarily quasi integrable w.r.t. every $\mathbb{P}^{[h_t]}_{\mu, u_t(h_t)}$, with $u_t(h_t) \in U^{[h_t]}$ and μ the marginal probability measure on the t-th coordinate of H induced by $\mathbb{P}_{\nu,u}$. Here ν is our fixed chosen starting distribution, and u is any control in U, the tail of which is $u_t(h_t)$.

However, for $\mathbb{P}_{\nu,u}$ - a.a. $h_t \in H_t$ the function $r^{[h_t]}$ is quasi integrable w.r.t. $\mathbb{P}^{[h_t]}_{\mu,u_t(h_t)}$, since otherwise the quasi integrability of r w.r.t. $\mathbb{P}_{\nu,u}$ would be violated.

Beside the tail $u_t(h_t)$ of a control $u \in U$, we define the *head of a control u up to time t* $\in T$ as a function ${}_tu:\{\tau \in T | \tau < t\} \times H \times A \to [0,1]$ with ${}_tu(\tau,h,\cdot) = u(\tau,h,\cdot)$, the restriction of u to $\{\tau \in T | \tau < t\} \times H \times A$.

The family of processes $(\Sigma^{[h_t]} \mid h_t \in H_t,\ t \in T)$

has the semigroup property, i.e.

$(\Sigma^{[h_t]})^{[h'_\tau]} = \Sigma^{[h_t h'_\tau]}$ for all $h_t \in H_t$, $h'_\tau \in H_\tau^{[h_t]}$.

We will prove this property only for the probability measures of the process. Choosing $h_t \in H_t$ and $h'_\tau \in H_\tau^{[h_t]}$, we have

$$(\mathbb{P}^{[h_t]})^{[h'_\tau]}_{x'_\tau,(u_t(h_t))_\tau(h'_\tau)} = \mathbb{P}^{[h_t]}_{h'_\tau,u_t(h_t)} = \mathbb{P}_{h_t h'_\tau,u} = \mathbb{P}^{[h_t h'_\tau]}_{x'_\tau,u_\tau(h_t h'_\tau)}.$$

The semigroup property can be proved even more easily for the other components of the process.

For each $t \in T$ the *value of a control u, given* $h_t \in H_t$ is a function v_t: $H_t \times U \to \mathbb{R}$ with

$$v_t(h_t,u) = \begin{cases} E_{h_t,u}\ r(H) & \text{if this integral exists,} \\ -\infty & \text{otherwise.} \end{cases}$$

For each $t \in T$ the *value given* $h_t \in H_t$ is a function w_t: $H_t \to \mathbb{R}$ with

$$w_t(h_t) = \sup_{u\in U} v_t(h_t,u).$$

Again in accordance with these definitions we have for each $t \in T$ and for a fixed starting distribution ν

$$v_t(H_t,u) = E_{\nu,u}^{F_t} r(H) \quad \mathbb{P}_{\nu,u} - \text{a.s.}$$

and $$w_t(H_t) = \sup_{u\in U} v_t(H_t,u) = \sup_{u\in U} E_{\nu,u}^{F_t} r(H).$$

Analogous to the discrete-time situation we formulate the concept of ν-optimality.

4.1.1. DEFINITION. A control $u^* \in U$ is called *ν-optimal* iff

$$w_t(H_t) = v_t(H_t,u^*) \quad \mathbb{P}_{\nu,u^*} - \text{a.s.} \quad \text{for all } t \in T.$$

Without loss of generality we assume from now on, that $t_0 = 0$.

We want to make one more remark in this section. In chapter 2 we have introduced the model of the decision process in the same way as is done in Hinderer (1970). In chapter 4 we have replaced the sets of admissible actions by more directly formulated restrictions on the set of controls. For the rest, we have build up the model along the same lines as in chapter 2. However, an alternative approach to the model is also possible, and perhaps even more transparent.

Such a set-up should start with a description of the sample space H, together with a set of probability measures $\mathbb{P}_{h_t,u}$, with u an element of a set of indices U. These measures satisfy the requirements given above. In order to have the possibility to "concatenate" measures $\mathbb{P}_{h_\tau,u'}$ for $\tau > t$ and $\mathbb{P}_{h_\tau,u''}$ for $\tau \le t$, t fixed, we assume the requirements for the set U to be satisfied. Then we introduce the function r together with its requirements. This defines the process Σ.

4.2. GENERAL UTILITY

The concept of ν-optimality will be characterized in this section. Again this is done by means of ν-conserving and ν-equalizing strategies. The derivation of this characterization is similar to that in the discrete-time case (cf. the proof of theorem 2.2.4).

4.2.1. DEFINITION. A control $u^* \in U$ is called *ν-conserving* iff for all $t_1, t_2 \in T$ with $t_2 \geq t_1$

$$w_{t_1}(H_{t_1}) = E_{\nu,u^*}^{F_{t_1}} w_{t_2}(H_{t_2}) \quad \mathbb{P}_{\nu,u^*} - \text{a.s.}$$

(It is supposed in this definition that the right-hand side is well defined). So u^* is ν-conserving iff $(w_t(H_t),\ t \in T)$ is a (continuous-time) martingale w.r.t. $\mathbb{P}_{\nu,u^*}$.

4.2.2. DEFINITION. A control $u^* \in U$ is called *ν-equalizing* iff

$$\lim_{t\to\infty} E_{\nu,u^*}[w_t(H_t) - v_t(H_t,u^*)] = 0.$$

4.2.3. THEOREM. A necessary and sufficient condition for the ν-optimality of a control $u^* \in U$ is, that u^* is ν-conserving and ν-equalizing.

PROOF. Suppose u^* is ν-optimal. Using the ν-optimality we get for each pair $t_1, t_2 \in T$ with $t_2 \geq t_1$

$$w_{t_1}(H_{t_1}) = v_{t_1}(H_{t_1},u^*) = E_{\nu,u^*}^{F_{t_1}} v_{t_2}(H_{t_2},u^*) = E_{\nu,u^*}^{F_{t_1}} w_{t_2}(H_{t_2}) \quad \mathbb{P}_{\nu,u^*} - \text{a.s.}$$

So u^* is ν-conserving. Also

$$E_{\nu,u^*}[w_t(H_t) - v_t(H_t,u^*)] = 0,$$

by the ν-optimality. Letting $t \to \infty$ we see, that u^* is ν-equalizing. So ν-conservingness and ν-equalizingness are necessary for ν-optimality.

Now suppose u^* is ν-conserving and ν-equalizing, then for all $\tau,\tau' \in T$ with $\tau' > \tau$

$$E_{\nu,u^*}\, w_\tau(H_\tau) = E_{\nu,u^*}\, w_{\tau'}(H_{\tau'}) = \lim_{t\to\infty} E_{\nu,u^*}\, w_t(H_t) = \lim_{t\to\infty} E_{\nu,u^*}\, v_t(H_t,u^*) =$$

$$= \lim_{t\to\infty} E_{\nu,u^*}\, E^{F_t}_{\nu,u^*}\, r(H) = E_{\nu,u^*}\, E^{F_\tau}_{\nu,u^*}\, r(H) = E_{\nu,u^*}\, v_\tau(H_\tau,u^*).$$

And since $v_\tau(H_\tau,u^*) \le w_\tau(H_\tau)$ $\mathbb{P}_{\nu,u^*}$ - a.s., it follows that $v_\tau(H_\tau,u^*) = w_\tau(H_\tau)$ $\mathbb{P}_{\nu,u^*}$ - a.s. for all $\tau \in T$. □

4.3. RECURSIVE UTILITY

In this section we consider the continuous-time analogues of theorems 3.2.3 and 3.2.5. The only proof we give for the characterization, is the analogue of the first proof of corollary 3.2.6. We start with a definition of recursiveness. Let $\zeta_t\colon H \to H$ be defined for each $t \in T$ as the function, that maps a history h into the tail of h beginning with (x_t,a_t).

4.3.1. DEFINITION. (i) The process Σ is called *separable* iff for all $t \in T$, all histories $h_t \in H_t$ and all controls $u \in U$ there exists a control $u^* \in U$, such that for all $h''_t \in H_t$ and all $x'_t \in X$

$$\mathbb{P}^{[h_t]}_{x'_t,u_t(h_t)} = \mathbb{P}^{[h''_t]}_{x'_t,u^*(h''_t)}$$

for any $h_t \in H_t$. The utility r is called *recursive* iff for each $t,\tau \in T$ with $\tau \ge t$ there exist functions $\theta^{[t]}_\tau$, $\chi^{[t]}_\tau$ and $r^{[t]}$, such that

$$r^{[0]} = r,$$

$$r^{[t]}(h) = \theta^{[t]}_\tau(h_\tau) + \chi^{[t]}_\tau(h_\tau)\, r^{[\tau]}(\zeta_{\tau-t}(h))$$

for all $h \in H^{[t]}$ and h_τ the restriction of h to $H^{[t]}_\tau$ with

$$\theta_\tau^{[t]} : H_\tau^{[t]} \to \mathbb{R},$$

$$\chi_\tau^{[t]} : H_\tau^{[t]} \to \mathbb{R}^+,$$

$$r^{[t]} : H^{[t]} \to \mathbb{R},$$

both $\theta_\tau^{[t]}$ and $\chi_\tau^{[t]}$ measurable and integrable, and $r^{[t]}$ measurable and quasi integrable. To ensure the uniqueness of the decomposition of $r^{[t]}$, we define $\chi_\tau^{[t]}(h_\tau) = 0$ iff $r^{[\tau]}(h') =$ constant for each $h' \in H^{[\tau]}$ with its first component x'_τ equal to the last component x_τ of $h_\tau \in H_\tau^{[t]}$.

If the process Σ is separable and if it has a recursive utility, then for each $h_t \in H_t$ the proces $\Sigma^{[h_t]}$ depends on h_t rather weakly, since the dependence is mainly on t. Note that actually the function $r^{[h_t]}$ is replaced by $r^{[t]}$. The controls are allowed to depend on h_t, but from the proof of lemma 3.1.3 it follows immediately that the set of measures $\mathbb{P}^{[h_t]}_{x'_t,u_t(h_t)}$ only depends on h_t through t. Therefore it is possible to speak a bit loosely about the <u>t-delayed process</u> $\Sigma^{[t]}$ instead of $\Sigma^{[h_t]}$. So the superscript [t] is used instead of $[h_t]$. Compare also the remarks after lemma 3.1.4.

The following lemma is the analogue of lemma 3.1.4.

4.3.2. LEMMA. If r is a recursive utility, then for all $t \in T$ and $h_t \in H_t$ we have

$$v_t^{[0]}(h_t,u) = \theta_t^{[0]}(h_t) + \chi_t^{[0]}(h_t)\ v_t^{[t]}(x_t,u_t(h_t)),$$

$$w_t^{[0]}(h_t) = \theta_t^{[0]}(h_t) + \chi_t^{[0]}(h_t)\ w_t^{[t]}(x_t).$$

PROOF. Completely analogous to the proof of lemma 3.1.4. □

The next step in the framework of this section is an analogue of lemma 3.2.2.

4.3.3. LEMMA. If r is a recursive utility, then for each $h \in H$ and $\tau, t \in T$ with $\tau \le t$

(i) $\theta_t^{[0]}(h_t) = \theta_\tau^{[0]}(h_\tau) + \chi_\tau^{[0]}(h_\tau) \cdot \theta_t^{[\tau]}(\zeta_\tau(h_t))$,

(ii) $\chi_t^{[0]}(h_t) = \chi_\tau^{[0]}(h_\tau) \cdot \chi_t^{[\tau]}(\zeta_\tau(h_t))$.

PROOF. On the one hand

$$r^{[0]}(h) = \theta_t^{[0]}(h_t) + \chi_t^{[0]}(h_t)\, r^{[t]}(\zeta_t(h)),$$

and on the other hand

$$r^{[0]}(h) = \theta_\tau^{[0]}(h_\tau) + \chi_\tau^{[0]}(h_\tau)\, r^{[\tau]}(\zeta_\tau(h)) =$$

$$= \theta_\tau^{[0]}(h_\tau) + \chi_\tau^{[0]}(h_\tau) \quad \theta_t^{[\tau]}(\zeta_\tau(h_t)) + \chi_t^{[\tau]}(\zeta_\tau(h_t)) \cdot r^{[t]}(\zeta_t(h)).$$

Hence, independent of the choice of h,

$$[\theta_t^{[0]}(h_t) - \theta_\tau^{[0]}(h_\tau) - \chi_\tau^{[0]}(h_\tau)\ \theta_t^{[\tau]}(\zeta_\tau(h_t))] +$$

$$+ [\chi_t^{[0]}(h_t) - \chi_\tau^{[0]}(h_\tau)\ \chi_t^{[\tau]}(\zeta_\tau(h_t))]\ \ r^{[t]}(\zeta_t(h)) = 0.$$

Using the same arguments as in the proof of lemma 3.2.3 we obtain the result. □

Now we come to a characterization of ν-conservingness, in the context of a recursive utility.

4.3.4. THEOREM. If r is a recursive utility, then

$$w_t^{[t]}(X_t) = E_{\nu,u}^{F_t} [\theta_\tau^{[t]}(\zeta_t(H_\tau)) + \chi_\tau^{[t]}(\zeta_t(H_\tau)) + w_\tau^{[\tau]}(X_\tau)]$$

$\mathbb{P}_{\nu,u}$ - a.s. for all $t, \tau \in T$ with $\tau \ge t$, is a necessary and sufficient condition for the control $u \in U$ to be ν-conserving.

PROOF. The proof is exactly the same as the proof of theorem 3.2.3, except for the fact that lemmas 4.3.2 and 4.3.3 are used instead of lemmas 3.1.4 and 3.2.2, respectively. □

In the continuous-time case we have also the concept of a ν-vanishing tail.

4.3.5. DEFINITION. The utility r is called ν-*tail vanishing* iff it is recursive, and for all $u \in \mathcal{U}$

$$\lim_{t\to\infty} E_{\nu,u}\, \chi_t^{[0]}(H_t)\; v_t^{[t]}(X_t, u_t(H_t)) = 0$$

(or, equivalently $\underline{\lim}_{t\to\infty} E_{\nu,u}\, \theta_t^{[0]}(H_t) = E_{\nu,u}\, v_0(H_0,u))$.

This leads to the following theorem.

4.3.6. THEOREM. If r is a ν-tail vanishing utility and $u \in \mathcal{U}$, then the following two assertions are equivalent

(i) $\lim_{t\to\infty} E_{\nu,u}\, [w_t(H_t) - v_t(H_t,u)] \begin{smallmatrix}=\\>\end{smallmatrix} 0,$

(ii) $\lim_{t\to\infty} E_{\nu,u}\, \chi_t^{[0]}\, (H_t)\; w_t^{[t]}\, (X_t) \begin{smallmatrix}=\\>\end{smallmatrix} 0.$

These formulae should both be read with equality, or both with strict inequality.

PROOF. See the proof of theorem 3.2.5. □

Hence in the C/G/G/1 process the analogues of corollary 3.2.6 (a reformulation of the characterization of ν-optimality) and corollary 3.2.7 (the property anne) hold.

So by now we have actually proved our remark in the introduction of chapter 4, that the principle of optimality as it occurs e.g. in Rishel (1970), Davis and Varaiya (1973), Boel and Varaiya (1977), is precisely our concept of conservingness. For so far as in these papers the time horizon is assumed to be finite, it follows that all controls are equalizing, so conservingness suffices for optimality. A similar example (with a finite time horizon) can be found in Bather and Chernoff (1967), whose lemma 3.1 states that conservingness implies optimality.
Let us discuss some examples with $\mathcal{T} = [0,\infty)$, i.e. with an infinite time horizon.

EXAMPLE A: (cf. The model in Boel and Varaiya (1977)). Suppose the process Σ to be separable, and let ν be such that $\mathbb{P}_{\nu,u}(X_0 = x_0) = 1$ for all $u \in U$ and a fixed $x_0 \in X$. Assume furthermore the existence of a (jointly) measurable nonpositive function $\rho\colon X \times A \to \mathbb{R}^-$, called the *instantaneous reward* or *reward density*, such that for all $u \in U$ and for $\mathbb{P}_{\nu,u}$ - a.a. $h \in H$ the utility r in the point h can be written as

$$r(h) = \int_0^\infty \rho(x_\tau, a_\tau)\, d\tau \; .$$

It is supposed, that for all $u \in U$

$$v_0(x_0,u) = E_{x_0,u} \int_0^\infty \rho(X_\tau, A_\tau)\, d\tau > -\infty.$$

This expression is well defined by the measurability and quasi integrability of r.

It is easy to see, that r is recursive and that for all $t, \tau \in T$ with $\tau \geq t$,

$$\theta_\tau^{[t]}(h_\tau) = \int_t^\tau \rho(x_\sigma, a_\sigma)\, d\sigma, \quad r^{[t]}(h) = \int_t^\infty \rho(x_\sigma, a_\sigma)\, d\sigma \quad \text{and} \quad \chi_\tau^{[t]} = 1.$$

Since by the monotonicity of the functions $\theta_t^{[0]}$, $t \in T$, in a fixed point $h \in H$,

$$\lim_{t\to\infty} E_{\nu,u}\, \theta_t^{[0]}(H_t) = E_{\nu,u} \int_0^\infty \rho(X_\tau, A_\tau)\, d\tau = E_{\nu,u}\, v_0(X_0,u),$$

it follows that r is also ν-tail vanishing.

Using the nonpositivity of ρ and the fact that $\underline{\lim}_{t\to\infty} E_{\nu,u}\, w_t^{[t]}(X_t) \geq 0$ by theorem 4.3.6 we conclude that

$$\lim_{t\to\infty} E_{\nu,u}\, w_t^{[t]}(X_t) = 0 \text{ for every } u \in \mathcal{U}.$$

So each $u \in \mathcal{U}$ is ν-equalizing. Hence $u^* \in \mathcal{U}$ is ν-optimal iff for all $t_1, t_2 \in \mathsf{T}$ with $t_1 \leq t_2$

$$\text{(i)} \quad w_{t_1}^{[t_1]}(X_{t_1}) = E_{\nu,u^*}^{F_{t_1}} \int_{t_1}^{t_2} \rho(X_\tau, A_\tau)\,d\tau + E_{\nu,u}^{F_{t_1}}\, w_{t_2}^{[t_2]}(X_{t_2}) \quad \mathbb{P}_{\nu,u^*}\text{ - a.s.}$$

This result is the analogue of the second assertion of theorem 4.1. in Boel and Varaiya (1977). It is also the analogue of the essential negative case, discussed in example 3.3.D.
Actually, we could have made ρ dependent on $t \in \mathsf{T}$. This is worked out in the next example.

EXAMPLE B: In fact, this example covers the situation described in Doshi (1976). Again we assume Σ to be separable. Let $\alpha > 0$ be a so-called discount rate and let ν be a fixed starting distribution. Suppose we have an instantaneous reward ρ: $\mathsf{T} \times \mathsf{X} \times \mathsf{A} \to \mathbb{R}$, which is a jointly measurable function satisfying $|\rho| \leq M < \infty$, such that for $t \in \mathsf{T}$, $u \in \mathcal{U}$ and for $\mathbb{P}_{\nu,u}$-a.a. $h \in \mathsf{H}$ the utility r in the point h can be written as

$$r(h) = \int_0^\infty e^{-\alpha\tau}\, \rho(\tau, x_\tau, a_\tau)\,d\tau =$$

$$= \int_0^t e^{-\alpha\tau}\, \rho(\tau, x_\tau, a_\tau)\,d\tau + e^{-\alpha t} \int_t^\infty e^{-\alpha(\tau-t)}\, \rho(\tau, x_\tau, a_\tau)\,d\tau.$$

So r is recursive with

$$\theta_{t_2}^{[t_1]}(\zeta_{t_1}(h)) = \int_{t_1}^{t_2} e^{-\alpha(\tau-t_1)}\,\rho(\tau, x_\tau, a_\tau)\,d\tau$$

and $\chi_{t_2}^{[t_1]}(\zeta_{t_1}(h)) = e^{-\alpha(t_2-t_1)}$.

Since $|\rho| \le M$, the functions $r^{[t]}$ are uniformly bounded: $|r^{[t]}| \le \frac{M}{\alpha}$. From $|v_t^{[t]}(X_t,u)| \le \frac{M}{\alpha}$ $\mathbb{P}_{\nu,u}$-a.s. for all $u \in \mathsf{U}$, it follows that

$$\lim_{t\to\infty} |E_{\nu,u}\ \chi_t^{[0]}(H_t)\ v_0^{[t]}(X_t,u_t(H_t))| \le \lim_{t\to\infty} e^{-\alpha t}\frac{M}{\alpha} = 0,$$

i.e. r is ν-tail vanishing. Since also $|w_t^{[t]}(X_t)| \le \frac{M}{\alpha}$ $\mathbb{P}_{\nu,u}$ - a.s., it follows from theorem 4.3.6, that every $u \in \mathsf{U}$ is ν-equalizing. Hence $u^* \in \mathsf{U}$ is ν-optimal iff

$$\text{(i)}\quad w_{t_1}^{[t_1]}(X_{t_1}) = E_{\nu,u^*}^{F_{t_1}} \int_{t_1}^{t_2} e^{-\alpha(\tau-t_1)}\ \rho(\tau,X_\tau,A_\tau)d\tau + e^{-\alpha(t_2-t_1)}\ w_{t_2}^{[t_2]}(X_{t_2})$$

$$\mathbb{P}_{\nu,u^*} - \text{a.s.}$$

Also we know, that u^* is ν-optimal iff for all $t \in \mathsf{T}$

$$\text{(ii)}\quad w_t^{[t]}(X_t) = v_t^{[t]}(X_t,u_t^*(H_t)) \quad \mathbb{P}_{\nu,u^*} - \text{a.s.}$$

(One part of the assertion follows from lemma 4.2.2 and lemma 4.3.2 with the subsequent remark, the other part is trivial).
Suppose the decision process satisfies conditions, that enable us the use of theorem 1.7 in Dynkin (1965) (see Doshi (1976) for a precise formulation of all conditions required). One of these conditions is, that u is a Markovian control, i.e. that u_t does not depend on the history before time t. Dynkin's result says, that for each $t \in \mathsf{T}$ the function $v_t^{[t]}$ is the unique solution of

$$\text{(iii)}\quad \alpha\ v(x_t,u_t) = \rho(t,x_t,u(t,x_t)) + (A_{u(t,\cdot)}\ v(\cdot,u_t))\ (x_t).$$

Here $A_{u(t,\cdot)}$ is the infinitesimal operator of the process at time t, determined by u. Now suppose u* is a Markovian ν-optimal control. Then from (ii) and (iii) it follows, that for all $t \in T$

$$\text{(iv)} \quad \alpha\, w_t^{[t]}(X_t) = \rho(t,X_t,u^*(t,X_t)) = (A_{u^*(t,\cdot)}\, w_t^{[t]})\ (X_t) \quad \mathbb{P}_{\nu,u^*} - \text{a.s.}$$

On the other hand, supposing (iv) to hold for a Markovian control u*, it follows from (iii) that (ii) is fulfilled, so u* is ν-optimal. Hence (iv) is a necessary and sufficient condition for ν-optimality of a Markovian control u*. Note that (iv) is the infinitesimal analogue of (i).

EXAMPLE C: Finally, we want to make some remarks about the average reward criterion. Let Σ be separable and let ν be a fixed starting distribution. Let ρ be a jointly measurable function $\rho\colon X \times A \to \mathbb{R}$, such that for all $u \in U$ and for $\mathbb{P}_{\nu,u}$ - a.a. $h \in H$ the utility r in the point h can be written as

$$r(h) = \lim_{t\to\infty} \frac{1}{t} \int_0^t \rho(x_\tau, u_\tau)\, d\tau .$$

As in the discrete-time case, r is recursive with $\theta_\tau^{[t]} = 0$ and $\chi_\tau^{[t]} = 1$ for $t,\tau \in T, \tau \geq t$, but r is not ν-tail vanishing. Then the conservingness means, that the process does not move towards a less favourable recurrent subset of the state space. Equalizingness means, that the player actually receives the gain corresponding to the most favourable recurrent subset of the state space he can reach. This formulation seems to be rather helpful in interpreting e.g. the optimality condition in formula 16 of de Leve, Federgruen and Tijms (1977).

CHAPTER 5

THE D (AND C)/G/G/2 PROCESS WITH A ZERO-SUM UTILITY

The main difference between the process in this chapter and that studied in the previous three chapters, is that here the decision process is controlled by two players. Furthermore, these players have opposite aims, or formulated alternatively, the process has a zero-sum utility.
It is well known that methods used to compute the optimal value, and to determine the optimal strategy in two-person zero-sum games, have much in common with the methods used in one-player games. (See e.g. Shapley (1953) and van der Wal and Wessels (1976). The survey of Parthasarathy and Stern (1976) contains many other interesting references.) Here we show that the characterization of optimal strategies in the two-player game is also strongly akin to the characterization in the one-player game. On the other hand, this characterization in connection with alternative notions of conservingness, that are not suitable for a characterization, gives more insight in the optimality concept itself. Moreover, some of the alternative notions of conservingness lead to characterizations of other, stronger optimality concepts such as subgame perfectness, persistent optimality and tail optimality.
In section 1 we describe the model of the more general D/G/G/n process (recall that n is the number of players). In section 2 and 3 we restrict ourselves to the case n = 2 with a zero-sum utility. Several types of optimality are introduced and characterized in section 2, and a number of counterexamples clarify the differences between these concepts. In section 3 these results are transferred to the situation with a recursive utility. Furthermore an analogue of the optimality principle (corollary 3.1.5) can be found in this section. In section 4 the continuous-time case is treated.

5.1. THE D/G/G/n PROCESS

The general D/G/G/n process, where n is the cardinal number of the set of players, looks very much like the D/G/G/1 process. We denote the set of players by $\mathbb{N}_n$. Note that for a finite n the set $\mathbb{N}_n = \{\ell \in \mathbb{N} \mid 0 \le \ell < n\}$.
The D/G/G/n process is defined by the tuple

$$(T,(X,\mathcal{X}),((A^{(\ell)},\mathcal{A}^{(\ell)}) \mid \ell \in \mathbb{N}_n) , (L_t^{(\ell)} \mid \ell \in \mathbb{N}_n, t \in T),$$
$$(p_t \mid t \in T), (r^{(\ell)} \mid \ell \in \mathbb{N}_n))$$

together with a set of requirements.
Here again $T = \{0,1,\dots\}$ is the time space, $(X,\mathcal{X})$ the measurable state space and $(p_t \mid t \in T)$ the family of transition functions. But now there are n measurable action spaces $(A^{(\ell)}, \mathcal{A}^{(\ell)})$, n families $(L_t^{(\ell)}, t \in T)$ and n utility functions $r^{(\ell)}$, $\ell \in \mathbb{N}_n$. $A^{(\ell)}$ is the action space for player ℓ and $r^{(\ell)}$ is his utility. Each $L_t^{(\ell)}$ is a subset of $\times_{k=0}^{t}(X \times A^{(\ell)})$, and if $(x_0,a_0^{(\ell)},\dots,x_t,a_t^{(\ell)}) \in L_t^{(\ell)}$, then $a_t^{(\ell)}$ is called an admissible action in $(x_0,a_0^{(\ell)},\dots,x_t)$ for player ℓ. In a similar way as before we suppose that $L_t^{(\ell)} \in \otimes_{k=0}^{t}(\mathcal{X} \otimes \mathcal{A}^{(\ell)})$ and that the h_t-section $L_{t\,h_t}^{(\ell)}$ of $L_t^{(\ell)}$ is nonempty for all $h_t = (x_0,a_0^{(\ell)},\dots,x_t)$, $\ell \in \mathbb{N}_n$.

To describe the behaviour of the process properly, we introduce $(A,\mathcal{A}) := (A^{(0)} \times \dots \times A^{(n-1)},\mathcal{A}^{(0)} \otimes \dots \otimes \mathcal{A}^{(n-1)})$. The sample space $(H,\mathcal{H})$ is defined as $(X \times A \times X \times A \times\dots, \mathcal{X} \otimes \mathcal{A} \otimes \dots)$ and $(H_t,\mathcal{H}_t)$, the space of histories up to time t, is defined as $(X \times A \times X \times A \times\dots \times X, \mathcal{X} \otimes \mathcal{A} \otimes\dots\otimes \mathcal{X})$ with t+1 factors X and $\mathcal{X}$ and t factors A and $\mathcal{A}$. Each p_t from the set of transition functions is a transition probability from $(X \times A \times\dots\times X \times A, \mathcal{X} \otimes \mathcal{A} \otimes\dots\otimes \mathcal{X} \otimes \mathcal{A})$, with t+1 factors $X,\mathcal{X}$, A and $\mathcal{A}$, into $(X,\mathcal{X})$. The second part of the transition mechanism of the process is prescribed by a (*simultaneous*) *strategy* $\pi = (\pi^{(0)}, \pi^{(1)},\dots,\pi^{(n-1)})$. Here each $\pi^{(\ell)}$ is a *strategy for player* ℓ, and it can be written as $\pi^{(\ell)} = (\pi_0^{(\ell)}, \pi_1^{(\ell)},\dots)$ where for each $t \in T$ the function $\pi_t^{(\ell)}$ is a transition probablity from $(H_t,\mathcal{H}_t)$ into $(A,\mathcal{A})$, with the extra condition that for all $h_t = (x_0,a_0,\dots,x_t) \in H_t$ the probability measure $\pi_t^{(\ell)}(h_t,\cdot)$ is concentrated on the set of admissible actions for player ℓ in $h_t \in H_t$. We also write π as $(\pi_0,\pi_1,\dots)$ with $\pi_t = (\pi_t^{(0)}, \pi_t^{(1)},\dots,\pi_t^{(n-1)})$. It should be noted, that π_t selects only cylinder sets in $\mathcal{A}$. The set of all simultaneous strategies π is denoted by Π, and the set of all strategies $\pi^{(\ell)}$ for player ℓ is denoted by $\Pi^{(\ell)}$.

As in chapter 2 the Ionescu Tulcea theorem provides the construction of a suitable probability measure $\mathbb{P}_{x_0,\pi}$ on the sample space, for each $x_0 \in X$ and $\pi \in \Pi$. This $\mathbb{P}_{x_0,\pi}$ is the uniquely determined probability measure for the process which starts in x_0, with a transition mechanism prescribed by π and $(p_t \mid t \in T)$. As in chapter 2, we have a unique probability measure $\mathbb{P}_{h_t,\pi}$ for each $h_t \in H_t$ and $\pi \in \Pi$, and a probability measure $\mathbb{P}_{\nu,\pi}$ for each starting distribution ν on $(X,\mathcal{X})$. From now on ν is assumed to be fixed.

Defining L_t as $\{(h_t,a) \mid h_t \in H_t,\ a = (a^{(0)},a^{(1)},\ldots,a^{(n-1)}) \in A$ with $(h_t,a^{(\ell)}) \in L_t^{(\ell)}\}$, and defining $\hat{H}$ as $\bigcap_{k=0}^{\infty} (L_k \times X \times A \times X \times A \times \ldots)$. we may apply theorem 2.1.1 to conclude that $\mathbb{P}_{x_0,\pi}(\hat{H}) = 1$ for every $x_0 \in X$ and $\pi \in \Pi$.

We also assume, that for each $\ell \in \mathbb{N}_n$ the function $r^{(\ell)} : H \to \mathbb{R}$ is Borel measurable and quasi integrable with respect to each $\mathbb{P}_{\nu,\pi}$. The symbols $H, H_t, X_t, A_t, F_t, \mathbb{E}_{x_0,\pi}, \mathbb{E}_{\nu,\pi}, \mathbb{E}_{h_t,\pi}$ have a similar meaning as before.

The *value of strategy* π *for player* ℓ, *given* h_t is defined as a function $v_t^{(\ell)} : H_t \times \Pi \to \mathbb{R}$ with

$$v_t^{(\ell)}(h_t,\pi) = \begin{cases} \mathbb{E}_{h_t,\pi}\, r^{(\ell)}(H) & \text{if this integral exists,} \\ -\infty & \text{otherwise.} \end{cases}$$

We will use the notation $(\pi^*;\ \ell: \pi^{(\ell)})$ for a strategy that is obtained from $\pi^* \in \Pi$ by replacing the component $\pi^{*(\ell)}$ by $\pi^{(\ell)} \in \Pi^{(\ell)}$. The *value for player* ℓ, *given* h_t *and given* π^* *for the other players* is a function $\psi_t^{(\ell)} : H_t \times \Pi \to \mathbb{R}$ satisfying

$$\psi_t^{(\ell)}(h_t,\pi^*) = \sup_{\pi^{(\ell)} \in \Pi^{(\ell)}} \mathbb{E}_{h_t,(\pi^*;\ell:\pi^{(\ell)})}\, r^{(\ell)}(H).$$

5.1.1. DEFINITION. A simultaneous strategy $\pi^* \in \Pi$ is called *ν-optimal* (or a *ν-equilibrium strategy*) iff for all $\ell \in \mathbb{N}_n$ and all $t \in T$

$$\psi_t^{(\ell)}(H_t,\pi^*) = v_t^{(\ell)}(H_t,\pi^*) \qquad \mathbb{P}_{\nu,\pi^*} - \text{a.s.}$$

We emphasize that the above type of optimality is precisely the well known Nash optimality for all time instances $t \in T$, with respect to a fixed starting distribution ν.

5.2. THE D/G/G/2 PROCESS WITH A GENERAL ZERO-SUM UTILITY

In this section and the next one we will study the D/G/G/n process with $n = 2$ and $r^{(0)} = -r^{(1)}$. This process is called a *two-person* (or *two-player*) *game*, and because of the condition $r^{(0)} + r^{(1)} = 0$ the process is said to have a *zero-sum utility*. The latter condition implies that the two players have opposite aims, since a gain for the one player is a loss for the other. For this special situation we will change our notation a little: π denotes only a strategy for player 0, and ρ denotes a strategy for player 1; a simultaneous strategy is denoted by (π,ρ), and never by π as in the general D/G/G/n case; $r := r^{(0)}$; $v_t := v_t^{(0)}$ for all $t \in T$, and instead of $v_t(h_t,(\pi,\rho))$ we write $v_t(h_t,\pi,\rho)$; for all $t \in T$ we define $\varphi_t\colon H_t \times \Pi^{(0)} \to \mathbb{R}$ by $\varphi_t(h_t,\pi) = \psi_t^{(1)}(h_t,(\pi,\rho))$ and $\psi_t\colon H_t \times \Pi^{(1)} \to \mathbb{R}$ by $\psi_t(h_t,\rho) = \psi_t^{(0)}(h_t,(\pi,\rho))$. This means that for all $t \in T$

$$\varphi_t(h_t,\pi^*) = \inf_{\rho\in\Pi^{(1)}} E_{h_t,(\pi^*,\rho)}\, r(H).$$

Apparently, ν-optimality of a strategy $(\pi^*,\rho^*) \in \Pi$ can be reformulated as

$$v_t(H_t,\pi,\rho) \le v_t(H_t,\pi^*,\rho^*) \le v_t(H_t,\pi^*,\rho) \quad \mathbb{P}_{\nu,(\pi^*,\rho^*)} - \text{a.s.}$$

for all $\pi \in \Pi^{(0)}$ and $\rho \in \Pi^{(1)}$ and $t \in T$. Now we may apply a well known standard reasoning (cf. the proof of lemma 5.2.1) to see, that ν-optimality of $(\pi^*,\rho^*) \in \Pi$ implies that for $\mathbb{P}_{\nu,(\pi^*,\rho^*)}$ - almost all $h_t \in H_t$.

$$\sup_\pi \inf_\rho v_t(h_t,\pi,\rho) = \inf_\rho \sup_\pi v_t(h_t,\pi,\rho) = v_t(h_t,\pi^*,\rho^*).$$

The reasoning is the following: for $\mathbb{P}_{\nu,(\pi^*,\rho^*)}$ - a.a. $h_t \in H_t$

5.2.0.1. $v_t(h_t,\pi^*,\rho^*) = \inf_{\rho} v_t(h_t,\pi^*,\rho) \le \sup_{\pi} \inf_{\rho} v_t(h_t,\pi,\rho) \le$

$$\le \sup_{\pi} \inf_{\rho} \sup_{\pi} v_t(h_t,\pi,\rho) = \inf_{\rho} \sup_{\pi} v_t(h_t,\pi,\rho) \le \sup_{\pi} v_t(h_t,\pi,\rho^*) =$$

$$= v_t(h_t,\pi^*,\rho^*).$$

An important role in the sequel is played by the functions $w_t : H_t \to \mathbb{R}$, $t \in T$, defined by

$$w_t(h_t) = \sup_{\pi} \inf_{\rho} v_t(h_t,\pi,\rho).$$

We call w_t the *saddle given h_t*. This name may seen a bit misleading, since w_t is usually called a saddle function iff

$$w_t(h_t) = \sup_{\pi} \inf_{\rho} v_t(h_t,\pi,\rho) = \inf_{\rho} \sup_{\pi} v_t(h_t,\pi,\rho).$$

In general such a saddle function does not exist. Note however that our saddle is really an extension of the usual concept, and moreover it is always well defined.

The following results will be derived in this section. A characterization of ν-optimality is given in terms of ν-conservingness and ν-equalizingness. These concepts are defined by use of the functions φ_t and ψ_t. We also introduce alternative notions of conservingness and equalizingness formulated in terms of the functions w_t. These alternative concepts reveal some weak aspects of the ν-optimality, but they are not very useful for a characterization of ν-optimality. However, at the same time these new concepts lead to stronger optimality concepts, which have those weak aspects to a lesser degree. The weakest of these optimality concepts is persistent optimality, introduced in Groenewegen (1976), and the strongest is subgame perfectness, introduced in Selten (1965) as perfectness, and reintroduced in Selten (1975) as subgame perfectness. Another concept called tail optimality, is stronger than persistent optimality and weaker than subgame perfectness. All these relatively strong types of optimality can be characterized by means of different types of conservingness and equalizingness, all formulated in terms of the saddles w_t. These characterizations will be derived in this section, with the exception of persistent optimality, that will be treated in the next section.

5.2.1. LEMMA. If $(\pi^*,\rho^*) \in \Pi$ is ν-optimal, then for all $t \in T$

$$v_t(H_t,\pi^*,\rho^*) = \varphi_t(H_t,\pi^*) = \psi_t(H_t,\rho^*) = w_t(H_t) \quad \mathbb{P}_{\nu,(\pi^*\rho^*)} \text{ - a.s.}$$

PROOF. The result follows directly from formula 5.2.0.1. □

In order to formulate theorem 5.2.3 we need an extra assumption.

5.2.2. ASSUMPTION. Let ν be a fixed starting distribution and let $\rho^* \in \Pi^{(1)}$ and $\pi^* \in \Pi^{(0)}$ be arbitrarily chosen. It is assumed, that for all $\pi \in \Pi^{(0)}$ the functions $\psi_t(\cdot,\rho^*)$ are $\mathbb{P}_{\nu,(\pi,\rho^*)}$ - almost equal to a measurable function, and that for all $\rho \in \Pi^{(1)}$ the functions $\varphi_t(\cdot,\pi^*)$ are $\mathbb{P}_{\nu,(\pi^*,\rho)}$ - almost equal to a measurable function. Moreover it is assumed, that for all probability measures μ on H_t, $t \in T$, all $\varepsilon > 0$ and all $m \in \mathbb{R}$ **firstly** there exist strategies $\pi',\pi'' \in \Pi^{(0)}$ such that

$$v_t(H_t,\pi',\rho^*) > \psi_t(H_t,\rho^*) - \varepsilon \quad \mu \text{ - a.s. on } \{h_t \in H_t \mid \psi_t(h_t,\rho^*) < \infty\},$$

$$v_t(H_t,\pi'',\rho^*) > m \qquad \mu \text{ - a.s. on } \{h_t \in H_t \mid \psi_t(h_t,\rho^*) = \infty\},$$

and secondly there exist strategies $\rho',\rho'' \in \Pi^{(1)}$ such that

$$v_t(H_t,\pi^*,\rho') < \varphi_t(H_t,\pi^*) + \varepsilon \quad \mu \text{ - a.s. on } \{h_t \in H_t \mid \varphi_t(h_t,\pi^*) > -\infty\},$$

$$v_t(H_t,\pi^*,\rho'') < m \qquad \mu \text{ - a.s. on } \{h_t \in H_t \mid \varphi_t(h_t,\pi^*) = -\infty\}.$$

Now we give the analogue of theorem 2.2.1. We need this analogue for the derivation of the characterization of tail optimality and persistent optimality, at least in this chapter. In chapter 6, however, we derive a somewhat different characterization of both these optimality concepts, without using theorem 5.2.3.

5.2.3. THEOREM. Let assumption 5.2.2 be satisfied. Then it holds that

$$\psi_t(H_t,\rho^*) \geq E^{F_t}_{\nu,(\pi,\rho^*)}\ \psi_{t+1}(H_{t+1},\rho^*) \quad \mathbb{P}_{\nu,(\pi,\rho^*)} \text{ - a.s. for all } \pi \in \Pi^{(0)},$$

$$\varphi_t(H_t,\pi^*) \leq E^{F_t}_{\nu,(\pi^*,\rho)}\ \varphi_{t+1}(H_{t+1},\pi^*) \quad \mathbb{P}_{\nu,(\pi^*,\rho)} \text{ - a.s. for all } \rho \in \Pi^{(1)},$$

i.e. the functions ψ_t and φ_t form a supermartingale and a submartingale respectively.

PROOF. Fixing $\rho^* \in \Pi^{(1)}$, we are in the situation of a D/G/G/1 process, with ψ_t as value function. Hence theorem 2.2.1 applies. On the other hand fixing $\pi^* \in \Pi^{(0)}$, we may apply the proof of theorem 2.2.1, but now with minimizing instead of maximizing, and with "greater than" instead of "less than". □

5.2.4. DEFINITION. A strategy $(\pi^*,\rho^*) \in \Pi$ is called ν-*conserving* iff for all $t \in T$

(i) $\varphi_t(H_t,\pi^*) = E^{F_t}_{\nu,(\pi^*,\rho^*)}\ \varphi_{t+1}(H_{t+1},\pi^*) \quad \mathbb{P}_{\nu,(\pi^*,\rho^*)}$ - a.s.

(ii) $\psi_t(H_t,\rho^*) = E^{F_t}_{\nu,(\pi^*,\rho^*)}\ \psi_{t+1}(H_{t+1},\rho^*) \quad \mathbb{P}_{\nu,(\pi^*,\rho^*)}$ - a.s.

i.e. both $\{\varphi_t(H_t,\pi^*) \mid t \in T\}$ and $\{\psi_t(H_t,\rho^*) \mid t \in T\}$ are martingales w.r.t. $\mathbb{P}_{\nu,(\pi^*,\rho^*)}$.

5.2.5. DEFINITION. A strategy $(\pi^*,\rho^*) \in \Pi$ is called ν-*equalizing* iff

(i) $\lim_{t\to\infty} E_{\nu,(\pi^*,\rho^*)} [\varphi_t(H_t,\pi^*) - v_t(H_t,\pi^*,\rho^*)] = 0$

(ii) $\lim_{t\to\infty} E_{\nu,(\pi^*,\rho^*)} [\psi_t(H_t,\rho^*) - v_t(H_t,\pi^*,\rho^*)] = 0$

5.2.6. THEOREM. A necessary and sufficient condition for ν-optimality of $(\pi^*,\rho^*) \in \Pi$ is that (π^*,ρ^*) is both ν-conserving and ν-equalizing.

PROOF. Suppose (π^*,ρ^*) is ν-optimal. Then π^* is ν-optimal in the D/G/G/1 process that arises from fixing ρ^*. Hence formulae 5.2.4 (ii) and 5.2.5 (ii) hold. Analogously 5.2.4 (i) and 5.2.5 (i) hold, so (π^*,ρ^*) is ν-conserving and ν-equalizing.
Suppose (π^*,ρ^*) is ν-conserving and ν-equalizing. Then by 5.2.4 (ii) and 5.2.5 (ii), the strategy π^* is ν-conserving and ν-equalizing in the D/G/G/1 process that arises from fixing ρ^*. Hence for all $\pi \in \Pi^{(0)}$ and $t \in T$

$$v_t(H_t,\pi,\rho^*) \le v_t(H_t,\pi^*,\rho^*) \qquad \mathbb{P}_{\nu,(\pi^*,\rho^*)} \text{ - a.s.}$$

And analogously for all $\rho \in \Pi^{(1)}$

$$v_t(H_t,\pi^*,\rho^*) \le v_t(H_t,\pi^*,\rho) \qquad \mathbb{P}_{\nu,(\pi^*,\rho^*)} \text{ - a.s.}$$

These two formulae together establish the ν-optimality of (π^*,ρ^*). □

Now we come to an alternative type of conservingness.

5.2.7. DEFINITION. A strategy $(\pi^*,\rho^*) \in \Pi$ is called ν-*saddle conserving* iff for all $t \in T$ and $(\pi,\rho) \in \Pi$

$$E^{F_t}_{\nu,(\pi,\rho^*)}\, w_{t+1}(H_{t+1}) \le w_t(H_t) \le E^{F_t}_{\nu,(\pi^*,\rho)}\, w_{t+1}(H_{t+1}) \quad \mathbb{P}_{\nu,(\pi^*,\rho^*)} \text{ - a.s.}$$

The next theorem is an analogue of lemma 7.1 in Davis and Varaiya (1973).

5.2.8. THEOREM. If $(\pi^*,\rho^*) \in \Pi$ is ν-optimal, then (π^*,ρ^*) is ν-saddle conserving, provided that assumption 5.2.2 is satisfied.

PROOF. Lemma 5.2.1 gives

$$w_t(H_t) = v_t(H_t,\pi^*,\rho^*) = \varphi_t(H_t,\pi^*) \quad \mathbb{P}_{\nu,(\pi^*,\rho^*)} \text{ - a.s.}$$

Since $\varphi_t(H_t,\pi^*)$ is a submartingale with respect to any $\mathbb{P}_{\nu,(\pi^*,\rho)}$, we have

$$w_t(H_t) \le E^{F_t}_{\nu,(\pi^*,\rho)}\ \varphi_{t+1}(H_{t+1},\pi^*) \le E^{F_t}_{\nu,(\pi^*,\rho)}\ w_{t+1}(H_{t+1}) \quad \mathbb{P}_{\nu,(\pi^*,\rho^*)} - \text{a.s.}$$

Analogously we have

$$w_t(H_t) \ge E^{F_t}_{\nu,(\pi,\rho^*)}\ \psi_{t+1}(H_{t+1},\rho^*) \ge E^{F_t}_{\nu,(\pi,\rho^*)}\ \inf_\rho \sup_\pi v_{t+1}(H_{t+1},\pi,\rho) \ge$$

$$\ge E^{F_t}_{\nu,(\pi,\rho^*)}\ \sup_\pi \inf_\rho v_{t+1}(H_{t+1},\pi,\rho) = E^{F_t}_{\nu,(\pi,\rho^*)}\ w_{t+1}(H_{t+1}) \quad \mathbb{P}_{\nu,(\pi,\rho^*)} - \text{a.s.}$$

□

However, ν-saddle conservingness together with ν-equalizingness is not sufficient for ν-optimality (see also Groenewegen and Wessels (1977)).

5.2.9. THEOREM. COUNTEREXAMPLE. Even if all strategies are ν-equalizing, then the ν-saddle conservingness of $(\pi^*,\rho^*) \in \Pi$ does not imply its ν-optimality.

PROOF. We introduce the following D/F/F/2 process (see theorem 5.2 in Groenewegen (1976)).

$$\Gamma_0 = \begin{array}{c|c|} & 1 \\ \hline 1 & 0+\Gamma_0 \\ \hline \end{array} \qquad \Gamma_1 = \begin{array}{c|c|} & 1 \\ \hline 1 & 0+\Gamma_0 \\ \hline \end{array} \qquad \Gamma_2 = \begin{array}{c|c|} & 1 \\ \hline 1 & 0+\Gamma_1 \\ \hline 2 & -10+\Gamma_0 \\ \hline \end{array} \qquad \Gamma_3 = \begin{array}{c|c|c|} & 1 & 2 \\ \hline 1 & 0+\Gamma_1 & 1+\Gamma_2 \\ \hline \end{array}$$

Here the block notation should be read as follows.

$$
\begin{array}{c|c|}
 & a^{(1)} \\
\hline
a^{(0)} & \begin{array}{c} \theta(i,a^{(0)},a^{(1)}) \\ + \\ \sum_{j\in X} p(i,a^{(0)},a^{(1)},j)\ \Gamma_j \end{array} \\
\hline
 & \Gamma_i
\end{array}
$$

In state i of a countable state space X a 'game' Γ_i is played, i.e. if in state i player 0 chooses action $a^{(0)}$ and player 1 chooses action $a^{(1)}$, then an immediate reward $\theta(i,a^{(0)},a^{(1)})$ is earned by player 0, and the system moves with probability $p(i,a^{(0)},a^{(1)},j)$ to state j.
The utility function is defined as $r(h) = \sum_{t=0}^{\infty} \theta(x_t,a_t^{(0)}, a_t^{(1)})$,
so r is the usual additive utility. The Markov strategies we study, will be given only for the states 2 and 3 in this order, e.g.

$$(\pi^*,\rho^*) := \begin{pmatrix} 1 & 2 & 2 & \ldots & , & 1 & 1 & \ldots \\ 1 & 1 & 1 & \ldots & , & 1 & 1 & \ldots \end{pmatrix}$$

Here the top row prescribes the actions to be chosen in state 2, and the bottom row prescribes the actions in state 3. Hence (π^*,ρ^*) prescribes that at each time in state 3 action 1 is chosen both by player 0 and player 1, and that in state 2 both players choose action 1 at time 0, but at all the other times player 0 chooses action 2 and player 1 chooses action 1. It can be verified very easily that (π^*,ρ^*) is ν-saddle conserving (with ν such that $\mathbb{P}_{\nu,(\pi^*,\rho^*)}(X_0 = 3) = 1$), since $w_t(H_t) = 0$ $\mathbb{P}_{\nu,(\pi^*,\rho^*)}$ - a.s.
For all $(\pi,\rho) \in \Pi$ the equality $w_t(h_t) = v_t(h_t,\pi,\rho)$ holds for $t \geq 2$, since for $t \geq 2$ the system is either in state 1 or in state 0 with probability 1. Hence all strategies are ν-equalizing.
Define

$$\rho' := \begin{pmatrix} 1 & 1 & \ldots \\ 2 & 2 & \ldots \end{pmatrix}$$

Then it is easy to see that

$$v_0(3,\pi^*,\rho') = -9 < v_0\ (3,\pi^*,\rho^*) = 0,$$

hence (π^*,ρ^*) is not ν-optimal. □

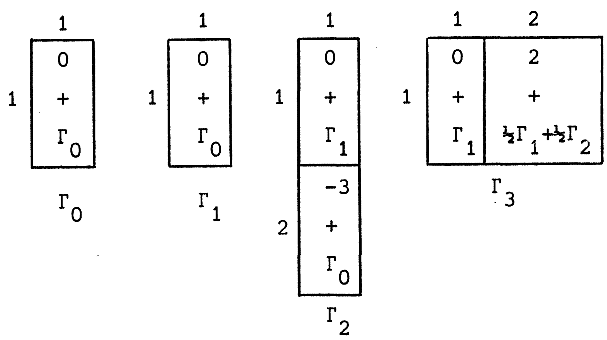
However, the following theorem states, that ν-optimality itself is not sufficient for a type of conservingness, that is somewhat stronger that ν-saddle conservingness.

5.2.10. THEOREM. COUNTEREXAMPLE. If a strategy $(\pi^*,\rho^*) \in \Pi$ is ν-optimal, then it is not necessarily true that for all $t \in T$ and $(\pi,\rho) \in \Pi$

$$w_t(H_t) \le E^{F_t}_{\nu,(\pi^*,\rho)}\ w_{t+1}(H_{t+1}) \qquad \mathbb{P}_{\nu,(\pi^*,\rho)} \text{ - a.s.}$$

and $$w_t(H_t) \ge E^{F_t}_{\nu,(\pi,\rho^*)}\ w_{t+1}(H_{t+1}) \qquad \mathbb{P}_{\nu,(\pi,\rho^*)} \text{ - a.s.}$$

PROOF. We introduce the following D/F/F/2 process, which is a variant of the counterexample 5.2.9.



Again $r(h) = \sum_{t=0}^{\infty} \theta(x_t, a_t^{(0)}, a_t^{(1)})$,

$$\mathbb{P}_{\nu,(\pi,\rho)}(X_0 = 3) = 1 \text{ for all } (\pi,\rho) \in \Pi,$$

$$(\pi^*,\rho^*) := \begin{pmatrix} 1\ 2\ 2\ \dots\ ,\ 1\ 1\ \dots \\ 1\ 1\ \dots\ ,\ 1\ 1\ \dots \end{pmatrix}$$

$$\rho' := \begin{pmatrix} 1\ 1\ \dots \\ 2\ 2\ \dots \end{pmatrix} .$$

(The strategies are only given for states 2 and 3 in this order). Then it is not difficult to see that (π^*,ρ^*) is indeed ν-optimal, but

$$w_1(H_1) \not\leq E^{F_1}_{\nu,(\pi^*,\rho')}\ w_2(H_2) \qquad \mathbb{P}_{\nu,(\pi^*,\rho')} \text{ - a.s.,}$$

since for $h_1 = (3,1,2,2)$ we have $\mathbb{P}_{\nu,(\pi^*,\rho')}(H_1 = h_1) = \frac{1}{2}$,

and $w_1(h_1) = 2 \not\leq E_{h_1,(\pi^*,\rho')}\ w_2(H_2) = 2 - 3 = -1$. □

On the other hand, the conserving property as formulated in the last theorem, together with equalizingness is not sufficient for optimality.

5.2.11. THEOREM. COUNTEREXAMPLE. If a strategy $(\pi^*,\rho^*) \in \Pi$ is ν-equalizing and if for all $t \in T$ and $(\pi,\rho) \in \Pi$

$$w_t(H_t) \leq E^{F_t}_{\nu,(\pi^*,\rho)}\ w_{t+1}(H_{t+1}) \qquad \mathbb{P}_{\nu,(\pi^*,\rho)} \text{ - a.s.}$$

and $$w_t(H_t) \geq E^{F_t}_{\nu,(\pi,\rho^*)}\ w_{t+1}(H_{t+1}) \qquad \mathbb{P}_{\nu,(\pi,\rho^*)} \text{ - a.s.,}$$

then (π^*,ρ^*) is not necessarily ν-optimal.

PROOF. We introduce the following D/F/F/2 process (see also theorem 3.2 in Groenewegen (1976)).

	1
1	$0 + \Gamma_0$

Γ_0

	1	2
1	$1 + \Gamma_0$	$1 + \Gamma_0$
2	$1 + \Gamma_0$	$0 + \Gamma_1$

Γ_1

$$r(h) = \sum_{t=0}^{\infty} \theta(x_t, a_t^{(0)}, a_t^{(1)}),$$

$$\mathbb{P}_{\nu,(\pi,\rho)}(X_0 = 1) = 1 \text{ for all } (\pi,\rho) \in \Pi.$$

The strategies are only given for state 1. It can be seen immediately that

$$(\pi^*,\rho^*) = (2\ 2\ \ldots,\ 1\ 1\ \ldots)$$

is ν-equalizing. In addition (π^*,ρ^*) has the property as formulated in theorem 5.2.11. But (π^*,ρ^*) is not ν-optimal, since

$$v_0(1,\pi^*,\rho^*) = 1 > v_0(1,\pi^*,2\ 2\ \ldots) = 0. \qquad \square$$

Now the following conclusions are obvious. Firstly, conservingness formulated in terms of the saddle function w_t seems not very useful in a characterization of optimality. Secondly, counterexample 5.2.10 shows that optimality of a strategy (π^*,ρ^*) does not imply for instance that π^* exploits the mistakes of player 1.
It so happens that optimality concepts can be defined in which the strategies of each player do exploit the mistakes of the counterplayer more or less, and which can be characterized in terms of the saddle function. One of these concepts is given now, two other concepts will be given later on, one at the end of this section and the other in the next section. For a good understanding of the definitions of these concepts, it is important to note, with respect to what pairs of strategies (π,ρ) we require an (in)equality to hold $\mathbb{P}_{\nu,(\pi,\rho)}$ - a.s.

5.2.12. DEFINITION. A strategy $(\pi^*,\rho^*) \in \Pi$ is called ν-*subgame perfect* iff for all $t \in T$ and $(\pi,\rho) \in \Pi$

$$v_t(H_t,\pi,{}_t\rho\rho^*(t;H_t)) \le w_t(H_t) \le v_t(H_t,{}_t\pi\pi^*(t;H_t),\rho) \quad \mathbb{P}_{\nu,(\pi,\rho)} \text{ - a.s.}$$

Corresponding to this new type of optimality, we define new conserving and equalizing properties. It will be proved that together they characterize subgame perfectness. We first give the new conservingness.

5.2.13. DEFINITION. A strategy $(\pi^*,\rho^*) \in \Pi$ is called ν-*overall saddling* iff for all $t \in T$ and $(\pi,\rho) \in \Pi$

$$E^{F_t}_{\nu,(\pi,{}_t\rho\rho^*(t;H_t))}\, w_{t+1}(H_{t+1}) \le w_t(H_t) \le E^{F_t}_{\nu,({}_t\pi\pi^*(t;H_t),\rho)}\, w_{t+1}(H_{t+1})$$

$$\mathbb{P}_{\nu,(\pi,\rho)} \text{ - a.s.}$$

We now give the new equalizingness.

5.2.14. DEFINITION. A strategy $(\pi^*,\rho^*) \in \Pi$ is called ν-*overall asymptotically definite* iff for all $t \in T$ and $(\pi,\rho) \in \Pi$

$$\lim_{\tau\to\infty} E^{F_t}_{\nu,(\pi,{}_t\rho\rho^*(t;H_t))}[\, w_{t+\tau}(H_{t+\tau}) - v_{t+\tau}(H_{t+\tau},\pi,{}_t\rho\rho^*(t;H_t))]^- = 0$$

and $$\lim_{\tau\to\infty} E^{F_t}_{\nu,({}_t\pi\pi^*(t;H_t),\rho)}[\, w_{t+\tau}(H_{t+\tau}) - v_{t+\tau}(H_{t+\tau},{}_t\pi\pi^*(t;H_t),\rho)]^+ = 0.$$

The names we have chosen, may seem a bit strange. The reason is that the names ν-saddling and ν-asymptotically definite have been reserved for concepts that will be defined later in this chapter.

5.2.15. THEOREM. A necessary and sufficient condition for ν-subgame perfectness of a strategy $(\pi^*,\rho^*) \in \Pi$ is, that (π^*,ρ^*) is both ν-overall saddling and ν-overall asymptotically definite.

PROOF. Suppose (π^*,ρ^*) is ν-subgame perfect. First we prove the ν-overall saddlingness.

$$w_t(H_t) \le v_t(H_t,{}_t\pi\pi^*(t;H_t),{}_{t+1}\rho\rho^*(t+1;H_{t+1})) =$$

$$= E^{F_t}_{\nu,({}_t\pi\pi^*(t;H_t),{}_{t+1}\rho\rho^*(t+1;H_{t+1}))}\, v_{t+1}(H_{t+1},{}_t\pi\pi^*(t;H_t),{}_{t+1}\rho\rho^*(t+1;H_{t+1})) =$$

$$= E^{F_t}_{\nu,({}_t\pi\pi^*(t;H_t),\rho)}\, w_{t+1}(H_{t+1}) \qquad \mathbb{P}_{\nu,(\pi,\rho)} \text{ - a.s.}$$

(In the first and the last step we used the subgame perfectness.)

Analogously

$$w_t(H_t) \geq E^{F_t}_{\nu,(\pi,{}_t\rho\rho^*(t;H_t))}\, w_{t+1}(H_{t+1}) \qquad \mathbb{P}_{\nu,(\pi,\rho)} \text{ - a.s.}$$

In order to show that (π^*,ρ^*) is ν-overall asymptotically definite, we note that for all $\tau \geq 0$

$$w_{t+\tau}(H_{t+\tau}) - v_{t+\tau}(H_{t+\tau},{}_t\pi\pi^*(t;H_t),\rho) \leq 0 \quad \mathbb{P}_{\nu,({}_t\pi\pi^*(t;H_t),\rho)} \text{ - a.s.}$$

Hence $\lim\limits_{\tau\to\infty} E^{F_t}_{\nu,({}_t\pi\pi^*(t;H_t),\rho)} [w_{t+\tau}(H_{t+\tau}) - v_{t+\tau}(H_{t+\tau},{}_t\pi\pi^*(t;H_t),\rho))]^+ = 0.$

And analogously

$$\lim_{\tau\to\infty} E^{F_t}_{\nu,(\pi,{}_t\rho\rho^*(t;H_t))} [w_{t+\tau}(H_{t+\tau}) - v_{t+\tau}(H_{t+\tau},\pi,{}_t\rho\rho^*(t;H_t))]^- = 0.$$

Now assume (π^*,ρ^*) is ν-overall saddling and ν-overal asymptotically definite. Using that (π^*,ρ^*) is ν-overall saddling we get that

$$w_t(H_t) \leq E^{F_t}_{\nu,({}_t\pi\pi^*(t;H_t),\rho)} w_{t+\tau}(H_{t+\tau}) \leq \lim_{\tau\to\infty} E^{F_t}_{\nu,({}_t\pi\pi^*(t;H_t),\rho)} w_{t+\tau}(H_{t+\tau}).$$

Using that (π^*,ρ^*) is ν-overall asymptotically definite, we may conclude that

$$w_t(H_t) \leq E^{F_t}_{\nu,({}_t\pi\pi^*(t;H_t),\rho)}\; v_{t+\tau}(H_{t+\tau},{}_t\pi\pi^*(t;H_t),\rho) = v_t(H_t,{}_t\pi\pi^*(t;H_t),\rho)$$

$$\mathbb{P}_{\nu,(\pi,\rho)} \text{ - a.s.,}$$

and analogously

$$w_t(H_t) \geq v_t(H_t,\pi,{}_t\rho\rho^*(t;H_t)) \qquad \mathbb{P}_{\nu,(\pi,\rho)} \text{ - a.s.} \qquad \square$$

In addition to this rather strong type of optimality we define and characterize two others types of optimality. For ν-tail optimality this will be done now. For ν-persistent optimality this is to be done in the next section.

5.2.16. DEFINITION. A strategy $(\pi^*,\rho^*) \in \Pi$ is called ν-*tail optimal* iff for all $t \in T$ and $(\pi,\rho) \in \Pi$

(i) $v_t(H_t, {}_t\pi^*\pi(t;H_t), {}_t\rho\rho^*(t;H_t)) \le w_t(H_t) \le v_t(H_t,\pi^*,\rho) \quad \mathbb{P}_{\nu,(\pi^*,\rho)}$ - a.s.

(ii) $v_t(H_t,\pi,\rho^*) \le w_t(H_t) \le v_t(H_t, {}_t\pi\pi^*(t;H_t), {}_t\rho^*\rho(t;H_t)) \quad \mathbb{P}_{\nu,(\pi,\rho^*)}$ - a.s.

It can be verified very easily that ν-subgame perfectness implies ν-tail optimality, and that ν-tail optimality implies ν-optimality.

5.2.17. DEFINITION. A strategy $(\pi^*,\rho^*) \in \Pi$ is called

(a) ν-*tail saddling* iff for all $t \in T$ and $(\pi,\rho) \in \Pi$

(i) $E^{F_t}_{\nu,({}_t\pi^*\pi(t;H_t),{}_t\rho\rho^*(t;H_t))} \psi_{t+1}(H_{t+1}, {}_t\rho\rho^*(t;H_t)) \le w_t(H_t) \le$

$\le E^{F_t}_{\nu,(\pi^*,\rho)} w_{t+1}(H_{t+1}) \qquad \mathbb{P}_{\nu,(\pi^*,\rho)}$ - a.s.

(ii) $E^{F_t}_{\nu,(\pi,\rho^*)} w_{t+1}(H_{t+1}) \le w_t(H_t) \le$

$\le E^{F_t}_{\nu,({}_t\pi\pi^*(t;H_t),{}_t\rho^*\rho(t;H_t))} \varphi_{t+1}(H_{t+1}, {}_t\pi\pi^*(t;H_t)) \qquad \mathbb{P}_{\nu,(\pi,\rho^*)}$ - a.s.

(b) ν-*asymptotically definite* iff for all $t \in T$ and $(\pi,\rho) \in \Pi$

(i) $\lim_{\tau\to\infty} E^{F_t}_{\nu,(\pi^*,\rho)} [\, w_{t+\tau}(H_{t+\tau}) - v_{t+\tau}(H_{t+\tau},\pi^*,\rho)]^+ = 0 \quad \mathbb{P}_{\nu,(\pi^*,\rho)}$ - a.s.

(ii) $\lim_{\tau\to\infty} E^{F_t}_{\nu,(\pi,\rho^*)} [\, w_{t+\tau}(H_{t+\tau}) - v_{t+\tau}(H_{t+\tau},\pi,\rho^*)]^- = 0 \quad \mathbb{P}_{\nu,(\pi,\rho^*)}$ - a.s.

5.2.18. THEOREM. Suppose assumption 5.2.2. holds. A necessary and sufficient condition for ν-tail optimality of a strategy $(\pi^*,\rho^*) \in \Pi$ is, that (π^*,ρ^*) is ν-tail saddling and ν-asymptotically definite.

PROOF. Assume (π^*,ρ^*) is ν-tail optimal. To show the ν-tail saddlingness, we first note that analogously to the proof of theorem 5.2.15 we have

$$w_t(H_t) \le v_t(H_t,\pi^*,{}_{t+1}\rho\rho^*(t+1;H_{t+1})) =$$

$$= E^{F_t}_{\nu,(\pi^*,{}_{t+1}\rho\rho^*(t+1;H_{t+1}))}\ v_{t+1}(H_{t+1},\pi^*,{}_{t+1}\rho\rho^*(t+1;H_{t+1})) =$$

$$= E^{F_t}_{\nu,(\pi^*,{}_{t+1}\rho\rho^*(t+1;H_{t+1}))}\ w_{t+1}(H_{t+1}) = E^{F_t}_{\nu,(\pi^*,\rho)}\ w_{t+1}(H_{t+1})$$

$$\mathbb{P}_{\nu,(\pi^*,\rho)} - \text{a.s.}$$

Next we note that ν-tail optimality implies, that

$$w_t(H_t) = v_t(H_t,\pi^*,{}_t\rho\rho^*(t;H_t)) = \psi_t(H_t,{}_t\rho\rho^*(t;H_t)) \quad \mathbb{P}_{\nu,(\pi^*,\rho)} - \text{a.s.}$$

Hence, using theorem 5.2.3, which applies by assumption 5.2.2, we may conclude

$$w_t(H_t) = \psi_t(H_t,{}_t\rho\rho^*(t;H_t)) \ge$$

$$\ge E^{F_t}_{\nu,({}_t\pi^*\pi(t;H_t),{}_t\rho\rho^*(t;H_t))}\ \psi_{t+1}(H_{t+1},{}_t\rho\rho^*(t;H)) \qquad \mathbb{P}_{\nu,(\pi^*,\rho)} - \text{a.s.}$$

Analogously we have

$$E^{F_t}_{\nu,(\pi,\rho^*)}\ w_{t+1}(H_{t+1}) \le w_t(H_t) \le$$

$$\le E^{F_t}_{\nu,({}_t\pi\pi^*(t;H_t),{}_t\rho^*\rho(t;H_t))}\ \varphi_{t+1}(H_{t+1},{}_t\pi\pi^*(t;H_t)) \qquad \mathbb{P}_{\nu,(\pi,\rho^*)} - \text{a.s.}$$

This establishes the ν-tail saddlingness.

Since for all $\tau \geq 0$ we have, that

$$w_{t+\tau}(H_{t+\tau}) - v_{t+\tau}(H_{t+\tau},\pi^*,\rho) \leq 0 \qquad \mathbb{P}_{\nu,(\pi^*,\rho)} \text{ - a.s.,}$$

it follows that

$$\lim_{\tau\to\infty} E^{F_t}_{\nu,(\pi^*,\rho)} [w_{t+\tau}(H_{t+\tau}) - v_{t+\tau}(H_{t+\tau},\pi^*,\rho)]^+ = 0 \ \mathbb{P}_{\nu,(\pi^*,\rho)} \text{ - a.s.}$$

This establishes the ν-asymptotic definiteness, since the other part of definition 5.2.17(b) can be proved in the same way.
Now suppose (π^*,ρ^*) to be ν-tail saddling and ν-asymptotically definite. We are to show, that (π^*,ρ^*) is ν-tail optimal. Firstly we have

$$w_t(H_t) \leq E^{F_t}_{\nu,(\pi^*,\rho)} w_{t+1}(H_{t+1}) \leq \lim_{\tau\to\infty} E^{F_t}_{\nu,(\pi^*,\rho)} w_{t+\tau}(H_{t+\tau}) \leq$$

$$\leq \lim_{\tau\to\infty} E^{F_t}_{\nu,(\pi^*,\rho)} v_{t+\tau}(H_{t+\tau},\pi^*,\rho) = v_t(H_t,\pi^*,\rho) \qquad \mathbb{P}_{\nu,(\pi^*,\rho)} \text{ - a.s.}$$

Secondly it follows, that

$$w_t(H_t) \geq E^{F_t}_{\nu,({}_t\pi^*\pi(t;H_t),{}_t\rho\rho^*(t;H_t))} \psi_{t+1}(H_{t+1},{}_t\rho\rho^*(t;H_t)) \geq$$

$$\geq E^{F_t}_{\nu,({}_t\pi^*\pi(t;H_t),{}_t\rho\rho^*(t;H_t)} v_{t+1}(H_{t+1},{}_t\pi^*\pi(t;H_t),{}_t\rho\rho^*(t;H_t)) =$$

$$= v_t(H_t,{}_t\pi^*\pi(t;H_t),{}_t\rho\rho^*(t;H_t)) \qquad \mathbb{P}_{\nu,(\pi^*,\rho)} \text{ - a.s.}$$

Analogously it holds, that

$$v_t(H_t,\pi,\rho^*) \leq w_t(H_t) \leq v_t(H_t,{}_t\pi\pi^*(t;H_t),{}_t\rho^*\rho(t;H_t)) \quad \mathbb{P}_{\nu,(\pi,\rho^*)} \text{ - a.s.}$$

which completes the proof. □

5.3. THE D/G/G/2 PROCESS WITH A RECURSIVE ZERO-SUM UTILITY

In this section the results of the previous section will be specialized to the recursive case.

Recursiveness in this section is just the same as it is in chapter 3, i.e. the D/G/G/2 process is t-separable for every $t \in T$, and there exist functions $\theta_\tau^{[t]}$, $\chi_\tau^{[t]}$ and $r^{[t]}$ with $t,\tau \in T$, such that $r^{[0]} = r$ and $r_\tau^{[t]}(h) = \theta_\tau^{[t]}(h_\tau) + \chi_\tau^{[t]}(h_\tau) r^{[\tau]}(\zeta^{\tau-t}(h))$ for all $h=(x_t,a_t,\ldots)=(h_\tau,a_\tau,x_{\tau+1},\ldots) \in \mathop{X}_{k=t}^{\infty}(X \times A)$. (For details see definitions 3.2.1, 3.1.1 and 3.1.2. Note that now (quasi) integrability is required with respect to each $\mathbb{P}_{\nu,(\pi,\rho)}$ instead of $\mathbb{P}_{\nu,\pi}$.)

From now on in this section r is supposed to be recursive. This has as an immediate consequence that lemma 3.2.2 applies, i.e. we have

5.3.1. LEMMA. If r is a recursive utility, then for all $t \in T$ and $h_t \in H_t$

$$\theta_\tau^{[0]}(h_\tau) = \sum_{k=1}^{\tau} \left[\prod_{\ell=1}^{k-1} \chi_\ell^{[\ell-1]}(x_{\ell-1},a_{\ell-1}^{(0)},a_{\ell-1}^{(1)},x_\ell) \right] \theta_k^{[k-1]}(x_{k-1},a_{k-1}^{(0)},a_{k-1}^{(1)},x_k),$$

$$\chi_\tau^{[0]}(h_\tau) = \prod_{k=1}^{\tau} \chi_k^{[k-1]}(x_{k-1},a_{k-1}^{(0)},a_{k-1}^{(1)},x_k).$$

Moreover, lemma 3.1.4 applies directly if ψ_t or φ_t is substituted for the value function w_t in chapter 2 and 3. Its proof can even be repeated for the saddle w_t of this chapter. As in chapter 3 we denote by $\psi_\tau^{[t]}$ (etc.) the function ψ_τ (etc) for the t-delayed process.

5.3.2. LEMMA. If r is a recursive utility, then for all $t \in T$, $h_t \in H_t$ and $(\pi,\rho) \in \Pi$

$$\psi_t(h_t,\pi) = \theta_t^{[0]}(h_t) + \chi_t^{[0]}(h_t)\ \psi_t^{[t]}(x_t,\pi(t;h_t)),$$

$$\varphi_t(h_t,\rho) = \theta_t^{[0]}(h_t) + \chi_t^{[0]}(h_t)\ \varphi_t^{[t]}(x_t,\rho(t;h_t)),$$

$$w_t(h_t) = \theta_t^{[0]}(h_t) + \chi_t^{[0]}(h_t)\ w_t^{[t]}(x_t),$$

$$v_t(h_t,\pi,\rho) = \theta_t^{[0]}(h_t) + \chi_t^{[0]}(h_t)\ v_t^{[t]}(x_t,\pi(t;h_t),\rho(t;h_t)).$$

So we are in a position to prove the following analogue of theorem 3.2.3.

5.3.3. THEOREM. If r is a recursive utility, then

(i) $\psi_t(h_t,\rho^*) \stackrel{=}{>} E_{h_t,(\pi^*,\rho^*)}\ \psi_{t+1}(H_{t+1},\rho^*)$ is equivalent to

$$\psi_t^{[t]}(x_t,\rho^*(t;h_t)) \stackrel{=}{>} E_{h_t,(\pi^*,\rho^*)}\ [\theta_{t+1}^{[t]}(X_t,A_t,X_{t+1}) + $$
$$+\ \chi_{t+1}^{[t]}(X_t,A_t,X_{t+1})\ \psi_{t+1}^{[t+1]}(X_{t+1},\rho^*(t+1;H_{t+1}))],$$

(ii) $\varphi_t(h_t,\pi^*) \stackrel{=}{<} E_{h_t,(\pi^*,\rho^*)}\ \varphi_{t+1}(H_{t+1},\pi^*)$ is equivalent to

$$\varphi_t^{[t]}(x_t,\pi^*(t;h_t)) \stackrel{=}{<} E_{h_t,(\pi^*,\rho^*)}\ [\theta_{t+1}^{[t]}(X_t,A_t,X_{t+1}) + $$
$$+\ \chi_{t+1}^{[t]}(X_t,A_t,X_{t+1})\ \varphi_{t+1}^{[t+1]}(X_{t+1},\pi^*(t+1;H_{t+1}))],$$

(iii) $w_t(h_t) \stackrel{\leq}{\stackrel{=}{>}} E_{h_t,(\pi^*,\rho^*)}\ w_{t+1}(H_{t+1})$ is equivalent to

$$w_t^{[t]}(x_t) \stackrel{\leq}{\stackrel{=}{>}} E_{h_t,(\pi^*,\rho^*)}\ [\theta_{t+1}^{[t]}(X_t,A_t,X_{t+1}) + \chi_{t+1}^{[t]}(X_t,A_t,X_{t+1})\,w_{t+1}^{[t+1]}(X_{t+1})].$$

PROOF. The proof is completely the same as the proof of theorem 3.2.3, except that lemma 5.3.1 and 5.3.2 are used instead of lemma 3.2.2 and 3.1.4. □

The property ν-tail vanishingness is essentially the same as in chapter 3, i.e. r is recursive and for all $(\pi,\rho) \in \Pi$

5.3.4.1. $$\lim_{t\to\infty} E_{\nu,(\pi,\rho)}\ \chi_t^{[0]}(H_t)\ v_t^{[t]}(X_t,\pi(t;H_t),\rho(t;H_t)) = 0\ .$$

Now the analogue of theorem 3.2.5 follows immediately.

5.3.5. THEOREM. If r is a ν-tail vanishing utility, then

(i) $$\lim_{t\to\infty} E_{\nu,(\pi,\rho)}\ [\psi_t(H_t,\rho) - v_t(H_t,\pi,\rho)] \overset{=}{>} 0$$ is equivalent to

$$\lim_{t\to\infty} E_{\nu,(\pi,\rho)}\ \chi_t^{[0]}(H_t)\ \psi_t^{[t]}(X_t,\rho(t;H_t)) \overset{=}{>} 0,$$

(ii) $$\lim_{t\to\infty} E_{\nu,(\pi,\rho)}\ [\varphi_t(H_t,\pi) - v_t(H_t,\pi,\rho)] \overset{=}{<} 0$$ is equivalent to

$$\lim_{t\to\infty} E_{\nu,(\pi,\rho)}\ \chi_t^{[0]}(H_t)\ \varphi_t^{[t]}(X_t,\pi(t;H_t)) \overset{=}{<} 0$$

(iii) $$\lim_{t\to\infty} E_{\nu,(\pi,\rho)}\ [w_t(H_t) - v_t(H_t,\pi,\rho)] \lesseqgtr 0$$ is equivalent to

$$\lim_{t\to\infty} E_{\nu,(\pi,\rho)}\ \chi_t^{[0]}(H_t)\ w_t^{[t]}(X_t) \lesseqgtr 0.$$

PROOF. Use the proof of theorem 3.2.5. □

Both theorems together lead in a straightforward way to the following result.

5.3.6. COROLLARY. Let r be a ν-tail vanishing utility. A necessary and sufficient condition for ν-optimality of $(\pi^*,\rho^*) \in \Pi$ is the validity of the following four assertions for all $t \in T$

5.3.6.1. $\psi_t^{[t]}(X_t,\rho^*(t;H_t)) = E_{\nu,(\pi^*,\rho^*)}^{F_t} [\theta_{t+1}^{[t]}(X_t,A_t,X_{t+1}) + $

$+ \chi_{t+1}^{[t]}(X_t,A_t,X_{t+1})\ \psi_{t+1}^{[t+1]}(X_{t+1},\rho^*(t+1;H_{t+1}))] \qquad \mathbb{P}_{\nu,(\pi^*,\rho^*)}$ - a.s.,

5.3.6.2. $\varphi_t^{[t]}(X_t,\pi^*(t;H_t)) = E_{\nu,(\pi^*,\rho^*)}^{F_t} [\theta_{t+1}^{[t]}(X_t,A_t,X_{t+1}) + $

$+ \chi_{t+1}^{[t]}(X_t,A_t,X_{t+1})\ \varphi_{t+1}^{[t+1]}(X_{t+1},\pi^*(t+1;H_{t+1}))] \qquad \mathbb{P}_{\nu,(\pi^*,\rho^*)}$ - a.s.

5.3.6.3. $\lim_{t\to\infty} E_{\nu,(\pi^*,\rho^*)}\ \chi_t^{[0]}(H_t)\ \psi_t^{[t]}(X_t,\rho^*(t;H_t)) = 0,$

5.3.6.4. $\lim_{t\to\infty} E_{\nu,(\pi^*,\rho^*)}\ \chi_t^{[0]}(H_t)\ \varphi_t^{[t]}(X_t,\pi^*(t;H_t)) = 0.$

To be able to characterize ν-subgame perfectness in a similar way as ν-optimality, we need the following concept.

5.3.7. DEFINITION. A recursive utility is called ν-*overall tail vanishing* iff for all $t \in T$ and $(\pi,\rho) \in \Pi$

5.3.7.1. $\lim_{\tau\to\infty} E_{\nu,(\pi,\rho)}^{F_t}\ \chi_\tau^{[0]}(H_\tau)\ v_\tau^{[\tau]}(X_\tau,\pi(\tau;H_\tau),\rho(\tau;H_\tau)) = 0$

$\mathbb{P}_{\nu,(\pi,\rho)}$ - a.s.

It should be noted that in general the ν-overall tail vanishing property does not imply the ν-tail vanishing property, since in formula 5.3.7.1 we have almost sure convergence, and in formula 5.3.4.1 we have convergence which is only slightly weaker than L^1 convergence. There exists a standard example of nonnegative functions f_n converging a.s. to a function f, but not in L^1 sense. By the nonnegativity of the f_n's this example shows at the same time, that a.s. convergence in general does not imply the "L^1 convergence in the above weaker sense". After corollary 5.3.9 we will discuss a sufficient condition for this implication to be true.

5.3.8. THEOREM. If r is a ν-overall tail vanishing utility, then

$$\lim_{\tau\to\infty} E^{F_t}_{\nu,(\pi,\rho)} [w_\tau(H_\tau) - v_\tau(H_\tau,\pi,\rho)] \overset{\le}{\underset{>}{=}} 0 \text{ is equivalent to}$$

$$\lim_{\tau\to\infty} E^{F_t}_{\nu,(\pi,\rho)} \chi^{[0]}_\tau(H_\tau)\, w^{[\tau]}_\tau(X_\tau) \overset{\le}{\underset{>}{=}} 0.$$

PROOF. Analogous to the proof of theorem 3.2.5. □

Applying theorem 5.3.3 and 5.3.8 we obtain the following result.

5.3.9. COROLLARY. Let r be a ν-overall tail vanishing utility. A necessary and sufficient condition for ν-subgame perfectness of $(\pi^*,\rho^*) \in \Pi$ is the validity of the following three assertions for all $t \in T$ and $(\pi,\rho) \in \Pi$

5.3.9.1. $$E^{F_t}_{\nu,(\pi,{}_t\rho\rho^*(t;H_t))} [\theta^{[t]}_{t+1}(X_t,A_t,X_{t+1}) + \chi^{[t]}_{t+1}(X_t,A_t,X_{t+1})\, w^{[t+1]}_{t+1}(X_{t+1})] \le$$

$$\le w^{[t]}_t(X_t) \le E^{F_t}_{\nu,({}_t\pi\pi^*(t;H_t),\rho)} [\theta^{[t]}_{t+1}(X_t,A_t,X_{t+1}) +$$

$$+ \chi^{[t+1]}_{t+1}(X_t,A_t,X_{t+1}) w^{[t+1]}_{t+1}(X_{t+1})] \qquad \mathbb{P}_{\nu,(\pi,\rho)} \text{ - a.s.}$$

5.3.9.2. $$\lim_{\tau\to\infty} E^{F_t}_{\nu,(\pi,{}_t\rho\rho^*(t;H_t))} \chi^{[0]}_\tau(H_\tau)\, w^{[\tau]}_\tau(X_\tau) \ge 0 \qquad \mathbb{P}_{\nu,(\pi,\rho)} \text{ - a.s.}$$

5.3.9.3. $$\lim_{\tau\to\infty} E^{F_t}_{\nu,({}_t\pi\pi^*(t;H_t),\rho)} \chi^{[0]}_\tau(H_\tau)\, w^{[\tau]}_\tau(X_\tau) \le 0 \qquad \mathbb{P}_{\nu,(\pi,\rho)} \text{ - a.s.}$$

It is well known (see e.g. Meyer (1966) section II.2) that uniform integrability is a necessary and sufficient condition for L^1 convergence of a sequence of random variables that converge almost surely. Suppose that r is ν-overall tail vanishing. Fix $t \in T$ and $(\pi,\rho) \in \Pi$. Then we have $\mathbb{P}_{\nu,(\pi,\rho)}$ - a.s. convergence (to zero) of

$$\lambda_{t,\tau}(H) := E^{F_t}_{\nu,(\pi,\rho)} \; \chi^{[0]}_{\tau}(H_\tau) \; v^{[\tau]}_{\tau}(X_\tau, \pi(\tau;H_\tau), \rho(\tau;H_\tau))$$

for $\tau \to \infty$. In order to have also L^1 convergence (with respect to the measure $\mathbb{P}_{\nu,(\pi,\rho)}$) of the functions $\lambda_{t,\tau}$, $\tau \geq 0$, it is necessary and sufficient that these functions are uniformly integrable, i.e.

$$\lim_{t\to\infty} \sup_{\tau\geq 0} \int_{\{|\lambda_{t,\tau}|>c\}} |\lambda_{t,\tau}(h)| \; \mathbb{P}_{\nu,(\pi,\rho)}(dh) = 0 \; .$$

This gives rise to the following theorem.

5.3.10. THEOREM. A strategy $(\pi^*,\rho^*) \in \Pi$ is ν-tail vanishing if it is ν-tail overall vanishing, provided that for every $(\pi,\rho) \in \Pi$ there exist real numbers M,N, such that

$$E_{\nu,(\pi,\rho)} \; r(H) = E_{\nu,(\pi,\rho)} \; [\sum_{\ell=1}^{\infty} \chi^{[0]}_{\ell=1}(H_{\ell-1}) \; \theta^{[\ell-1]}_{\ell}(X_{\ell-1}, A_{\ell-1}, X_\ell) + N \,]$$

and $$E_{\nu,(\pi,\rho)} \sum_{\ell=1}^{\infty} \chi^{[0]}_{\ell-1}(H_{\ell-1}) \; | \; \theta^{[\ell-1]}_{\ell}(X_{\ell-1}, A_{\ell-1}, X_\ell) \; | = M < \infty.$$

PROOF. Fix $(\pi,\rho) \in \Pi$. For all $\tau \geq 0$

$$|\lambda_{0,\tau}(H)| = |E^{F_t}_{\nu,(\pi,\rho)} \; \chi^{[0]}_{\tau}(H_\tau) \; v^{[\tau]}_{\tau}(X_\tau, \pi(\tau;H_\tau), \rho(\tau;H_\tau))| \leq$$

$$\leq E^{F_t}_{\nu,(\pi,\rho)} \sum_{\ell=1}^{\infty} \chi^{[0]}_{\ell-1}(H_{\ell-1}) \; |\theta^{[\ell-1]}_{\ell}(X_{\ell-1}, A_{\ell-1}, X_\ell)| \quad \mathbb{P}_{\nu,(\pi,\rho)} \text{ - a.s.}$$

This means that $E_{\nu,(\pi,\rho)} \; |\lambda_{0,\tau}(H)| \leq M < \infty$.

Hence $\int \sup_{\tau} |\lambda_{0,\tau}(h)| \; \mathbb{P}_{\nu,(\pi,\rho)}(dh) < \infty$.

So
$$0 = \lim_{c\to\infty} \sup_{\tau} \int_{\{\sup_{\tau}|\lambda_{0,\tau}|>c\}} \sup_{\tau} |\lambda_{0,\tau}(h)| \; \mathbb{P}_{\nu,(\pi,\rho)}(dh) \geq$$

$$\geq \lim_{c\to\infty} \sup_{\tau} \int_{\{|\lambda_{0,\tau}|>c\}} \sup_{\tau} |\lambda_{0,\tau}(h)| \; \mathbb{P}_{\nu,(\pi,\rho)}(dh) \geq$$

$$\geq \lim_{c\to\infty} \sup_{\tau} \int_{\{|\lambda_{0,\tau}|>c\}} |\lambda_{0,\tau}(h)| \; \mathbb{P}_{\nu,(\pi,\rho)}(dh) \geq 0.$$

This implies that the sequence of functions $\lambda_{0,\tau}$, $\tau \geq 0$ is uniformly integrable.
Hence the L^1 limit of $\lambda_{0,\tau}(H)$ exists, and is equal to zero since the $\mathbb{P}_{\nu,(\pi,\rho)}$ - a.s. limit is zero. So

$$0 = \lim_{\tau\to\infty} E_{\nu,(\pi,\rho)} \mid \lambda_{0,\tau}(H) - 0 \mid = \lim_{\tau\to\infty} E_{\nu,(\pi,\rho)} \lambda_{0,\tau}(H),$$

which is precisely the ν-tail vanishing property. □

It seems worthwile to note that, if for a discrete state space the condition for theorem 5.3.10 is satisfied for all degenerate starting distributions ν, and if moreover $N = 0$ for all $(\pi,\rho) \in \Pi$, then the utility has a so-called charge structure as introduced in Hordijk (1974) for the convergent dynamic programming situation.
For the ν-tail optimality a characterization can be given, similar to the characterization in theorem 5.3.9.

5.3.11. COROLLARY. Let r be a ν-overall tail vanishing utility, and let assumption 5.2.3 hold. A necessary and sufficient condition for ν-tail optimality of $(\pi^*,\rho^*) \in \Pi$ is the validity of the following inequalities for all $t \in T$ and $(\pi,\rho) \in \Pi$

5.3.11.1. $$E^{F_t}_{\nu,({}_t\pi^*\pi(t,H_t),{}_t\rho\rho^*(t,H_t))}[\theta^{[t]}_{t+1}(X_t,A_t,X_{t+1}) + \chi^{[t]}_{t+1}(X_t,A_t,X_{t+1})\ \psi^{[t+1]}_{t+1}(X_{t+1},\rho^*(t+1;H_{t+1}))] \le$$

$$\le w_0(X_t) \le E^{F_t}_{\nu,(\pi^*,\rho)}\ [\theta^{[t]}_{t+1}(X_t,A_t,X_{t+1}) + \chi^{[t]}_{t+1}(X_t,A_t,X_{t+1})\ w^{[t+1]}_{t+1}(X_{t+1})] \qquad \mathbb{P}_{\nu,(\pi^*,\rho)} \text{ - a.s.,}$$

5.3.11.2. $$E^{F_t}_{\nu,(\pi,\rho^*)}[\theta^{[t]}_{t+1}(X_t,A_t,X_{t+1}) + \chi^{[t]}_{t+1}(X_t,A_t,X_{t+1})\ w^{[t+1]}_{t+1}(X_{t+1})] \le$$

$$\le w^{[t]}_t(X_t) \le E^{F_t}_{\nu,({}_t\pi\pi^*(t;H_t),{}_t\rho^*\rho(t;H_t))}[\theta^{[t]}_{t+1}(X_t,A_t,X_{t+1}) + \chi^{[t]}_{t+1}(X_t,A_t,X_{t+1})\ \varphi^{[t+1]}_{t+1}(X_{t+1},\pi^*(t+1;H_{t+1}))] \quad \mathbb{P}_{\nu,(\pi,\rho^*)} \text{ - a.s.,}$$

5.3.11.3. $$\lim_{\tau\to\infty} E^{F_t}_{\nu,(\pi^*,\rho)}\ \chi^{[0]}_\tau(H_\tau)\ w^{[\tau]}_\tau\ (X_\tau) \le 0 \qquad \mathbb{P}_{\nu,(\pi^*,\rho)} \text{ - a.s.,}$$

5.3.11.4. $$\lim_{\tau\to\infty} E^{F_t}_{\nu,(\pi,\rho^*)}\ \chi^{[0]}_\tau(H_\tau)\ w^{[\tau]}_\tau(X_\tau) \ge 0 \qquad \mathbb{P}_{\nu,(\pi,\rho^*)} \text{ - a.s.}$$

PROOF. The reason that for this corollary we need an extra proof, whereas the statements of corollary 5.3.9 and 5.3.6 are immediate, is that here we need a result, that is a variation of theorem 5.3.3. The result is the following. If r is a recursive utility, then

$$w_t(h_t) \ge E_{h_t,(\pi,\rho)}\ \psi_{t+1}(H_{t+1},\rho) \text{ is equivalent to}$$

$$w^{[t]}_t(x_t) \ge E_{h_t,(\pi,\rho)}\ [\theta^{[t]}_{t+1}(X_t,A_t,X_{t+1}) + \chi^{[t]}_{t+1}(X_t,A_t,X_{t+1})\ \psi^{[t+1]}_{t+1},\rho(t+1;H_{t+1}))],$$

and analogously

$$w_t(h_t) \geq E_{h_t,(\pi,\rho)}\ \varphi_{t+1}(H_{t+1},\pi) \text{ is equivalent to}$$

$$w_t^{[t]}(X_t) \geq E_{h_t,(\pi,\rho)}\ [\theta_{t+1}^{[t]}(X_t,A_t,X_{t+1}) + \\ + \chi_{t+1}^{[t]}(X_t,A_t,X_{t+1})\ \varphi_{t+1}^{[t+1]}(X_{t+1},\pi(t+1;H_{t+1}))].$$

Its proof is precisely the proof of theorem 5.3.3. Now corollary 5.3.11 follows immediately from theorem 5.3.8 and 5.3.3, and the above result. □

In chapter 3 we have given two more proofs of the characterization of ν-optimality, one by the martingale approach, and another making use of the optimality principle. It is possible to repeat these proofs for the situation described here, but we shall confine ourselves to the proof, that for a recursive utility the analogue of the optimality principle holds.

5.3.12. THEOREM. If r is a recursive utility and $(\pi^*,\rho^*) \in \Pi$ is ν-optimal, then for all $t \in T$ $(\pi^*(t;h_t),\rho^*(t;h_t))$ is μ-optimal for $\mathbb{P}_{\nu,(\pi^*,\rho^*)}$ - a.a. $h_t \in H_t$, with $\mu = \mathbb{P}^{(2t+1)}_{\nu,(\pi^*,\rho^*)}$ the marginal probability on the (2t+1) - th coordinate of H, that is the state space at time t.

PROOF. From definition 5.1.1. (or also from lemma 5.2.1) we know that

$$v_t(H_t,\pi^*,\rho^*) = \varphi_t(H_t,\pi^*) \qquad \mathbb{P}_{\nu,(\pi^*,\rho^*)} \text{ - a.s.}$$

Hence by lemma 5.3.2

$$\theta_t^{[0]}(H_t) + \chi_t^{[0]}(H_t) v_t^{[t]}(X_t,\pi^*(t;H_t),\rho^*(t;H_t)) = \theta_t^{[0]}(H_t) + \\ + \chi_t^{[0]}(H_t)\varphi_t^{[t]}(X_t,\pi^*(t;H_t)) \qquad \mathbb{P}_{\nu,(\pi^*,\rho^*)} \text{ - a.s.}$$

In other words for all $\rho \in \Pi^{(1)}$

$$v_t^{[t]}(X_t,\pi^*(t;H_t),\rho^*(t;H_t)) = \varphi_t^{[t]}(X_t,\pi^*(t;H_t)) \leq v_t^{[t]}(X_t,\pi^*(t;H_t),\rho)$$

$$\mathbb{P}_{\nu,(\pi^*,\rho^*)} \text{ - a.s.} \qquad \square$$

5.3.13. DEFINITION. If the ν-optimality of $(\pi^*,\rho^*) \in \Pi$ implies that $(\pi^*(t;h_t),\rho^*(t;h_t))$ is μ-optimal in the process $\Sigma^{[h_t]}$ for $\mathbb{P}_{\nu,(\pi^*,\rho^*)}$ - a.a. $h_t \in H_t$, and with $\mu = \mathbb{P}^{(2t+1)}_{\nu,(\pi^*,\rho^*)}$ as defined above, then we say that *the optimality principle holds* for the two-person zero-sum game Σ.

It can be seen from the first inequality of formula 5.3.11.1 and from the last of 5.3.11.2 that ν-tail optimality of a strategy is a stronger form of the optimality principle.

Finally we want to introduce another concept of optimality, called ν-persistent optimality. This new concept seems to make sense, only if the D/G/G/2 process is separable and moreover *stationary*, and if the utility is recursive *in a stationary way*."Stationarity" of the process means, that the admissibility of an action does not depend on time t, and that the transition functions p_t do not depend on t either. With "recursiveness in a stationary way" we mean, that for all $t \in T$ the function $r^{[t]} = r$, which implies $\chi^{[t]}_{t+1} = \chi^{[0]}_1 =: \chi$ and $\theta^{[t]}_{t+1} = \theta^{[0]}_1 =: \theta$. Both stationarity conditions together imply, that $v^{[t]}_\tau = v_{\tau-t}$ and $w^{[t]}_\tau = w_{\tau-t}$, since the set of tails of strategies from time t on is equal to the set of strategies itself. These remarks should be kept in mind while reading the following definitions, especially when we allow a strategy π^* or ρ^* to prescribe actions from time t on instead of time 0. In the remainder of this section it is supposed, that these stationarity conditions are satisfied for the process as well as for the utility.

5.3.14. DEFINITION. A strategy $(\pi^*,\rho^*) \in \Pi$ is called

(a) ν-*persistently optimal* iff for all $t \in T$ and $(\pi,\rho) \in \Pi$

$$v_0(X_t,\pi(t;H_t),\rho^*) \le w_0(X_t) \le v_0(X_t,\pi^*(t;H_t),\rho(t;H_t)) \quad \mathbb{P}_{\nu,(\pi^*,\rho)}\text{-a.s.},$$

$$v_0(X_t,\pi(t;H_t),\rho^*(t;H_t)) \le w_0(X_t) \le v_0(X_t,\pi^*,\rho(t;H_t)) \quad \mathbb{P}_{\nu,(\pi,\rho^*)}\text{-a.s.}$$

(b) ν-*saddling* iff for all $t \in T$ and $(\pi,\rho) \in \Pi$

$$E^{F_t}_{\nu,({}_t\pi^*\pi(t;H_t),{}_t\rho\rho^*)} [\theta(X_t,A_t,X_{t+1}) +$$

$$+ \chi(X_t,A_t,X_{t+1})\, \psi_0(X_{t+1},\rho^*(1;(X_t,A_t,X_{t+1})))] \le$$

$$\leq w_0(X_t) \leq E^{F_t}_{\nu,(\pi^*,\rho)} [\theta(X_t,A_t,X_{t+1}) + \chi(X_t,A_t,X_{t+1}) w_0(X_{t+1})]$$

$$\mathbb{P}_{\nu,(\pi^*,\rho)} \text{ - a.s.,}$$

$$E^{F_t}_{\nu,(\pi,\rho^*)} [\theta(X_t,A_t,X_{t+1}) + \chi(X_t,A_t,X_{t+1}) w_0(X_{t+1})] \leq w_0(X_t) \leq$$

$$\leq E^{F_t}_{\nu,({}_t\pi\pi^*,{}_t\rho^*\rho(t;H_t))} [\theta(X_t,A_t,X_{t+1}) +$$

$$+ \chi(X_t,A_t,X_{t+1}) \varphi_0(X_{t+1},\pi^*(1;(X_t,A_t,X_{t+1})))] \qquad \mathbb{P}_{\nu,(\pi,\rho^*)} \text{ - a.s.}$$

5.3.15. THEOREM. Let assumption 5.2.3 hold, let r be a ν-overall tail vanishing utility, and let the stationarity conditions be satisfied both for the process and for the utility. A necessary and sufficient condition for ν-persistent optimality of a strategy $(\pi^*,\rho^*) \in \Pi$ is, that (π^*,ρ^*) is ν-saddling and that formulae 5.3.11.3 and 5.3.11.4 hold.

PROOF. From the stationarity of the process and the utility, from lemma 5.3.2 and from theorem 5.3.3 together with its extension needed in the proof of theorem 5.3.11, we may reformulate the assertion of theorem 5.3.15 as follows. Suppose assumption 5.2.2 holds. The validity of the inequalities

5.3.15.1. $v_t(H_t,{}_t\pi^*\pi(t;H_t),{}_t\rho\rho^*) \leq w_t(H_t) \leq v_t(H_t,\pi^*,\rho) \qquad \mathbb{P}_{\nu,(\pi^*,\rho)}$ - a.s.

5.3.15.2. $v_t(H_t,\pi,\rho^*) \leq w_t(H_t) \leq v_t(H_t,{}_t\pi\pi^*,{}_t\rho^*\rho(t;H_t)) \qquad \mathbb{P}_{\nu,(\pi,\rho^*)}$ - a.s.

is a necessary and sufficient condition for the validity of the following assertion and inequalities

5.3.15.3. (π^*,ρ^*) is ν-asymptotically definite,

5.3.15.4. $E^{F_t}_{\nu,({}_t\pi^*\pi(t;H_t),{}_t\rho\rho^*)} \psi^{[0]}_{t+1}(H_{t+1},{}_t\rho\rho^*) \leq w^{[0]}_t(H_t) \leq$

$$\leq E^{F_t}_{\nu,(\pi^*,\rho)} w^{[0]}_{t+1}(H_{t+1}) \qquad \mathbb{P}_{\nu,(\pi^*,\rho)} \text{ - a.s.}$$

5.3.15.5. $E^{F_t}_{\nu,(\pi,\rho^*)} w^{[0]}_{t+1}(H_{t+1}) \le w^{[0]}_t(H_t) \le$

$$\le E^{F_t}_{\nu,({}_t\pi\pi^*,{}_t\rho^*\rho(t;H_t))} \varphi^{[0]}_{t+1}(H_{t+1},{}_t\pi\pi^*) \qquad \mathbb{P}_{\nu,(\pi,\rho^*)} \text{ - a.s.}$$

The proof of this statement is completely analogous to the proof of theorem 5.2.18. □

REMARK. From the proof of this theorem it follows that formulae 5.3.15.1 and 5.3.15.2 may also serve as a definition of ν-persistent optimality, and that formulae 5.3.15.4 and 5.3.15.5 may serve as definition of ν-saddlingness.

We want to make one more remark about the ν-persistent optimality. Assuming r to be a ν-overall tail vanishing utility we may use the characterization of theorem 5.3.15. It is easy to see, that the first inequality in formula 5.3.15.1 expresses the optimality (for player 1) of ρ^* in $\mathbb{P}_{\nu,(\pi^*,\rho)}$ - almost all X_t, and that the last inequality in formula 5.3.15.2 expresses the optimality (for player 0) of π^* in $\mathbb{P}_{\nu,(\pi,\rho^*)}$ - almost all X_t. Recall, that the sets of admissible actions $L^{(\ell)}_t$ and the transition probabilities $p^{(\ell)}_t$ do not depend on t. Then, if (π^*,ρ^*) is not only ν-persistently optimal, but also μ-optimal for every starting distribution μ, the first inequality of 5.3.15.1 and the last of 5.3.15.2 are automatically satisfied. And this last condition holds e.g. if the state space is discrete and ν is a starting distribution that gives positive probability to every state. So in this particular situation ν-tail optimality implies ν-persistent optimality.

5.4. THE C/G/G/2 PROCESS WITH A ZERO-SUM UTILITY.

In this section we will extend the results of both previous sections to the continuous-time case. We will start with a description of the general C/G/G/n process. Then we will derive, as in chapter 4, characterizations of the various optimality concepts for the case n = 2 and a zero-sum utility. The case n ≥ 2 and a nonzero-sum utility will be treated in chapter 6. We will try to avoid duplications of descriptions and proofs as much as possible by referring to earlier proofs.

The general C/G/G/n process, with n a cardinal number, is defined as a tuple

$$\Sigma = (T, (X,\mathcal{X}), ((A^{(\ell)},\mathcal{A}^{(\ell)}) \mid \ell \in \mathbb{N}_n), (U^{(\ell)} \mid \ell \in \mathbb{N}_n),$$

$$(\mathbb{P}_{x_0,(u^{(0)},\ldots,u^{(n-1)})} \mid x_0 \in X, u^{(\ell)} \in U^{(\ell)}), (r^{(\ell)} \mid \ell \in \mathbb{N}_n)).$$

As before, T is the time space, $(X,\mathcal{X})$ the measurable state space, $(A^{(\ell)},\mathcal{A}^{(\ell)})$ the measurable action space for player ℓ, $U^{(\ell)}$ the set of controls for player ℓ, $\mathbb{P}_{x_0,(u^{(0)},\ldots,u^{(n-1)})}$ a probability measure on the sample space

$$(H,\mathcal{H}) = (\mathop{X}_{t\in T}\{X \times (A^{(0)} \times \ldots \times A^{(n-1)})\}, \bigotimes_{t\in T}\{\mathcal{X} \times (\mathcal{A}^{(0)} \otimes \ldots \otimes \mathcal{A}^{(n-1)})\}$$

and $r^{(\ell)}$ the utility function for player ℓ. We define by $(A,\mathcal{A}) = (A^{(0)} \times \ldots \times A^{(n-1)}, \mathcal{A}^{(0)} \otimes \ldots \otimes \mathcal{A}^{(n-1)})$, the set of simultaneous actions, and $\mathcal{F}_t$ is the σ-field generated by sets of type $\mathop{X}_{\tau\in T}(X_\tau \times A_\tau)$ with $X_\tau \in \mathcal{X}, A_\tau \in \mathcal{A}, X_\tau \times A_\tau$ unequal to $X \times A$ for finitely many $\tau \le t$ and $X_\tau = X$ if $\tau > t$ and $A_\tau = A$ if $\tau \ge t$. This implies $\mathcal{F}_{t_0} \subset \mathcal{F}_{t_1} \subset \ldots \subset \mathcal{H}$ for each sequence $t_0 < t_1 < \ldots$ in T.

Define $H_t = [\mathop{X}_{\tau\in T,\tau<t} (X \times A)] \times X$, and let $h \in H$. The symbol h_t denotes the truncation of h contained in H_t. The set of simultaneous controls is defined as $U = U^{(0)} \times \ldots \times U^{(n-1)}$. Each $u^{(\ell)} \in U^{(\ell)}$ is a function $u^{(\ell)}: T \times H \times \mathcal{A}^{(\ell)} \to [0,1]$ such that $u^{(\ell)}(t,\cdot,\cdot)$ is a transition probability from $(H_t,\mathcal{H}_t)$ into $(A^{(\ell)},\mathcal{A}^{(\ell)})$. Hence u is nonanticipative, i.e. $u(t,h',\cdot) = u(t,h'',\cdot)$ for all $t \in T$ and $h', h'' \in H$ with $h'_t = h''_t$.

each set $U^{(\ell)}$ is assumed to be closed under exchange of tails (cf. section 4.1). Let ν be an arbitrary starting distribution on $(X,\mathcal{X})$. We assume that $\mathbb{P}_{x,u}$ is measurable in x, and we define $\mathbb{P}_{\nu,u} = \int_X \mathbb{P}_{x,u}\, \nu(dx)$. Furthermore we assume the existence of a probability measure $\mathbb{P}_{h_t,u}$ on H for each $h_t \in H_t$ and $u \in U$, such that $\mathbb{P}_{h_t,u}$ is an $\mathcal{F}_t$-measurable function of h_t, and moreover that $\mathbb{P}_{h_t,u}(h_t \times A' \times \underset{\tau \in T}{\underset{\tau>t}{X}} (X \times A)) = u(t,h_t,A')$ for all $t \in T$, $u \in U$, $A' \in A$. We also suppose that $\mathbb{P}_{h_t,u}$ depends nonanticipatively on u, and that $\mathbb{P}_{h_t,u}$ satisfies the conditions (i) and (ii) in section 4.1.

The utility functions $r^{(\ell)}: H \to \mathbb{R}$ with $\ell \in \mathbb{N}_n$ are supposed to be measurable and quasi integrable w.r.t. $\mathbb{P}_{\nu,u}$ for all $u \in U$ and for ν fixed.

The symbols $E_{\nu,u}$, $E_{h_t,u}$, H, H_t, X_t and A_t are used in the same manner as before. Corresponding to the C/G/G/n process Σ, there exists for each $t \in T$ and $h_t \in H_t$ a C/G/G/n process $\Sigma^{[h_t]}$, that is called the t-delayed process. The description of $\Sigma^{[h_t]}$ is completely analogous to the description of the t-delayed process in section 4.1.

The *value of control u for player ℓ, given h_t* is a function $v_t^{(\ell)}: H_t \times U \to \mathbb{R}$ with

$$v_t^{(\ell)}(h_t,u) = E_{h_t,u}\, r^{(\ell)}(H).$$

Using the notation $(u^*;\ell:u^{(\ell)})$ for a control that is obtained from $u^* \in U$ by replacing the component $u^{*(\ell)}$ by $u^{(\ell)} \in U^{(\ell)}$, we define the *value for player ℓ given h_t and given u^* for the other players* as $\psi_t^{(\ell)}: H_t \times U \to \mathbb{R}$ with

$$\psi_t^{(\ell)}(h_t,u^*) = \begin{cases} \sup\limits_{u^{(\ell)} \in U^{(\ell)}} E_{h_t,(u^*;\ell:u^{(\ell)})}\, r^{(\ell)}(H) & \text{if this integral exists,} \\ -\infty & \text{otherwise} \end{cases}$$

5.4.1. DEFINITION. A control $u^* \in U$ is called *ν-optimal* (or a *ν-equilibrium control*) iff for all $\ell \in \mathbb{N}_n$ and all $t \in T$

$$\psi_t^{(\ell)}(H_t,u^*) = v_t^{(\ell)}(H_t,u^*) \qquad \mathbb{P}_{\nu,u^*} \text{ - a.s.}$$

In the remainder of this section we will restrict ourselves to the two-person zero-sum case: $n = 2$, $r^{(0)} + r^{(1)} = 0$. As in section 5.2 and 5.3 we define $r = r^{(0)}$, $v_t = v_t^{(0)}$, and the *saddle (function)* $w_t: H_t \to \mathbb{R}$

with $w_t(h_t) = \sup_{u^{(0)}} \inf_{u^{(1)}} v_t(h_t,u)$.

As before we suppose, that the expressions occurring in the definitions are well defined.

5.4.2. DEFINITION. A control $u^* \in U$ is called ν-*conserving* iff for all $t_1,t_2 \in T$ with $t_2 \geq t_1$

(i) $\psi_{t_1}^{(0)}(H_{t_1},u^*) = E_{\nu,u^*}^{F_{t_1}} \psi_{t_2}^{(0)}(H_{t_2},u^*)$ $\qquad \mathbb{P}_{\nu,u^*}$ - a.s.,

(ii) $\psi_{t_1}^{(1)}(H_{t_1},u^*) = E_{\nu,u^*}^{F_{t_1}} \psi_{t_2}^{(1)}(H_{t_2},u^*)$ $\qquad \mathbb{P}_{\nu,u^*}$ - a.s.

5.4.3. DEFINITION. A control $u^* \in U$ is called ν-*equalizing* iff

(i) $\lim_{t\to\infty} E_{\nu,u^*}[\psi_t^{(0)}(H_t,u^*) - v_t(H_t,u^*)] = 0$,

(ii) $\lim_{t\to\infty} E_{\nu,u^*}[\psi_t^{(1)}(H_t,u^*) - v_t(H_t,u^*)] = 0$.

5.4.4. THEOREM. A necessary and sufficient condition for ν-optimality of a control $u^* \in U$ is that u^* is ν-conserving and ν-equalizing.

PROOF. Combining the proof of theorem 5.2.6 with the characterization for the C/G/G/1 case (theorem 4.2.3) gives the result. □

As before we use the symbol $u_t^{(\ell)}(h_t)$ to denote the tail of control $u^{(\ell)}$ from time t on, given the history before time t. The symbol ${}_tu^{(\ell)}$ denotes the head of control $u^{(\ell)}$ before time t.
Again we suppose, that the expressions occurring in the definition are well defined.

5.4.5. DEFINITION. A control $u^* \in U$ is called

(i) ν-*subgame perfect* iff for all $u \in U$ and $t \in T$

$$v_t(H_t, u^{(0)}, {}_tu^{(1)}u_t^{*(1)}(H_t)) \le w_t(H_t) \le v_t(H_t, {}_tu^{(0)}u_t^{*(0)}(H_t), u^{(1)})$$

$$\mathbb{P}_{\nu,u} \text{ - a.s.}$$

(ii) ν- *overall saddling* iff for all $u \in \mathsf{U}$ and $t_1, t_2 \in \mathsf{T}$ with $t_2 \ge t_1$

$$E^{F_{t_1}}_{\nu,(u^{(0)}, {}_{t_1}u^{(1)}u^{*(1)}_{t_1}(H_{t_1}))} w_{t_2}(H_{t_2}) \le w_{t_1}(H_{t_1}) \le$$

$$\le E^{F_{t_1}}_{\nu,({}_{t_1}u^{(0)}u^{*(0)}_{t_1}(H_{t_1}), u^{(1)})} w_{t_2}(H_{t_2}) \qquad \mathbb{P}_{\nu,u} \text{ - a.s.,}$$

(iii) ν-*overall asymptotically definite* iff for all $u \in \mathsf{U}$ and $t \in \mathsf{T}$

$$\lim_{\tau\to\infty} E^{F_t}_{\nu,(u^{(0)}, {}_tu^{(1)}u_t^{*(1)}(H_t))} [w_{t+\tau}(H_{t+\tau}) - v_{t+\tau}(H_{t+\tau}, u^{(0)}, {}_tu^{(1)}u_t^{*(1)}(H_t))]^- = 0$$

$$\mathbb{P}_{\nu,u} \text{ - a.s.}$$

and

$$\lim_{\tau\to\infty} E^{F_t}_{\nu,({}_tu^{(0)}u_t^{*(0)}(H_t), u^{(1)})} [w_{t+\tau}(H_{t+\tau}) - v_{t+\tau}(H_{t+\tau}, {}_tu^{(0)}u_t^{*(0)}(H_t), u^{(1)})]^+ = 0$$

$$\mathbb{P}_{\nu,u} \text{ - a.s.}$$

5.4.6. THEOREM. A necessary and sufficient condition for ν-subgame perfectness of a control $u^* \in \mathsf{U}$ is that u^* is ν-overall saddling and ν-overall asymptotically definite.

PROOF. Completely analogous to the proof of theorem 5.2.15. □

For a control $u^* \in \mathsf{U}$ we can define ν-persistent optimality, ν-tail optimality, ν-saddlingness, ν-tail saddlingness and ν-asymptotical definiteness in the same way. But, since we did not formulate the continuous-time analogue of theorem 2.2.1, we cannot give proofs of the corresponding characterizations similar to the proofs of theorems 5.2.18 and 5.3.15. In chapter 6, however, we shall derive a slightly different characterization for a more general situation.

From now on we suppose r to be recursive, i.e. for all $t,\tau \in T$ with $\tau \geq t$ and $h \in H$

$$r^{[0]}(h) = r(h)$$

$$r^{[t]}(h) = \theta_{\tau}^{[t]}(h_{\tau}) + \chi_{\tau}^{[t]}(h\ )\ r^{[t]}(\zeta_{\tau-t}(h))$$

(for details see definition 4.3.1). This leads to the following statement. (As before we use the superscript [t] instead of $[h_t]$ to refer to the t-delayed process).

5.4.7. LEMMA. If r is a recursive utility, then for each $t \in T$, $u \in U$ and starting distribution ν

$$\psi_t^{(\ell)}(h_t,u) = \theta_t^{[0]}(h_t) + \chi_t^{[0]}(h_t)\ \psi_t^{[t](\ell)}(x_t,u_t(h_t)) \quad \text{for } \ell = 0 \text{ or } 1,$$

$$w_t(h_t) = \theta_t^{[0]}(h_t) + \chi_t^{[0]}(h_t)\ w_t^{[t]}(x_t),$$

$$v_t(h_t,u) = \theta_t^{[0]}(h_t) + \chi_t^{[0]}(h_t)\ v_t^{[t]}(x_t,u_t(h_t)).$$

PROOF. Combine the proofs of lemma 4.3.2 and 5.3.2. □

Now ν-conserving, ν-saddling, ν-tail saddling and ν-overall saddling can be characterized (for a recursive utility) in the same manner as is done in theorem 4.3.4.
If moreover r is ν-tail-vanishing (see definition 4.3.5, which must be read as if $u \in U$ is a simultaneous control), then the analogue of theorem 4.3.6 for the functions $\psi_t^{(0)}$ and $\psi_t^{(1)}$ instead of w_t follows immediately.
In the continuous-time case we define the ν-overall tail vanishing property of the utility r as: for all $t \in T$ and $u \in U$

$$\lim_{\tau\to\infty} E_{\nu,u}^{F_t}\ \chi_{\tau}^{[0]}(H_{\tau})v_{\tau}^{[\tau]}(X_{\tau},u_{\tau}(H_{\tau})) = 0 \qquad \mathbb{P}_{\nu,u}\text{ - a.s.}$$

The following result holds.

5.4.8. THEOREM. If r is a ν-overall tail vanishing utility, then a necessary and sufficient condition for u* ∈ U to be ν-overall asymptotically definite, is that

$$\lim_{\tau\to\infty} E^{F_t}_{\nu,(u^{(0)},{}_tu^{(1)}u_t^{*(1)}(H_t))} [w_\tau^{[\tau]}(X_\tau)-v_\tau^{[\tau]}(X_\tau,u_\tau^{(0)}(H_\tau),u_\tau^{*(1)}(H_\tau))]^- = 0 \qquad \mathbb{P}_{\nu,u}\text{ - a.s.}$$

and
$$\lim_{\tau\to\infty} E^{F_t}_{\nu,({}_tu^{(0)}u_t^{*(0)}(H_t),u^{(1)})} [w_\tau^{[\tau]}(X_\tau)-v_\tau^{[\tau]}(X_\tau,u_\tau^{*(0)}(H_\tau),u_\tau^{(1)}(H_\tau))]^+ = 0 \qquad \mathbb{P}_{\nu,u}\text{ - a.s.}$$

for all t ∈ T and u ∈ U.

PROOF. Use the proof of theorem 3.2.5. □

Now the continuous-time analogues of corollaries 5.3.6 and 5.3.9 are selfevident.

CHAPTER 6

THE D/G/G/n PROCESS AND THE C/G/G/n PROCESS

In this chapter we generalize the results of the previous chapter to the more general situation with n players (n may be any cardinal number) and a general (not necessarily zero-sum) utility. However, we still restrict ourselves to a noncooperative situation, and we only characterize Nash equilibria and extensions thereof. For the discrete as well as for the continuous-time case we have already discussed this model in section 5.1 and 5.4 respectively. Since for both cases our way of proving the results is very much alike, this chapter contains only one section, in which the results for both cases are derived simultaneously. Moreover, as already noted in chapter 4, the continuous-time case can be treated in such a way that the discrete-time case is covered by it. One new optimality concept is introduced in this chapter: semi subgame perfectness.

6.1. CHARACTERIZATIONS OF OPTIMALITY IN THE C (AND D)/G/G/n PROCESS

In this section we will denote the D/G/G/n process by the D-case, and the C/G/G/n process by the C-case. In order to avoid proving the same thing twice, for the D-case and for the C-case, we extend the notations and terminology of the C-case to the D-case. This means that from now on a strategy $\pi \in \Pi$ is also called a control $u \in U$. Now we may say in the D and C-case, see definition 5.1.1 and 5.4.1, that a control $u^* \in U$ is ν-optimal iff for all $\ell \in \mathbb{N}_n$ and all $t \in T$ (recall that $\mathbb{N}_n$ is the set of players)

6.1.0.1. $\psi_t^{(\ell)}(H_t,u^*) = v_t^{(\ell)}(H_t,u^*) \qquad \mathbb{P}_{\nu,u^*}$ - a.s.

6.1.1. DEFINITION. A control $u^* \in U$ is called

(i) ν-*conserving* iff for all $\ell \in \mathbb{N}_n$ and $t_1,t_2 \in T$ with $t_2 \geq t_1$

6.1.1.1. $\psi_{t_1}^{(\ell)}(H_{t_1},u^*) = E_{\nu,u^*}^{F_{t_1}} \psi_{t_2}^{(\ell)}(H_{t_2},u^*) \qquad \mathbb{P}_{\nu,u^*}$ - a.s.,

(ii) ν-*equalizing* iff for all $\ell \in \mathbb{N}_n$

6.1.1.2. $\lim_{t\to\infty} E_{\nu,u^*}[\psi_t^{(\ell)}(H_t,u^*) - v_t^{(\ell)}(H_t,u^*)] = 0.$

(As before, the expressions in the definition are supposed to be well defined.)

6.1.2. THEOREM. A necessary and sufficient condition for ν-optimality of a control $u^* \in U$ is that u^* is ν-conserving and ν-equalizing.

PROOF. Fixing ℓ and $u^{*(k)}$ for $k \neq \ell$, $k \in \mathbb{N}_n$, we are in the situation of a D(or C)/G/G/1 process with $\psi_t^{(\ell)}$ as value function.

Hence for this fixed ℓ formula 6.1.0.1 is equivalent to 6.1.1.1 and 6.1.1.2 together (theorem 5.2.6 and 5.4.4). So the theorem is proved. □

6.1.3. DEFINITION. A control $u^* \in U$ is called

(i) ν-*subgame perfect* iff for all $\ell \in \mathbb{N}_n$, $t \in T$ and $u \in U$

$$v_t^{(\ell)}(H_t, {}_t u u_t^*(H_t)) = \psi_t^{(\ell)}(H_t, {}_t u u_t^*(H_t)) \qquad \mathbb{P}_{\nu,u} \text{ - a.s.,}$$

(ii) ν-*overall conserving* iff for all $\ell \in \mathbb{N}_n$, $u \in U$ and $t_1, t_2 \in T$ with $t_2 \geq t_1$

$$\psi_{t_1}^{(\ell)}(H_{t_1}, {}_{t_1} u u_{t_1}^*(H_{t_1})) = E_{\nu,({}_{t_1} u u_{t_1}^*(H_{t_1}))}^{F_{t_1}} \psi_{t_2}^{(\ell)}(H_{t_2}, {}_{t_1} u u_{t_1}^*(H_{t_1}))$$

$$\mathbb{P}_{\nu,u} \text{ - a.s.,}$$

(iii) ν-*overall equalizing* iff for all $\ell \in \mathbb{N}_n$, $t \in T$ and $u \in U$

$$\lim_{\tau\to\infty} E_{\nu,{}_t u u_t^*(H_t)}^{F_t} [\psi_\tau^{(\ell)}(H_\tau, {}_t u u_t^*(H_t)) - v_\tau^{(\ell)}(H_\tau, {}_t u u_t^*(H_t))] = 0$$

$$\mathbb{P}_{\nu,u} \text{ - a.s.}$$

It is not difficult to see that for the two-person zero-sum case the ν-subgame perfectness of definition 6.1.3 is the same as that of definition 5.2.12 (resp. 5.4.5). Something similar can be shown for the other two concepts. We come back to this directly after the next theorem.

6.1.4. THEOREM. A necessary and sufficient condition for ν-subgame perfectness of a control $u^* \in U$ is that u^* is ν-overall conserving and ν-overall equalizing.

PROOF. The proof is essentially the same as the proof of theorem 2.2.4. Suppose u^* is ν-subgame perfect. Then for all $\ell \in \mathbb{N}_n$, $u \in U$ and $t_1, t_2 \in T$ with $t_2 \geq t_1$

$$\psi_{t_1}^{(\ell)}(H_{t_1}, {}_{t_1}uu_{t_1}^*(H_{t_1})) = v_{t_1}^{(\ell)}(H_{t_1}, {}_{t_1}uu_{t_1}^*(H_{t_1})) =$$

$$= E_{\nu, {}_{t_1}uu_{t_1}^*(H_{t_1})}^{F_{t_1}} v_{t_2}^{(\ell)}(H_{t_2}, {}_{t_1}uu_{t_1}^*(H_{t_1})) =$$

$$= E_{\nu, {}_{t_1}uu_{t_1}^*(H_{t_1})}^{F_{t_1}} \psi_{t_2}^{(\ell)}(H_{t_2}, {}_{t_1}uu_{t_1}^*(H_{t_1})) \qquad \mathbb{P}_{\nu,u} - \text{a.s.}$$

This establishes the ν-overall conservingness. And since for all $t, \tau \in T$ with $\tau \geq t$

$$\psi_\tau^{(\ell)}(H_\tau, {}_tuu_t^*(H_t)) - v_\tau^{(\ell)}(H_\tau, {}_tuu_t^*(H_t)) = 0 \qquad \mathbb{P}_{\nu,u} - \text{a.s.}$$

the ν-overall equalizingness follows immediately.

Now suppose u^* is ν-overall conserving and ν-overall equalizing. Then for all $\ell \in \mathbb{N}_n$, $u \in U$ and $t, \tau \in T$ with $\tau \geq t$

$$\psi_t^{(\ell)}(H_t, {}_tuu_t^*(H_t)) = \lim_{\tau\to\infty} E_{\nu, {}_tuu_t^*(H_t)}^{F_t} \psi_\tau^{(\ell)}(H_\tau, {}_tuu_t^*(H_t)) =$$

$$= \lim_{\tau\to\infty} E^{F_t}_{\nu,{}_t uu^*_t(H_t)}\; v^{(\ell)}_\tau(H_\tau, {}_t uu^*_t(H_t)) = v^{(\ell)}_t(H_t, {}_t uu^*_t(H_t)) \quad \mathbb{P}_{\nu,u} - \text{a.s.}$$

This proves the theorem. □

It may be noted that for the two-person zero-sum case the above characterization seems different from the characterization given in theorem 5.2.15 (resp. 5.4.6). Nevertheless they are actually the same, since for the two-person zero-sum case the ν-subgame perfectness implies that for all $t \in T$, $u \in U$ and $\ell \in \{0,1\}$

$$\psi^{(\ell)}_t(H_t, {}_t uu^*_t(H_t)) = w_t(H_t) \qquad \mathbb{P}_{\nu,u} - \text{a.s.}$$

Now we come to another type of optimality: semi subgame perfectness. It can also be found in Couwenbergh (1977) where it is called semi persistency. We prefer this new name, since this concept has to do not so much with persistency, as with a kind of optimality for all subgames of one player, as is shown in the proof of theorem 6.1.6.

6.1.5. DEFINITION. A control $u^* \in U$ is called

(i) ν-*semi subgame perfect* iff for all $\ell \in \mathbb{N}_n$, $t \in T$ and $u \in U$

6.1.5.1. $\psi^{(\ell)}_t(H_t, u^*) = v^{(\ell)}_t(H_t, (u^*; \ell : {}_t u^{(\ell)} u^{*(\ell)}_t(H_t)))$

$$\mathbb{P}_{\nu,(u^*;\ell:u^{(\ell)})} - \text{a.s.},$$

(ii) ν-*strongly conserving* iff for all $\ell \in \mathbb{N}_n$, $u \in U$ and $t_1, t_2 \in T$ with $t_2 \geq t_1$

6.1.5.2. $\psi^{(\ell)}_{t_1}(H_{t_1}, u^*) = E^{F_{t_1}}_{\nu,(u^*;\ell:{}_{t_1}u^{(\ell)}u_{t_1}^{*(\ell)}(H_t))}\; \psi^{(\ell)}_{t_2}(H_{t_2}, u^*)$

$$\mathbb{P}_{\nu,(u^*;\ell:\, u^{(\ell)})} - \text{a.s.}$$

(iii) ν-*strongly equalizing* iff for all $\ell \in \mathbb{N}_n$, $t \in T$ and $u \in U$

$$6.1.5.3.\quad \lim_{\tau\to\infty} E^{F_t}_{\nu,(u^*;\ell:{}_tu^{(\ell)}u_t^{*(\ell)}(H_t))}[\psi_\tau^{(\ell)}(H_\tau,u^*) + \\ - v_\tau^{(\ell)}(H_\tau,(u^*;\ell:{}_tu^{(\ell)}u_t^{*(\ell)}(H_t)))] = 0 \quad \mathbb{P}_{\nu,(u^*;\ell:u^{(\ell)})} - \text{a.s.}$$

6.1.6. THEOREM. A necessary and sufficient condition for ν-semi subgame perfectness of a control $u^* \in U$ is that u^* is ν-strongly conserving and ν-strongly equalizing.

PROOF. Fixing ℓ and $u^{*(k)}$ for $k \neq \ell$, $k \in \mathbb{N}_n$, we are in the situation of a D (or C)/G/G/1 process. The value function of this process equals $\psi_t^{(\ell)}$. This means that formula 6.1.5.1 actually expresses the ν-subgame perfectness of $u^{*(\ell)}$ in this D (or C)/G/G/1 process, formula 6.1.5.2 expresses the ν-overall conserving property of $u^{*(\ell)}$, and formula 6.1.5.3 expresses the ν-overall equalizing property of $u^{*(\ell)}$. Hence, by theorem 6.1.4 it follows that for our fixed ℓ formula 6.1.5.1 is equivalent to 6.1.5.2 and 6.1.5.3 together. This proves the theorem. □

It can be seen immediately that ν-subgame perfectness implies ν-semi subgame perfectness, that ν-semi subgame perfectness implies ν-optimality, and that neither of the reverse implications holds.
For the two-person zero-sum case ν-tail optimality implies ν-semi subgame perfectness, since for instance formula 5.2.16. (ii) implies

$$v_t^{(0)}(H_t,\pi,\rho^*) \le v_t^{(0)}(H_t,{}_t\pi\pi^*(t;H_t),\rho^*) = \psi_t^{(0)}(H_t,\pi,\rho^*) \quad \mathbb{P}_{\nu,(\pi,\rho^*)} - \text{a.s.}$$

That the converse is not true, follows from the following counterexample.

6.1.7. THEOREM. COUNTEREXAMPLE. ν-Semi subgame perfectness is not sufficient for ν-tail optimality.

PROOF. Consider the following D/F/F/2 process.

Γ_0:

	1
1	$0 + \Gamma_0$

Γ_1:

	1
1	$0 + \Gamma_0$
2	$-5 + \Gamma_0$

Γ_2:

	1	2
1	$0 + \Gamma_0$	$5 + \Gamma_0$

Γ_3:

	1	2
1	$0 + \Gamma_0$	$10 + \Gamma_1$
2	$-10 + \Gamma_2$	$0 + \Gamma_0$

$$r(h) = \sum_{t=0}^{\infty} \theta(x_t, a_t^{(0)}, a_t^{(1)}), \quad \mathbb{P}_{\nu,(\pi,\rho)}(X_0=3)=1 \text{ for all } (\pi,\rho) \in \Pi.$$

The strategies for player 0 are only given for state 1 and 3 in this order, and for player 1 only for state 2 and 3 in this order.

Define $\pi^* = \begin{pmatrix} 1 & 2 & 2 & \dots \\ 1 & 1 & \dots & \end{pmatrix} \quad \rho^* = \begin{pmatrix} 1 & 2 & 2 & \dots \\ 1 & 1 & \dots & \end{pmatrix}$

Then it is easy to verify that (π^*,ρ^*) is ν-semi subgame perfect but not ν-tail optimal. □

Only under rather stringent conditions we can prove that ν-semi subgame perfectness implies ν-persistent optimality as defined for the case n=2, see definition 5.3.14. These conditions are: the process is separable and stationary, r is recursive in a stationary way, the (ν-semi subgame perfect) strategy (π^*,ρ^*) is μ-optimal for all degenerate starting distributions μ (so far these conditions without the stationarity assumptions are precisely the conditions under which ν-tail optimality implies ν-persistent optimality for the case n=2), and moreover for all $t \in T$, $u \in U$ and $\ell \in \{0,1\}$

$$\psi_0^{(\ell)}(X_t,u^*) = \psi_0^{(\ell)}(X_t,u_t^*(H_t)) \qquad \mathbb{P}_{\nu,(u^*;\ell:u^{(\ell)})} \text{ - a.s.}$$

(This $\psi_0^{(\ell)}$ plays exactly the same role as the function v in the proof of theorem 2.4.1 in Couwenbergh (1977).) Under these conditions, the assumption that (π^*,ρ^*) is ν-semi subgame perfect, gives for $\ell=0$

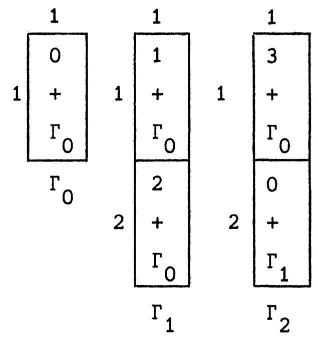
$$v_0^{(0)}(X_t,(u^{(0)},u_t^{*(1)}(H_t)) \le v_0^{(0)}(X_t,u_t^*(H_t)) = \psi_0^{(0)}(X_t,u_t^*(H_t)) =$$

$$= \psi_0^{(0)}(X_t,u^*) = v_0^{(0)}(X_t,u^*) \le v_0^{(0)}(X_t,(u^{*(0)},u^{(1)})) \quad \mathbb{P}_{\nu,(u^{(0)},u^{*(1)})}\text{- a.s.}$$

This together with the analogous result for $\ell=1$ yields the ν-persistent optimality of (π^*,ρ^*).
That the converse is not true can be seen from the following counterexample given in Couwenbergh (1977).

6.1.8. THEOREM. COUNTEREXAMPLE. ν-Persistent optimality does not imply ν-semi subgame perfectness.

PROOF. Consider the following D/F/F/1 process.



$$r(h) = \sum_{t=0}^{\infty} \theta(u_t,a_t^{(0)},a_t^{(1)}), \quad \mathbb{P}_{\nu,(\pi,\rho)}(X_0=2) = 1 \text{ for all } (\pi,\rho) \in \Pi.$$

The strategies are only given for state 1 and 2 in this order.

Defining $(\pi^*,\rho^*) = \begin{pmatrix} 1\ 1\ \ldots\ ,\ 1\ 1\ \ldots \\ 2\ 1\ 1\ \ldots,\ 1\ 1\ \ldots \end{pmatrix}$,

it is easy to verify that (π^*,ρ^*) is ν-persistently optimal but not ν-semi subgame perfect. □

Now we will formulate our results for a recursive utility, but only verbally. (Note that in the two previous counterexamples the utility is recursive and ν-(overall) tail vanishing).
Recursiveness and the ν-(overall) tail vanishing property are defined as

before. This gives immediately that theorem 5.3.3, 5.3.5 and their C-case analogues also hold for $\psi_t^{(\ell)}$ instead of ψ_t, and that theorem 5.3.8 and its C-case analogue also hold for $\psi_t^{(\ell)}$ instead of w_t. But then the analogues of corollary 5.3.6, 5.3.9 and their C-case counterparts are selfevident, even for the ν-semi subgame perfectness.

In this chapter we have not yet discussed the extensions of the persistent optimality and the tail optimality. It turns out that also these concepts can be generalized and characterized. First we will do this for the ν-persistent optimality . The following definition can be read, as though we had presupposed that the process is separable and stationary, and that the utility is recursive in a stationary way(cf. definition 5.3.14 and its foregoing remarks). Yet we prefer a general definition here, corresponding to formulae 5.3.15.1 and 5.3.15.2, since it is more suitable for the derivation of the characterization theorem. Moreover, from the remark directly after the proof of theorem 5.3.15 we know, that formulae 5.3.15.1 and 5.3.15.2 may serve indeed as definition of ν-persistent optimality for the case n=2.

6.1.9. DEFINITION. A control $u^* \in \mathsf{U}$ is called

(i) ν-*persistently optimal* iff for all $k,\ell \in \mathbb{N}_n$, $u \in \mathsf{U}$ and $t_1,t_2 \in \mathsf{T}$ with $t_2 \geq t_1$

$$v_{t_2}^{(\ell)}(H_{t_2},(u^*;k:{}_{t_1}u^{(k)}u^{*(k)})) = \psi_{t_2}^{(\ell)}(H_{t_2},(u^*;k:{}_{t_1}u^{(k)}u^{*(k)})) \qquad \mathbb{P}_{\nu,(u^*;k:u^{(k)})}\text{ - a.s.,}$$

(ii) ν-*persistently conserving* iff for all $k,\ell \in \mathbb{N}_n$, $u \in \mathsf{U}$ and $t_1,t_2 \in \mathsf{T}$ with $t_2 \geq t_1$

$$\psi_{t_1}^{(\ell)}(H_{t_1},(u^*;k:{}_{t_1}u^{(k)}u^{*(k)})) =$$

$$= E_{\nu,(u^*;k:{}_{t_1}u^{(k)}u^{*(k)})}^{F_{t_1}}\ \psi_{t_2}^{(\ell)}(H_{t_2},(u^*;k:{}_{t_1}u^{(k)}u^{*(k)})) \qquad \mathbb{P}_{\nu,(u^*;k:u^{(k)})}\text{ - a.s.}$$

(iii) ν-*persistently equalizing* iff for all $k,\ell \in \mathbb{N}_n$, $t \in \mathsf{T}$ and $u \in \mathsf{U}$

$$\lim_{\tau\to\infty} E^{F_t}_{\nu,(u^*;k:{}_tu^{(k)}u^{*(k)})} [\psi_\tau^{(\ell)}(H_\tau,(u^*;k:{}_tu^{(k)}u^{*(k)})) +$$

$$- v_\tau^{(\ell)}(H_\tau(u^*;k:{}_tu^{(k)}u^{*(k)}))] = 0 \qquad \mathbb{P}_{\nu,(u^*;k:u^{(k)})} - \text{a.s.}$$

6.1.10. THEOREM. A necessary and sufficient condition for ν-persistent optimality of a control $u^* \in U$ is, that u^* is ν-persistently conserving and ν-persistently equalizing.

PROOF. We can repeat the proof of theorem 6.1.4 for this situation. □

We will conclude this chapter by introducing and characterizing ν-tail optimality.

6.1.11. DEFINITION. A control $u^* \in U$ is called

(i) ν-*tail optimal* iff for all $k,\ell \in \mathbb{N}_n$, $u \in U$ and $t_1,t_2 \in T$ with $t_2 \geq t_1$

$$v_{t_2}^{(\ell)}(H_{t_2},(u^*;k:{}_{t_1}u^{(k)}u_{t_1}^{*(k)}(H_{t_1}))) = \psi_{t_2}^{(\ell)}(H_{t_2},(u^*;k:{}_{t_1}u^{(k)}u_{t_1}^{*(k)}(H_{t_1})))$$

$$\mathbb{P}_{\nu,(u^*;k:u^{(k)})} - \text{a.s.}$$

(ii) ν-*tail conserving* iff for all $k,\ell \in \mathbb{N}_n$, $u \in U$ and $t_1,t_2 \in T$ with $t_2 \geq t_1$

$$\psi_{t_1}^{(\ell)}(H_{t_1},(u^*;k:{}_{t_1}u^{(k)}u_{t_1}^{*(k)}(H_{t_1}))) =$$

$$= E^{F_{t_1}}_{\nu,(u^*;k:{}_{t_1}u^{(k)}u_{t_1}^{*(k)}(H_{t_1}))} \psi_{t_2}^{(\ell)}(H_{t_2},(u^*;k:{}_{t_1}u^{(k)}u_{t_1}^{*(k)}(H_{t_1})))$$

$$\mathbb{P}_{\nu,(u^*;k:u^{(k)})} - \text{a.s.}$$

(iii) ν-*tail equalizing* iff for all $k,\ell \in \mathbb{N}_n$, $t \in T$ and $u \in U$

$$\lim_{\tau\to\infty} E^{F_t}_{\nu,(u^*;k:{}_tu^{(k)}u_t^{*(k)}(H_t))} [\psi_\tau^{(\ell)}(H_\tau,(u^*;k:{}_tu^{(k)}u_t^{*(k)}(H_t))) +$$

$$- v_\tau^{(\ell)}(H_\tau,(u^*;k:{}_tu^{(k)}u_t^{*(k)}(H_t)))] = 0 \qquad \mathbb{P}_{\nu,(u^*;k:u^{(k)})}\text{- a.s.}$$

6.1.12. THEOREM. A necessary and sufficient condition for ν-tail optimality of a control $u^* \in U$ is, that u^* is ν-tail conserving and ν-tail equalizing.

PROOF. We can repeat the proof of theorem **6.1.4 for this situation.** □

It follows directly from the definitions, that ν-tail optimality is implied by the concept of ν-subgame perfectness and implies ν-semi subgame perfectness. The remarks, made in section 5.3 (after the proof of corollary 5.3.11 and 5.3.12) about the relation between ν-tail optimality and ν-persistent optimality, also apply here. We omit the special form of the characterization of ν-tail optimality and ν-persistent optimality for a recursive utility, since it can be derived in a straightforward way.

NOTATIONS

a	(simultaneous) action, typical elem. of A
$a^{(\ell)}$	action for player ℓ, typical elem. of $A^{(\ell)}$
a_t	action at time t, typical elem. of A
A	space of (simultaneous) actions
$A^{(\ell)}$	action space for player ℓ
$\mathcal{A}, \mathcal{A}^{(\ell)}$	σ-field on A, $A^{(\ell)}$ resp.
A_t	action at time t (random var.)
$E_{h_t,u}, E_{h_t,\pi}, E_{x,u}, E_{x,\pi}, E_{\nu,u}, E_{\nu,\pi}$	expectation operator w.r.t. $\mathbb{P}_{h_t,u}, \mathbb{P}_{h_t,\pi}$, $\mathbb{P}_{x,u}, \mathbb{P}_{x,\pi}, \mathbb{P}_{\nu,u}, \mathbb{P}_{\nu,\pi}$ resp.
$E_{\nu,u}^{\mathcal{F}_t}, E_{\nu,\pi}^{\mathcal{F}_t}$	conditional expectation w.r.t. $\mathcal{F}_t$ corresponding to $E_{\nu,u}, E_{\nu,\pi}$ resp.
$\mathcal{F}_t$	σ-field on H, generated by H_t
h	history, typical elem. of H
h_t	history up to time t, typical elem. of H_t
$h_t h'_\tau$	concatenation of h_t and h'_τ, such that x_t has disappeared
H	space of histories or sample space
H_t	space of histories up to time t
$H^{[h_t]}$	space of histories in the t-delayed process $\Sigma^{[h_t]}$ $(\tau \geq t)$
$\mathcal{H}, \mathcal{H}_t, \mathcal{H}^{[h_t]}$	σ-field on $H, H_t, H^{[t]}$ resp.
H	history (random var.)
H_t	history up to time t (random var.)
K_t	space of histories before time t+1
$\mathcal{K}_t$	σ-field on K_t
L_t	subset of K_t, determining the admissible actions
$L_t^{(\ell)}$	set that determines the admissible actions for player ℓ

$L_\tau^{[h_t]}$	set that determines the admissible actions in the t-delayed process $\Sigma^{[h_t]}$
$L_{t\,h_t}$	h_t-section of L_t
$\mathbb{N}$	$\{0,1,2,...\}$
$\mathbb{N}_n$	set of players with cardinal number n
p_t	transition probability from K_t to the "next" state space
$\mathbb{P}_{h_t,u}$	probability measure on H, given history h_t and control u
$\mathbb{P}_{h_t,\pi}$	probability measure on H, given history h_t and strategy π
$\mathbb{P}_{x,u}$	probability measure on H, given starting state x and control u
$\mathbb{P}_{x,\pi}$	probability measure on H, given starting state x and strategy π
$\mathbb{P}_{\nu,u}$	probability measure on H for a starting distribution ν and control u
$\mathbb{P}_{\nu,\pi}$	probability measure on H for a starting distribution ν and strategy π
$\mathbb{P}_{h_\tau,u}^{[h_t]}$	probability measure on $H^{[h_t]}$, given h_t and u
r	utility (function) ; $r: H \to \mathbb{R}$
$r^{(\ell)}$	utility for player ℓ; $r^{(\ell)}: H \to \mathbb{R}$
$r^{[h_t]}$	utility in the t-delayed process $\Sigma^{[h_t]}$
$r^{[t]}$	part of the decomposition of a recursive utility r
t	time instant, typical elem. of T
T	time space
$T^{[h_t]}, T^{[t]}$	time space in the t-delayed process $\Sigma^{[h_t]}$
u	(simultaneous) control, typical elem. of U
$u^{(\ell)}$	control for player ℓ, typical elem. of $U^{(\ell)}$
${}_tu$	head of a control u before time t
$u_t(h_t)$	tail of a control u from time t on given history h_t

$(u^*;\ell;u^{(\ell)})$	simultaneous control, obtained from $u^* \in U$ by replacing $u^{*(\ell)}$ by $u^{(\ell)}$
${}_t uu_t^*(H_t)$	(simultaneous) control, before time t equal to u and from time t on equal to u^*.
U	set of (simultaneous) controls
$U^{(\ell)}$	set of controls for player ℓ
$U^{[h_t]}$	set of (simultaneous) controls in the t-delayed process $\Sigma^{[h_t]}$
v_t	value of a strategy (or control); $v_t: H \times \Pi \to \mathbb{R}$ or $v_t: H_t \times U \to \mathbb{R}$
$v_\tau^{[t]}$	value of a strategy (or control) in the t-delayed process $\Sigma^{[t]}$
$v_t^{(\ell)}$	value of a simultaneous strategy (or control) for player ℓ
w_t	value (function), if there is 1 player; saddle (function), if there are 2 players; $w_t: H_t \to \mathbb{R}$
$w_\tau^{[t]}$	value (saddle) function in the t-delayed process $\Sigma^{[t]}$
x	state, typical elem. of X
x_t	state at time t, typical elem. of X
X	state space
$\mathcal{X}$	σ-field on X.
X_t	state at time t (random var.)
Γ_i	game in state i
ζ	shift (discrete time: one step to the right)
ζ^t	shift (discrete time: t steps to the right)
ζ_t	shift (continuous time: a length t to the right)
$\theta_\tau^{[t]}$	part of the decomposition of a recursive utility r
ν	starting distribution
π	1 player: strategy; 2 players 0-sum: strategy for player 0; otherwise: simultaneous strategy

π_t	transition probability from H_t to A_t
$\pi^{(\ell)}$	strategy for player ℓ
${}_t\pi$	head of a strategy π before time t
$\pi(t;h_t)$	tail of a strategy π from time t on, given history h_t
${}_t\pi\pi^*(t;H_t)$	strategy, before time t equal to π, from time t on equal to π^*
(π,ρ)	simultaneous strategy for the 2-person 0-sum case
$(\pi^*;\ell:\pi^{(\ell)})$	simultaneous strategy obtained from π^* by replacing $\pi^{*(\ell)}$ by $\pi^{(\ell)}$
Π	set of (simultaneous) strategies
$\Pi^{(\ell)}$	set of strategies for player ℓ
$\Pi^{[h_\tau]}$	set of strategies in the t-delayed process $\Sigma^{[h_t]}$
ρ	strategy for player 1 in the 2-player 0-sum case
Σ	decision process
$\Sigma^{[h_t]}$	t-delayed process, given h_t, corresponding to Σ
$\Sigma^{[t]}$	t-delayed process, corresponding to Σ (not dependent on h_t)
τ	time instant, typical elem. of T
φ_t	value for player 1 given a strategy for player 0 in the 2-person 0-sum case; $\varphi_t\colon H_t \times \Pi^{(0)} \to \mathbb{R}$ or $\varphi_t\colon H_t \times U^{(0)} \to \mathbb{R}$
$\chi_\tau^{[t]}$	part of the decomposition of a recursive utility r
ψ_t	value for player 0 given a strategy for player 1 in the 2-person 0-sum case; $\psi_t\colon H_t \times \Pi^{(1)} \to \mathbb{R}$ or $\psi_t\colon H_t \times U^{(1)} \to \mathbb{R}$
$\psi_t^{(\ell)}$	value for player ℓ given a strategy for the other players; $\psi_t^{(\ell)}\colon H_t \times \Pi \to \mathbb{R}$ or $\psi_t^{(\ell)}\colon H_t \times U \to \mathbb{R}$

INDEX

REFERENCES

AUMANN, R.J., *Mixed and behaviour strategies in infinite extensive games,* Advances in game theory; ed. by M. Dresher, L.S. Shapley and A.W. Tucker. Princeton: Princeton University Press, 1964. 627-650. (Ann. Math. Studies; 52).

BATHER, J. & H. CHERNOFF, *Sequential decisions in the control of a space-ship* (finite fuel), J. Appl. Probability 4, (1967), 584-604.

BELLMAN, R., *Dynamic programming,* Princeton: Princeton University Press, 1957.

BENEŠ, V., *Existence of optimal stochastic control laws,* SIAM J. Control 9 (1971), 446-475.

BERKOVITZ, L.D., *Optimal control theory,* New York etc.: Springer, 1974 (Appl. Math. Sci.; 12).

BLACKWELL, D., *Discounted dynamic programming,* Ann. Math. Statist. 36, (1965), 226-235.

BLACKWELL, D., *On stationary policies,* J. Roy. Statist. Soc., Ser. A. 133 (1970), 33-38.

BLACKWELL, D., D. FREEDMAN and M. ORKIN, *The optimal reward operator in dynamic programming,* Ann. Probability 2 (1974), 926-941.

BOEL, R. & P. VARAIYA, *Optimal control of jump processes,* SIAM J. Control Optimization 15 (1977), 92-119.

BOEL, R., P. VARAIYA and E. WONG, *Martingales on jump processes II: Applications,* SIAM J. Control 13 (1975), 1021-1061.

COUWENBERGH, H.A.M., *Characterization of strong (Nash) equilibrium points in Markov games,* Eindhoven: Eindhoven University of Technology, Dept. of Math., 1977 (Memorandum COSOR; 77-09).

DAVIS, M.H.A. & P. VARAIYA, *Dynamic programming conditions for partially observable stochastic systems,* SIAM J. Control 11 (1973), 226-261.

DOSHI, B.T., *Continuous time control of Markov processes on an arbitrary state space: discounted rewards,* Ann. Statist. 4 (1976), 1219-1235.

DUBINS, L.E. & L.J. SAVAGE, *How to gamble if you must: inequalities for stochastic processes,* New York: Mc.Graw-Hill, 1965.

DYNKIN, E.B., *Markov processes I,* Berlin etc.: Springer, 1965.

FURUKAWA, N. & S. IWAMOTO, *Markovian decision processes with recursive reward function,* Bull. Math. Statist. 15 (1973), 79-91.

GAVISH, B. & P.J. SCHWEITZER, *An optimality principle for Markovian decision processes,* J. Math. Anal. Appl. 54, (1976), 173-184.

GIRSANOV, I.V., *On transforming a certain class of stochastic process by absolutely continuous substitution of measures,* Theor. Probability Appl. 5 (1960), 285-301.

GROENEWEGEN, L.P.J., *Convergence results related to the equalizing property in a Markov decision process,* Eindhoven: Eindhoven University of Technology, Dept. of Math., 1975 (Memorandum COSOR; 75-18).

GROENEWEGEN, L.P.J., *Markov games: properties of and conditions for optimal strategies,* Eindhoven: Eindhoven University of Technology, Dept. of Math., 1976 (Memorandum COSOR; 76-24).

GROENEWEGEN, L.P.J. & K.M. VAN HEE, *Markov decision processes and quasi-martingales,* Recent developments in statistics; ed. by J.R. Barra et al. Amsterdam: North-Holland Publ. Comp., 1977. 453-459.

GROENEWEGEN, L.P.J. & J. WESSELS, *On the relation between optimality and saddle-conservation in Markov games,* Dynamische Optimierung; Tagungsband des Sonderforschungsbereiches 72. Bonn: Math. Inst. Universität Bonn, 1977, (Bonner Mathematische Schriften; 98). 19-31.

HEE, K.M. VAN, A. HORDIJK & J. VAN DER WAL, *Successive approximations for convergent dynamic programming,* Markov decision theory; ed. by H.C. Tijms and J. Wessels, Amsterdam: Mathematisch Centrum, 1977 (Math. Centre Tracts; 93), 183-211.

HINDERER, K., *Foundations of non-stationary dynamic programming with discrete time parameter,* Berlin etc.: Springer, 1970 (Lecture notes in Operations Res. and Math. Econ.; 33).

HINDERER, K., *Instationäre dynamische Optimierung bei schwachen Voraussetzungen über die Gewinnfunktionen,* Abh. Math. Sem. Univ. Hamburg 36 (1971), 208-223.

HORDIJK, A., *Dynamic programming and Markov potential theory*, Amsterdam: Mathematisch Centrum, 1974 (Math. Centre Tracts; 51).

KERTZ, R.P. & D.C. NACHMAN, *Optimal plans for discrete-time non-stationary dynamic programming with general total reward function I: the topology of weak convergence case*, Atlanta: Georgia Institute of Technology, College of industrial management, 1977 (Tech. rep. MS-77-1).

KREPS, D.M., *Decision problems with expected utility criteria, 1: upper and lower convergent utility*, Math. Operations Res. 2 (1977), 45-53.

LEVE, G. DE, A. FEDERGRUEN & H.C. TIJMS, *A general Markov decision method, I: model and techniques*, Adv. Appl. Prob. 9 (1977), 296-315.

MANDL, P., *Estimation and control in Markov chains*, Advances in Appl. Probability 6 (1974), 40-60.

MEYER, P.A., *Probabilitê et potentiel*, Paris: Hermann, 1966.

NEVEU, J., *Mathematical foundations of the calculus of probability*, San Francisco: Holden-Day, 1965.

NUNEN, J.A.E.E. VAN, *Contracting Markov decision processes*, Amsterdam: Mathematisch Centrum, 1976 (Math. Centre Tracts; 71).

PARTHASARATHY, T. & M. STERN, *Markov games - a survey*, Chicago: University of Illinois at Chicago Circle, 1976 (Tech. rep.).

RIEDER, U., *On optimal policies and martingales in dynamic programming*, J. Appl. Probability 13 (1976), 507-518.

RISHEL, R., *Necessary and sufficient dynamic programming conditions for continuous time stochastic optimal control*, SIAM J. Control 8 (1970), 559-571.

SCHÄL, M., *On dynamic programming: compactness of the space of policies*, Stochastic Process Appl. 3 (1975), 345-364.

SELTEN, R., *Spieltheoretischen Behandlung eines Oligopolmodells mit Nachfrageträgheit*, Z. gesammte Staatswissenschaft 121 (1965), 301-324; 667-684.

SELTEN, R., *Reexamination of the perfectness concept for equilibrium points in extensive games*, Internat. J. Game Theory 4 (1975), 25-55.

SHAPLEY, L.S., *Stochastic games,* Proc. Nat. Acad. Sci. U.S.A. 39 (1953), 1095-1100.

SHREVE, S.E., *Dynamic programming in complete separable spaces,* Urbana: University of Illinois, 1977 (Tech. resp. R-755; UILU-ENG 77-2202).

STRAUCH, R., *Negative dynamic programming,* Ann. Math. Statist. 37 (1966), 871-889.

STRIEBEL, C., *Optimal control of discrete time stochastic systems,* Berlin etc.: Springer, 1975 (Lecture Notes in Econ. and Math. Systems; 110).

SUDDERTH, W.D., *On the Dubins and Savage characterization of optimal strategies,* Ann. Math. Statist. 43 (1972), 498-507.

WAGNER, D.H., *Survey of measurable selection theorems,* SIAM J. Control Optimization 15 (1977), 859-903.

WAL, J. VAN DER & J. WESSELS, *Successive approximation methods for Markov games,* Markov decision theory; ed. by H.C. Tijms and J. Wessels, Amsterdam: Mathematisch Centrum, 1977 (Math. Centre Tracts; 93) 39-55.

TITLES IN THE SERIES MATHEMATICAL CENTRE TRACTS

(An asterisk before the MCT number indicates that the tract is under preparation).

A leaflet containing an order form and abstracts of all publications mentioned below is available at the Mathematisch Centrum, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands. Orders should be sent to the same address.

MCT 1 T. VAN DER WALT, *Fixed and almost fixed points*, 1963. ISBN 90 6196 002 9.

MCT 2 A.R. BLOEMENA, *Sampling from a graph*, 1964. ISBN 90 6196 003 7.

MCT 3 G. DE LEVE, *Generalized Markovian decision processes, part I: Model and method*, 1964. ISBN 90 6196 004 5.

MCT 4 G. DE LEVE, *Generalized Markovian decision processes, part II: Probabilistic background*, 1964. ISBN 90 6196 005 3.

MCT 5 G. DE LEVE, H.C. TIJMS & P.J. WEEDA, *Generalized Markovian decision processes, Applications*, 1970. ISBN 90 6196 051 7.

MCT 6 M.A. MAURICE, *Compact ordered spaces*, 1964. ISBN 90 6196 006 1.

MCT 7 W.R. VAN ZWET, *Convex transformations of random variables*, 1964. ISBN 90 6196 007 X.

MCT 8 J.A. ZONNEVELD, *Automatic numerical integration*, 1964. ISBN 90 6196 008 8.

MCT 9 P.C. BAAYEN, *Universal morphisms*, 1964. ISBN 90 6196 009 6.

MCT 10 E.M. DE JAGER, *Applications of distrubutions in mathematical physics*, 1964. ISBN 90 6196 010 X.

MCT 11 A.B. PAALMAN-DE MIRANDA, *Topological semigroups*, 1964. ISBN 90 6196 011 8.

MCT 12 J.A.Th.M. VAN BERCKEL, H. BRANDT CORSTIUS, R.J. MOKKEN & A. VAN WIJNGAARDEN, *Formal properties of newspaper Dutch*, 1965. ISBN 90 6196 013 4.

MCT 13 H.A. LAUWERIER, *Asymptotic expansions*, 1966, out of print; replaced by MCT 54.

MCT 14 H.A. LAUWERIER, *Calculus of variations in mathematical physics*, 1966. ISBN 90 6196 020 7.

MCT 15 R. DOORNBOS, *Slippage tests*, 1966. ISBN 90 6196 021 5.

MCT 16 J.W. DE BAKKER, *Formal definition of programming languages with an application to the definition of ALGOL 60*, 1967. ISBN 90 6196 022 3.

MCT 17 R.P. VAN DE RIET, *Formula manipulation in ALGOL 60, part 1*, 1968. ISBN 90 6196 025 8.

MCT 18 R.P. VAN DE RIET, *Formula manipulation in ALGOL 60, part 2*, 1968. ISBN 90 6196 038 X.

MCT 19 J. VAN DER SLOT, *Some properties related to compactness*, 1968. ISBN 90 6196 026 6.

MCT 20 P.J. VAN DER HOUWEN, *Finite difference methods for solving partial differential equations*, 1968. ISBN 90 6196 027 4.

MCT 21 E. WATTEL, *The compactness operator in set theory and topology*, 1968. ISBN 90 6196 028 2.

MCT 22 T.J. DEKKER, *ALGOL 60 procedures in numerical algebra, part 1*, 1968. ISBN 90 6196 029 0.

MCT 23 T.J. DEKKER & W. HOFFMANN, *ALGOL 60 procedures in numerical algebra, part 2*, 1968. ISBN 90 6196 030 4.

MCT 24 J.W. DE BAKKER, *Recursive procedures*, 1971. ISBN 90 6196 060 6.

MCT 25 E.R. PAËRL, *Representations of the Lorentz group and projective geometry*, 1969. ISBN 90 6196 039 8.

MCT 26 EUROPEAN MEETING 1968, *Selected statistical papers, part I*, 1968. ISBN 90 6196 031 2.

MCT 27 EUROPEAN MEETING 1968, *Selected statistical papers, part II*, 1969. ISBN 90 6196 040 1.

MCT 28 J. OOSTERHOFF, *Combination of one-sided statistical tests*, 1969. ISBN 90 6196 041 X.

MCT 29 J. VERHOEFF, *Error detecting decimal codes*, 1969. ISBN 90 6196 042 8.

MCT 30 H. BRANDT CORSTIUS, *Exercises in computational linguistics*, 1970. ISBN 90 6196 052 5.

MCT 31 W. MOLENAAR, *Approximations to the Poisson, binomial and hypergeometric distribution functions*, 1970. ISBN 90 6196 053 3.

MCT 32 L. DE HAAN, *On regular variation and its application to the weak convergence of sample extremes*, 1970. ISBN 90 6196 054 1.

MCT 33 F.W. STEUTEL, *Preservation of infinite divisibility under mixing and related topics*, 1970. ISBN 90 6196 061 4.

MCT 34 I. JUHÁSZ, A. VERBEEK & N.S. KROONENBERG, *Cardinal functions in topology*, 1971. ISBN 90 6196 062 2.

MCT 35 M.H. VAN EMDEN, *An analysis of complexity*, 1971. ISBN 90 6196 063 0.

MCT 36 J. GRASMAN, *On the birth of boundary layers*, 1971. ISBN 90 6196 064 9.

MCT 37 J.W. DE BAKKER, G.A. BLAAUW, A.J.W. DUIJVESTIJN, E.W. DIJKSTRA, P.J. VAN DER HOUWEN, G.A.M. KAMSTEEG-KEMPER, F.E.J. KRUSEMAN ARETZ, W.L. VAN DER POEL, J.P. SCHAAP-KRUSEMAN, M.V. WILKES & G. ZOUTENDIJK, *MC-25 Informatica Symposium* 1971. ISBN 90 6196 065 7.

MCT 38 W.A. VERLOREN VAN THEMAAT, *Automatic analysis of Dutch compound words*, 1971. ISBN 90 6196 073 8.

MCT 39 H. BAVINCK, *Jacobi series and approximation*, 1972. ISBN 90 6196 074 6.

MCT 40 H.C. TIJMS, *Analysis of (s,S) inventory models*, 1972. ISBN 90 6196 075 4.

MCT 41 A. VERBEEK, *Superextensions of topological spaces*, 1972. ISBN 90 6196 076 2.

MCT 42 W. VERVAAT, *Success epochs in Bernoulli trials (with applications in number theory)*, 1972. ISBN 90 6196 077 0.

MCT 43 F.H. RUYMGAART, *Asymptotic theory of rank tests for independence*, 1973. ISBN 90 6196 081 9.

MCT 44 H. BART, *Meromorphic operator valued functions*, 1973. ISBN 90 6196 082 7.

MCT 45 A.A. BALKEMA, *Monotone transformations and limit laws* 1973. ISBN 90 6196 083 5.

MCT 46 R.P. VAN DE RIET, *ABC ALGOL, A portable language for formula manipulation systems, part 1: The language*, 1973. ISBN 90 6196 084 3.

MCT 47 R.P. VAN DE RIET, *ABC ALGOL, A portable language for formula manipulation systems, part 2: The compiler*, 1973. ISBN 90 6196 085 1.

MCT 48 F.E.J. KRUSEMAN ARETZ, P.J.W. TEN HAGEN & H.L. OUDSHOORN, *An ALGOL 60 compiler in ALGOL 60, Text of the MC-compiler for the EL-X8*, 1973. ISBN 90 6196 086 X.

MCT 49 H. KOK, *Connected orderable spaces*, 1974. ISBN 90 6196 088 6.

MCT 50 A. VAN WIJNGAARDEN, B.J. MAILLOUX, J.E.L. PECK, C.H.A. KOSTER, M. SINTZOFF, C.H. LINDSEY, L.G.L.T. MEERTENS & R.G. FISKER (eds), *Revised report on the algorithmic language ALGOL 68*, 1976. ISBN 90 6196 089 4.

MCT 51 A. HORDIJK, *Dynamic programming and Markov potential theory*, 1974. ISBN 90 6196 095 9.

MCT 52 P.C. BAAYEN (ed.), *Topological structures*, 1974. ISBN 90 6196 096 7.

MCT 53 M.J. FABER, *Metrizability in generalized ordered spaces*, 1974. ISBN 90 6196 097 5.

MCT 54 H.A. LAUWERIER, *Asymptotic analysis, part 1*, 1974. ISBN 90 6196 098 3.

MCT 55 M. HALL JR. & J.H. VAN LINT (eds), *Combinatorics, part 1: Theory of designs, finite geometry and coding theory*, 1974. ISBN 90 6196 099 1.

MCT 56 M. HALL JR. & J.H. VAN LINT (eds), *Combinatorics, part 2: Graph theory, foundations, partitions and combinatorial geometry*, 1974. ISBN 90 6196 100 9.

MCT 57 M. HALL JR. & J.H. VAN LINT (eds), *Combinatorics, part 3: Combinatorial group theory*, 1974. ISBN 90 6196 101 7.

MCT 58 W. ALBERS, *Asymptotic expansions and the deficiency concept in statistics*, 1975. ISBN 90 6196 102 5.

MCT 59 J.L. MIJNHEER, *Sample path properties of stable processes*, 1975. ISBN 90 6196 107 6.

MCT 60 F. GÖBEL, *Queueing models involving buffers*, 1975. ISBN 90 6196 108 4.

*MCT 61 P. VAN EMDE BOAS, *Abstract resource-bound classes, part 1*, ISBN 90 6196 109 2.

*MCT 62 P. VAN EMDE BOAS, *Abstract resource-bound classes, part 2*, ISBN 90 6196 110 6.

MCT 63 J.W. DE BAKKER (ed.), *Foundations of computer science*, 1975. ISBN 90 6196 111 4.

MCT 64 W.J. DE SCHIPPER, *Symmetric closed categories*, 1975. ISBN 90 6196 112 2.

MCT 65 J. DE VRIES, *Topological transformation groups 1 A categorical approach*, 1975. ISBN 90 6196 113 0.

MCT 66 H.G.J. PIJLS, *Locally convex algebras in spectral theory and eigenfunction expansions*, 1976. ISBN 90 6196 114 9.

*MCT 67 H.A. LAUWERIER, *Asymptotic analysis, part 2*, ISBN 90 6196 119 X.

MCT 68 P.P.N. DE GROEN, *Singularly perturbed differential operators of second order*, 1976. ISBN 90 6196 120 3.

MCT 69 J.K. LENSTRA, *Sequencing by enumerative methods*, 1977. ISBN 90 6196 125 4.

MCT 70 W.P. DE ROEVER JR., *Recursive program schemes: Semantics and proof theory*, 1976. ISBN 90 6196 127 0.

MCT 71 J.A.E.E. VAN NUNEN, *Contracting Markov decision processes*, 1976. ISBN 90 6196 129 7.

MCT 72 J.K.M. JANSEN, *Simple periodic and nonperiodic Lamé functions and their applications in the theory of conical waveguides*, 1977. ISBN 90 6196 130 0.

MCT 73 D.M.R. LEIVANT, *Absoluteness of intuitionistic logic*, 1979. ISBN 90 6196 122 X.

MCT 74 H.J.J. TE RIELE, *A theoretical and computational study of generalized aliquot sequences*, 1976. ISBN 90 6196 131 9.

MCT 75 A.E. BROUWER, *Treelike spaces and related connected topological spaces*, 1977. ISBN 90 6196 132 7.

MCT 76 M. REM, *Associations and the closure statement*, 1976. ISBN 90 6196 135 1.

MCT 77 W.C.M. KALLENBERG, *Asymptotic optimality of likelihood ratio tests in exponential families*, 1977. ISBN 90 6196 134 3.

MCT 78 E. DE JONGE & A.C.M. VAN ROOIJ, *Introduction to Riesz spaces*, 1977. ISBN 90 6196 133 5.

MCT 79 M.C.A. VAN ZUIJLEN, *Empirical distributions and rank statistics*, 1977. ISBN 90 6196 145 9.

MCT 80 P.W. HEMKER, *A numerical study of stiff two-point boundary problems*, 1977. ISBN 90 6196 146 7.

MCT 81 K.R. APT & J.W. DE BAKKER (eds), *Foundations of computer science II*, part 1, 1976. ISBN 90 6196 140 8.

MCT 82 K.R. APT & J.W. DE BAKKER (eds), *Foundations of computer science II*, part 2, 1976. ISBN 90 6196 141 6.

MCT 83 L.S. BENTHEM JUTTING, *Checking Landau's "Grundlagen" in the AUTOMATH system*, 1979. ISBN 90 6196 147 5.

MCT 84 H.L.L. BUSARD, *The translation of the elements of Euclid from the Arabic into Latin by Hermann of Carinthia (?) books vii-xii*, 1977. ISBN 90 6196 148 3.

MCT 85 J. VAN MILL, *Supercompactness and Wallman spaces*, 1977. ISBN 90 6196 151 3.

MCT 86 S.G. VAN DER MEULEN & M. VELDHORST, *Torrix I, A programming system for operations on vectors and matrices over arbitrary fields and of variable size*. 1978. ISBN 90 6196 152 1.

*MCT 87 S.G. VAN DER MEULEN & M. VELDHORST, *Torrix II*, ISBN 90 6196 153 X.

MCT 88 A. SCHRIJVER, *Matroids and linking systems*, 1977. ISBN 90 6196 154 8.

MCT 89 J.W. DE ROEVER, *Complex Fourier transformation and analytic functionals with unbounded carriers*, 1978. ISBN 90 6196 155 6.

*MCT 90 L.P.J. GROENEWEGEN, *Characterization of optimal strategies in dynamic games*, . ISBN 90 6196 156 4.

MCT 91 J.M. GEYSEL, *Transcendence in fields of positive characteristic*, 1979. ISBN 90 6196 157 2.

MCT 92 P.J. WEEDA, *Finite generalized Markov programming*, 1979. ISBN 90 6196 158 0.

MCT 93 H.C. TIJMS & J. WESSELS (eds), *Markov decision theory*, 1977. ISBN 90 6196 160 2.

MCT 94 A. BIJLSMA, *Simultaneous approximations in transcendental number theory*, 1978. ISBN 90 6196 162 9.

MCT 95 K.M. VAN HEE, *Bayesian control of Markov chains*, 1978. ISBN 90 6196 163 7.

MCT 96 P.M.B. VITÁNYI, *Lindenmayer systems: Structure, languages, and growth functions*, 1980. ISBN 90 6196 164 5.

*MCT 97 A. FEDERGRUEN, *Markovian control problems; functional equations and algorithms*, . ISBN 90 6196 165 3.

MCT 98 R. GEEL, *Singular perturbations of hyperbolic type*, 1978. ISBN 90 6196 166 1.

MCT 99 J.K. LENSTRA, A.H.G. RINNOOY KAN & P. VAN EMDE BOAS, *Interfaces between computer science and operations research*, 1978. ISBN 90 6196 170 X.

MCT 100 P.C. BAAYEN, D. VAN DULST & J. OOSTERHOFF (eds), *Proceedings bicentennial congress of the Wiskundig Genootschap, part 1*, 1979. ISBN 90 6196 168 8.

MCT 101 P.C. BAAYEN, D. VAN DULST & J. OOSTERHOFF (eds), *Proceedings bicentennial congress of the Wiskundig Genootschap, part 2*, 1979. ISBN 90 6196 169 6.

MCT 102 D. VAN DULST, *Reflexive and superreflexive Banach spaces*, 1978. ISBN 90 6196 171 8.

MCT 103 K. VAN HARN, *Classifying infinitely divisible distributions by functional equations*, 1978. ISBN 90 6196 172 6.

MCT 104 J.M. VAN WOUWE, *Go-spaces and generalizations of metrizability*, 1979. ISBN 90 6196 173 4.

*MCT 105 R. HELMERS, *Edgeworth expansions for linear combinations of order statistics*, . ISBN 90 6196 174 2.

MCT 106 A. SCHRIJVER (ed.), *Packing and covering in combinatorics*, 1979. ISBN 90 6196 180 7.

MCT 107 C. DEN HEIJER, *The numerical solution of nonlinear operator equations by imbedding methods*, 1979. ISBN 90 6196 175 0.

MCT 108 J.W. DE BAKKER & J. VAN LEEUWEN (eds), *Foundations of computer science III*, part 1, 1979. ISBN 90 6196 176 9.

MCT 109 J.W. DE BAKKER & J. VAN LEEUWEN (eds), *Foundations of computer science III*, part 2, 1979. ISBN 90 6196 177 7.

MCT 110 J.C. VAN VLIET, *ALGOL 68 transput*, part I: *Historical review and discussion of the implementation model*, 1979. ISBN 90 6196 178 5.

MCT 111 J.C. VAN VLIET, *ALGOL 68 transput*, part II: *An implementation model*, 1979. ISBN 90 6196 179 3.

MCT 112 H.C.P. BERBEE, *Random walks with stationary increments and renewal theory*, 1979. ISBN 90 6196 182 3.

MCT 113 T.A.B. SNIJDERS, *Asymptotic optimality theory for testing problems with restricted alternatives*, 1979. ISBN 90 6196 183 1.

MCT 114 A.J.E.M. JANSSEN, *Application of the Wigner distribution to harmonic analysis of generalized stochastic processes*, 1979. ISBN 90 6196 184 X.

MCT 115 P.C. BAAYEN & J. VAN MILL (eds), *Topological Structures II*, part 1, 1979. ISBN 90 6196 185 5.

MCT 116 P.C. BAAYEN & J. VAN MILL (eds), *Topological Structures II*, part 2, 1979. ISBN 90 6196 186 6.

MCT 117 P.J.M. KALLENBERG, *Branching processes with continuous state space*, 1979. ISBN 90 6196 188 2.

MCT 118 P. GROENEBOOM, *Large deviations and asymptotic efficiencies*, 1980. ISBN 90 6196 190 4.

MCT 119 F.J. PETERS, *Sparse matrices and substructures, with a novel implementation of finite element algorithms*, 1980. ISBN 90 6196 192 0.

MCT 120 W.P.M. DE RUYTER, *On the asymptotic analysis of large-scale ocean circulation*, 1980. ISBN 90 6196 192 9.

MCT 121 W.H. HAEMERS, *Eigenvalue techniques in design and graph theory*, 1980. ISBN 90 6196 194 7.

MCT 122 J.C.P. BUS, *Numerical solution of systems of nonlinear equations*, 1980. ISBN 90 6196 195 5.

MCT 123 I. YUHÁSZ, *Cardinal functions in topology - ten years later*, 1980. ISBN 90 6196 196 3.

MCT 124 R.D. GILL, *Censoring and stochastic integrals*, 1980. ISBN 90 6196 197 1.

MCT 125 R. EISING, *2-D systems, an algebraic approach*, 1980. ISBN 90 6196 198 X.

MCT 126 G. VAN DER HOEK, *Reduction methods in nonlinear programming*, 1980. ISBN 90 6196 199 8.

MCT 127 J.W. KLOP, *Combinatory reduction systems*, 1980. ISBN 90 6196 200 5.

MCT 128 A.J.J. TALMAN, *Variable dimension fixed point algorithms and triangulations*, 1980. ISBN 90 6196 201 3.

MCT 129 G. VAN DER LAAN, *Simplicial fixed point algorithms*, 1980. ISBN 90 6196 202 1.

MCT 130 P.J.W. TEN HAGEN et al., *ILP Intermediate language for pictures*, 1980. ISBN 90 6196 204 8.

MCT 131 R.J.R. BACK, *Correctness preserving program refinements: Proof theory and applications*, 1980. ISBN 90 6196 207 2.

MCT 132 H.M. MULDER, *The interval function of a graph*, 1980. ISBN 90 6196 208 0.

MCT 133 C.A.J. KLAASSEN, *Statistical performance of location estimators*, 1981. ISBN 90 6196 209 9.

MCT 134 J.C. VAN VLIET & H. WUPPER (eds), *Proceedings international conference on ALGOL 68*, 1981. ISBN 90 6196 210 2.

MCT 135 J.A.G. GROENENDIJK, T.M.V. JANSSEN & M.J.B. STOKHOF (eds), *Formal methods in the study of language*, part I, 1981. ISBN 90 6196 211 0.

MCT 136 J.A.G. GROENENDIJK, T.M.V. JANSSEN & M.J.B. STOKHOF (eds), *Formal methods in the study of language*, part II, 1981. ISBN 90 6196 213 7.