CWI Tract

14

Analytic spaces and dynamic
programming:
a measure-theoretic approach

J.G.F. Thiemann

ACKNOWLEDGEMENTS

CONTENTS

INTRODUCTION

Analytic topological spaces are used in dynamic programming in order to avoid certain measurability problems. In this monograph we give a measure-theoretic alternative for these spaces serving the same purpose. Before giving an overview of the contents, we briefly describe the measurability problems encountered in dynamic programming and indicate how they have been solved.

Consider a system that passes through a sequence of states in the course of time and suppose that a controller can influence each of the transitions of the system to a new state by taking certain actions. At each transition the new state of the system depends on the old state and on the action chosen by the controller in a stochastic way; it is the probability distribution of the new state that is determined by the old state and the action, rather than the new state itself. Also, for each realization of this process, i.e. for each sequence of states and actions, a utility is defined, that is, a number representing the desirability of the realization. Now the controller tries to choose his actions such as to maximize the expected utility. In general, this implies that the action to be chosen at each instant of time depends on the state the system is in at that time. The sequence of these choice functions, one for each transition of the system, is called a strategy.

When the state space, i.e. the set of states the system can be in, and the action space, i.e. the set of actions available to the controller, are finite or countably infinite, and when only finitely many transitions of the system are considered, no difficulties arise in defining the expectation of the utility. However, when these spaces are uncountable, this is no longer the case. For the expected utility to be definable it seems necessary that the state space and the action space are measurable spaces and that the choice functions in a strategy as well as the utility are measurable functions. In the construction of strategies yielding maximal expected utility one meets measurability problems that can be described in their simplest form as follows: If f is a measurable function of two variables, then, in general, $\sup_y f(x,y)$ is not a measurable function of x. Also, when the

supremum is attained for each x, there does not necessarily exist a measurable function φ such that $\sup_y f(x,y) = f(x,\varphi(x))$.

Blackwell was among the first who paid attention to these problems. He took Borel spaces as state and action spaces, a measurable utility and measurable strategies. Later on, in a paper by Blackwell, Freedman, and Orkin, this formalism was generalized: analytic state and action spaces, a semianalytic utility and analytically measurable strategies. Another generalization is due to Shreve, who took Borel spaces again, a semianalytic utility and universally measurable strategies (for the papers in question see the references). In the last two formalisms the measurability problems mentioned earlier do no longer appear.

The use of Borel spaces and analytic spaces is unsatisfactory in so far as topological conditions are imposed on the system in order to avoid difficulties that are measure-theoretic by nature. Moreover, the theory of analytic spaces is far from trivial, and quite remote from the things one expects when turning to dynamic programming. In this monograph we shall develop, within a purely measure-theoretic framework, those parts of the theory of analytic spaces that are material for dynamic programming, and we shall show how they can be applied. The formalism for dynamic programming treated in the following is quite general: the utility need not be the sum of single-step utilities but may be an arbitrary function of the realization, while the restrictions imposed on the choice of the actions do not concern the actions themselves but rather the probabilities on the action space on which the choice is based.

The monograph is divided into three chapters. In Chapter I the measure-theoretic prerequisites are introduced, the main topics being universal measurability and the Souslin operation. This chapter is self-contained: apart from the Radon-Nikodym theorem and the martingale convergence theorem only elementary measure theory is required. In the second chapter analytic measurable spaces are introduced. The defining property of these spaces is common to all analytic topological spaces, and our measure-theoretic approach is therefore a generalization of the topological one. Also for this chapter no a priori knowledge is needed, and, consequently, it may serve as an introduction to the subject. However, only those topics are treated that are needed for the applications in Chapter III, or that serve a good conception. The final chapter is devoted to dynamic programming. Although familiarity with this subject is not needed for the understanding of this chapter, it will certainly add to its appreciation. So,

this part of the monograph should not be taken as introductory. The topics treated have been chosen so as to give the reader a good impression of the use of analytic measurable spaces. Consequently, results that are based on, say, a particular structure of the utility or on the choice of particular strategies are not considered.

Although there are some new results in this monograph and the purely measure-theoretic approach as such may be considered as new, on many occasions our line of reasoning was inspired by arguments found in the literature; our main sources were the books of Bertsekas & Shreve, Christensen, Hinderer, and Hoffmann-Jørgensen (see references).

CHAPTER I
MEASURE-THEORETIC PREREQUISITES

The measure-theoretic prerequisites needed for the understanding of
the theory of analytic spaces and their applications are collected in this
chapter. In its first section only rather well-known facts are recalled
(often with an indication of a proof) and some notations are introduced.
In the second section sets of probabilities are equipped with a structure
such as to make the theory of measurable spaces apply to them. The, perhaps
less familiar, subjects of universal measurability and Souslin sets are
treated from scratch in the sections 3 and 4, respectively. Also, in
section 4, the first new result appears, viz. Proposition 4.8. In section 5
a nontopological compactness notion is introduced and applied to probabili-
ties. Finally, in section 6, we derive a result on measurability of inte-
grals that play a central role in Chapter III.

§ 1. Preliminaries

1)   The terms: set, collection, and class all stand for the same thing.
Which of them is used depends merely on the role played by them in the
argument. A set is called *countable* when it is finite or countably
infinite.

Let E be a set. Then for every pair A,B of subsets of E the complement
(with respect to E) of A is denoted by A^c and the difference $A \cap B^c$ of A
and B by $A \setminus B$. When A is a collection of subsets of E then
$A_c := A \cup \{A^c \mid A \in A\}$. Moreover, A_d (A_s, A_δ, A_σ, respectively) denotes
the collection of those subsets of E that are finite intersections
(finite unions, countable intersections, countable unions, respectively)
of members of A, while σA stands for the σ-algebra generated by A, i.e.,
the smallest σ-algebra of subsets of E that contains A. Collections like
$(A_s)_\delta$ will be simply denoted by $A_{s\delta}$ etc. For each collection A of sub-
sets of some set we have $A_{sd} = A_{ds}$: the inclusion $A_{sd} \subset A_{ds}$ is obvious,
whereas the reverse inclusion follows from this by complementation and
de Morgan's rule.

2)   DEFINITION. A collection A of sets is called a *Dynkin class* when

i) $\forall_{A,B \in A}$ $[A \supset B \Rightarrow A \setminus B \in A]$

ii) if $(A_n)_{n \in \mathbb{N}}$ is an increasing sequence in A then $\bigcup_{n \in \mathbb{N}} A_n \in A$.

For Dynkin classes we have the following proposition due to Dynkin, a proof of which can be found in [Ash] Theorem 4.1.2, and in [Cohn] Theorem 1.6.1.

PROPOSITION 1.1. *Let A be a collection of subsets of a set E such that $A_d = A$ and $E \in A$. Then $\sigma(A)$ is the smallest Dynkin class containing A.*

3)   Next we collect some facts on mappings. Let φ be a mapping of a set E into a set F. For every subset A of E we define $\varphi A := \{\varphi x \mid x \in A\}$, which is a subset of F. By φ^{-1} we denote the mapping of the collection of all subsets of F into the collection of all subsets of E defined by $\varphi^{-1}B := \{x \in E \mid \varphi x \in B\}$. In particular we have $\varphi^{-1}F = E$ and $\varphi^{-1}\emptyset = \emptyset$. Note that we did not suppose φ to be injective or surjective. Note also that φ^{-1} does not map points of F on points of E, so φ^{-1} should not be considered as an inverse of φ in the usual sense.

For every collection B of subsets of F we define $\varphi^{-1}B := \{\varphi^{-1}B \mid B \in B\}$, which is a collection of subsets of E. Mappings like φ^{-1} will be used extensively in the sequel, so we list some of their properties.

4)   The mapping φ^{-1} commutes with the set-theoretic operations union, intersection and complementation, i.e. for any pair A,B of subsets of F we have $\varphi^{-1}(A \cup B) = (\varphi^{-1}A) \cup (\varphi^{-1}B)$, $\varphi^{-1}(A \cap B) = (\varphi^{-1}A) \cap (\varphi^{-1}B)$ and $\varphi^{-1}(A^c) = (\varphi^{-1}A)^c$. More generally, we have, for every collection B of subsets of F, the equalities $\varphi^{-1}(\cup B) = \cup(\varphi^{-1}B)$ and $\varphi^{-1}(\cap B) = \cap(\varphi^{-1}B)$. As a simple consequence we have, for every collection B of subsets of F, that $\varphi^{-1}(B_\rho) = (\varphi^{-1}B)_\rho$, where ρ stands for any of the subscripts s, σ, d, δ or c.

The properties of φ^{-1} considered so far are simple consequences of the definition of φ^{-1}. Slightly more involved is the proof of the following property: when B is a collection of subsets of F, then $\varphi^{-1}(\sigma B) = \sigma(\varphi^{-1}B)$. To prove this we observe that σB is closed under the formation of countable unions and under complementation and that the

same holds for the collection $\varphi^{-1}(\sigma B)$, due to the properties of φ^{-1}
mentioned before. So $\varphi^{-1}(\sigma B)$ is a σ-algebra which, moreover, contains
$\varphi^{-1}B$. Hence $\varphi^{-1}(\sigma B) \supset \sigma(\varphi^{-1}B)$. On the other hand, the collection
$\{C \subset F \mid \varphi^{-1}C \in \sigma(\varphi^{-1}B)\}$ contains B and it is closed under the formation
of countable unions and under complementation. It therefore contains
$\sigma(B)$ and, consequently, we have $\varphi^{-1}(\sigma B) \subset \sigma(\varphi^{-1}B)$.

   Each of the properties of φ^{-1} mentioned above expresses the fact that
φ^{-1} commutes with a certain set-theoretic operation on collections of
sets.

5)   A particular case of the foregoing is forming the trace of a collec-
tion of sets. Let F be a set, let B be a collection of subsets of F, and
let E be a subset of F. The *trace* $B|E$ of B on E is defined to be the
collection $\{B \cap E \mid B \in B\}$ of subsets of E. Obviously we have $B|E = \varphi^{-1}B$,
where $\varphi: E \to F$ is the identity on E. From this it follows for instance
that the trace of a σ-algebra is again a σ-algebra.

6)   The main object of interest in this monograph will be measurable
spaces, i.e. pairs (E,E) where E is a set and E is a σ-algebra of subsets
of E. Subsets of E that belong to E will be called *measurable* subsets of
(E,E). To simplify the notation we shall, as a rule, not mention the
σ-algebra of a measurable space when confusion is unlikely. A subset E
of a measurable space (F,F) will always be supposed to be endowed with
the σ-algebra $F|E$; when we want to stress the measurable-space structure
of E we call E a *subspace* of F rather than a subset of F.

7)   A product of measurable spaces will always be supposed to be equipped
with the product σ-algebra. We recall some facts on this subject.

   Let $((F_i, F_i))_{i \in I}$ be a family of measurable spaces, let F be the
Cartesian product $\Pi_{i \in I} F_i$ of the family $(F_i)_{i \in I}$, and, for each $i \in I$,
let π_i be the i-th *coordinate* on F, i.e. the mapping $\pi_i: F \to F_i$ defined
by $\pi_i(x) := x_i$. The *product* $\otimes_{i \in I} F_i$ of the family $(F_i)_{i \in I}$ of σ-algebras
is by definition the smallest σ-algebra F on F such that, for each $i \in I$,
the mapping $\pi_i: (F,F) \to (F_i, F_i)$ is measurable.

   The measurable space $(\Pi_{i \in I} F_i, \otimes_{i \in I} F_i)$ will be denoted by $\Pi_{i \in I}(F_i, F_i)$
or simply by $\Pi_{i \in I} F_i$ and it is called the *product* of the family
$((F_i, F_i))_{i \in I}$ of measurable spaces.

8)   As is clear from its definition, the product σ-algebra $\otimes_{i \in I} F_i$ is generated by the collection $B := \cup_{i \in I}(\pi_i^{-1} F_i)$ and therefore by the collection B_d. The members of the collection B_d are called *measurable cylinders*; they are characterized by the fact that they can be written as $\Pi_{i \in I} A_i$, where $A_i \in F_i$ for all $i \in I$ and $A_i = F_i$ for all but a finite number of $i \in I$.

9)   Let (E, E) be another measurable space and let $\varphi: E \to \Pi_{i \in I} F_i$. Then

$$\varphi^{-1} \otimes_{i \in I} F_i = \varphi^{-1} \sigma \cup_{i \in I}(\pi_i^{-1} F_i) = \sigma \cup_{i \in I} \varphi^{-1}(\pi_i^{-1} F_i) = \sigma \cup_{i \in I}(\pi_i \circ \varphi)^{-1} F_i ,$$

so $\varphi^{-1} \otimes_{i \in I} F_i \subset E$ iff $\forall_{i \in I} [(\pi_i \circ \varphi)^{-1} F_i \subset E]$, hence φ is measurable iff $\forall_{i \in I} [\pi_i \circ \varphi$ is measurable]. This property can be phrased as: "A mapping into a product space is measurable iff its coordinates are measurable".

10)   When all spaces in a family $(F_i)_{i \in I}$ are identical to a space F, then the product space $\Pi_{i \in I} F_i$ may be denoted by F^I. In particular we have the spaces $\mathbb{N}^{\mathbb{N}}$ of all sequences of positive integers; each $n \in \mathbb{N}^{\mathbb{N}}$ equals the sequence $(n_1, n_2, n_3, \ldots)$ of its coordinates.

When only two measurable spaces, say (F_1, F_1) and (F_2, F_2), are involved the product space is denoted by $(F_1, F_1) \times (F_2, F_2)$ and the product σ-algebra by $F_1 \otimes F_2$. Cylinders are usually called *rectangles* in this case. For any two collections A_1 and A_2 of subsets of F_1 and F_2, respectively, the collection $\{A_1 \times A_2 \mid A_1 \in A_1$ and $A_2 \in A_2\}$ is denoted by $A_1 \times A_2$.

11)   Let (E, E) be a measurable space. A *probability* on (E, E) is a mapping $\mu: E \to [0,1]$ such that $\mu(E) = 1$ and such that $\mu(\cup_{n \in \mathbb{N}} A_n) = \Sigma_{n \in \mathbb{N}} \mu(A_n)$ for each sequence $(A_n)_{n \in \mathbb{N}}$ of mutually disjoint members of E. When μ is a probability on (E, E) and when φ is a measurable mapping of (E, E) into a measurable space (F, F), then $\mu \circ \varphi^{-1}$ is a probability on (F, F). As a special case of this we have the following. Let μ be a probability on a product space $\Pi_{i \in I}(F_i, F_i)$. Then, for each $I' \subset I$, the *marginal* of μ on $\Pi_{i \in I'}(F_i, F_i)$ is defined to be the probability $\mu \circ \pi^{-1}$, where π is the projection of $\Pi_{i \in I} F_i$ onto $\Pi_{i \in I'} F_i$.

Frequently used probabilities are those concentrated at one point. Let (E, E) be a measurable space and let $x \in E$. Then the probability δ_x

on E is defined by $\delta_x(A) = 1_A(x)$ and it is called *the probability con-
centrated* at x. Note that {x} is not supposed to be a measurable subset
of E. Obviously, for each measurable function f on E we have
$$\int_E f \, d\delta_x = f(x).$$

12) A subset A of a set E is said to separate a pair x,y of distinct
points of E when either $x \in A$ and $y \notin A$ or $x \notin A$ and $y \in A$. A collection
A of subsets of E is called *separating* if each pair of distinct points
of E is separated by a member of A. A measurable space (E,E) is called
separated when E is separating and it is called *countably separated* when
some countable subclass of E is separating. When the σ-algebra E is
generated by a subclass A, then a pair x,y of distinct points of E is
separated by some member of E iff it is separated by some member of A,
because the collection of subsets of E that do not separate x and y is a
σ-algebra and it therefore contains E as soon as it contains A.

Identification of points of a measurable space that are not separated
by measurable sets yields a separated measurable space whose σ-algebra
is isomorphic to the σ-algebra of the original space. Such an identifica-
tion therefore is inessential in many situations. We do not, however,
restrict ourselves to separated spaces as this turns out to be incon-
venient.

13) A measurable space (E,E) is called *countably generated* when the σ-
algebra E is generated by some countable subclass. When the σ-algebra E
of a measurable space is generated by a (not necessarily countable)
subclass A and when E_0 is a countably generated sub-σ-algebra of E, then
there exists a countable subclass A_0 of A such that $E_0 \subset \sigma(A_0)$. This is
a simple consequence of the fact that the collection $\cup\{\sigma(A_0) \mid A_0$ count-
able subclass of $A\}$ is a σ-algebra which contains A and therefore E.

14) By $\mathbb{N}$ we denote the set $\{1,2,3,\ldots\}$ of positive integers equipped
with the σ-algebra of all its subsets. $\mathbb{R}$ denotes the set of real
numbers endowed with the σ-algebra generated by the collection of all
intervals; the members of this σ-algebra are called *Borel sets*. $\overline{\mathbb{R}}$ is
the set $\mathbb{R} \cup \{-\infty,\infty\}$ equipped with the usual ordering and with the σ-
algebra generated by the collection of intervals [a,b] $(a,b \in \overline{\mathbb{R}})$.
Addition and multiplication of $\mathbb{R}$ are extended on $\overline{\mathbb{R}}$ in the usual way
subject to the convention that $\infty + (-\infty) = -\infty$ and $0 \cdot (-\infty) = 0 \cdot \infty = 0$.

Note that in $\overline{\mathbb{R}}$ multiplication by -1 is not distributive over addition: $(-1)\cdot(\infty + (-\infty)) \neq (-1)\cdot\infty + (-1)\cdot(-\infty)$.

By a *function* on a set E we mean a mapping of E into $\overline{\mathbb{R}}$. A function will be called *positive* when its values belong to $[0,\infty]$. For any set A the function 1_A is defined by: $1_A(x) = 1$ when $x \in A$ and $1_A(x) = 0$ when $x \notin A$. Mappings will often be denoted by their argument-value pairs separated by the symbol $\mapsto$. Example: $x \mapsto x^2 + 1$ ($x \in \mathbb{R}$) denotes the mapping defined on $\mathbb{R}$ that assigns to x the value $x^2 + 1$.

§ 2. Spaces of probabilities

In this section the set of all probabilities on a measurable space is equipped with a certain structure, which makes it a measurable space with some desirable properties. These spaces of probabilities play a predominant role in the sequel; not only do many notions and results find their most natural formulation in terms of this structure, but also results bearing upon individual probabilities often can be derived more easily when these probabilities are considered as members of the measurable space of all probabilities on a certain space.

There are several ways to provide the set M of all probabilities on a measurable space (E,E) with a σ-algebra. One way is to endow this set with the metric corresponding to the total-variation norm, which turns M into a metric space (see [Neveu] section IV.1). Next, one may equip M with the σ-algebra generated by the topology of this metric space. A second method applies when the σ-algebra E itself is generated by some topology T on E. In this case one may consider the smallest topology on M such that the mapping $\mu \mapsto {}_E\!\int f\,d\mu$ is continuous on M for every bounded continuous function f on (E,T). This, so-called weak, topology again can be used to generate a σ-algebra on M. In the latter procedure continuity may be replaced by upper semicontinuity; this results in a third construction of a σ-algebra on M (see [Topsøe]).

In our approach, however, we directly define a σ-algebra on M without first constructing a topology.

DEFINITION. Let (E,E) be a measurable space. The set of all probabilities on E will be denoted by $\tilde{E}(E)$. For every $A \in E$ the function $\mu \mapsto \mu(A)$ maps

$\widetilde{E}(E)$ into the measurable space $\mathbb{R}$ and $\widetilde{E}$ is defined to be the smallest σ-algebra on $\widetilde{E}(E)$ with respect to which all these functions are measurable.

The measurable space $(\widetilde{E}(E), \widetilde{E})$ will also be denoted by $(E,E)^{\sim}$. When no confusion can arise both the set $\widetilde{E}(E)$ and the space $(\widetilde{E}(E), \widetilde{E})$ will be denoted by just $\widetilde{E}$.

The definition of $\widetilde{E}$ can also be phrased: $\widetilde{E}$ is the smallest σ-algebra on $\widetilde{E}(E)$ such that, for every $A \in E$, the function $\mu \mapsto {}_E\!\int 1_A d\mu$ is measurable. We shall see later on (Proposition 6.1) that this is equivalent to the measurability of $\mu \mapsto {}_E\!\int f d\mu$ for all bounded measurable functions f on (E,E). So, our construction of the σ-algebra $\widetilde{E}$ bears some resemblance to the constructions discussed above and, in fact, coincides with them in some special cases (see [Bertsekas & Shreve] Proposition 7.25).

The set $\{\mu \mapsto \mu(A) \mid A \in E\}$ of functions on $\widetilde{E}(E)$ in terms of which the σ-algebra $\widetilde{E}$ is defined can be considerably reduced:

PROPOSITION 2.1. *Let* (E,E) *be a measurable space and let* A *be a subclass of* E *that generates* E *and is closed under formation of finite intersections. Then* $\widetilde{E}$ *is the smallest σ-algebra on* $\widetilde{E}$ *such that for every* A $\in$ A *the function* $\mu \mapsto \mu(A)$ *on* $\widetilde{E}$ *is measurable.*

PROOF. Let B be the smallest σ-algebra on $\widetilde{E}$ such that for each A $\in$ A the function $\mu \mapsto \mu(A)$ is measurable on $(\widetilde{E}, B)$. It follows from the definition of $\widetilde{E}$ and from A $\subset$ E that $B \subset \widetilde{E}$.

Now the sets A $\in$ E for which the function $\mu \mapsto \mu(A)$ is measurable on $(\widetilde{E}, B)$ constitute a Dynkin class D containing A $\cup$ {E}. As the collection A $\cup$ {E} is closed under the formation of finite intersections it follows from Proposition 1.1 that $D \supset \sigma(A \cup \{E\}) = E$. So for every A $\in$ E the function $\mu \mapsto \mu(A)$ is measurable on $(\widetilde{E}, B)$, which implies that $B \supset \widetilde{E}$. $\qquad\square$

The following proposition is a frequently used tool to check measurability of mappings into spaces of probabilities.

PROPOSITION 2.2. *Let* (E,E) *be a measurable space and let* A *be a subclass of* E *that generates* E *and is closed under formation of finite intersections. Then a mapping* φ: F $\to$ $\widetilde{E}$ *of an arbitrary measurable space* F *into* $\widetilde{E}$ *is measurable iff for every* A $\in$ A *the function* $x \mapsto \varphi(x)(A)$ *is measurable on* F.

PROOF. The "only if" part is a trivial consequence of the definition of $\tilde{E}$. So, for every $A \in A$, let $\psi_A: \tilde{E} \to \mathbb{R}$ be the mapping $\mu \mapsto \mu(A)$ and suppose that the mapping $\psi_A \circ \varphi$ is measurable. Denoting by F and R the σ-algebras of F and $\mathbb{R}$, respectively, by Proposition 2.1 we have

$$\tilde{E} = \sigma \bigcup_{A \in A} (\psi_A^{-1} R)$$

and, consequently,

$$\varphi^{-1} \tilde{E} = \sigma \bigcup_{A \in A} \varphi^{-1} (\psi_A^{-1} R) = \sigma \bigcup_{A \in A} [(\psi_A \circ \varphi)^{-1} R] \subset F ,$$

because $(\psi_A \circ \varphi)^{-1} R \subset F$ for every $A \in A$. So φ is measurable. $\square$

In particular, the collection A may be the whole of E. This will be the case in most applications of Proposition 2.2. Note the similarity to the measurability property of mappings into product spaces mentioned in the preliminaries. In fact, a similar result holds for mappings into any measurable space whose σ-algebra is generated by a set of mappings.

We now give some applications of the foregoing proposition, which will be used in the sequel.

EXAMPLES.

1) Let (E, E) be a measurable space and let $\delta: E \to \tilde{E}$ be the mapping that maps each point x of E onto the probability δ_x concentrated at that point. Then δ is measurable, because for each $A \in E$ the number $\delta_x(A)$ equals $1_A(x)$, which is a measurable function of x.

2) Let φ be a measurable mapping of a measurable space (E, E) into a measurable space (F, F). Then, for each probability μ on E, $\mu \circ \varphi^{-1}$ is a probability on F. Moreover, $\mu \circ \varphi^{-1}$ depends measurably on μ, i.e. $\mu \mapsto \mu \circ \varphi^{-1}$ is a measurable mapping of $\tilde{E}$ into $\tilde{F}$, because for each $B \in F$ we have $(\mu \circ \varphi^{-1})(B) = \mu(\varphi^{-1} B)$ and, since $\varphi^{-1} B \in E$, the last expression is a measurable function of μ.

3)   When, in example 2), φ is taken to be a projection in a product space
     (see preliminaries 11) then it follows that the marginals of a probabil-
     ity on a product space depend measurably on that probability.

4)   When, in example 2), $(F,\mathcal{F})$ is taken to be $(E,\mathcal{E}_0)$ with $\mathcal{E}_0 \subset \mathcal{E}$, and
     when φ is taken to be the identity on E, then for each probability μ on
     E the probability $\mu \circ \varphi^{-1}$ is the restriction of μ to the sub-σ-algebra $\mathcal{E}_0$
     of $\mathcal{E}$. So the restriction of a probability to a sub-σ-algebra depends
     measurably on that probability.

5)   Another example is the product of probabilities. Let E and F be
     measurable spaces. Then the product $\mu \times \nu$ of a probability μ on E and a
     probability ν on F depends measurably on μ and ν simultaneously, i.e.
     $(\mu,\nu) \mapsto \mu \times \nu$ is a measruable mapping of $\widetilde{E} \times \widetilde{F}$ into $(E \times F)^{\sim}$. To prove
     this we apply Proposition 2.2, taking for $\mathcal{A}$ the collection of measurable
     rectangles of the product space E $\times$ F.

6)   The last example we consider is the transition probability. A transi-
     tion probability from a measurable space $(E,\mathcal{E})$ to a measurable space
     $(F,\mathcal{F})$ is a function p on E $\times$ F such that for every B $\in \mathcal{F}$ the function
     $x \mapsto p(x,B)$ is measurable and for every x $\in$ E the function $B \mapsto p(x,B)$ is
     a probability on $\mathcal{F}$. Clearly, p can be identified with a measurable
     mapping of E into $\widetilde{F}$. In fact, this will be the way by which transition
     probabilities will be introduced in section 9.

     Certain properties of a measurable space E are inherited by the space
$\widetilde{E}$, as is illustrated by the following proposition. Other examples will be
given later.

PROPOSITION 2.3. *Let* $(E,\mathcal{E})$ *be a measurable space. Then* $\widetilde{\mathcal{E}}$ *separates the
points of* $\widetilde{E}$. *When* $(E,\mathcal{E})$ *is countably generated, then* $(\widetilde{E},\widetilde{\mathcal{E}})$ *is countably
generated and countably separated.*

PROOF. Let $\mu_1,\mu_2 \in \widetilde{E}$ and $\mu_1 \neq \mu_2$. Then $\mu_1(A) < \mu_2(A)$ for some A $\in \mathcal{E}$. Hence
$\{\mu \in \widetilde{E} \mid \mu(A) \geq \mu_2(A)\}$ is a member of $\widetilde{\mathcal{E}}$ that separates μ_1 and μ_2. Now let
$(E,\mathcal{E})$ be countably generated and let $\mathcal{C}$ be a countable generating subclass
of $\mathcal{E}$. Then by Proposition 2.1 the σ-algebra $\widetilde{\mathcal{E}}$ is generated by the functions
$\mu \mapsto \mu(A)$ $(A \in \mathcal{C}_d)$ and, hence, by the countable collection

$\{\{\mu \in \widetilde{E} \mid \mu(A) \geq r\} \mid r \in Q, A \in C_d\}$, where Q denotes the set of rational numbers. As $(\widetilde{E}, \widetilde{E})$ is separated, this countable generating collection must be separating as well. □

Finally we remark that the theory which was dealt with in this section can easily be extended to bounded measures, not necessarily probabilities.

§ 3. Universal measurability

Let (E, E) be a measurable space and let μ be a probability on E. By the completion E_μ of E with respect to μ we mean the collection of subsets A of E for which there exist sets $B_1, B_2 \in E$ (depending on A) such that $B_1 \subset A \subset B_2$ and $\mu(B_1) = \mu(B_2)$. It is well known (see [Cohn] Proposition 1.5.1) that E_μ is a σ-algebra containing E, and that μ can be extended to a probability on E_μ. So, as far as μ is concerned, there is not much difference between the spaces (E, E) and (E, E_μ).

Now, the completion E_μ of E depends on the probability μ and one may ask whether there exists a σ-algebra which can be looked upon as a kind of completion of E for all probabilities on E simultaneously.

The subject of this section is to prove that such a completion does indeed exist, and that the measurability notion associated with it has some nice stability properties.

The usefulness of this generalized measurability concept will become evident in the subsequent sections where certain, not necessarily measurable, sets and mappings emerge which turn out to be measurable in this generalized sense.

DEFINITION. Let (E, E) be a measurable space and for each probability μ on E let E_μ be the completion of E with respect to μ.

A subset of E is called *universally measurable* if it belongs to E_μ for every probability μ on E. The collection of all universally measurable subsets of (E, E) is denoted by $U(E)$ and is called the *universal completion* of E.

PROPOSITION 3.1. *Let* (E,E) *be a measurable space. Then* $U(E)$ *is a σ-algebra containing* E. *When* A *is a σ-algebra of subsets of* E *such that* $E \subset A \subset U(E)$, *then every probability on* E *can be uniquely extended to a probability on* A.

PROOF. We have $U(E) = \cap_\mu E_\mu$, where the intersection is taken over all probabilities μ on E. Consequently $U(E)$ is the intersection of a collection of σ-algebras and therefore it is a σ-algebra itself. Obviously, $E \subset U(E)$.

Now let A be a σ-algebra such that $E \subset A \subset U(E)$ and let μ be a probability on E. Then μ can be extended to E_μ and, hence, to the sub-σ-algebra A of E_μ. Let μ' be an extension of μ to A and let $A \in A$. Then $A \in E_\mu$, so there exist $B_1, B_2 \in E$ such that $B_1 \subset A \subset B_2$ and $\mu(B_1) = \mu(B_2)$. This, however, implies $\mu(B_1) = \mu'(B_1) \leq \mu'(A) \leq \mu'(B_2) = \mu(B_2)$, so $\mu'(A)$ is uniquely determined by μ. From the arbitrariness of A it follows that μ' is the unique extension of μ on A. $\qquad\qquad\square$

The inclusion $E \subset U(E)$ can be strict as will be seen later (see Proposition 4.11).

When confusion is unlikely we shall denote both a probability on E and its extension to $U(E)$ by the same symbol.

The term "completion" for the collection $U(E)$ is justified by the following proposition.

PROPOSITION 3.2. *Let* (E,E) *be a measurable space. Then* $U(U(E)) = U(E)$, *i.e. in the measurable space* $(E,U(E))$ *every universally measurable subset is measurable.*

PROOF. Let $A \in U(U(E))$. We shall prove that $A \in U(E)$, i.e. that $A \in E_\mu$ for every probability μ on E. So, let μ be a probability on E and let its extensions to $U(E)$ and to $U(U(E))$ be denoted by μ as well. Since $A \in U(U(E)) \subset (U(E))_\mu$, there exists a set $B_1 \in U(E)$ such that $B_1 \subset A$ and $\mu(B_1) = \mu(A)$. From $B_1 \in U(E) \subset E_\mu$ it follows that there exists a set $C_1 \in E$ such that $C_1 \subset B_1$ and $\mu(C_1) = \mu(B_1)$, and hence such that $C_1 \subset A$ and $\mu(C_1) = \mu(A)$. By an analogous argument there exists a set $C_2 \in E$ such that $A \subset C_2$ and $\mu(A) = \mu(C_2)$. So $A \in E_\mu$. $\qquad\qquad\square$

As has been argued earlier, identification of points of a measurable space that are not separated by measurable sets is inessential in many cases. This is also the case with regard to universal completion. In fact,

the following proposition implies that the result of universal completion and identification of points does not depend on the order in which these two operations have been performed.

PROPOSITION 3.3. *A pair of points of a measurable space is separated by the measurable sets if it is separated by the universally measurable sets.*

PROOF. Let x and y be points in a measurable space (E, E) that are separated by the collection $U(E)$ of universally measurable sets. Then the probabilities δ_x and δ_y, which are defined on the σ-algebra of all subsets of E, do not coincide on $U(E)$. By Proposition 3.1 this implies that these probabilities do not coincide on E either and that therefore the points x and y are separated by E.   □

We now turn to the measurability of mappings with respect to σ-algebras of universally measurable sets.

DEFINITION. A mapping φ of a measurable space E into a measurable space F is called *universally measurable* when for each measurable subset B of F the set $\varphi^{-1}B$ is a universally measurable subset of E.

So, universal measurability with respect to the σ-algebras E and F is the same as measurability with respect to $U(E)$ and F. Consequently a mapping of a measurable space into a product space is universally measurable iff its coordinates are universally measurable. Also in Proposition 2.2 the words "measurable" can be replaced by "universally measurable". Universal measurability with respect to E and F is equivalent also with measurability with respect to $U(E)$ and $U(F)$, as is stated in the following proposition.

PROPOSITION 3.4. *Let* E *and* F *be measurable spaces, let* $\varphi: E \to F$ *be universally measurable and* B *a universally measurable subset of* F. *Then* $\varphi^{-1}B$ *is a universally measurable subset of* E.

PROOF. Let E be the σ-algebra of E. Due to Proposition 3.2 it is sufficient to prove that $\varphi^{-1}B$ belongs to $U(U(E))$ or, equivalently, that $\varphi^{-1}B$ belongs to $(U(E))_\nu$ for every probability ν on $U(E)$.

Let therefore ν be the probability on $U(E)$. Then $\nu \circ \varphi^{-1}$ is a probability on F. Now, B is a universally measurable subset of F, so there

exist measurable subsets B_1 and B_2 of F such that $B_1 \subset B \subset B_2$ and $(\nu \circ \varphi^{-1})B_1 = (\nu \circ \varphi^{-1})B_2$, and hence such that $\varphi^{-1}B_1 \subset \varphi^{-1}B \subset \varphi^{-1}B_2$ and $\nu(\varphi^{-1}B_1) = \nu(\varphi^{-1}B_2)$. Since $\varphi^{-1}B_1$ and $\varphi^{-1}B_2$ belong to $U(E)$ by the universal measurability of φ, this implies that $\varphi^{-1}B \in (U(E))_\nu$. $\qquad\square$

COROLLARY 3.5. *A composition of universally measurable mappings is universally measurable.*

As a particular kind of universally measurable mappings we have the universally measurable functions. Let (E,E) be a measurable space, μ a probability on E, and f a positive universally measurable function on E. Since f is measurable with respect to the σ-algebra $U(E)$ and as μ is uniquely extendible to $U(E)$, we can define $\int f\,d\mu$ to be the integral of f with respect to the measure space $(E,U(E),\mu)$. This integral is the unique extension, as a σ-additive functional, of the integral of positive measurable functions.

The last proposition in this section bears upon universal measurability in spaces of probabilities. Recall that the σ-algebra $\widetilde{E}$ on the space $\widetilde{E}$ of all probabilities on a measurable space (E,E) has been defined such that $\mu \mapsto \mu(A)$ is a measurable function on $\widetilde{E}$ for every measurable subset A of E.

PROPOSITION 3.6. *Let E be a measurable space and A a universally measurable subset of E. Then the function $\mu \mapsto \mu(A)$ is universally measurable on $\widetilde{E}$.*

PROOF. We have to show that $\mu \mapsto \mu(A)$ is measurable with respect to $(\widetilde{E})_\nu$ for every probability ν on $\widetilde{E}$. So let ν be a probability on $\widetilde{E}$ and let λ be defined on the σ-algebra E of E by $\lambda(B) := \int_{\widetilde{E}} \mu(B)\nu(d\mu)$. Then λ is easily seen to be a probability. As A is universally measurable and therefore belongs to E_λ, there exist sets $B_1, B_2 \in E$ such that $B_1 \subset A \subset B_2$ and $\lambda(B_2 \setminus B_1) = 0$. Consequently, by the definition of λ we have

$$0 = \lambda(B_2 \setminus B_1) = \int_{\widetilde{E}} \mu(B_2 \setminus B_1)\nu(d\mu) = \int_{\widetilde{E}} [\mu(B_2) - \mu(B_1)]\nu(d\mu) .$$

We also have for all $\mu \in \widetilde{E}$ the inequalities $\mu(B_1) \leq \mu(A) \leq \mu(B_2)$. So $\mu(A) = \mu(B_2)$ for ν-almost all $\mu \in \widetilde{E}$. Now $B_2 \in E$, so $\mu(B_2)$ is a measurable function of μ. Therefore the function $\mu \mapsto \mu(A)$ is measurable with respect to $(\widetilde{E})_\nu$. $\qquad\square$

As is easily seen, Proposition 3.6 is equivalent to the inclusion $[U(E)]^\sim \subset U(\tilde{E})$. When the σ-algebra E consists of $\emptyset$ and E only, then $\tilde{E}$ consists of only one probability and, consequently, this inclusion in fact is an equality. In all other cases, however, the inclusion is strict, as will be proved in section 4 (see the remark following Proposition 4.10).

§ 4. Souslin sets and Souslin functions

Let A be a class of subsets of a set E. In general there is no simple construction principle by which the members of $\sigma(A)$ can be obtained from those of A. In this section a construction principle, the Souslin operation, is considered which, when applied to a class A meeting certain conditions, yields a class of sets which contains $\sigma(A)$. Moreover, the class obtained is not too large, since it is itself contained in the σ-algebra of universally measurable sets derived from $\sigma(A)$.

Another instance where the Souslin operation appears concerns projections in product spaces. Let S be a measurable subset of a product space $E \times F$. Then the projection $\{x \in E \mid \exists_{y \in F} (x,y) \in S\}$ of S on E is in general not a measurable subset of E. It can, however, be obtained from the measurable subsets of E by application of the Souslin operation in many cases.

For the interesting history of the Souslin operation we refer to [Hoffmann-Jørgensen] Chapter II, § 11.

Recall that $\mathbb{N}^{\mathbb{N}}$ is the space of all (infinite) sequences of positive integers and that each $n \in \mathbb{N}^{\mathbb{N}}$ equals the (infinite) sequence $(n_1, n_2, n_3, \ldots)$ of its coordinates.

DEFINITION. Let A be a collection of sets and let Ψ be the set of all finite sequences of positive integers. A *Souslin scheme* on A is a family $(A_\psi)_{\psi \in \Psi}$, such that $\forall_{\psi \in \Psi} A_\psi \in A$. The *kernel* of a Souslin scheme $(A_\psi)_{\psi \in \Psi}$ is the set

$$\bigcup_{n \in \mathbb{N}^{\mathbb{N}}} \bigcap_{k \in \mathbb{N}} A_{(n_1, \ldots, n_k)} .$$

The collection consisting of the kernels of all Souslin schemes on a collection A is called the *Souslin class* generated by A; it is denoted by $S(A)$. The operation S, i.e. the mapping $A \mapsto S(A)$, is called the *Souslin operation*. When (E, E) is a measurable space then the members of $S(E)$ will be called *Souslin (sub)sets* of (E, E).

It follows directly from the definition of S that for each collection A and each $S \in S(A)$ there exists a countable subclass A_0 of A such that $S \in S(A_0)$. Also the implication $A \subset B \Rightarrow S(A) \subset S(B)$ and the inclusion $A \subset S(A)$ are obvious, but more can be said:

PROPOSITION 4.1. *Let A be a collection of subsets of some set. Then*

i) $A_\sigma \subset S(A)$ *and* $A_\delta \subset S(A)$,

ii)   *when* $A_c \subset S(A)$ *(in particular when A is closed under complementation) then* $\sigma(A) \subset S(A)$.

PROOF.

i)   Let $(B_i)_{i \in \mathbb{N}}$ be a sequence in A and let $A_{n_1, \ldots, n_k} := B_{n_1}$ (or B_k).
Then

$$\underset{n \in \mathbb{N}}{\bigcup}^{\mathbb{N}} \underset{k \in \mathbb{N}}{\bigcap} A_{n_1, \ldots, n_k} = \underset{i \in \mathbb{N}}{\bigcup} B_i \quad (\text{or } \underset{i \in \mathbb{N}}{\bigcap} B_i) .$$

So $A_\sigma \subset S(A)$ and $A_\delta \subset S(A)$.

ii)   To prove ii) we merely observe that by i) the collection
$\{S \in S(A) \mid S^c \in S(A)\}$ is a sub-σ-algebra of $S(A)$ which contains A. $\square$

It follows from ii) in the foregoing proposition that $\sigma(A) \subset S(A_c)$ for every collection A of subsets of some set. So, the members of a σ-algebra can be obtained from the members of a generating class by complementation followed by the Souslin operation.

Like most of the set-theoretic operations considered up to now the Souslin operation is idempotent:

PROPOSITION 4.2. *Let A be a collection of sets. Then $S(S(A)) = S(A)$.*

PROOF. The inclusion $S(A) \subset S(S(A))$ is obvious. To prove the reverse inclusion, let $A \in S(S(A))$. Then A can be written as

$$A = \underset{n \in \mathbb{N}}{\bigcup}^{\mathbb{N}} \underset{k \in \mathbb{N}}{\bigcap} A_{n_1, \ldots, n_k} ,$$

where $A_{n_1, \ldots, n_k} \in S(A)$ for all n and k. Also, for each $n \in \mathbb{N}^{\mathbb{N}}$ and $k \in \mathbb{N}$, we can write

$$A_{n_1,\ldots,n_k} = \bigcup_{m\in\mathbb{N}} \bigcap_{\ell\in\mathbb{N}} A_{n_1,\ldots,n_k;m_1,\ldots,m_\ell} ,$$

with $A_{n_1,\ldots,n_k;m_1,\ldots,m_\ell} \in \mathcal{A}$ for all m and ℓ. Consequently,

$$A = \bigcup_{n\in\mathbb{N}^{\mathbb{N}}} \bigcap_{k\in\mathbb{N}} \bigcup_{m\in\mathbb{N}^{\mathbb{N}}} \bigcap_{\ell\in\mathbb{N}} A_{n_1,\ldots,n_k;m_1,\ldots,m_\ell}$$

where all indexed sets belong to $\mathcal{A}$. This equality is equivalent to

$$A = \bigcup_{n\in\mathbb{N}^{\mathbb{N}}} \bigcup_{f:\mathbb{N}\to\mathbb{N}^{\mathbb{N}}} \bigcap_{k\in\mathbb{N}} \bigcap_{\ell\in\mathbb{N}} A_{n_1,\ldots,n_k;f(k)_1,\ldots,f(k)_\ell}$$

because for every family $(B_{k,m})_{k\in\mathbb{N},\, m\in\mathbb{N}^{\mathbb{N}}}$ of sets we have

$$x \in \bigcap_{k\in\mathbb{N}} \bigcup_{m\in\mathbb{N}^{\mathbb{N}}} B_{k,m} \quad\Leftrightarrow$$

$$\forall_{k\in\mathbb{N}} \exists_{m\in\mathbb{N}^{\mathbb{N}}} x \in B_{k,m} \quad\Leftrightarrow$$

$$\exists_{f:\mathbb{N}\to\mathbb{N}^{\mathbb{N}}} \forall_{k\in\mathbb{N}} x \in B_{k,f(k)} \quad\Leftrightarrow$$

$$x \in \bigcup_{f:\mathbb{N}\to\mathbb{N}^{\mathbb{N}}} \bigcap_{k\in\mathbb{N}} B_{k,f(k)} .$$

Next, in the expression for A we combine the two unions into one. To this end let $\alpha: \mathbb{N} \to \mathbb{N}$ and $\beta: \mathbb{N} \to \mathbb{N}$ be such that the mapping $n \mapsto (\alpha(n),\beta(n))$ is a surjection of $\mathbb{N}$ onto $\mathbb{N} \times \mathbb{N}$. Then for each sequence $n \in \mathbb{N}^{\mathbb{N}}$ and for each mapping $f: \mathbb{N} \to \mathbb{N}^{\mathbb{N}}$ there exists a (non-unique) mapping $g: \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, such that

$$n_i = \alpha(g(i,1)) \quad \text{and} \quad f(i)_j = \beta(g(i,j)) \qquad (i,j \in \mathbb{N}) .$$

Therefore we can write A in the form

$$A = \bigcup_{g:\mathbb{N}\times\mathbb{N}\to\mathbb{N}} \bigcap_{(k,\ell)\in\mathbb{N}\times\mathbb{N}} A_{\alpha(g(1,1)),\ldots,\alpha(g(k,1));\beta(g(k,1)),\ldots,\beta(g(k,\ell))}$$

where we have also combined the two intersections.

Our next step is the transformation of the set $\mathbb{N} \times \mathbb{N}$, appearing twice, into $\mathbb{N}$. Let $\gamma: \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ be a bijection such that

(1) $\forall_{i,j,i',j' \in \mathbb{N}} [i+j < i'+j' \Rightarrow \gamma(i,j) < \gamma(i',j')]$.

Since γ is injective, for each mapping g: $\mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ there exists a
sequence h $\in \mathbb{N}^{\mathbb{N}}$ such that $g(i,j) = h_{\gamma(i,j)}$ $(i,j \in \mathbb{N})$. As a consequence
we have

$$A = \bigcup_{h \in \mathbb{N}^{\mathbb{N}}} \bigcap_{(k,\ell) \in \mathbb{N} \times \mathbb{N}} A_{\alpha(h_{\gamma(1,1)}), \ldots, \alpha(h_{\gamma(k,1)}); \beta(h_{\gamma(k,1)}), \ldots, \beta(h_{\gamma(k,\ell)})} .$$

It follows from (1) that the numbers $\gamma(1,1),\ldots,\gamma(k,\ell)$ appearing in the
multi-index do not exceed $\gamma(k,\ell)$. So, the multi-index is a function F of
the coordinates $h_1,h_2,\ldots,h_{\gamma(k,\ell)}$ of h only. The function F itself depends
on k and ℓ only and therefore on $\gamma(k,\ell)$ only, because γ is injective. As γ
is bijective, we can replace (k,ℓ) by $\gamma(k,\ell)$ in the intersection thus ob-
taining

$$A = \bigcup_{h \in \mathbb{N}^{\mathbb{N}}} \bigcap_{j \in \mathbb{N}} A_{F_j(h_1,\ldots,h_j)} ,$$

for suitably defined index functions F_j on $\mathbb{N}^j$ $(j \in \mathbb{N})$. Consequently, A
belongs to $S(A)$. □

As a simple consequence of the two foregoing propositions a Souslin
class is closed under the formation of countable unions and countable
intersections. In general however, a Souslin class is not closed under
complementation and therefore is not a σ-algebra.

The following two propositions express the fact that the Souslin
operation commutes with certain other operations.

PROPOSITION 4.3. *Let* φ *be a mapping of a set* E *into a set* F *and let* $\mathcal{B}$ *be a
collection of subsets of* F. *Then* $\varphi^{-1}(S(\mathcal{B})) = S(\varphi^{-1}\mathcal{B})$.

PROOF. Since φ^{-1} commutes with unions and intersections, for each Souslin
scheme B on $\mathcal{B}$ we have

$$\varphi^{-1}\left(\bigcup_{n \in \mathbb{N}^{\mathbb{N}}} \bigcap_{k \in \mathbb{N}} B_{n_1,\ldots,n_k} \right) = \bigcup_{n \in \mathbb{N}^{\mathbb{N}}} \bigcap_{k \in \mathbb{N}} \left(\varphi^{-1} B_{n_1,\ldots,n_k} \right)$$

Hence the result. □

PROPOSITION 4.4. *Let* F *be a set,* E *a subset of* F *and* B *a collection of subsets of* F. *Then*

i) $S(B)|E = S(B|E)$

ii)   *when, in addition,* $E \in S(B)$, *then*

$$S(B)|E = \{S \in S(B) \mid S \subset E\} \ .$$

PROOF.

i)   Let $\varphi: E \to F$ be the identity on E. Then the result follows from the preceding proposition.

ii)   Let S be a subset of E. When $S \in S(B)|E$, then for some $S' \in S(B)$ we have $S = S' \cap E \in [S(B)]_d = S(B)$. When, on the other hand, $S \in S(B)$, then $S = S \cap E \in S(B)|E$.   □

The last set-theoretic property of Souslin classes we mention concerns product σ-algebras.

PROPOSITION 4.5. *Let* (E,E) *and* (F,F) *be measurable spaces, let* A *be a collection of subsets of* E *such that* $E \subset S(A)$, *and let* B *be a collection of subsets of* F *such that* $F \subset S(B)$. *Then* $E \otimes F \subset S(A \times B)$. *In particular,* $E \otimes F \subset S(E \times F)$.

PROOF. As a simple consequence of the definition of the Souslin operation we have, for each $B \in B$,

$$S(A) \times \{B\} = \{S \times B \mid S \in S(A)\} = S(\{A \times B \mid A \in A\}) \subset S(A \times B) \ .$$

Consequently, $S(A) \times B \subset S(A \times B)$.

Interchanging the role of A and B in the foregoing argument we get $A \times S(B) \subset S(A \times B)$ and, applying this argument to the collections A and $S(B)$ instead of A and B, we get $S(A) \times S(B) \subset S(A \times S(B))$. The last two inclusions yield:

$$E \times F \subset S(A) \times S(B) \subset S(A \times S(B)) \subset SS(A \times B) = S(A \times B) \ .$$

Now the complement of any measurable rectangle in $E \times F$ is the union of two such rectangles, so

$$(E \times F)_c \subset (E \times F)_s \subset [S(A \times B)]_s = S(A \times B) \ .$$

It now follows from the second part of Proposition 4.1 that

$$E \otimes F = \sigma(E \times F) \subset S(A \times B) \ .$$ $\qquad\qquad\Box$

Beside the set-theoretic properties mentioned before, Souslin classes have some useful measure-theoretic features. To start with, we have the following relation to universally measurable sets.

PROPOSITION 4.6. *Let* (E,E) *be a measurable space. Then each Souslin set is universally measurable and the σ-algebra of universally measurable sets is closed under the Souslin operation, i.e.* $S(E) \subset U(E) = SU(E)$.

The proof of this proposition will be combined with the proof of Proposition 5.2.

From the foregoing sections we know that for every (universally) measurable set A in a measurable space (E,E) the function $\mu \mapsto \mu(A)$ is (universally) measurable on $\tilde{E}$. What can be said of this function when A is a Souslin set? Of course, in general measurability with respect to $S(\tilde{E})$ is not defined, as $S(\tilde{E})$ may fail to be a σ-algebra.

In order to describe the dependence of $\mu(A)$ on μ when A is a Souslin set, we introduce functions which closely resemble the measurable functions.

DEFINITION. Let (E,E) be a measurable space. A function $f: E \to \overline{\mathbb{R}}$ is called a *Souslin function* if $\{x \in E \mid f(x) > a\} \in S(E)$ for each $a \in \mathbb{R}$.

The class of Souslin functions on a measurable space is closed under certain operations as stated in the following proposition. However, when f is a Souslin function, the function $-f$ need not be one, because the complement of a Souslin set may not be a Souslin set.

Note that in the following proposition we use the conventions $\infty - \infty = -\infty$ and $\pm\infty \cdot 0 = 0$ (see Preliminaries 14).

PROPOSITION 4.7.

i)   *Let* $(f_n)_{n \in \mathbb{N}}$ *be a sequence of Souslin functions on a measurable space. Then* $\sup_n f_n$, $\inf_n f_n$, $\limsup_n f_n$ *and* $\liminf_n f_n$ *are Souslin functions as well.*

ii)   *Let f and g be Souslin functions on a measurable space. Then f + g*
also is a Souslin function. When in addition f and g are positive,
then fg is a Souslin function as well.

PROOF. Let the functions be defined on a measurable space $(E, \mathcal{E})$.

i)   For every $a \in \mathbb{R}$ we have

$$\{x \in E \mid \sup_n f_n(x) > a\} = \bigcup_{n \in \mathbb{N}} \{x \in E \mid f_n(x) > a\} \in S(E)_\sigma = S(E)$$

and

$$\{x \in E \mid \inf_n f_n(x) > a\} =$$

$$= \bigcup_{m \in \mathbb{N}} \bigcap_{n \in \mathbb{N}} \{x \in E \mid f_n(x) \geq a + \frac{1}{m}\} \in S(E)_{\delta\sigma} = S(E) \ .$$

So $\sup_n f_n$ and $\inf_n f_n$ are Souslin functions. From this and the
equalities $\limsup_n f_n = \inf_n \sup_{m \geq n} f_m$ and $\liminf_n f_n = \sup_n \inf_{m \geq n} f_m$
the remainder of i) follows.

ii)   For every $a \in \mathbb{R}$ we have

$$\{f+g > a\} = \bigcup_{r \in Q} \{f > r\} \cap \{g > a-r\} \in S(E)_{d\sigma} = S(E) \ ,$$

where Q is the set of rational numbers. So $f + g$ is a Souslin function.
For f and g positive and $a \geq 0$ we have

$$\{fg > a\} = \bigcup_{\substack{r \in Q \\ r > 0}} \{f > r\} \cap \{g > \frac{a}{r}\} \in S(E)$$

which implies that fg is a Souslin function.   □

Recall that we introduced Souslin functions in order to describe the
function $\mu \mapsto \mu(A)$ for Souslin sets A.

PROPOSITION 4.8. *Let A be a Souslin set of a measurable space E. Then*
$\mu \mapsto \mu(A)$ *is a Souslin function on $\tilde{E}$.*

The proof of this proposition will be combined with the proof of
Proposition 5.2.

We conclude this section with some results that will not be used in the rest of this monograph. We first consider the inclusions $E \subset S(E) \subset U(E)$ and $[U(E)]^{\sim} \subset U(\widetilde{E})$, as given in Proposition 4.6 and the remark following Proposition 3.6, and in particular the question whether these inclusions are strict.

PROPOSITION 4.9. *Let* (E,E) *be a countably generated measurable space that can be mapped measurably onto the product space* $\mathbb{N}^{\mathbb{N}}$. *Then* $S(E)$ *is not closed under complementation.*

PROOF. Let $A = \{A_1, A_2, \ldots\}$ be a countable generating subclass of E that is closed under complementation. Then by Proposition 4.1, $E = \sigma(A) \subset S(A) \subset \subset S(E)$, so $S(E) = S(A)$.

Let Ψ be the set of all finite sequences of positive integers and let F be the product space $\mathbb{N}^{\Psi}$. Moreover, let the subset U of $E \times F$ be defined by

$$U := \bigcup_{n \in \mathbb{N}^{\mathbb{N}}} \bigcap_{k \in \mathbb{N}} \bigcup_{\ell \in \mathbb{N}} [A_{\ell} \times \{y \in F \mid y_{(n_1, \ldots, n_k)} = \ell\}] .$$

Then U is a Souslin subset of $E \times F$. Also, for each $y \in F$, we find for the "y-section" of U:

$$\{x \in E \mid (x,y) \in U\} = \bigcup_{n \in \mathbb{N}^{\mathbb{N}}} \bigcap_{k \in \mathbb{N}} A_{y_{(n_1, \ldots, n_k)}} .$$

So for each $S \in S(A)$ we have $S = \{x \in E \mid (x,y) \in U\}$ for some $y \in F$.

Now Ψ is countably infinite, so F is isomorphic to $\mathbb{N}^{\mathbb{N}}$ and, therefore, there exists a measurable surjection $\varphi\colon E \to F$. The mapping $\chi\colon E \to E \times F$, defined by $\chi(x) = (x, \varphi(x))$, is measurable, because its coordinates are, and $\chi^{-1}U$ is a Souslin subset of E by Proposition 4.3.

It will be sufficient to prove, that the complement C of $\chi^{-1}U$ is not a Souslin set of E. We reason by contradiction. Suppose $C \in S(E)$. Then $C \in S(A)$ and therefore $C = \{x \in E \mid (x, y_0) \in U\}$ for some $y_0 \in F$. Since φ is surjective, $y_0 = \varphi(x_0)$ for some $x_0 \in E$. Now we have the following equivalences:

$$x_0 \in C \Leftrightarrow (x_0, y_0) \in U \Leftrightarrow (x_0, \varphi(x_0)) \in U \Leftrightarrow x_0 \in \chi^{-1}U \Leftrightarrow x_0 \notin C ,$$

which is a contradiction. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

PROPOSITION 4.10. *Let* (E, E) *be a measurable space and let* $E \neq \{\emptyset, E\}$. *Then* $S(\tilde{E})$ *is not closed under complementation.*

PROOF. Let $A \in E \setminus \{\emptyset, E\}$, and let $x \in A$ and $y \notin A$. Then $\delta_x(A) = 1 \neq 0 = \delta_y(A)$, so $\delta_x \neq \delta_y$. Consequently,

$$I := \{\lambda \delta_x + (1-\lambda)\delta_y \mid \lambda \in (0,1]\}$$

is a subspace of $\tilde{E}$ which is isomorphic to the interval $(0,1]$.

Next consider the mapping

$$n \mapsto \sum_{k=1}^{\infty} 2^{-(n_1 + \ldots + n_k)}$$

of $\mathbb{N}^{\mathbb{N}}$ onto $(0,1]$. As this mapping is an isomorphism, the space $(0,1]$ can be mapped measurably onto $\mathbb{N}^{\mathbb{N}}$.

The foregoing implies that I can be mapped measurably onto $\mathbb{N}^{\mathbb{N}}$ and, as a consequence of the foregoing proposition, the Souslin class of I is not closed under complementation. As the Souslin class of I is the trace on I of the Souslin class $S(\tilde{E})$ of $\tilde{E}$, the class $S(\tilde{E})$ is not closed under complementation either. $\qquad\square$

Now let (E, E) be a measurable space such that $E \neq \{\emptyset, E\}$. Then $U(E) \neq \{\emptyset, E\}$ and it follows from the foregoing proposition applied to the space $(E, U(E))$ that $S(U(E)^{\sim})$ is not closed under complementation and that the inclusion $U(E)^{\sim} \subset S(U(E)^{\sim})$ is therefore strict. From the inclusion $U(E)^{\sim} \subset U(\tilde{E})$ mentioned after Proposition 3.6 and from Proposition 4.6 we deduce

$$S(U(E)^{\sim}) \subset S(U(\tilde{E})) = U(\tilde{E}) .$$

The inclusion $U(E)^{\sim} \subset U(\tilde{E})$ is therefore strict.

As an interesting consequence of Proposition 4.9 we have:

PROPOSITION 4.11. *Let* B *be the σ-algebra of Borel subsets of* $\mathbb{R}$. *Then the inclusions* $B \subset S(B) \subset U(B)$ *are strict and the σ-algebra* $U(B)$ *is not countably generated.*

PROOF. The space $(\mathbb{R}, B)$ can be mapped measurably onto its subspace $(0,1]$, e.g. by the mapping $x \mapsto (1 + x^2)^{-1}$, and the space $(0,1]$ can be mapped

measurably onto $\mathbb{N}^{\mathbb{N}}$ (see the proof of Proposition 4.10). So, $(\mathbb{R}, B)$ can be mapped measurably onto $\mathbb{N}^{\mathbb{N}}$ and, of course, the same holds for $(\mathbb{R}, U(B))$.

Since B is countably generated, by Proposition 4.9 the collection $S(B)$ is not closed under complementation and therefore it is not a σ-algebra. Consequently, the inclusions $B \subset S(B) \subset U(B)$ are strict.

Next suppose that $U(B)$ is countably generated. Then it follows from Proposition 4.9 applied to the space $(\mathbb{R}, U(B))$ that $S(U(B))$ is not closed under complementation, which contradicts Proposition 4.6. So $U(B)$ is not countably generated. $\qquad\square$

Beside the σ-algebra of universally measurable sets there is another σ-algebra of subsets of a measurable space which has equally nice properties. The remainder of this section is devoted to this σ-algebra. The result will not be used in the sequel.

For any measurable space (E, E) let $L(E)$ be the smallest σ-algebra of subsets of E that contains E and is closed under the Souslin operation S. Following Bertsekas and Shreve we call $L(E)$ the limit-σ-algebra of (E, E) and we call its members limit measurable subsets of (E, E) (see [Bertsekas & Shreve] p. 292).

All properties of the σ-algebra of universally measurable sets derived up to now are shared by the σ-algebra of limit measurable sets, as will be proved presently.

Since the role of universally measurable sets and universally measurable mappings in the remainder of this monograph is based entirely on these common properties, the adjective "universally measurable" can be replaced everywhere by "limit measurable" without affecting the validity of the results.

A property of limit σ-algebras that is possibly not shared by all universal completions is the equality

$$L(E) = \bigcup_{E_0} L(E_0) \; ,$$

where the union is over all countably generated sub-σ-algebras E_0 of E. To prove this equality we merely note that the righthand side is a σ-algebra which is closed under S.

To prove the analogy claimed above we first observe that, due to Proposition 4.6, for every measurable space (E,E) we have $L(E) \subset U(E)$. From this inclusion the analogues of the Propositions 3.1 and 3.3 for the σ-algebra of limit measurable sets easily follow. As to Proposition 3.4, we remark that for a limit measurable mapping $\varphi\colon (E,E) \to (F,F)$ the collection $\{B \subset F \mid \varphi^{-1}B \in L(E)\}$ is a σ-algebra which is closed under S and which contains F. The analogue of Corollary 3.5 is a simple consequence of the analogue of Proposition 3.4 again, while the analogue of Proposition 3.2 follows directly from the definition of L.

The proof of the analogue of Proposition 3.6 is slightly more laborious. Let (E,E) be a measurable space and B the collection consisting of those members A of $U(E)$ for which the function $\mu \mapsto \mu(A)$ on $\tilde{E}$ is limit measurable, i.e. measurable with respect to $L(\tilde{E})$. B is easily seen to be a Dynkin class. Zorn's lemma implies that among the subclasses of B that contain E and are closed under the formation of finite intersections, there is a maximal one, say A. By Dynkin's theorem (Proposition 1.1) we have $\sigma(A) \subset B$ and the maximality of A therefore implies that $\sigma(A) = A$, i.e. that A is a σ-algebra.

We now consider the space (E,A). Since $E \subset A \subset U(E)$, it follows from Proposition 3.1 that the probabilities on E and those on A can be identified in an obvious way, so the spaces $(E,E)^{\sim}$ and $(E,A)^{\sim}$ are composed of the same set $\tilde{E}$ of probabilities. Now $A \subset B$, so for every $A \in A$ the function $\mu \mapsto \mu(A)$ on $\tilde{E}$ is measurable with respect to $L(\tilde{E})$ and consequently $\tilde{A} \subset L(\tilde{E})$. It follows from Prposition 4.8, applied to the space (E,A), that for every $A \in S(A)$ we have $\bigvee_{a \in \mathbb{R}} \{\mu \in \tilde{E} \mid \mu(A) > a\} \in S(\tilde{A})$ and, as $S(\tilde{A}) \subset SL(\tilde{E}) = L(\tilde{E})$, also that $\mu \mapsto \mu(A)$ is limit measurable on $\tilde{E}$.

Due to the maximality of A again we conclude from this that $S(A) = A$. So A is a σ-algebra which contains E and which is closed under S and it therefore contains $L(E)$.

We thus have proved the analogue of Proposition 3.6, that for every limit measurable subset A of E the function $\mu \mapsto \mu(A)$ on $\tilde{E}$ is limit measurable.

The proof of Proposition 4.11 applies also to the limit-σ-algebra of the space $\mathbb{R}$ and this σ-algebra is therefore not countably generated.

§ 5. Semicompact classes

In this section we introduce the concept of a semicompact class. "Countably compact" might have been a more suitable adjective for these classes, because their defining property is precisely the set-theoretic feature exhibited by the class of closed subsets of a countably compact topological space. However, the term "semicompact" is the most usual one.

The usefulness of semicompact classes lies in the fact that σ-additivity of functions defined on algebras of sets can be deduced from certain approximation properties of semicompact classes (see [Neveu] Proposition I.6.2; note that in this reference the term "compact" is used instead of "semicompact").

DEFINITION. A collection A of sets is said to possess the *finite intersection property* if every finite subcollection of A has a nonempty intersection. A collection A of sets is called *semicompact* if every countable subcollection of A which possesses the finite intersection property has a nonempty intersection.

PROPOSITION 5.1. *When C is a semicompact collection of sets, then $C_{s\delta}$ is semicompact as well.*

PROOF. Let $\{B_n \mid n \in \mathbb{N}\}$ be a subcollection of C_s having the finite intersection property. We first prove that there exists a subcollection $\{C_n \mid n \in \mathbb{N}\}$ of C having the finite intersection property and such that $\forall_{n \in \mathbb{N}} \ C_n \subset B_n$.

We proceed by recursion. Let $n \in \mathbb{N}$ and suppose that $C_1, \ldots, C_{n-1} \in C$ have been defined such that the collection $B' := \{C_1, \ldots, C_{n-1}, B_n, B_{n+1}, \ldots\}$ has the finite intersection property.

As $B_n \in C_s$, we have $B_n = \bigcup_{m=1}^{p} C'_m$ for some $p \in \mathbb{N}$ and $C'_1, \ldots, C'_p \in C$. Now for some $m \in \{1, \ldots, p\}$ the collection $B' \cup \{C'_m\}$ has the finite intersection property: if not, then for each $m \in \{1, \ldots, p\}$ there exists a finite subcollection B'_m of B' such that $(\cap \ B'_m) \cap C'_m = \emptyset$. Consequently, $\bigcup_{m=1}^{p} B'_m$ is a finite subcollection of B' whose intersection does not meet $\bigcup_{m=1}^{p} C'_m$ and this contradicts the finite intersection property of B'. So there exists a set $C_n \in C$ such that $B' \cup \{C_n\}$, and therefore also $\{C_1, \ldots, C_{n-1}, C_n, B_{n+1}, \ldots\}$, has the finite intersection property.

The sequence $(C_n)_{n \in \mathbb{N}}$ constructed in this way has the desired properties. These properties and the semicompactness of C now imply that $\cap_{n \in \mathbb{N}} B_n \supset \cap_{n \in \mathbb{N}} C_n \neq \emptyset$. Finally it follows from the arbitrariness of the sequence $(B_n)_{n \in \mathbb{N}}$ that C_s is semicompact.

Next, let B be a countable subclass of $C_{s\delta}$ such that $\cap B = \emptyset$. Then $B = \{\cap_{m \in \mathbb{N}} C_{nm} \mid n \in \mathbb{N}\}$ for a suitable choice of $C_{nm} \in C_s$. Hence $\cap_{(n,m) \in \mathbb{N}^2} C_{nm} = \cap B = \emptyset$ and, as C_s is semicompact, $\cap_{(n,m) \in I} C_{nm} = \emptyset$ for some finite subset I of $\mathbb{N}^2$. This implies that B has a finite subclass the intersection of which is empty. From the arbitrariness of B it follows that $C_{s\delta}$ is semicompact.   □

Let λ be Lebesgue measure on $\mathbb{R}$. It is well known that every Lebesgue-measurable subset A of $\mathbb{R}$ can be approximated from the inside by compact sets in the following sense:

$$\lambda(A) = \sup \{\lambda(B) \mid B \subset A \text{ and } B \text{ compact}\} .$$

This property of the Lebesgue measure is called inner regularity. The following proposition shows that all probabilities on a measurable space are inner regular with respect to suitably chosen collections of subsets of the space.

PROPOSITION 5.2. *Let* (E,E) *be a measurable space and let* C *be a collection of subsets of* E *such that* $C \subset E \subset S(C)$. *Then*

$$\mu(A) = \sup \{\mu(B) \mid B \subset A \text{ and } B \in C_{s\delta}\}$$

for every probability μ *on* E *and for every universally measurable subset* A *of* E.

PROOF. As announced, the proofs of the Propositions 4.6 and 4.8 and of the present proposition will be given simultaneously. For pairs $(n_1,\dots,n_\ell), (m_1,\dots,m_\ell)$ of sequences (of length ℓ) of positive integers we define $(n_1,\dots,n_\ell) \leq (m_1,\dots,m_\ell)$ to mean $n_j \leq m_j$ $(j = 1,\dots,\ell)$.

Let $A \in S(E)$. Then by the definition of S we can write

$$A = \bigcup_{n \in \mathbb{N}^\mathbb{N}} \bigcap_{k \in \mathbb{N}} A_{n_1,\dots,n_k} ,$$

where the sets $A_{n_1,\dots,n_k}$ belong to E. For each finite sequence $(m_1,\dots,m_\ell)$

of positive integers we define

$$B_{m_1,\ldots,m_\ell} := \cup \, \{ \underset{k\le\ell}{\cap} A_{n_1,\ldots,n_k} \mid (n_1,\ldots,n_\ell) \le (m_1,\ldots,m_\ell) \} \ .$$

Since both the intersection and the union occurring in this definition are finite, the sets $B_{m_1,\ldots,m_\ell}$ belong to E. Also for each sequence m of positive integers the sequence $(B_{m_1,\ldots,m_\ell})_{\ell\in\mathbb{N}}$ of sets is decreasing and, as we shall show, $\cap_{\ell\in\mathbb{N}} B_{m_1,\ldots,m_\ell} \subset A$. To prove this inclusion, let $x \in \cap_{\ell\in\mathbb{N}} B_{m_1,\ldots,m_\ell}$. Then by the definition of the sets $B_{m_1,\ldots,m_\ell}$ we have

$$\forall_{\ell\in\mathbb{N}} \ \exists_{(n_1,\ldots,n_\ell)\le(m_1,\ldots,m_\ell)} \ [x \in \underset{k\le\ell}{\cap} A_{n_1,\ldots,n_k}]$$

and, consequently,

$$(*) \quad \left(\left\{ n \in \mathbb{N}^{\mathbb{N}} \mid (n_1,\ldots,n_\ell) \le (m_1,\ldots,m_\ell) \text{ and } x \in \underset{k\le\ell}{\cap} A_{n_1,\ldots,n_k} \right\} \right)_{\ell\in\mathbb{N}}$$

is a decreasing sequence of nonempty subsets of $\mathbb{N}^{\mathbb{N}}$, each of which belongs to the collection A_{ds}, where

$$A := \{ \{ n \in \mathbb{N}^{\mathbb{N}} \mid n_p = q \} \mid p,q \in \mathbb{N} \} \ .$$

Now A is semicompact, as is easily seen, and A_{ds} is semicompact as well, because $A_{ds} = A_{sd} \subset A_{s\delta}$. The sequence $(*)$ therefore has a nonempty intersection, i.e. there exists an $n \in \mathbb{N}^{\mathbb{N}}$ such that

$$\forall_{\ell\in\mathbb{N}} \ x \in \underset{k\le\ell}{\cap} A_{n_1,\ldots,n_k} \ ,$$

and therefore such that

$$x \in \underset{k\in\mathbb{N}}{\cap} A_{n_1,\ldots,n_k} \ .$$

So $x \in A$, and the inclusion has been proved.

Now let μ be a probability on E and let $a \in \mathbb{R}$. Suppose that

$$(1) \qquad \mu^*(A) \ge a \ ,$$

where μ^* is the outer measure corresponding to μ. Then, as will be

28

demonstrated, for every $i \in \mathbb{N}$ there exists an $m \in \mathbb{N}^{\mathbb{N}}$ such that for every $\ell \in \mathbb{N}$

$$\mu^*(\cup \{ \bigcap_{k \in \mathbb{N}} A_{n_1,\ldots,n_k} \mid n \in \mathbb{N}^{\mathbb{N}} \text{ with } (n_1,\ldots,n_\ell) \leq (m_1,\ldots,m_\ell)\}) > \mu^*(A) - \frac{1}{i} ,$$

so

$$(2) \qquad \forall_{i \in \mathbb{N}} \, \exists_{m \in \mathbb{N}^{\mathbb{N}}} \, \forall_{\ell \in \mathbb{N}} \, [\mu(B_{m_1,\ldots,m_\ell}) > a - \frac{1}{i}] .$$

The existence of such an m for every i follows, by induction on ℓ, from the continuity of μ^* on increasing sequences of sets, and from the relations

$$\cup \{ \bigcap_{k \in \mathbb{N}} A_{n_1,\ldots,n_k} \mid n \in \mathbb{N}^{\mathbb{N}} \text{ with } (n_1,\ldots,n_{\ell-1}) \leq (m_1,\ldots,m_{\ell-1})\} =$$

$$= \lim_{m_\ell \to \infty} \cup \{ \bigcap_{k \in \mathbb{N}} A_{n_1,\ldots,n_k} \mid n \in \mathbb{N}^{\mathbb{N}} \text{ with } (n_1,\ldots,n_\ell) \leq (m_1,\ldots,m_\ell)\} .$$

Statement (2) is equivalent to

$$(2') \qquad \mu \in \bigcap_{i \in \mathbb{N}} \, \bigcup_{m \in \mathbb{N}^{\mathbb{N}}} \, \bigcap_{\ell \in \mathbb{N}} \, \{\nu \in \widetilde{E} \mid \nu(B_{m_1,\ldots,m_\ell}) > a - \frac{1}{i}\} .$$

For each $m \in \mathbb{N}^{\mathbb{N}}$ the sequence $(B_{m_1,\ldots,m_\ell})_{\ell \in \mathbb{N}}$ is decreasing and therefore

$$\mu(\bigcap_{\ell \in \mathbb{N}} B_{m_1,\ldots,m_\ell}) = \lim_{\ell \to \infty} \mu(B_{m_1,\ldots,m_\ell}) .$$

Together with (2) this implies

$$\forall_{i \in \mathbb{N}} \, \exists_{m \in \mathbb{N}^{\mathbb{N}}} \, [\mu(\bigcap_{\ell \in \mathbb{N}} B_{m_1,\ldots,m_\ell}) \geq a - \frac{1}{i}] ,$$

so

$$\sup_{m \in \mathbb{N}^{\mathbb{N}}} \mu(\bigcap_{\ell \in \mathbb{N}} B_{m_1,\ldots,m_\ell}) \geq a .$$

As $A \supset \bigcap_{\ell \in \mathbb{N}} B_{m_1,\ldots,m_\ell}$ for every $m \in \mathbb{N}^{\mathbb{N}}$, we can conclude

$$(3) \qquad \mu_*(A) \geq \sup_{m \in \mathbb{N}^{\mathbb{N}}} \mu(\bigcap_{\ell \in \mathbb{N}} B_{m_1, \ldots, m_\ell}) \geq a \, ,$$

where μ_* is the inner measure corresponding to μ.

Now (3) is easily seen to imply (1). So, for every $\mu \in \widetilde{E}$ and $a \in \mathbb{R}$, the statements (1), (2), (2') and (3) are equivalent. From these equivalences we shall now derive the proofs for the Propositions 4.6, 4.8, and 5.2.

i) Taking $a := \mu^*(A)$ in the equivalent statements (1) and (3), we get $\mu_*(A) \geq \mu^*(A)$. So A belongs to E_μ. As μ is arbitrary, this implies that A is universally measurable. The arbitrariness of A, in turn, implies that $S(E) \subset U(E)$. Applying this inclusion to the space $(E, U(E))$ instead of (E, E) we get $SU(E) \subset UU(E)$, so $SU(E) = U(E)$. Thus the proof of Proposition 4.6 is complete. Note that the existence of a collection C, as mentioned in Proposition 5.2, does not restrict the space (E, E), since we always have $E \subset E \subset S(E)$.

ii) The equivalence of (1) and (2') together with the universal measurability of A implies that the set $\{\nu \in \widetilde{E} \mid \nu(A) \geq a\}$ equals the right-hand side of (2'), which is a Souslin subset of $\widetilde{E}$. From the arbitrariness of a we conclude that

$$\{\nu \in \widetilde{E} \mid \nu(A) > b\} = \bigcup_{i \in \mathbb{N}} \{\nu \in \widetilde{E} \mid \nu(A) \geq b + \tfrac{1}{i}\} \in S(\widetilde{E})$$

for every $b \in \mathbb{R}$. So $\nu \mapsto \nu(A)$ is a Souslin function, thus proving Proposition 4.8.

iii) To prove Proposition 5.2 we take $a := \mu^*(A)$ once more. It then follows from the equivalence of (1) and (3) that

$$\mu(A) = \sup_{m \in \mathbb{N}^{\mathbb{N}}} \mu(\bigcap_{\ell \in \mathbb{N}} B_{m_1, \ldots, m_\ell}) \, .$$

Now C is a subclass of E such that $E \subset S(C)$, and hence, such that $A \subset S(E) \subset SS(C) = S(C)$. The sets $A_{n_1, \ldots, n_k}$ may therefore be supposed to belong to C, and the sets $\bigcap_{\ell \in \mathbb{N}} B_{m_1, \ldots, m_\ell}$ can be assumed to belong to $C_{ds\delta}$. Since $C_{ds\delta} = C_{sd\delta} = C_{s\delta}$, the equality appearing in the proposition has been proved for Souslin sets A and, consequently, for measurable sets A. For universally measurable sets A the equality follows from this and the mere definition of universal measurability. $\square$

DEFINITION. A positive function μ defined on a collection A of sets is called *additive* (σ-*additive*) when, for each finite (countable) subclass A_0 of A consisting of pairwise disjoint sets, and such that $\cup\, A_0 \in A$, the equality $\mu(\cup\, A_0) = \sum_{A \in A_0} \mu(A)$ holds.

When μ is additive then obviously $\mu(\emptyset) = 0$.

PROPOSITION 5.3. *Let A be a semicompact algebra of subsets of a set. Then every positive additive function on A is σ-additive.*

PROOF. Let $\{A_n \mid n \in \mathbb{N}\}$ be a countable collection of pairwise disjoint members of A such that $\cup_{n \in \mathbb{N}} A_n \in A$. Then the collection $\{\cup_{n=m}^{\infty} A_n \mid m \in \mathbb{N}\}$ has empty intersection. As a subcollection of A the collection is semicompact, and consequently $\cup_{n=m}^{\infty} A_n = \emptyset$ for some $m \in \mathbb{N}$, which in turn implies that only finitely many members of $\{A_n \mid n \in \mathbb{N}\}$ are nonempty. The foregoing implies that additivity and σ-additivity are equivalent for positive functions defined on A. $\qquad\Box$

Next we introduce the auxiliary space $\mathbb{D}$, the so-called Cantor space. The role played by $\mathbb{D}$ strongly resembles the role of the space $\mathbb{R}$ in many measure-theoretic arguments. In fact, $\mathbb{D}$ and $\mathbb{R}$ can be shown to be isomorphic measurable spaces ([Bertsekas & Shreve] Proposition 7.16). The space $\mathbb{D}$ however is better suited to our needs.

DEFINITION. The measurable space $\mathbb{D}$ is the product space $\Pi_{n \in \mathbb{N}} (D_n, \mathcal{D}_n)$, where, for each $n \in \mathbb{N}$, $D_n := \{0,1\}$ and $\mathcal{D}_n$ is the σ-algebra consisting of all (four) subsets of D_n.

Note that for any countable set I the spaces $\mathbb{D}^I$ and $\mathbb{D}$ are isomorphic. As another useful property of $\mathbb{D}$ we have:

PROPOSITION 5.4. *The σ-algebra of $\mathbb{D}$ is generated by a countable semicompact algebra.*

PROOF. Let $C := \{\{x \in \mathbb{D} \mid x_i = j\} \mid i \in \mathbb{N}, j = 0,1\}$. Then C is generating, countable and semicompact. By Proposition 5.1 the same holds for C_{sd}. As C is closed under complementation, the collection C_{sd} is an algebra. $\qquad\Box$

§ 6. Measurability of integrals

A well-known theorem in measure theory says that the integral $\int f(x,y)\mu(dy)$ of a measurable function of two variables is a measurable function of x. In this section we shall prove that this integral depends measurably on μ as well, and that similar results are valid for universally measurable functions and Souslin functions. Before proving these measurability properties we introduce a generalization of the integral, which will turn out to be convenient later on.

A universally measurable function f defined on a measurable space E can be written in the form $f^+ - f^-$, where f^+ and f^- are the positive universally measurable functions on E defined by

$$f^+(x) = \max \{f(x),0\} \quad \text{and} \quad f^-(x) = \max \{-f(x),0\} \ .$$

When μ is a probability on E then f is called *quasi-integrable* with respect to μ when at least one of the integrals $\int f^+ d\mu$ and $\int f^- d\mu$ is finite. In this case one defines

$$\int f\,d\mu := \int f^+ d\mu - \int f^- d\mu \ .$$

As is easily seen, a universally measurable function defined on a measurable space is quasi-integrable with respect to every probability on that space only if it is bounded from above or from below. In order not to be forced to consider bounded functions only or to demand quasi-integrability in advance every time, we generalize the integral concept such as to be applicable even to functions that are not quasi-integrable.

DEFINITION. For a universally measurable function f and a probability μ on a measurable space we define $\underline{\int} f\,d\mu$ to be equal to $\int f\,d\mu$ if f is quasi-integrable with respect to μ and equal to $-\infty$ in the other case.

Using the convention $\infty - \infty = -\infty$ and $\pm \infty \cdot 0 = 0$ (see Preliminaries 14) we conclude that $\underline{\int} f\,d\mu$ depends on f in an additive and positively homogeneous manner:

$$\underline{\int} (f+g)d\mu = \underline{\int} f\,d\mu + \underline{\int} g\,d\mu \quad \text{and} \quad \underline{\int} \lambda f\,d\mu = \lambda \underline{\int} f\,d\mu \quad \text{for } \lambda > 0 \ .$$

In particular we have

$$\int_{-} f \, d\mu = \int_{-} f^{+} d\mu - \int^{-} f^{-} d\mu = \int^{-} f^{+} d\mu - \int_{-} f^{-} d\mu \ .$$

In general however, $\int -f \, d\mu$ differs from $- \int f \, d\mu$. Also, Fubini's theorem cannot be generalized.

We now turn to the measurability properties of the integral defined above.

PROPOSITION 6.1. *Let f be a universally measurable (or measurable, or Souslin) function on a measurable space E. Then the function $\mu \mapsto \int f \, d\mu$ on $\widetilde{E}$ is universally measurable (or measurable, or Souslin, respectively).*

PROOF. We prove the statement on Souslin functions. We first consider the case that f is positive. Then f is the limit of the increasing sequence

$$\left(2^{-n} \sum_{m=1}^{\infty} 1_{\{f > m \, 2^{-n}\}} \right)_{n \in \mathbb{N}}$$

of positive functions, so

$$\int f \, d\mu = \lim_{n} 2^{-n} \sum_{m=1}^{\infty} \mu\{f > m \, 2^{-n}\} \quad \text{for each } \mu \in \widetilde{E} \ .$$

For each $m, n \in \mathbb{N}$ the set $\{f > m \, 2^{-n}\}$ is a Souslin subset of E, which implies by Proposition 4.8 that $\mu\{f > m \, 2^{-n}\}$ is a Souslin function of μ. Applying Proposition 4.7 we conclude that $\int f \, d\mu$ also is a Souslin function of μ. As any bounded function differs only a constant from a positive one this result also holds for bounded Souslin functions f.

Now let f be an arbitrary Souslin function, and for each $m, n \in \mathbb{N}$ let the function f_{m}^{n} be defined on E by

$$f_{m}^{n}(x) := \min \{n, \max\{-m, f(x)\}\} \ .$$

Note that f_{m}^{n} is the function f truncated at the values $+n$ and $-m$. For each $m, n \in \mathbb{N}$, f_{m}^{n} is a bounded Souslin function and

$$\int_{-} f \, d\mu = \lim_{n} \lim_{m} \int f_{m}^{n} d\mu \ .$$

It now follows from Proposition 4.7 that $\int f \, d\mu$ is a Souslin function of μ.

The statement on (universally) measurable functions f can be proved in a similar way: instead of Proposition 4.8 use the fact that $\mu(A)$ depends (universally) measurable on μ for every (universally) measurable set A, i.e. Proposition 3.6 and the definition of $\tilde{E}$. $\qquad\qquad \square$

PROPOSITION 6.2. *Let* E *and* F *be measurable spaces and* f *a universally measurable (or measurable, or Souslin) function on* E × F. *Then the functions* $y \mapsto f(x,y)$ *(x ∈ E) on* F *and the function* $(x,\mu) \mapsto \int f(x,y)\mu(dy)$ *on* E × $\tilde{F}$ *are universally measurable (or measurable, or Souslin, respectively).*

PROOF. We prove the statement on Souslin functions; the other cases can be treated in a similar way.

Let x ∈ E. Then the mapping $y \mapsto (x,y)$ of F into E × F is measurable, because its coordinates $y \mapsto x$ and $y \mapsto y$ are measurable. Now $y \mapsto f(x,y)$ is the composition of this measurable mapping and the Souslin function f and therefore it is a Souslin function itself.

For every x ∈ E, $\mu \in \tilde{F}$ we have

$$\int f^+(x,y) \, \mu(dy) = \iint f^+(x',y) \, \delta_x(dx')\mu(dy) = \int f^+ \, d(\delta_x \times \mu)$$

and similarly for f^-. Consequently $\int f(x,y)\mu(dy)$ equals $\int f \, d(\delta_x \times \mu)$, which is a Souslin function of $\delta_x \times \mu$ by Proposition 6.1. The result now follows from the fact that $\delta_x \times \mu$ depends measurably on (x,μ) (see the examples 1 and 5 following Proposition 2.2), and that the composition of a measurable mapping and a Souslin function is a Souslin function. $\qquad \square$

CHAPTER II
ANALYTIC SPACES

We have now arrived at the main topic of this monograph: a measure-
theoretic treatment of analytic spaces. The importance of analytic topo-
logical spaces (for a definition see [Hoffmann-Jørgensen], Ch. III § 1) in
dynamic programming is due to certain properties of the σ-algebras generated
by their topologies. In our measure-theoretic approach we have taken these
properties as a starting point, and, in fact, we have chosen one of them as
the defining property of the class of measurable spaces to be studied. This
class, therefore, is an extension of the class of analytic topological
spaces as far as the measure-theoretic structure is concerned.

The facts on analytic spaces collected in the first section of this
chapter should give the reader a good impression of such spaces. But, since
we have always kept in mind the applications of chapter III, the treatment
of the subject can only be considered complete in connection with these
applications. In the second section, the one on separating classes, count-
ably generated analytic spaces are considered, and their relation to topo-
logical analytic spaces is made clear. In the third and final section of
this chapter probabilities on analytic spaces are treated, and, in particu-
lar, sets of these probabilities that will play a role in the applications.

§ 7. Analytic spaces

DEFINITION. A measurable space F is called *analytic* if for every measurable
space E and every Souslin subset S of E × F:

i)   The projection S_E of S on E is a Souslin subset of E,

ii) S contains the graph of a universally measurable mapping of the
     subspace S_E of E into F.

The reader should note that, in the above definition, it is not the
graph of the mapping that is supposed to be universally measurable but the
mapping itself.

A σ-algebra E of subsets of a set E is called analytic, when the
space (E, E) is analytic.

The usefulness of analytic spaces for dynamic programming is mainly due to the property expressed in the following proposition, as will become evident in Chapter III.

PROPOSITION 7.1 (Exact selection theorem). *Let* E *be a measurable space,* F *an analytic space,* S *a Souslin subspace of* E × F *every section* $S_x := \{y \in F \mid (x,y) \in S\}$ *of which is nonempty, and* f *a Souslin function on* S. *Further let* $g: E \to \overline{\mathbb{R}}$ *be defined by* $g(x) = \sup_{y \in S_x} f(x,y)$ *and let*

$$T := \{x \in E \mid \exists_{y \in S_x} \; g(x) = f(x,y)\} \; .$$

Then g *is a Souslin function and* T *is universally measurable.*

Moreover, for every universally measurable function h *on* E *such that*

$$h(x) \begin{cases} < g(x) & \textit{if } -\infty < g(x), \\ = g(x) & \textit{else,} \end{cases}$$

a universally measurable mapping $\varphi: E \to F$ *exists the graph of which is contained in* S *and such that for each* $x \in E$

$$f(x,\varphi(x)) \begin{cases} = g(x) & \textit{if } x \in T, \\ > h(x) & \textit{else.} \end{cases}$$

PROOF. Let $\mathcal{E}$ and $\mathcal{F}$ be the σ-algebras of E and F, respectively. For every $a \in \mathbb{R}$ the set $\{x \in E \mid g(x) > a\}$ is the projection on E of the Souslin subset $\{(x,y) \in S \mid f(x,y) > a\}$ of E × F, and therefore it is a Souslin subset of E, since F is analytic. This implies that g is a Souslin function.

Next, let

$$A := \{(x,y) \in S \mid f(x,y) = g(x)\}$$

and

$$B := \{(x,y) \in S \mid f(x,y) > h(x)\} \; .$$

Then

$$A = S \cap [\bigcap_{r \in Q} (\{r < f\} \cup (\{g \leq r\} \times F))]$$

and

$$B = S \cap [\underset{r \in Q}{\cup} (\{r<f\} \cap (\{h \leq r\} \times F))] \, ,$$

where Q is the set of rational numbers. Now, for each $r \in Q$ we have

$$\{r<f\} \in S(E \otimes F) \subset S(U(E) \otimes F) \, ,$$

$$\{g \leq r\} \times F \in U(E) \times F \in S(U(E) \otimes F) \, ,$$

and similarly for $\{h \leq r\} \times F$. Consequently, A and B belong to $S(U(E) \otimes F)$.

We now consider the product space $(E, U(E)) \times (F, F)$. From the analyticity of F we conclude that the projection A_E of A on E is universally measurable with respect to $U(E)$ and that there exists a mapping α of A_E into F that is universally measurable with respect to $U(E) | A_E$ and whose graph is contained in A. Similarly, the projection B_E of B on E is universally measurable with respect to $U(E)$ and there exists a mapping β of B_E into F that is universally measurable with respect to $U(E) | B_E$ and whose graph is contained in B. Moreover, it follows from the definitions of A and B that $A_E \cup B_E = E$.

Now $T = A_E$, so T is universally measurable with respect to $U(E)$ and therefore with respect to E by Proposition 3.2.

Finally, the mapping $\varphi : E \to F$ defined by

$$\varphi(x) \quad \begin{array}{ll} = \alpha(x) & \text{if } x \quad A_E \\ = \beta(x) & \text{else} \end{array}$$

is easily seen to have the desired properties. □

For f a Souslin function defined on $E \times F$ taking the values 0 and 1 only, the exact selection theorem reduces to the definition of analyticity of F. Consequently, the validity of the exact selection theorem characterizes analytic spaces.

We shall show that the class of analytic spaces is closed under the constructions for measurable spaces most commonly used and that it therefore contains many of the measurable spaces encountered in practice. To begin with, we shall prove that the space $\mathbb{D}$, introduced in section 5, is analytic.

PROPOSITION 7.2. *The space* $\mathbb{D}$ *is analytic.*

PROOF. Let E be a measurable space and S a Souslin subset of $E \times \mathbb{D}$. For each $k \in \mathbb{N}$ let the partition P_k of $\mathbb{D}$ be defined by

$$P_k := \{\{y \in \mathbb{D} \mid y_i = n_i \ (i = 1,\ldots,k)\} \mid n \in \{0,1\}^k\}$$

and let $P := \cup_{k \in \mathbb{N}} P_k$. Then, for each $k \in \mathbb{N}$ and for each $P \in P_k$, P^c is the union of $2^k - 1$ members of P_k. So $P_c \subset P_s \subset S(P)$. Since, moreover, the σ-algebra $\mathcal{D}$ of $\mathbb{D}$ is generated by P, we have $\mathcal{D} \subset S(P)$ as a consequence of Proposition 4.1. When E is the σ-algebra of E, then, by Proposition 4.5, $E \otimes \mathcal{D} \subset S(E \times P)$ and, therefore, $S(E \otimes \mathcal{D}) \subset S(E \times P)$. Consequently, by the definition of S the set S can be written as

$$S = \bigcup_{n \in \mathbb{N}^{\mathbb{N}}} \bigcap_{k \in \mathbb{N}} (A_{n_1,\ldots,n_k} \times B_{n_1,\ldots,n_k}) \ ,$$

where the A's are measurable subsets of E and the B's belong to P.

For each n and k the set $B_{n_1,\ldots,n_k}$ is contained in a member of P_k or equal to a finite union of members of P_k; in any case $B_{n_1,\ldots,n_k}$ can be written as a countable union, say $\cup_{\ell \in \mathbb{N}} B^\ell_{n_1,\ldots,n_k}$, of members of P each of which is contained in some member of P_k. Hence we have

$$S = \bigcup_{n \in \mathbb{N}^{\mathbb{N}}} \bigcap_{k \in \mathbb{N}} \left(A_{n_1,\ldots,n_k} \times \bigcup_{\ell \in \mathbb{N}} B^\ell_{n_1,\ldots,n_k} \right) .$$

From this we deduce:

$$S = \bigcup_{n \in \mathbb{N}^{\mathbb{N}}} \bigcap_{k \in \mathbb{N}} \bigcup_{\ell \in \mathbb{N}} \left(A_{n_1,\ldots,n_k} \times B^\ell_{n_1,\ldots,n_k} \right) =$$

$$= \bigcup_{n \in \mathbb{N}^{\mathbb{N}}} \bigcup_{m \in \mathbb{N}^{\mathbb{N}}} \bigcap_{k \in \mathbb{N}} \left(A_{n_1,\ldots,n_k} \times B^{m_k}_{n_1,\ldots,n_k} \right) =$$

$$= \bigcup_{n \in \mathbb{N}^{\mathbb{N}}} \bigcap_{k \in \mathbb{N}} \left(A_{\alpha(n_1),\ldots,\alpha(n_k)} \times B^{\beta(n_k)}_{\alpha(n_1),\ldots,\alpha(n_k)} \right) ,$$

where $n \mapsto (\alpha(n),\beta(n))$ is some surjection of $\mathbb{N}$ onto $\mathbb{N} \times \mathbb{N}$. Hence

$$S = \bigcup_{n \in \mathbb{N}^{\mathbb{N}}} \bigcap_{k \in \mathbb{N}} (A'_{n_1,\ldots,n_k} \times B'_{n_1,\ldots,n_k}) ,$$

where, for each n and k, $A'_{n_1,\ldots,n_k}$ is a measurable subset of of E, and $B'_{n_1,\ldots,n_k}$ is a member of P that is contained in some member of P_k. This

38

implies that for each $n \in \mathbb{N}^{\mathbb{N}}$ $\bigcap_{k \in \mathbb{N}} B'_{n_1,\ldots,n_k}$ contains at most one point of $\mathbb{D}$.

For each n and k we now define

$$A''_{n_1,\ldots,n_k} \quad \begin{cases} = A'_{n_1,\ldots,n_k} & \text{if } \bigcap_{\ell=1}^{k} B'_{n_1,\ldots,n_\ell} \neq \emptyset , \\ = \emptyset & \text{else.} \end{cases}$$

Then

$$S = \bigcup_{n \in \mathbb{N}^{\mathbb{N}}} \bigcap_{k \in \mathbb{N}} (A''_{n_1,\ldots,n_k} \times B'_{n_1,\ldots,n_k}) .$$

Moreover, as P is semicompact, for each $n \in \mathbb{N}^{\mathbb{N}}$ we have the implications:

$$\bigcap_{k \in \mathbb{N}} A''_{n_1,\ldots,n_k} \neq \emptyset \;\Rightarrow\; \forall_{k \in \mathbb{N}} \; A''_{n_1,\ldots,n_k} \neq \emptyset \;\Rightarrow\;$$

$$\forall_{k \in \mathbb{N}} \; \bigcap_{\ell=1}^{k} B'_{n_1,\ldots,n_\ell} \neq \emptyset \;\Rightarrow\; \bigcap_{k \in \mathbb{N}} B'_{n_1,\ldots,n_k} \neq \emptyset .$$

We now show that the projection S_E of S on E is a Souslin subset of E. Let $x \in S_E$. Then there is some $y \in \mathbb{D}$ such that $(x,y) \in S$ and hence

$$x \in \bigcup_{n \in \mathbb{N}^{\mathbb{N}}} \bigcap_{k \in \mathbb{N}} A''_{n_1,\ldots,n_k} .$$

On the other hand, when x belongs to the last set, then there exists $n \in \mathbb{N}^{\mathbb{N}}$ such that $x \in \bigcap_{k \in \mathbb{N}} A''_{n_1,\ldots,n_k}$, which implies $\bigcap_{k \in \mathbb{N}} B'_{n_1,\ldots,n_k} \neq \emptyset$, and hence, the existence of $y \in \mathbb{D}$ with $y \in \bigcap_{k \in \mathbb{N}} B'_{n_1,\ldots,n_k}$. Then

$$(x,y) \in \bigcap_{k \in \mathbb{N}} (A''_{n_1,\ldots,n_k} \times B'_{n_1,\ldots,n_k}) \subset S ,$$

so the point x belongs to S_E. The foregoing implies that S_E equals $\bigcup_n \bigcap_k A''_{n_1,\ldots,n_k}$ and, consequently, S_E is a Souslin subset of E.

Next we show that S contains the graph of a universally measurable mapping $\varphi\colon S_E \to \mathbb{D}$. For each $n \in \mathbb{N}^{\mathbb{N}}$ and $k \in \mathbb{N}$ we define

$$A^*_{n_1,\ldots,n_k} := \Big(\bigcup_{m \in \mathbb{N}^{\mathbb{N}}} \bigcap_{\ell \in \mathbb{N}} A''_{n_1,\ldots,n_k,m_1,\ldots,m_\ell} \Big) \cap A''_{n_1,\ldots,n_k} .$$

Then

$$S_E \subset \bigcup_{i \in \mathbb{N}} A_i^* \quad \text{and} \quad A_{n_1, \ldots, n_k}^* \subset \bigcup_{i \in \mathbb{N}} A_{n_1, \ldots, n_k, i}^*$$

for each $n \in \mathbb{N}^{\mathbb{N}}$ and $k \in \mathbb{N}$. Now let $x \in S_E$ and let $p \in \mathbb{N}^{\mathbb{N}}$ be such that, for each $k \in \mathbb{N}$, p_k is the smallest positive integer for which $x \in A_{p_1, \ldots, p_k}^*$. Then

$$\bigcap_{k \in \mathbb{N}} A_{p_1, \ldots, p_k}^{\prime\prime} \supset \bigcap_{k \in \mathbb{N}} A_{p_1, \ldots, p_k}^* \supset \{x\} \neq \emptyset$$

and hence $\bigcap_{k \in \mathbb{N}} B_{p_1, \ldots, p_k}^{\prime} \neq \emptyset$. This implies that the last intersection contains precisely one point, say $\varphi(x)$, of $\mathbb{D}$. Moreover, we have

$$(x, \varphi(x)) \in \bigcap_{k \in \mathbb{N}} A_{p_1, \ldots, p_k}^{\prime\prime} \times B_{p_1, \ldots, p_k}^{\prime} \subset S .$$

The graph of the mapping $\varphi: S_E \to \mathbb{D}$ defined above is therefore contained in S.

To prove the universal measurability of φ, let $k \in \mathbb{N}$ and $B \in P_k$. Let N be the set of those $n \in \mathbb{N}^k$ for which $B_{n_1, \ldots, n_k}^{\prime} \subset B$. Then the following four statements are equivalent:

$$\varphi(x) \in B , \quad B_{p_1, \ldots, p_k}^{\prime} \subset B , \quad (p_1, \ldots, p_k) \in N ,$$

$$x \in \bigcup_{n \in N} \bigcap_{\ell=1}^{k} \left(A_{n_1, \ldots, n_\ell}^* \setminus \bigcup_{i=1}^{n_\ell - 1} A_{n_1, \ldots, n_{\ell-1}, i}^* \right) .$$

As each A^* is a Souslin subset of E and N is countable, $\varphi^{-1} B$ belongs to the σ-algebra generated by the Souslin subsets of S_E and, hence, is universally measurable. Now the result follows from $\mathcal{D} = \sigma(\bigcup_{k \in \mathbb{N}} P_k)$. $\square$

The analyticity of a large class of measurable spaces will be deduced from the analyticity of $\mathbb{D}$. Material in these deductions is Proposition 7.5, where mappings of analytic spaces into $\mathbb{D}$ are considered. First we give some preparatory results on mappings in general.

PROPOSITION 7.3. *Let E and F be measurable spaces, let F be countably separated and let φ: E $\to$ F be measurable. Then the graph of φ is a measurable subset of* E $\times$ F.

PROOF. Let A be a countable collection of measurable subsets of F that separates the points of F. Then for every $(x,y) \in$ E $\times$ F we have:

$$(x,y) \notin \text{graph } \varphi \quad \Leftrightarrow \quad y \neq \varphi(x) \quad \Leftrightarrow$$

$$\exists_{A\in A} \; [(\varphi(x) \in A \land y \notin A) \lor (\varphi(x) \notin A \land y \in A)] \quad \Leftrightarrow$$

$$(x,y) \in \bigcup_{A\in A} \; [(\varphi^{-1}A \times A^c) \cup ((\varphi^{-1}A)^c \times A)] \; .$$

So the graph of φ is the complement of a countable union of measurable subsets of E $\times$ F, and therefore is measurable itself. □

DEFINITION. Let (E,E) and (F,F) be measurable spaces and let φ: E $\to$ F. A *right inverse* of φ is a mapping ψ: φE $\to$ E such that $\varphi \circ \psi$ is the identity on φE. The mapping φ is called *strictly measurable* if $E = \varphi^{-1}F$. The mapping φ is called an *isomorphic embedding* when it is strictly measurable and injective.

Note that every mapping has a right inverse, but that this inverse is unique only when the mapping is injective.

PROPOSITION 7.4.

i)   *Let (E,E) and (F,F) be measurable spaces and let φ: E $\to$ F be strictly measurable. Then $S(E) = \varphi^{-1}S(F)$ and every right inverse of φ is measurable with respect to $F|(\varphi E)$.*

ii) *Let E be a countably generated measurable space. Then there exists a strictly measurable mapping φ: E $\to$ $\mathbb{D}$.*

PROOF.

i) $S(E) = S(\varphi^{-1}F) = \varphi^{-1}S(F)$ because of Proposition 4.3. Let ψ be a right inverse of φ. For every measurable subset A of E there exists a measurable subset B of F such that $A = \varphi^{-1}B$ and, hence, such that

$$\psi^{-1}A = \psi^{-1}(\varphi^{-1}B) = (\varphi \circ \psi)^{-1}B = B \cap (\varphi E) \ .$$

Consequently, ψ is measurable with respect to $F|(\varphi E)$.

ii)   Let $\{C_n \mid n \in \mathbb{N}\}$ be a countable collection generating the σ-algebra E of E. Moreover, let $\mathcal{D}$ be the σ-algebra of $\mathbb{D}$, and for each $n \in \mathbb{N}$ let $D_n := \{y \in \mathbb{D} \mid y_n = 1\}$. Define $\varphi: E \to \mathbb{D}$ by $(\varphi(x))_n := 1_{C_n}(x)$ $(n \in \mathbb{N}, x \in E)$. Then for every $n \in \mathbb{N}$ we have $C_n = \varphi^{-1}D_n$ and consequently $\sigma\{C_n \mid n \in \mathbb{N}\} = \varphi^{-1}\sigma\{D_n \mid n \in \mathbb{N}\}$ or equivalently $E = \varphi^{-1}\mathcal{D}$. So φ is strictly measurable. $\square$

The characterization of analytic spaces given in the following proposition is a useful alternative to the defining one. It will be used frequently in the rest of this chapter.

PROPOSITION 7.5. *A measurable space* F *is analytic iff for every measurable mapping* $\varphi: F \to \mathbb{D}$ *the range of* φ *is a Souslin subset of* $\mathbb{D}$ *and* φ *has a universally measurable right inverse.*

PROOF. Let F be analytic and $\varphi: F \to \mathbb{D}$ measurable. As $\mathbb{D}$ is countably separated, the graph of φ is a measurable subset of $F \times \mathbb{D}$ by Proposition 7.3 and, hence, a Souslin subset of $F \times \mathbb{D}$. The range of φ is the projection on $\mathbb{D}$ of the graph of φ and therefore is a Souslin subset of $\mathbb{D}$. Moreover, the graph of φ contains the graph of a universally measurable mapping of φF into F, which obviously is a right inverse of φ.

Conversely, let (F, F) be a measurable space having the property stated in the proposition. We shall show that F obeys the definition of analyticity. To this end let (E, E) be a measurable space, and let S be a Souslin subset of $E \times F$. Then by Propositions 4.2 and 4.5 we have $S \in S(E \otimes F) \subset SS(E \times F) = S(E \times F)$, i.e. S is the kernel of a Souslin scheme on $E \times F$. As each Souslin scheme involves only countably many sets, we even have $S \in S(E \times C)$ for some countable subclass C of F.

By Proposition 7.4 a mapping $\varphi: F \to \mathbb{D}$ exists that is strictly measurable with respect to $\sigma(C)$, and, therefore, measurable with respect to F. When $\psi: E \times F \to E \times \mathbb{D}$ is defined by $\psi(x,y) = (x, \varphi(y))$, then ψ is easily seen to be strictly measurable with respect to $E \otimes \sigma(C)$. It now follows from Proposition 7.4 that a Souslin subset S' of $E \times \mathbb{D}$ exists such that $S = \psi^{-1}S'$, and hence, such that $\psi S = S' \cap (E \times \varphi F)$. The properties of F imply that φF is a Souslin subset of $\mathbb{D}$. So $E \times \varphi F$ is a Souslin subset of

E × ID and this in turn implies that ψS is the intersection of two Souslin subsets of E × ID and, consequently, a Souslin set itself.

From the definition of ψ it now follows that the projection S_E of S on E equals the projection of ψS on E. The latter set, however, is a Souslin subset of E, due to the analyticity of ID.

It also follows from the analyticity of ID that there exists a universally measurable mapping α: S_E → ID the graph of which is contained in ψS. Further, by the properties of F, there exists a universally measurable right inverse β of φ. The mapping β ∘ α: S_E → F therefore is universally measurable, and for every x ∈ S_E we have

$$\psi(x, (\beta \circ \alpha)x) = (x, (\varphi \circ \beta \circ \alpha)x) = (x, \alpha x) \in \psi S \subset S' .$$

Hence,

$$\text{graph}(\beta \circ \alpha) \subset \psi^{-1} S' = S .$$

The foregoing implies that F satisfies the defining conditions for analyticity. □

We are now able to prove the stability properties of the class of analytic spacec announced earlier.

PROPOSITION 7.6. *The class of analytic spaces is closed under the formation of*
i)   *measurable images,*
ii) *Souslin subspaces,*
iii) *product spaces.*

PROOF.

i)   Let a measurable space F be the image of an analytic space E under a measurable mapping ψ. We shall prove that the characterization of analytic spaces given in the preceeding proposition applies to F. So, let φ be a measurable mapping of F into ID. Then φ ∘ ψ is a measurable mapping of the analytic space E into ID and, by the preceeding proposition, its range is a Souslin subset of ID, while it has a universally measurable right inverse χ. As the range of φ equals the range of φ ∘ ψ and as ψ ∘ χ is a right inverse of φ, the range of φ is a Souslin subset of ID and φ has a universally measurable right

inverse. From this and the arbirariness of φ it follows that F is analytic.

ii) To prove the statement on Souslin subspaces we use the definition of analyticity. Let F be a Souslin subspace of an analytic space F'. Further, let E be a measurable space and let S be a Souslin subset of E × F. Then, by Proposition 4.4, S is the intersection of E × F and some Souslin subset of E × F'. Now E × F itself is a Souslin subset of E × F', so S is a Souslin subset of E × F' as well. From the analyticity of F' we deduce that the projection S_E of S on E is a Souslin subset of E, and that S contains the graph of a universally measurable mapping of S_E into F'. The range of this mapping is contained in F and it can, therefore, be considered as a universally measurable mapping of S_E into F.

The foregoing implies that F is analytic.

iii) Let $(F_i, F_i)_{i \in I}$ be a family of analytic spaces and let $F := \Pi_{i \in I} F_i$. For each measurable mapping $\varphi: F \to \mathbb{D}$ we shall show that the range of φ is a Souslin subset of $\mathbb{D}$ and that φ has a universally measurable right inverse. These facts imply that F is analytic.

Since the σ-algebra $\mathcal{D}$ of $\mathbb{D}$ is countably generated, so is the sub-σ-algebra $\varphi^{-1}\mathcal{D}$ of F. So $\varphi^{-1}\mathcal{D}$ is contained in a sub-σ-algebra of F that is generated by countably many rectangles in $\Pi_{i \in I} F_i$. Consequently, there exists a countable subset J of I and, for each $i \in I$, a countably generated sub-σ-algebra F_i^0 of F_i such that $\forall_{i \in I \setminus J} F_i^0 = \{\emptyset, F_i\}$ and such that $\varphi^{-1}\mathcal{D} \subset \otimes_{i \in I} F_i^0$.

We shall now show that the mapping $\varphi: F \to \mathbb{D}$ can be decomposed into a mapping $\psi: F \to \mathbb{D}^J$ and a mapping $\chi: \psi F \to \mathbb{D}$. To this end let, for each $j \in J$, the mapping $\psi_j: F_j \to \mathbb{D}$ be strictly measurable with respect to F_j^0 (such mappings exist by Proposition 7.4) and let $\psi: F \to \mathbb{D}^J$ be defined by $\psi(x)_j := \psi_j(x_j)$ $(j \in J, x \in F)$. Now suppose that x and x' are points of F such that $\psi(x) = \psi(x')$. Then for each $j \in J$ we have $\psi_j(x_j) = \psi_j(x_j')$, so x_j and x_j' are not separated by $\psi_j^{-1}\mathcal{D}$, hence not by F_j^0. Since F_i^0 is trivial for $i \notin J$, x and x' are not separated by $\otimes_{i \in I} F_i^0$ and therefore they are not separated by the smaller class $\varphi^{-1}\mathcal{D}$ either. As $\mathcal{D}$ separates the point of $\mathbb{D}$, this implies that $\varphi(x) = \varphi(x')$. The arbitrariness of x and x' now implies that $\varphi = \chi \circ \psi$ for some mapping $\chi: \psi F \to \mathbb{D}$.

Next we show that the range of ψ is an analytic space and that ψ has a universally measurable right inverse. Let $j \in J$. Since (F_j, F_j) is analytic and ψ_j is measurable with respect to F_j, the range of ψ_j is a Souslin subset, S_j say, of $\mathbb{D}$ and ψ_j has a right inverse, say β_j, that is universally measurable with respect to F_j. Moreover, when π_j is the j-th coordinate of $\mathbb{D}^J$, then $\pi_j^{-1} S_j$ is a Souslin subset of $\mathbb{D}^J$ by Proposition 4.3. The foregoing now implies that

$$\psi F = \prod_{j \in J} \psi_j F_j = \prod_{j \in J} S_j = \bigcap_{j \in J} (\pi_j^{-1} S_j) ,$$

so ψF is a countable intersection of Souslin sets and, therefore, a Souslin set itself. Since J is countable, the space $\mathbb{D}^J$ is isomorphic to $\mathbb{D}$ and therefore it is analytic. Consequently ψF, being a Souslin subset of $\mathbb{D}^J$, is analytic by the result ii) proved above.

To construct a universally measurable right inverse $\beta: \psi F \to F$ of ψ we merely choose a point $a \in F$ and define

$$[\beta(y)]_i \quad \begin{aligned} &= \beta_i(y_i) \quad \text{if } i \in J \\ &= a_i \quad\quad \text{if } i \in I \setminus J \end{aligned} \quad (i \in I, \ y \in \psi F) .$$

The universal measurability of β follows from the fact that for each coordinate π_i' of $\prod_{i \in I} F_i$ we have

$$\pi_i' \circ \beta \quad \begin{aligned} &= \beta_i \circ \pi_i \quad \text{if } i \in J \\ &= a_i \quad\quad \text{if } i \in I \setminus J , \end{aligned}$$

and from the universal measurability of the β_i's (see remark preceeding Proposition 3.4).

Finally, we turn to the mapping $\chi: \psi F \to \mathbb{D}$. It follows from the definition of $\psi: F \to \mathbb{D}^J$ that ψ is strictly measurable with respect to $\otimes_{i \in I} F_i^0$. Since we also have

$$\psi^{-1}(\chi^{-1}\mathcal{D}) = (\chi \circ \psi)^{-1}\mathcal{D} = \varphi^{-1}\mathcal{D} \subset \otimes_{i \in I} F_i^0 ,$$

it follows that $\chi^{-1}\mathcal{D}$ is contained in the σ-algebra of ψF. So χ is measurable. As the domain ψF of χ has been proved to be analytic, the

range of χ is a Souslin subset of ID and χ has a universally measurable right inverse, say γ.

Now the range of φ equals the range of χ, and β ∘ γ is easily seen to be a universally measurable right inverse of φ. From the arbitrariness of φ it now follows that the space $\Pi_i\ (F_i, F_i)$ is analytic. □

COROLLARY 7.7. *Let* (F,F) *be an analytic space and let* G *be a sub-σ-algebra of* F. *Then the space* (F,G) *is analytic as well.*

The construction principles mentioned in Proposition 7.6, and the analyticity of ID enable us to obtain a large variety of analytic spaces. We give an example.

PROPOSITION 7.8. *The space* IR *is analytic.*

PROOF. Let φ: ID → [0,1] be defined by

$$\varphi(x) = \sum_{n=1}^{\infty} x_n \cdot 2^{-n} .$$

Then φ is a surjection. Moreover, φ is measurable, because for each n ∈ IN the number x_n depends measurably on x. So, [0,1] is a measurable image of ID and, therefore, it is analytic. The same holds for the measurable subspace (0,1) of [0,1] and, consequently, for the space IR, which is the image of (0,1) under the measurable mapping

$$x \mapsto \frac{1}{x} + \frac{1}{x-1} .$$ □

When the σ-algebras E and F of two measurable spaces (E,E) and (F,F) are isomorphic, the spaces (E,E) and (F,F) themselves are not necessarily isomorphic. In many situations, however, this distinction is immaterial. Note, for instance, that the spaces (Ẽ,Ẽ) and (F̃,F̃) are isomorphic if the σ-algebras E and F are isomorphic. It follows from these considerations and the following proposition that it is often possible to consider a countably generated analytic space to be a Souslin subspace of ID.

PROPOSITION 7.9. *Let* (F,F) *be a countably generated analytic space. Then* F *is isomorphic to the σ-algebra of some Souslin subspace of* ID.

PROOF. It follows from Proposition 7.4 that there is a strictly measurable mapping φ: F → ID. Proposition 7.5 implies that the range of φ is a Souslin subset of ID. Hence the result. □

If (F,F) is a measurable space, G a sub-σ-algebra of F and μ a probability on G, then μ is not necessarily extendable to a probability on F. When such an extension happens to exist, it need not be unique. For analytic spaces, however, things are not too bad:

PROPOSITION 7.10. *Let* (F,F) *be an analytic space and let* G *be a countably generated sub-σ-algebra of* F. *Then there exists a universally measurable mapping* φ: (F,G)~ → (F,F)~ *such that for every probability* μ *on* G *the probability* φ(μ) *is an extension of* μ *to* F.

PROOF. By Proposition 7.4 there exists a mapping ψ: F → ID that is (strictly) measurable with respect to G, and hence, measurable with respect to F. Since (F,F) is analytic, by Proposition 7.5 there is a right inverse χ of ψ that is universally measurable with respect to F.

Now let φ: (F,G)~ → (F,F)~ be defined by φ(μ) := μ ∘ ψ⁻¹ ∘ χ⁻¹.
For every A ∈ F the set χ⁻¹A is a universally measurable subset of the space ψF. Consequently, the set ψ⁻¹(χ⁻¹A) is a universally measurable subset of (F,G) by Proposition 3.4, which in turn implies, by Proposition 3.6, that μ(ψ⁻¹(χ⁻¹A)) depends universally measurably on μ ∈ F̃(G). So, from the definitions of (F,F)~ and universal measurability it follows that the mapping φ is universally measurable.

Finally, we shall show that for every probability μ on G the probability φ(μ) is an extension of μ to F. From the strict measurability of ψ it follows that G = ψ⁻¹D, where D is the σ-algebra of ID. So what we have to prove is that φ(μ) and μ coincide on ψ⁻¹D. Therefore, let B ∈ D. Then

$$\varphi(\mu)(\psi^{-1}B) = (\mu\circ\psi^{-1}\circ\chi^{-1})(\psi^{-1}B) = \mu((\psi^{-1}(\psi\circ\chi)^{-1}B) = \mu(\psi^{-1}B) \,,$$

because ψ ∘ χ is the identity on ψF. □

The mapping φ in the foregoing proposition is a (nonunique) univer-
sally measurable right inverse of the measurable mapping $\rho: (F,F)^{\sim} \to (F,G)^{\sim}$
defined by $(\rho\mu)(G) = \mu(G)$ $(G \in G)$, i.e. the mapping corresponding to
restriction to G (see example 4 following Proposition 2.2).

The remaining part of this section consists of the proof of Proposi-
tion 7.13, which states that analyticity of a space F implies analyticity
of the space $\tilde{F}$. We start with two lemmas.

LEMMA 7.11. $\tilde{\mathbb{D}}$ *is analytic.*

PROOF. Let $\mathring{A}$ be a countable semicompact algebra generating the σ-algebra of
$\mathbb{D}$, let the mapping $\varphi: \tilde{\mathbb{D}} \to [0,1]^{\mathring{A}}$ be defined by $(\varphi\mu)_A = \mu(A)$, and let M be
the subspace of $[0,1]^{\mathring{A}}$ consisting of all additive functions $\nu: \mathring{A} \to [0,1]$
for which $\nu(\mathbb{D}) = 1$.

Since $\mathring{A}$ is a semicompact algebra, every member of M is σ-additive by
Proposition 5.3, and therefore is uniquely extendable to a probability on
$\sigma(\mathring{A})$. This, however, implies that φ is injective and that the range of φ
equals M. The mapping φ even is an isomorphic embedding because the σ-
algebra of $\tilde{\mathbb{D}}$ is generated by the mappings $\mu \mapsto \mu(A)$ $(A \in \mathring{A})$, that is, by
the mappings $\mu \mapsto (\varphi\mu)_A$ $(A \in \mathring{A})$. It is therefore sufficient to prove that M
is analytic.

Now $[0,1]^{\mathring{A}}$ is analytic, because $[0,1]$ is, and M is a measurable
subset of $[0,1]^{\mathring{A}}$, as it can be written as a countable intersection of
measurable sets, e.g.

$$\bigcap_{(A,B)} \{\nu \in [0,1]^{\mathring{A}} \mid \nu(A) + \nu(B) = \nu(A + B) \text{ and } \nu(\mathbb{D}) = 1\} ,$$

where the intersection is over the countable set of disjoint pairs (A,B) of
members of $\mathring{A}$. Hence M is analytic. □

LEMMA 7.12. *Let* F *be a countably generated analytic space. Then* $\tilde{F}$ *is*
analytic.

PROOF. By Proposition 7.9 the σ-algebra of F is isomorphic to the σ-algebra
of some Souslin subspace, S say, of $\mathbb{D}$, and, consequently, the spaces $\tilde{F}$ and
$\tilde{S}$ are isomorphic. It will therefore be sufficient to prove that $\tilde{S}$ is
analytic.

To this end let us define the mapping φ: $\tilde{S}$ → $\tilde{\mathbb{D}}$ by $(\varphi\mu)(B) := \mu(B \cap S)$ $(B \in \mathcal{D})$, where $\mathcal{D}$ is the σ-algebra of $\mathbb{D}$. When $\mu,\mu' \in \tilde{S}$ are such that $\varphi\mu = \varphi\mu'$, then μ and μ' coincide on the σ-algebra $\mathcal{D}|S$ of S and, therefore, are identical. So φ is injective. Moreover, since the σ-algebras of $\tilde{S}$ and $\tilde{\mathbb{D}}$ are generated by the functions $\mu \mapsto \mu(B \cap S)$ and $\nu \mapsto \nu(B)$ $(B \in \mathcal{D})$, respectively, the mapping φ is an isomorphic embedding. So $\tilde{S}$ is isomorphic to the range of φ and what remains to be proved is the analyticity of this range.

Now let $\mu \in \tilde{S}$. For every $B \in \mathcal{D}$ such that $S \subset B$ we have $(\varphi\mu)(B) =$ $= \mu(B \cap S) = \mu(S) = 1$, so $(\varphi\mu)(S) = 1$. On the other hand, let $\nu \in \tilde{\mathbb{D}}$ be such that $\nu(S) = 1$. When we define μ on $\mathcal{D}|S$ by $\mu(A) := \nu(A)$, then μ is a probability and $\varphi\mu = \nu$. So the range of φ is equal to the subset $\{\nu \in \tilde{\mathbb{D}} \mid \nu(S) = 1\}$ of $\tilde{\mathbb{D}}$. Since this subset can be written as

$$\bigcap_{n \in \mathbb{N}} \{\nu \in \tilde{\mathbb{D}} \mid \nu(S) > 1 - n^{-1}\} \,,$$

and since $\nu \mapsto \nu(S)$ is a Souslin function on $\tilde{\mathbb{D}}$ by Proposition 4.8, the range of φ is a Souslin subset of $\tilde{\mathbb{D}}$. From the analyticity of $\tilde{\mathbb{D}}$ together with Proposition 7.6 we now deduce that the range of φ is analytic. □

PROPOSITION 7.13. *If a measurable space* F *is analytic, then* $\tilde{F}$ *is analytic as well.*

PROOF. We shall prove that $\tilde{F}$ has the characterizing property of analytic spaces given in Proposition 7.5. Therefore, let φ: $(F,F)\tilde{} \to (\mathbb{D},\mathcal{D})$ be measurable. Since $\mathcal{D}$ is countably generated, $\varphi^{-1}\mathcal{D}$ is a countably generated sub-σ-algebra of $\tilde{F}$. From the definition of $\tilde{F}$ it follows that $\tilde{F}$ is genrated by the collection C of subsets of $\tilde{F}(F)$ of the form $\{\mu \in \tilde{F}(F) \mid \mu(A) \in B\}$, where $A \in F$ and where B is a measurable subset of $\mathbb{R}$. As a consequence, there exists a countable subclass C_0 of C such that $\varphi^{-1}\mathcal{D} \subset \sigma(C_0)$. Hence, by the definition of C, there exists a countable subclass A of F such that $\varphi^{-1}\mathcal{D}$ is contained in the σ-algebra generated by the functions $\mu \mapsto \mu(A)$ $(A \in A)$ on $\tilde{F}(F)$.

Now let G be the sub-σ-algebra of F generated by A. Suppose that μ' and μ'' are probabilities on F that coincide on G. Then μ and μ'' coincide on A and, therefore, they are not separated by the σ-algebra on $\tilde{F}(F)$ generated by the functions $\mu \mapsto \mu(A)$ $(A \in A)$. As a consequence μ' and μ'' are not separated by the smaller σ-algebra $\varphi^{-1}\mathcal{D}$ either. So, $\varphi\mu'$ and $\varphi\mu''$ are not

separated by $\mathcal{D}$ and, therefore, are identical, since $\mathcal{D}$ separates the points of $\mathbb{D}$.

The foregoing implies that we can write $\varphi = \chi \circ \psi$, where $\psi: (F,F)^\sim \rightarrow (F,G)^\sim$ is the mapping corresponding to restriction to G, i.e. $(\psi\mu)(A) := \mu(A)$ for all $A \in G$ and $\mu \in \widetilde{F}(F)$, and where χ maps $(F,G)^\sim$ into $(\mathbb{D}, \mathcal{D})$. As G is a sub-σ-algebra of F and since (F,F) is analytic, the space (F,G) is analytic as well by Corollary 7.7. Moreover, G is countably generated, so by Proposition 7.10 and the remark following it the mapping ψ is surjective and has a universally measurable right inverse.

From the definitions of $\widetilde{G}$ and ψ it follows that $\psi^{-1}\widetilde{G}$ is the σ-algebra on $\widetilde{F}(F)$ that is generated by the functions $\mu \mapsto \mu(A)$ $(A \in G)$; it therefore contains the σ-algebra $\varphi^{-1}\mathcal{D}$. Consequently, we have $\psi^{-1}\chi^{-1}\mathcal{D} = (\chi \circ \psi)^{-1}\mathcal{D} = = \varphi^{-1}\mathcal{D} \subset \psi^{-1}\widetilde{G}$ and, since ψ is surjective, $\chi^{-1}\mathcal{D} \subset \widetilde{G}$. So χ is measurable.

Due to Lemma 7.12, the space $(F,G)^\sim$ is analytic, because (F,G) is a countably generated analytic space. Together with the measurability of χ this implies, by Proposition 7.5, that the range of χ is a Souslin subset of $\mathbb{D}$ and that χ has a universally measurable right inverse.

Since ψ is surjective, the range of φ equals the range of χ and hence is a Souslin subset of $\mathbb{D}$. Moreover, the composition of the right inverses of χ and ψ mentioned above is a universally measurable right inverse of φ. $\square$

§ 8. Separating classes

We shall now prove the so-called first separation theorem for analytic spaces. From this theorem it follows that the two conditions in the definition of analytic spaces are not independent when countably generated or countably separated spaces are considered. Moreover, with the aid of this theorem it is possible to characterize those analytic spaces whose σ-algebras are generated by an analytic topology.

DEFINITION. Let A and B be collections of sets. Then B is said to *separate* A, if for every disjoint pair $A_1, A_2 \in A$ there exists a disjoint pair $B_1, B_2 \in B$ such that $A_1 \subset B_1$ and $A_2 \subset B_2$.

This terminology is consistent with the notion of a separating collection of subsets of a set introduced earlier, since such a collection is characterized by the fact that it separates the collection of singletons.

For the proof of the first separation theorem we need:

LEMMA 8.1. *Let* (E,E) *be a measurable space,* $\{S_i^1 \mid i \in \mathbb{N}\}$ *and* $\{S_i^2 \mid i \in \mathbb{N}\}$ *countable collections of subsets of* E, $S^1 := \bigcup_{i \in \mathbb{N}} S_i^1$ *and* $S^2 := \bigcup_{i \in \mathbb{N}} S_i^2$. *When the pair* $\{S^1, S^2\}$ *is not separated by* E, *then for some* $i,j \in \mathbb{N}$ *the pair* $\{S_i^1, S_j^2\}$ *is not separated by* E *either.*

PROOF. We argue by contradiction. Let the pair $\{S_i^1, S_j^2\}$ be separated by E for every $i,j \in \mathbb{N}$. Then for every $i,j \in \mathbb{N}$ there exists $M_{ij} \in E$ such that $S_i^1 \subset M_{ij}$ and $S_j^2 \subset M_{ij}^c$. Consequently, S^1 is contained in the set $\bigcup_{i \in \mathbb{N}} \bigcap_{j \in \mathbb{N}} M_{ij}$ and S^2 in the compliment of this set. So the pair $\{S^1, S^2\}$ is separated by E. □

PROPOSITION 8.2 (First separation theorem). *Let* (E,E) *be an analytic space. Then* E *separates* $S(E)$.

PROOF.

i)   We first consider the case that E equals $\mathbb{D}$. Let $\mathcal{D}$ be the σ-algebra of $\mathbb{D}$ and let A be a countable semicompact algebra generating $\mathcal{D}$ (cf. Proposition 5.4). As A is closed under complementation, by Proposition 4.1 we have $\sigma(A) \subset S(A)$. Moreover, $\mathcal{D} = \sigma(A)$, so $S(\mathcal{D}) \subset S(S(A)) = S(A)$. Now let $S^1, S^2 \in S(\mathcal{D})$. Then

$$S^r = \bigcup_{n \in \mathbb{N}^{\mathbb{N}}} \bigcap_{k \in \mathbb{N}} A^r_{n_1, \ldots, n_k} \qquad (r = 1,2)$$

for certain sets $A^r_{n_1, \ldots, n_k}$ belonging to A. For $r \in \{1,2\}$ and for every finite sequence $m_1, \ldots, m_p$ of positive integers we define

$$S^r_{m_1, \ldots, m_p} := \bigcup \{ \bigcap_{k \in \mathbb{N}} A^r_{n_1, \ldots, n_k} \mid n \in \mathbb{N}^{\mathbb{N}} \text{ and } \forall_{\ell \le p} \ n_\ell = m_\ell \} .$$

Then

$$S^r = \bigcup_{i \in \mathbb{N}} S^r_i \quad \text{and} \quad S^r_{m_1, \ldots, m_p} = \bigcup_{i \in \mathbb{N}} S^r_{m_1, \ldots, m_p, i} .$$

Let us suppose that the pair $\{S^1, S^2\}$ is not separated by $\mathcal{D}$. It then follows by induction on p, and by use of Lemma 8.1 at each step, that there exist $m^1, m^2 \in \mathbb{N}^{\mathbb{N}}$ such that for each p the pair

$$\left\{ S^r_{m^r_1,\ldots,m^r_p} \mid r = 1,2 \right\}$$

is not separated by $\mathcal{D}$. As

$$S^r_{m^r_1,\ldots,m^r_p} \subset \bigcap_{k\le p} A^r_{m^r_1,\ldots,m^r_k} \in \mathcal{D}$$

for every p and r, this implies that for every p the pair

$$\left\{ \bigcap_{k\le p} A^r_{m^r_1,\ldots,m^r_k} \mid r = 1,2 \right\}$$

is not disjoint. It now follows from the semicompactness of A that the set

$$\left(\bigcap_{k\in\mathbb{N}} A^1_{m^1_1,\ldots,m^1_k} \right) \cap \left(\bigcap_{k\in\mathbb{N}} A^2_{m^2_1,\ldots,m^2_k} \right)$$

is nonempty, and hence, that S^1 and S^2 are not disjoint. The foregoing implies that any disjoint pair of members of $S(\mathcal{D})$ is separated by $\mathcal{D}$.

ii) Next let $(E,\mathcal{E})$ be a countably generated analytic space. Then, by Proposition 7.9, the space $(E,\mathcal{E})$ may be supposed to be a Souslin subspace of $(\mathbb{D},\mathcal{D})$ and consequently $S(\mathcal{E}) \subset S(\mathcal{D})$ by Proposition 4.4. As $S(\mathcal{D})$ is separated by $\mathcal{D}$, so is $S(\mathcal{E})$. Since each member of $S(\mathcal{E})$ is contained in E, the collection $S(\mathcal{E})$ is separated by the trace $\mathcal{E}$ of $\mathcal{D}$ on E as well.

iii) Finally, let $(E,\mathcal{E})$ be an arbitrary analytic space and let $S^1, S^2 \in S(\mathcal{E})$. Then there exists a countable subclass C of $\mathcal{E}$ such that $S^1, S^2 \in S(C) \subset \subset S(\sigma(C))$ (see the remark following the definition of S). Applying ii) to the space $(E, \sigma(C))$ we conclude that the pair S^1, S^2 is separated by $\sigma(C)$, and hence, also by the larger class $\mathcal{E}$. $\qquad\square$

COROLLARY 8.3. *If $(E,\mathcal{E})$ is an analytic space, then*

$$\mathcal{E} = \{ A \subset E \mid A \in S(\mathcal{E}) \text{ and } A^c \in S(\mathcal{E}) \} \ ,$$

and hence, $\mathcal{E}$ is the largest σ-algebra contained in $S(\mathcal{E})$.

The following proposition is a generalization of the "only if"-part of Proposition 7.5.

PROPOSITION 8.4. *Let φ be a measurable mapping of an analytic space F into a countably separated measurable space G. Then the range of φ is a Souslin subset of G and φ has a universally measurable right inverse. When, in addition, φ is injective, then φ is an isomorphic embedding.*

PROOF. To prove the first part of the proposition we can repeat the first part of the proof of Proposition 7.5, with $\mathbb{D}$ replaced by G.

Next, let φ be injective. Let A be a measurable subset of F. We shall prove that φA is a measurable subset of φF. By Proposition 7.6, A is an analytic subspace of F, which is mapped by φ into the countably separated space φF. So, by what has already been proved above, φA is a Souslin subset of the space φF. A similar reasoning holds for the set A^c. Since φ is injective, $\varphi(A)$ and $\varphi(A^c)$ are complementary Souslin subsets of φF. From Proposition 7.6 we deduce that φF is an analytic space and, by Corollary 8.3, this implies that φA is measurable. As A is arbitrarily chosen, it follows that φ is an isomorphic embedding. □

PROPOSITION 8.5. *Let (E,E) be an analytic space. Then any countable separating subclass of E generates E.*

PROOF. Let C be a countable separating subclass of E. Then $(E,\sigma(C))$ is countably separated and the identity on E is a measurable injection of (E,E) onto $(E,\sigma(C))$. The preceding proposition now implies that E equals $\sigma(C)$. □

Proposition 8.5 implies an extremal property of the σ-algebra of a countably separated analytic space: no smaller σ-algebra is countably separated, no larger one is analytic.

In the definition of analytic spaces as well as in some propositions of the last two sections and in their proofs we can distinguish two parts: one bearing on Souslin sets and another one having to do with universally measurable mappings. Moreover, inspection of the proofs mentioned reveals that the first part is in a sense independent of the second one and, after skipping everything bearing on universally measurable mappings in definitions as well as in propositions and proofs, we are left with a theory

about a class of spaces that is possibly larger than the class of analytic spaces.

As has been remarked at the end of Section 4, another possible modification of the theory is obtained by substituting limit measurable sets and mappings for universally measurable ones. The following proposition implies that these three theories coincide for measurable spaces that are countably generated or countably separated.

PROPOSITION 8.6. *Let* F *be a measurable space that is countably generated or countably separated. Then the following statements are equivalent:*

i)   F *is analytic.*

ii)   *For every measurable space* E *and every Souslin subset* S *of* E × F, *the projection of* S *on* E *is a Souslin subset of* E.

iii)   *For every measurable mapping* φ: F → D *the range of* φ *is a Souslin subset of* D.

PROOF. The proof consists of two parts.

Part I. We first consider the case that F is countably generated. We shall prove the implications i) $\Rightarrow$ ii) $\Rightarrow$ iii) $\Rightarrow$ i).

i) $\Rightarrow$ ii). This follows from the definition of analyticity.

ii) $\Rightarrow$ iii). Let φ: F → D be a measurable mapping. Since D is countably separated, the graph of φ is a measurable subset of F × D by Proposition 7.3. As a consequence of ii) the projection on D of this graph, being the range of φ, is a Souslin subset of D.

iii) $\Rightarrow$ i). Let φ: F → D be measurable. By Proposition 7.5 it is sufficient to prove that φ has a universally measurable right inverse.

Let ψ: F → D be strictly measurable and let η: ψF → F be a measurable right inverse of ψ; the existence of such mappings follows from Proposition 7.4. Moreover, it follows from iii), applied to ψ, that the range ψF of ψ is a Souslin subspace of D and hence, that ψF is analytic.

Now we show that $\varphi = \varphi \circ \eta \circ \psi$. To this end let x and x' be points of F such that $\psi x = \psi x'$. Then ψx and $\psi x'$ are not separated by the σ-algebra $\mathcal{D}$ of D and, therefore, x and x' are not separated by the σ-algebra $\psi^{-1}\mathcal{D}$ of F. As $\varphi^{-1}\mathcal{D} \subset \psi^{-1}\mathcal{D}$, by the measurability of φ, the points x and x' are not separated by $\varphi^{-1}\mathcal{D}$ either, and therefore φx and $\varphi x'$ are not separated by $\mathcal{D}$.

54

So $\varphi x = \varphi x'$. In particular we can take $x' := (\eta \circ \psi)x$. Then $\psi x' = \psi x$, and
hence, $\varphi x = \varphi x' = (\varphi \circ \eta \circ \psi)x$. The arbitrariness of x now implies $\varphi = \varphi \circ \eta \circ \psi$.

Next consider the mapping $\varphi \circ \eta \colon \psi F \to \mathbb{D}$. It follows from the equality
last obtained that the range of $\varphi \circ \eta$ equals the range of φ. Moreover, $\varphi \circ \eta$
is measurable, since φ and η are, and ψF is analytic. So, by Proposition 7.5
$\varphi \circ \eta$ has a universally measurable right inverse, $\xi \colon \varphi F \to \psi F$ say. It follows
that the mapping $\eta \circ \xi$ is a universally measurable right inverse of φ.

Part II. Next let F be countably separated, and suppose that at least one
of the statements i), ii) or iii) is true.

Let F be the σ-algebra of F, C a countable separating subclass of F,
$A \in F$, and A the σ-algebra generated by the collection $C \cup \{A\}$. Then A is a
sub-σ-algebra of F, so each (hence, at least one) of the statements i), ii)
or iii) that holds for (F,F) also holds for (F,A). As A is countably gener-
ated, it follows from Part I that the three statements are equivalent for
(F,A). So each of them (in particular i)) holds for (F,A), i.e. (F,A) is
analytic. It now follows from Proposition 8.5 that the countable separating
class C generates A and therefore that $A \in \sigma(C)$.

Since A is an arbitrarily chosen member of F, we have $F \subset \sigma(C)$, and
hence, $F = \sigma(C)$. So (F,F) is countably generated and, by Part I, the three
statements are equivalent for (F,F). As at least one of them is true, they
all are. $\square$

The measurable spaces appearing in many applications are countably
generated or, when not, one often is interested only in a function defined
on such a space, to which again there corresponds a countably generated
σ-algebra. So there would have been no serious loss of applicability had we
incorporated in the definition of analytic spaces the condition that such a
space were countably generated. This however would have been at the cost of
stability of the class of analytic spaces under the formation of measurable
images and product spaces. A similar remark can be made about separation:
a nonseparated measurable space is not essentially different from a sepa-
rated one, but the property of being separated is not conserved under the
formation of measurable images.

We conclude this section with a characterization of those measurable
spaces whose σ-algebra is generated by an analytic topology. This charac-
terization will not be used in the sequel. The σ-algebra of a measurable

space is generated by an analytic topology iff that space is isomorphic to
a Souslin subset of the measurable space [0,1] (cf. [Hoffman-Jørgensen]
Ch. III §2, theorems 3 and 4). Since the measurable spaces [0,1] and $\mathbb{D}$
are isomorphic ([Bertsekas & Shreve] Proposition 7.16) the latter condition
is equivalent to the space being isomorphic to a Souslin subspace of $\mathbb{D}$. It
now follows from Propositions 7.2, 7.6 and 7.9 that this in turn is equiva-
lent to the condition that the space is countably separated and analytic.

Most of our results concerning countably generated analytic spaces
are known to be valid for analytic topological spaces. So these results are
new only in that they have been derived by measure-theoretic means only.

§ 9. Probabilities on analytic spaces

This section is devoted to what may be called Kolmogorov's theorem
for analytic spaces and to the decomposition of probabilities defined on
product spaces into a marginal probability and a transition probability.
In both topics semicompact classes play a central role and we start with a
proposition concerning them.

PROPOSITION 9.1. Let (F, F) be a countably generated analytic space. Then
there exists a semicompact subclass C of F such that for every $\mu \in \tilde{F}$ and
$A \in F$:

$$\mu(A) = \sup \{\mu(C) \mid C \in C \text{ and } C \subset A\} .$$

PROOF. As F is isomorphic to the σ-algebra of some Souslin subspace of $\mathbb{D}$
(cf. Proposition 7.9), (F, F) itself may be supposed to be a Souslin sub-
space of $\mathbb{D}$. By Proposition 5.4 the σ-algebra $\mathcal{D}$ of $\mathbb{D}$ is generated by a
semicompact algebra A and, as A is closed under complementation, we have
$\mathcal{D} \subset S(A)$ by Proposition 4.1. It now follows from Proposition 5.2 that, for
every Souslin subset S of $\mathbb{D}$ and for every $\mu \in \tilde{\mathbb{D}}$, we have

$$\mu(S) = \sup \{\mu(C) \mid C \in A_{s\delta} \text{ and } C \subset S\} .$$

Specializing to sets S contained in F and probabilities μ concentrated on
F we get

$$\mu(S) = \sup \{\mu(C) \mid C \in C \text{ and } C \subset S\} ,$$

56

where $C := \{C \in A_{s\delta} \mid C \subset F\}$. Moreover C is a subclass of $A_{s\delta}$ and $A_{s\delta}$ is semicompact because A is (Proposition 5.1). Consequently C is semicompact as well. Finally, we observe that every member of F is the intersection of F with some member of $\mathcal{D}$, and hence, is a Souslin subset of $\mathbb{D}$ that is contained in F. $\qquad\qquad\square$

PROPOSITION 9.2. *Let* $(F_i)_{i \in I}$ *be a family of analytic spaces and for every finite subset* I' *of* I *let* $\mu_{I'}$ *be a probability on* $\Pi_{i \in I'} F_i$, *such that, for every pair* I',I" *of finite subsets of* I *with* $I' \subset I''$, *the probability* $\mu_{I''}$ *is the marginal of* $\mu_{I''}$ *(corresponding to the projection of* $\Pi_{i \in I''} F_i$ *onto* $\Pi_{i \in I'} F_i$*). Then there exists a unique probability on* $\Pi_{i \in I} F_i$ *having the probabilities* $\mu_{I'}$ *(I' finite subset of* I*) as its marginals.*

PROOF. Let $F_{I'} := \Pi_{i \in I'} F_i$ for each subset I' of I, and let A be the collection of measurable cylinders of F_I. We now define $\mu: A \to \mathbb{R}$ as follows. For every $A \in A$ there exists a finite subset I' of I and a measurable cylinder A' of $F_{I'}$ such that $A = A' \times F_{I \setminus I'}$. Next we define $\mu(A) := \mu_{I'}(A')$. The definition of $\mu(A)$ is easily seen to be independent of the particular choice of I'. Moreover, the function μ thus defined is the only function on A whose marginals coincide with the probabilities $\mu_{I'}$. As A generates the σ-algebra of F_I, all that remains to be proved is σ-additivity of μ on A (see [Neveu] Proposition I.6.1).

Let C be a subclass of A consisting of a countable collection pairwise disjoint sets together with their union. Then there exists a family $(C_i)_{i \in I}$, with C_i a countable collection of measurable subsets of F_i for each $i \in I$, and such that C is contained in the semialgebra B of measurable cylinders of the space $\Pi_{i \in I}(F_i, \sigma(C_i))$. B is contained in A and the restriction of μ to B is the unique function on B whose marginals coincide with the restriction of the measures $\mu_{I'}$ to the spaces $\Pi_{i \in I'}(F_i, \sigma(C_i))$.

Now for each $i \in I$ the space $(F_i, \sigma(C_i))$ is countably generated and analytic, being a measurable image of the space F_i with the original σ-algebra. Therefore each of the σ-algebras $\sigma(C_i)$ contains a semicompact class as described in Proposition 9.1. These facts, however, imply ([Neveu] Theorem III.3) that μ is σ-additive on B, and hence, on C. From the arbitrariness of C it now follows that μ is σ-additive on A. $\qquad\square$

DEFINITION. Let E and F be measurable spaces. A *transition probability from E to F* is a mapping of E into $\tilde{F}$.

When p is such a transition probability, $x \in E$, and B a universally measurable subset of F, then we shall usually write $p(x;B)$ instead of $p(x)(B)$ in order to enhance readability.

PROPOSITION 9.3. *Let E and F be measurable spaces, let ν be a probability on E and let p be a universally measurable transition probability from E to F. Then there exists a unique probability on E × F, denoted by $\nu \times p$, such that*

$$\int\limits_{E\times F} f\,d(\nu\times p) = \int\limits_{E}\int\limits_{F} f(x,y)p(x;dy)\nu(dx)$$

for every universally measurable $f: E \times F \to \overline{\mathbb{R}}_+$

PROOF. By Proposition 6.2, $\int_F f(x,y)p(x)(dy)$ depends universally measurably on $(x,p(x))$ and therefore on x, because $p(x)$ depends universally measurably on x. So the repeated integral is well defined. Moreover it is easily seen to be nonnegative, to depend σ-additively on f, and to assume the value 1 for $f = 1_{E\times F}$. From this the result follows. $\qquad\square$

The notation introduced in Proposition 9.3 is consistent with the notation of product probabilities when constant transition probabilities are identified with their value. Note that for constant transition probabilities Proposition 9.3 reduces to the theorem of Tonelli (see [Cohn] Proposition 5.2.1).

In general, not every probability μ on a product space E × F can be written as the product of its marginal on E and a transition probability from E to F. When, however, the space F meets certain conditions, then such a decomposition of μ is possible, as is stated in the following proposition. This proposition, together with the "Exact selection theorem" (Proposition 7.1) forms the basis for the applications in Chapter III.

PROPOSITION 9.4. *Let E be a measurable space, F a countably generated analytic space, μ a probability on E × F and ν the marginal probability of μ on E. Then there exists a measurable transition probability p from E to F such that $\mu = \nu \times p$.*

58

PROOF. As, by Proposition 7.9, the σ-algebra of F is isomorphic to the σ-algebra of some Souslin subspace of $\mathbb{D}$, we can suppose F itself to be a Souslin subspace of $\mathbb{D}$. Let E and F be the σ-algebras of E and F, respectively, and let B be a countable semicompact algebra that generates the σ-algebra $\mathcal{D}$ of $\mathbb{D}$.

For every $B \in B$ the mapping $A \mapsto \mu(A \times (B \cap F))$ is a bounded nonnegative measure on E that is absolutely continuous with respect to ν. Hence by the Radon-Nikodym theorem ([Neveu] Proposition IV.1.4) there exists a measurable function $f_B: E \to \mathbb{R}_+$ such that for every $A \in E$

$$(*) \qquad \mu(A \times (B \cap F)) = \int_A f_B(x)\nu(dx) \ .$$

Now for every $A \in E$ $\mu(A \times (B \cap F))$ depends additively on B and equals $\nu(A)$ for $B = \mathbb{D}$. As a consequence of $(*)$ the same holds for the integral in $(*)$. Due to the countability of B this implies that there is a ν-null subset N of E such that for every $x \in E\backslash N$ the function $B \mapsto f_B(x)$ is additive, and $f_{\mathbb{D}}(x) = 1$. Redefining the functions f_B on N by $f_B(x) = f_B(x_0)$, where x_0 is some point in $E \backslash N$, we can suppose $N = \emptyset$ and still have $(*)$ for every $A \in E$, $B \in B$. As B is a semicompact algebra, for each $x \in E$ the function $B \mapsto f_B(x)$ is even σ-additive, and therefore it is uniquely extendible to a probability $B \mapsto \bar{f}_B(x)$ on $\mathcal{D}$. The functions $\bar{f}_B$, with $B \in \mathcal{D}$, are measurable again, because the sets $B \in \mathcal{D}$ for which $\bar{f}_B$ is measurable constitute a Dynkin class (see Preliminaries 2) containing the algebra B. Moreover we have $\mu(A \times (B \cap F)) = \int_A \bar{f}_B(x)\nu(dx)$ for all $A \in E$ and $B \in \mathcal{D}$, because both members of this equality, taken as functions of B, are measures on $\mathcal{D}$, and by $(*)$ these measures coincide on the generating algebra B.

Next consider the probability λ on $\mathcal{D}$ defined by $\lambda(B) := \mu(E \times (B \cap F))$. Then $\lambda^*(F) = 1$ and, as F is universally measurable, $\lambda_*(F) = 1$ as well. Consequently, there exists a set $G \in \mathcal{D}$ such that $G \subset F$ and $\lambda(G) = 1$. From the definition of λ it now follows that $\mu(E \times (F\backslash G)) = 0$.

Finally, for each $x \in E$ and each $B \in F$ we define $p(x)(B) := \bar{f}_{B \cap G}(x)$. It is easily seen that p is a measurable transition probability from E to F. Moreover, for every $A \in E$ and $B \in F$ we have

$$\mu(A \times B) = \mu(A \times B \cap G) + \mu(A \times B \cap (F\backslash G)) =$$

$$= \int_A \bar{f}_{B \cap G}(x)\nu(dx) = \int_A p(x)(B)\nu(dx) = (\nu \times p)(A \times B) \ ,$$

because $\mu(A \times B \cap (F \setminus G)) \leq \mu(E \times F \setminus G) = 0$. So μ and $\nu \times p$ coincide on the measurable rectangles of $E \times F$, which implies that they are equal. □

In the preceding proposition, the analyticity of F may be replaced by the weaker condition that a semicompact approximating subclass of $\tilde{F}$ exists for the marginal of μ on F (in the sense of Proposition 9.1; see [Pfanzagl & Pierlo] Section 7).

The decomposition of a probability on a product space into a marginal probability and a transition probability is, in general, not unique:

PROPOSITION 9.5. *Let* E *be a measurable space,* F *a countably generated measurable space,* $\nu \in \tilde{E}$, *and* $p, p': E \to \tilde{F}$ *universally measurable. Then* $\nu \times p = \nu \times p'$ *iff* $p(x) = p'(x)$ *for ν-almost all* $x \in E$.

PROOF. The "if"-part is a trivial consequence of the definition of $\nu \times p$ and $\nu \times p'$, so let $\nu \times p = \nu \times p'$. Let $\mathcal{B}$ be a countable algebra generating the σ-algebra of F. For every $B \in \mathcal{B}$ and every universally measurable subset A of E we have

$$\int_A p(x;B)\nu(dx) = (\nu \times p)(A \times B) = (\nu \times p')(A \times B) = \int_A p'(x;B)\nu(dx) \, ,$$

and hence, $p(x;B) = p'(x;B)$ for ν-almost all $x \in E$. It follows from the countability of $\mathcal{B}$ that for ν-almost $x \in E$ the measures $p(x)$ and $p'(x)$ coincide on $\mathcal{B}$ and therefore that they are equal. □

The nonuniqueness of the transition probability appearing in the decomposition of a probability μ enables us to make this transition probability depend on μ in a decent way:

PROPOSITION 9.6. *Let* E *be a countably generated measurable space and* F *a countably generated analytic space. Then there exists a universally measurable mapping* $p: (E \times F)^\sim \times E \to \tilde{F}$ *such that every probability* μ *on* $E \times F$ *is equal to the product* $\nu \times q$ *of its marginal* ν *on* E *and a transition probability* q *from* E *to* F *defined by* $q(x) := p(\mu, x)$. *When* $F = \mathbb{D}$, *then* p *may be supposed to be measurable.*

60

PROOF. As the σ-algebra of F is isomorphic to the σ-algebra of some Souslin subspace of $\mathbb{D}$, we may suppose F itself to be a Souslin subspace of $\mathbb{D}$. Let E and F be the σ-algebras of E and F, respectively, let $A = \{A_n \mid n \in \mathbb{N}\}$ be a countable algebra generating E and let B be a countable semicompact algebra generating the σ-algebra D of $\mathbb{D}$. For each $n \in \mathbb{N}$ let G_n be the partition of E generated by $\{A_1, A_2, \ldots, A_n\}$.

Now let μ be a probability on $E \times F$, ν its marginal on E, and $B \in B$. We define a sequence $(f_n)_{n \in \mathbb{N}}$ of functions on E by

$$f_n(x) := \sum \{\mu(G \times B \cap F)\mu(G \times F)^{-1} 1_G(x) \mid G \in G_n \text{ and } \mu(G \times F) > 0\} .$$

Let $m, n \in \mathbb{N}$ and $m \le n$. Then G_n is a refinement of G_m, so for each $A \in \sigma(G_m)$ we have

$$\int_A f_n d\nu = \sum \left\{ \int_G f_n d\nu \mid G \in G_n \text{ and } G \subset A \right\} =$$

$$= \sum \{\mu(G \times B \cap F) \mid G \in G_n \text{ and } G \subset A\} = \mu(A \times B \cap F) .$$

Moreover each of the functions f_n is bounded by 1. This implies that the sequence (f_n) is a martingale relative to the σ-algebras $\sigma(G_n)$ and the measure ν, and that the martingale convergence theorem applies (see [Chow & Teicher] §7.4 Theorem 2 ii)). So, if we define $f := \limsup_n f_n$, then

$$\int_A f d\nu = \int_A f_n d\nu = \mu(A \times (B \cap F))$$

for every $n \in \mathbb{N}$ and $A \in \sigma(G_n)$. Consequently, for every $A \in \bigcup_{n \in \mathbb{N}} \sigma(G_n) = A$, we have $\int_A f d\nu = \mu(A \times (B \cap F))$ and, as both members of this equality depend σ-additively on A, this equality holds for every $A \in \sigma(A) = E$.

The function f introduced above depends on μ and B, and we now make this dependence explicit by defining $f: (E \times F)^{\sim} \times E \times B \to \overline{\mathbb{R}}$ as

$$f(\mu, x, B) := \limsup_n \sum_{G \in G_n} \mu(G \times (B \cap F))h(\mu(G \times F))1_G(x) ,$$

where h is the measurable function on $\mathbb{R}$ defined by $h(t) = t^{-1}$ if $t \ne 0$ and $h(0) = 0$. The foregoing then implies that, for every $\mu \in (E \times F)^{\sim}$, $A \in E$ and $B \in B$, we have

$(*)$ $\qquad \displaystyle\int_A f(\mu,x,B)\nu(dx) = \mu(A \times (B \cap F))$,

where ν is the marginal of μ on E. Moreover, it follows directly from the definition of f, that $f(\mu,x,B)$ is a measurable function of (μ,x) for every $B \in \mathcal{B}$.

We now turn to the dependence on B. In general $f(\mu,x,B)$, taken as a function of B, is not a probability on $\mathcal{B}$, but a nonessential redefinition of f suffices to make it one. Let N be the set of pairs $(\mu,x) \in (E{\times}F)^{\sim} \times E$ for which the function $B \mapsto f(\mu,x,B)$ is not additive on $\mathcal{B}$ or for which $f(\mu,x,\mathbb{D}) \neq 1$. Then

$$N = \bigcup_{(B_1,B_2)} \{(\mu,x) \in (E{\times}F)^{\sim} \times E \mid f(\mu,x,B_1) + f(\mu,x,B_2) \neq f(\mu,x,B_1 \cup B_2)$$

$$\text{or } f(\mu,x,\mathbb{D}) \neq 1\} \ ,$$

where the union is over the countable set of disjoint pairs (B_1,B_2) of members of $\mathcal{B}$. The measurability of f mentioned above now implies that N is a measurable subset of $(E{\times}F)^{\sim} \times E$. Moreover, for every $\mu \in (E{\times}F)^{\sim}$ the section $\{x \in E \mid (\mu,x) \in N\}$ of N is a ν-null set, where ν is the marginal of μ on E: Let $B_1,B_2 \in \mathcal{B}$ be disjoint. It then follows from $(*)$, that

$$\int_A [f(\mu,x,B_1) + f(\mu,x,B_2) - f(\mu,x,B_1 \cup B_2)]\nu(dx) = 0$$

for every $A \in E$, and hence, that the integrand is a ν-null function. Similarly $f(\mu,x,\mathbb{D}) - 1$ is a ν-null function of x. Together with the definition of N this implies the result on the sections of N.

Now let $(\mu_0,x_0) \notin N$ and, for each $B \in \mathcal{B}$, redefine $f(\cdot,\cdot,B)$ on N by setting it equal to $f(\mu_0,x_0,B)$. Then, for every $(\mu,x) \in (E{\times}F)^{\sim} \times E$, $f(\mu,x,B)$ is an additive function of B on $\mathcal{B}$ and $f(\mu,x,\mathbb{D}) = 1$. Moreover, for every $B \in \mathcal{B}$, $f(\mu,x,B)$ still is a measurable function of (μ,x), and for every μ, A and B the equality $(*)$ still holds.

As $\mathcal{B}$ is a semicompact algebra, for every (μ,x) the function $B \mapsto f(\mu,x,B)$ is even σ-additive on $\mathcal{B}$ by Proposition 5.3, and it therefore is uniquely extendible to a probability $p(\mu,x)$ on $\mathcal{U}(\mathcal{D})$. It follows from Proposition 2.2 applied to the subclass $\mathcal{B}$ of $\mathcal{D}$, and from Proposition 3.6, that $p(\mu,x,B)$ depends (universally) measurable on (μ,x) for every (universally) measurable subset B of $\mathbb{D}$. Also, it follows from $(*)$, that for

62

every $\mu \in (E \times F)^\sim$, $A \in E$ and $B \in B$ we have:

(**) $\int_A p(\mu,x)(B)\nu(dx) = \mu(A \times (B \cap F))$

and, as both members of this equality depend σ-additively on A as well as on B, the equality holds for every (universally) measurable subset A of E and B of $\mathbb{D}$.

When $F = \mathbb{D}$, then the foregoing implies that p is a measurable mapping as mentioned in the proposition. When $F \neq \mathbb{D}$ for each $(\mu,x) \in (E \times F)^\sim \times E$ we consider the restriction of $p(\mu,x)$ to the σ-algebra F of F. Such a restriction is a σ-additive function on F, but it need not be a probability on F, since it may fail to attain the value 1 at F. Now let

$M := \{(\mu,x) \in (E \times F)^\sim \times E \mid p(\mu,x)(F) \neq 1\}$

Then M is universally measurable and, taking $(\mu_1,x_1) \notin M$, we redefine p on M by setting it equal to $p(\mu_1,x_1)$, which makes p a universally measurable mapping of $(E \times F)^\sim \times E$ into $\tilde{F}$. Moreover, this redefinition of p does not affect the validity of (**), because for every $\mu \in (E \times F)^\sim$ we have

$\int_A [p(\mu,x)(F) - 1]\nu(dx) = \mu(A \times F) - \nu(A) = 0$

for every universally measurable subset A of E, and hence, the integrand is a ν-null function. All this now implies, that the (redefined) mapping p meets the requirements of the proposition.   □

In the following two propositions we consider conditional expectations of functions defined on analytic spaces. Let p be a measurable transition probability from a measurable space E to a measurable space F, and for every positive measurable function f on F let the function φf on E be defined by $(\varphi f)(x) := \int f(y)p(x;dy)$. Then, by Proposition 6.1, φf is measurable. Moreover, for each $x \in E$, $(\varphi f)(x)$ depends σ-additively on f. These facts, together with the decomposition of probabilities given in Proposition 9.4, enable us to construct regular versions of conditional expectations. For general information see [Chow & Teicher] Section 7.2 and, in particular, Theorem 3; in fact, our next proposition is a generalization of this theorem.

PROPOSITION 9.7. *Let* (E,E) *be a measurable space,* μ *a probability on* E *and* F, G *sub-σ-algebras of* E *such that* (E,G) *is a countably generated analytic space. Then there exists an* F-*measurable transition probability* p *from* (E,F) *to* (E,G) *such that, for every positive* G-*measurable function* f *on* E, *the function* y ↦ ∫ f(x)p(y;dx) *is a version of the conditional expectation of* f *with respect to* F *and* μ.

PROOF. Let the measurable mapping ψ: (E,E) → (E,F) × (E,G) be defined by ψ(x) := (x,x). Then μ ∘ ψ^{-1} is a probability on F ⊗ G whose marginal on F equals μ. As (E,G) is countably generated and analytic, there exists a measurable transition probability p from (E,F) to (E,G) such that μ ∘ ψ^{-1} = μ × p. Now for every A ∈ F and B ∈ G we have

$$\int_A 1_B d\mu = \mu(A \cap B) = (\mu \circ \psi^{-1})(A \times B) = (\mu \times p)(A \times B) =$$

$$= \int_A p(y;B)\mu(dy) = \int_A \int_E 1_B(x)p(y;dx)\mu(dy) .$$

So, for every A ∈ F and for every positive G-measurable function f on E, the equality

$$\int_A f d\mu = \int_A \int_E f(x)p(y;dx)\mu(dy)$$

holds. ☐

PROPOSITION 9.8. *Let* (E,E) *be a countably generated analytic space,* μ *a probability on* E, *and* F *a sub-σ-algebra of* E. *Then there exists an* F-*measurable transition probability* p *from* (E,F) *to* (E,E) *such that, for every positive* E-*measurable function* f, *the function* y ↦ ∫ f(x)p(y;dx) *is a version of the conditional expectation of* f *with respect to* F *and* μ.

PROOF. Apply Proposition 9.7 taking G := E. ☐

Some properties of probabilities on a measurable space give rise to measurable, or Souslin, subsets of the space of all probabilities. In the rest of this section we shall discuss some of these. The reader may think the properties considered somewhat peculiar; their significance will become evident in Chapter III.

PROPOSITION 9.9. *Let* E *and* F *be countably generated measurable spaces, let* F *be analytic and let* G *be a measurable (Souslin) subset of* E × F̃. *Then the probabilities on* E × F *that can be decomposed into their marginal on* E *and a universally measurable transition probability whose graph is contained in* G *constitute a measurable (Souslin) subset of* (E × F)~.

PROOF. Let E, F and $\mathcal{D}$ be the σ-algebras of E, F and ID, respectively. As F is a countably generated analytic space, F is isomorphic to the σ-algebra of some Souslin subspace of ID by Proposition 7.9. So F itself may be supposed to be a Souslin subspace of ID.

For the proof we need some mappings, and we start by introducing them. Let φ: F̃ → ĨD be defined by $(\varphi\mu)(B) := \mu(B \cap F)$ $(B \in \mathcal{D})$. Since $F = \{B \cap F \mid B \in \mathcal{D}\}$, the mapping φ is strictly measurable. Consequently, the mapping $(x,\mu) \mapsto (x,\varphi\mu)$ of E × F̃ into E × ĨD is strictly measurable as well, and, therefore, there exists a measurable (Souslin) subset G' of E × ĨD such that

(1) $$\forall_{(x,\mu) \in E \times \tilde{F}} \; [(x,\mu) \in G \Leftrightarrow (x,\varphi\mu) \in G'] \; .$$

Next let ψ: (E×F)~ → (E×ID)~ be defined by $(\psi\mu)(C) := \mu(C \cap (E \times F))$ $(C \in E \otimes \mathcal{D})$. As $E \otimes F = \{C \cap (E \times F) \mid C \in E \otimes \mathcal{D}\}$, the mapping ψ is (strictly) measurable. Also, for each $\nu \in \tilde{E}$ and for every universally measurable q: E → F̃, we have

$$[\psi(\nu \times q)](A \times B) = (\nu \times q)[(A \times B) \cap (E \times F)] =$$

$$= (\nu \times q)[A \times (B \cap F)] = \int_A q(x)(B \cap F)\nu(dx) =$$

$$= \int_A \varphi(q(x))(B)\nu(dx) = [\nu \times (\varphi \circ q)](A \times B)$$

for every $A \in E$ and $B \in \mathcal{D}$, and hence

(2) $$\psi(\nu \times q) = \nu \times (\varphi \circ q) \; .$$

Further let p: (E×ID)~ × E → ĨD be measurable and such that for each $\nu \in \tilde{E}$ and for every universally measurable q: E → ĨD

(3) $$p(\nu \times q, x) = q(x) \quad \text{for } \nu\text{-almost every } x \in E.$$

Such a mapping p exists by Propositions 9.5 and 9.6.

Finally, let $\chi: (E \times \mathbb{D})^{\sim} \to (E \times \widetilde{\mathbb{D}})^{\sim}$ be defined by

$$(4) \qquad \int_{E \times \widetilde{\mathbb{D}}} f\, d(\chi\mu) := \int_{E \times \mathbb{D}} f(x, p(\mu, x))\mu(d(x,y)) \ ,$$

for each positive measurable function f on $E \times \widetilde{\mathbb{D}}$. In (4), the integral on the right hand side depends measurably on μ by Proposition 6.2 and this implies that the mapping χ is measurable. Note that (4) is equivalent to

$$(5) \qquad \int_{E \times \widetilde{\mathbb{D}}} f\, d(\chi\mu) = \int_{E} f(x, p(\mu, x))\nu(dx) \ ,$$

where ν is the marginal of μ on E.

Now let μ be a probability on $E \times F$ and let $\nu \times q$ be a decomposition of μ into its marginal ν on E and a universally measurable transition probability from E to F. Then we have

$$[(\chi \circ \psi)\mu](G') = [\chi(\psi(\nu \times q))](G') \overset{(2)}{=}$$

$$= [\chi(\nu \times (\varphi \circ q))](G') = \int_{E \times \widetilde{\mathbb{D}}} 1_{G'}\, d\chi(\nu \times (\varphi \circ q)) \overset{(5)}{=}$$

$$= \int_{E} 1_{G'}(x, p(\nu \times (\varphi \circ q), x))\nu(dx) \overset{(3)}{=}$$

$$= \int_{E} 1_{G'}(x, (\varphi \circ q)x)\nu(dx) \overset{(1)}{=} \int_{E} 1_{G}(x, q(x))\nu(dx) \ .$$

So, when the graph of q is contained in G, then $1_G(x, q(x)) = 1$ for ν-almost all $x \in E$, and hence, $[(\chi \circ \psi)\mu](G') = 1$. Suppose. on the other hand, that the latter equality holds. Then $(x, q(x)) \in G$ for ν-almost all $x \in E$. Now we may assume that a universally measurable mapping $q_0: E \to \widetilde{F}$ exists whose graph is contained in G; otherwise the proposition is trivially true. When we redefine $q(x) := q_0(x)$ for those $x \in E$ for which $(x, q(x)) \notin G$, then the graph of q will be contained in G and the equality $\mu = \nu \times q$ will still hold.

The foregoing implies that the probabilities μ on $E \times F$ that can be decomposed into their marginal on E and a universally measurable transition probability whose graph is contained in G, can be characterized by the equality $[(\chi \circ \psi)\mu](G') = 1$. Therefore, by the measurability of $\chi \circ \psi$, they constitute a measurable (Souslin) subset of $(E \times F)^{\sim}$. ☐

A particularly useful example of the measurable (Souslin) subset of $E \times \tilde{F}$ occurring in the foregoing proposition is expressed in

PROPOSITION 9.10. *Let* E *and* F *be measurable spaces,* S *a measurable (Souslin) subset of* $E \times F$, *and for every* $x \in E$ *let* $S_x := \{y \in F \mid (x,y) \in S\}$. *Then* $\{(x,\lambda) \in E \times \tilde{F} \mid \lambda(S_x) = 1\}$ *is a measurable (Souslin) subset of* $E \times \tilde{F}$.

PROOF. For every $(x,\lambda) \in E \times \tilde{F}$ we have

$$\lambda(S_x) = \int_{E \times F} 1_S(x',y')(\delta_x \times \lambda)(d(x',y')) \ .$$

It follows from Proposition 6.1 that the integral is a measurable (Souslin) function of $\delta_x \times \lambda$, and the examples 1 and 5 following Proposition 2.2 imply that $\delta_x \times \lambda$ depends measurably on (x,λ). From this the result follows. ☐

In the applications of analytic spaces in Chapter III we shall encounter convex linear combinations of probabilities. The following propositions express the fact that certain measurability properties of a set M of probabilities are preserved when we add to M all convex linear combinations of its members.

DEFINITION. Let M be a set of probabilities on a measurable space E. A probability μ on E is called a *countable convex combination* of members of M if there exist a countable family $(\mu_m)_{m \in I}$ of probabilities in M, and a countable family $(\alpha_m)_{m \in I}$ of numbers in $(0,1]$, such that $\Sigma_{m \in I} \alpha_m = 1$ and $\Sigma_{m \in I} \alpha_m \mu_m = \mu$. The set of countable convex combinations of members of M is called the *σ-convex hull* of M. M is called *σ-convex*, if M equals its σ-convex hull.

PROPOSITION 9.11. *Let* E *and* F *be measurable spaces,* G *a subset of* E × F̃ *and* M *the set of probabilities on* E × F *that can be decomposed into their marginal on* E *and a universally measurable transition probability whose graph is contained in* G. *When each of the sections* {λ ∈ F̃ | (x,λ) ∈ G} (x ∈ E) *of* G *is* σ*-convex, then* M *is* σ*-convex as well.*

PROOF. Let $\mu = \sum_{m\in I} \alpha_m \mu_m$ be a countable convex combination of probabilities on E × F such that for each m ∈ I the probability μ_m can be decomposed into its marginal ν_m on E and a universally measurable transition probability p_m whose graph is contained in G. Since the marginal ν of μ on E equals $\sum_{m\in I} \alpha_m \nu_m$, for each m ∈ I the probability ν_m is absolutely continuous with respect to ν, and by the Radon–Nikodym theorem there exists a positive measurable function f_m on E such that $d\nu_m = f_m \, d\nu$. Clearly $\sum_{m\in I} \alpha_m f_m(x) = 1$ for ν-almost all x ∈ E and, as the functions f_m are determined up to a ν-null function, this equality may be supposed to hold for each x ∈ E.

Now let the transition probability p: E → F̃ be defined by

$$p(x) := \sum_{m\in I} \alpha_m f_m(x) p_m(x) \ .$$

Then the graph of p is contained in G, since the sections of G are σ-convex. Moreover, for every measurable rectangle A × B ⊂ E × F we have

$$\mu(A \times B) = \sum_{m\in I} \alpha_m \mu_m(A \times B) = \sum_{m\in I} \alpha_m \int_A p_m(x;B) \nu_m(dx) =$$

$$= \int_A \sum_{m\in I} \alpha_m p_m(x;B) f_m(x) \nu(dx) = \int_A p(x;B) \nu(dx) =$$

$$= (\nu \times p)(A \times B) \ ,$$

so μ = ν × p. Hence μ ∈ M. □

PROPOSITION 9.12. *Let* E *be a measurable space,* F *a countably generated analytic space and* S *a Souslin subset of* E × F̃. *Moreover let* S̄ *be the subset of* E × F̃ *such that for every* x ∈ E *the section* {λ ∈ F̃ | (x,λ) ∈ S̄} *of* S̄ *equals the* σ*-convex hull of* {λ ∈ F̃ | (x,λ) ∈ S}. *Then* S̄ *is again a Souslin subset of* E × F̃.

PROOF

i)   We first consider the case that $E = \mathbb{D}$. Let

$$H := [0,1]^{\mathbb{N}} \times \mathbb{D} \times \widetilde{F}^{\mathbb{N}}$$

and

$$H_0 := \{(\alpha,x,y) \in H \mid \sum_{m \in \mathbb{N}} \alpha_m = 1 \text{ and } \forall_{m \in \mathbb{N}} (x,y_m) \in S\} \ .$$

It follows from the properties of product spaces (see Preliminaries 8 and 9) that, for each $m \in \mathbb{N}$, α_m and (x,y_m) depend measurably on $(\alpha,x,y) \in H$. So, by Proposition 4.3, H_0 is a Souslin subspace of H. As H is a product of analytic spaces, it follows from Proposition 7.6 (applied twice) that H_0 is analytic. Now let $\chi\colon H_0 \to \mathbb{D} \times \widetilde{F}$ be defined by

$$\chi(\alpha,x,y) = (x, \sum_{m \in \mathbb{N}} \alpha_m y_m) \ .$$

Then χ is measurable and the range of χ is $\overline{S}$. Since by Proposition 2.3 $\widetilde{F}$ is countably separated, $\mathbb{D} \times \widetilde{F}$ is countably separated as well. So, by Proposition 8.4, $\overline{S}$ is a Souslin subset of $\mathbb{D} \times \widetilde{F}$.

ii)   Next let E be an arbitrary measurable space and let E and F be the σ-algebras of E and F, respectively. By Propositions 4.2 and 4.5 we have $S \in S(E \otimes \widetilde{F}) \subset SS(E \times \widetilde{F}) = S(E \times \widetilde{F})$, and from this it follows (see the remark following the definition of the Souslin operation) that $S \in S(A \times \widetilde{F})$ for some countable subclass A of E. So S is a Souslin subset of $(E,E_0) \times \widetilde{F}$, where $E_0 := \sigma(A)$, which is a countably generated sub-σ-algebra of E.

Let $\varphi\colon (E,E_0) \to \mathbb{D}$ be strictly measurable (see Proposition 7.4) and let $\psi\colon (E,E_0) \times \widetilde{F} \to \mathbb{D} \times \widetilde{F}$ be defined by $\psi(x,\lambda) = (\varphi(x),\lambda)$. Then ψ is strictly measurable as well, and by Proposition 7.4 $S = \psi^{-1}T$ for some Souslin subset T of $\mathbb{D} \times \widetilde{F}$.

Finally, let $\overline{T}$ be the subset of $\mathbb{D} \times \widetilde{F}$ such that for every $x \in \mathbb{D}$ the section $\{\lambda \in \widetilde{F} \mid (x,\lambda) \in \overline{T}\}$ equals the σ-convex hull of $\{\lambda \in \widetilde{F} \mid (x,\lambda) \in T\}$. Then $\overline{S} = \psi^{-1}\overline{T}$ by the definition of ψ, and $\overline{T}$ is a Souslin subset of $\mathbb{D} \times \widetilde{F}$ by i). So $\overline{S}$ is a Souslin subset of $E \times \widetilde{F}$ by Proposition 4.3.   □

COROLLARY 9.13. *Let* F *be a countably generated analytic space and let* S *be a Souslin subset of* $\widetilde{F}$. *Then the* σ-*convex hull of* S *is a Souslin subset of* $\widetilde{F}$.

In the foregoing (Example 1, following Proposition 2.2) we introduced probabilities δ_x concentrated at a point x. A generalization of this kind of probability is given in:

DEFINITION. A probability is called *deterministic* when it attains the values 0 and 1 only.

In general, a deterministic probability need not be concentrated at one point. Also, neither the set of deterministic probabilities nor the set of probabilities concentrated at a point need be a measurable subset of the space of all probabilities. However, for countably generated spaces we have:

PROPOSITION 9.14. *Let* E *be a countably generated measurable space. Then*

i) *each deterministic probability on* E *is concentrated at one point;*

ii) *the set of deterministic probabilities on* E *is a measurable subset of* $\widetilde{E}$.

PROOF. Let A be a countable algebra generating the σ-algebra E of E.

i) Let $\mu \in \widetilde{E}$ be deterministic. The collection $A' := \{A \in A \mid \mu(A) = 1\}$ is countable, so $\mu(\cap A') = 1$, hence $\cap A' \neq \emptyset$. Now let $x \in \cap A'$. Then the probabilities μ and δ_x coincide on A and therefore on E. So $\mu = \delta_x$.

ii) Let $\mu \in \widetilde{E}$ be such that $\mu(A) \in \{0,1\}$ for every $A \in A$. Then $\{B \in E \mid \mu(B) \in \{0,1\}\}$ is a σ-algebra that contains A, and hence, equals E; so μ is deterministic. The set of all deterministic probabilities therefore equals

$$\bigcap_{A \in A} \{\mu \in \widetilde{E} \mid \mu(A) \in \{0,1\}\} ,$$

which is a countable intersection of measurable subsets of $\widetilde{E}$ and therefore measurable itself. $\square$

PROPOSITION 9.15. *Let* E *and* F *be countably generated measurable spaces, and let* F *be analytic. Then the probabilities on* E × F *that can be decomposed into their marginal on* E *and a universally measurable transition probability that attains deterministic values only constitute a measurable subset of* (E × F)$^\sim$.

PROOF. Let D be the set of deterministic probabilities on F, and let G := E × D. By Proposition 9.14 D is a measurable subset of $\tilde{F}$, so G is a measurable subset of E × $\tilde{F}$. Application of Proposition 9.9 now gives the desired result. $\qquad\square$

The final subject of this section is Ionescu-Tulcea's theorem; in fact a generalization of that theorem, since the transition probabilities appearing in it need not be measurable. Our proof is a modification of the proofs in [Neveu] section V.1.

LEMMA 9.16. *Let* E *and* F *be measurable spaces and let* p *be a (universally) measurable transition probability from* E *to* F. *Then* $\nu \mapsto \nu \times p$ *is a (universally) measurable mapping of* $\tilde{E}$ *into* (E × F)$^\sim$.

PROOF. Let A × B be a measurable rectangle in E × F. Then

$$(\nu \times p)(A \times B) = \int_E 1_A(x)p(x;B)\nu(dx) \ .$$

Since the integrand is a (universally) measurable function of x, by Proposition 6.1 the integral depends (universally) measurably on ν. The result follows from Proposition 2.2 applied to the collection of measurable rectangles of E × F. $\qquad\square$

LEMMA 9.17. *Let* E *and* F *be measurable spaces, let* μ *be a probability on* E × F *and let* $\mu^{(1)}$ *and* $\mu^{(2)}$ *be the marginals of* μ *on* E *and* F, *respectively. When* $\mu^{(1)}$ *or* $\mu^{(2)}$ *is deterministic, then* $\mu = \mu^{(1)} \times \mu^{(2)}$.

PROOF. Suppose that $\mu^{(1)}$ is deterministic. Let A × B be a measurable rectangle of E × F. When $\mu^{(1)}(A) = 0$, then $\mu(A \times B) \leq \mu(A \times F) = \mu^{(1)}(A) = 0$, so

$$\mu(A \times B) = 0 = \mu^{(1)}(A)\mu^{(2)}(B) = (\mu^{(1)} \times \mu^{(2)})(A \times B) \ .$$

When $\mu^{(1)}(A) = 1$, then $\mu(A^c \times B) \leq \mu(A^c \times F) = \mu^{(1)}(A^c) = 0$, so $\mu(A^c \times B) = 0$ and therefore

$$\mu(A \times B) = \mu(E \times B) = \mu^{(2)}(B) = \mu^{(1)}(A)\mu^{(2)}(B) = (\mu^{(1)} \times \mu^{(2)})(A \times B) \ .$$

So $\mu = \mu^{(1)} \times \mu^{(2)}$, because this equality holds on the class of measurable rectangles.   □

PROPOSITION 9.18 (Ionescu-Tulcea's theorem). *Let* $(E_1, E_2, \ldots)$ *be a finite or infinite sequence of measurable spaces and for each* n *let* p_n *be a (universally) measurable transition probability from* $E_1 \times \ldots \times E_n$ *to* E_{n+1}. *Then*

i)   *for each probability* ν *on* E_1 *there is a unique probability* μ_ν *on* $\Pi_{n \geq 1} E_n$ *such that*

$$\mu_\nu^{(1)} = \nu \quad \text{and, for each } n, \quad \mu_\nu^{(n+1)} = \mu_\nu^{(n)} \times p_n \ ,$$

where $\mu_\nu^{(n)}$ *denotes the marginal of* μ_ν *on* $E_1 \times \ldots \times E_n$

ii)   *there is a unique (universally) measurable transition probability* p *from* E_1 *to* $\Pi_{n \geq 2} E_n$ *such that, for each probability* ν *on* E_1,
$\mu_\nu = \nu \times p$.

PROOF. We prove the statement for infinite sequences; the case of finite sequences is contained in this.

Let A be the algebra of measurable subsets of the space $\Pi_{n \geq 1} E_n$ that depend on only finitely many coordinates. Let $x_1 \in E_1$ and let us define the function P on A as follows: For every $m \in \mathbb{N}$ and for every $A \in A$ depending on the first m coordinates only

(1) $P(A) := (\ldots(\delta_{x_1} \times p_1) \times \ldots \times p_m)(\varphi_m A) \ ,$

where φ_m is the projection of $\Pi_{n \geq 1} E_n$ onto $\Pi_{n \leq m+1} E_n$. Note that for each A this definition does not depend on the particular choice for m. Clearly, P is additive on A; we shall show that P is σ-additive on A, or, equivalently, that P is continuous at $\emptyset$. We argue by contradiction.

So let $(A_m)_{m \in \mathbb{N}}$ be a decreasing sequence in A satisfying $\cap_{m \in \mathbb{N}} A_m = \emptyset$, and suppose that $\lim_{m \to \infty} P(A_m) > 0$. Without loss of generality we suppose that for each $m \in \mathbb{N}$ the set A_m depends on the first m coordinates only, so

$$(2) \qquad \lim_{m \to \infty} (\dots(\delta_{x_1} \times p_1) \times \dots \times p_m)(\varphi_m A_m) = \lim_{m \to \infty} P(A_m) > 0 .$$

We shall prove, by induction on n, that an $x \in \Pi_{n \geq 1} E_n$ exists such that for each $n \in \mathbb{N}$

$$(3) \qquad \limsup_{m \to \infty} (\dots(\delta_{(x_1,\dots,x_n)} \times p_n) \times \dots \times p_m)(\varphi_m A_m) > 0 .$$

For $n = 1$, (3) is a consequence of (2). Also, for each $n \in \mathbb{N}$ by the definition of $\delta_{(x_1,\dots,x_n)} \times p_n$ the inequality (3) can be written as

$$\limsup_{m \to \infty} \int_{E_{n+1}} [(\dots(\delta_{(x_1,\dots,x_{n+1})} \times p_{n+1}) \times \dots \times p_m)(\varphi_m A_m)] p_n(x_1,\dots,x_n; dx_{n+1}) > 0 ,$$

and it follows from Fatou's lemma (see [Ash] 1.6.8) that this inequality remains valid when the order of integration and formation of the limit is changed. Hence

$$\limsup_{m \to \infty} (\dots(\delta_{(x_1,\dots,x_{n+1})} \times p_{n+1}) \times \dots \times p_m)(\varphi_m A_m) > 0$$

for some $x_{n+1} \in E_{n+1}$.

Let $n \in \mathbb{N}$. As $(A_m)_{m \in \mathbb{N}}$ is a decreasing sequence and since for each m the set A_m depends on the first m coordinates only, for each $m \geq n$ we have

$$(\dots(\delta_{(x_1,\dots,x_n)} \times p_n) \times \dots \times p_m)(\varphi_m A_m) \leq$$

$$\leq (\dots(\delta_{(x_1,\dots,x_n)} \times p_n) \times \dots \times p_m)(\varphi_m A_n) =$$

$$= \delta_{(x_1,\dots,x_n)}(\varphi_{n-1} A_n) = 1_{A_n}(x) .$$

Together with (3) this implies that $1_{A_n}(x) > 0$, and hence, $x \in A_n$. Since n is arbitrary, we conclude that $x \in \cap_{n \in \mathbb{N}} A_n$, which contradicts our assumption $\cap_{n \in \mathbb{N}} A_n = \emptyset$.

So P is σ-additive on A. Since the algebra A generates the σ-algebra of the space $\Pi_{n\geq 1}\ E_n$, P can be extended to a probability on that space, denoted again by P.

We now make explicit the dependence of P on x_1 by writing $P(x_1)$ instead of P. It follows from (1) by repeated application of Lemma 9.16 and by the measurability of δ that $P(x_1)(A)$ depends (universally) measurably on x_1 for each A ∈ A. Together with Proposition 2.2 this implies that the probability $P(x_1)$ depends (universally) measurably on x_1, and hence, by example 3 of section 2 that the marginal $p(x_1)$ of $P(x_1)$ on $E := \Pi_{n\geq 2}\ E_n$ depends (universally) measurably on x_1 as well. So p is a (universally) measurable transition probability from E_1 to E. Moreover, it follows from (1) that δ_x is the marginal of P(x) on E_1, so by Lemma 9.17 $P(x) = \delta_x \times p(x)$ $(x \in E_1)$.

Now let ν be a probability on E_1. Then for each m ∈ IN and for each A ∈ A depending on the first m coordinates only

$$(\nu \times p)(A) = \int_{E_1} \int_E 1_A(x,y)p(x;dy)\nu(dx) =$$

$$= \int_{E_1} \int_{E_1} \int_E 1_A(x',y)p(x;dy)\delta_x(dx')\nu(dx) =$$

$$= \int_{E_1} (\delta_x \times p(x))(A)\nu(dx) = \int_{E_1} P(x)(A)\nu(dx) =$$

$$= \int_{E_1} (...(\delta_x \times p_1) \times ... \times p_m)(\varphi_m A)\nu(dx) =$$

$$= (...(\nu \times p_1) \times ... \times p_m)(\varphi_m A) ,$$

so the marginal of ν × p on $E_1 \times ... \times E_{m+1}$ is $(...(\nu \times p_1) \times ... \times p_m)$. Hence, the probability ν × p satisfies i) and, since ν is arbitrary, the transition probability p satisfies ii).

Finally, the uniqueness of μ_ν in i) follows from the fact that μ_ν is completely defined by its restriction to the generating algebra A, hence, by its marginals. The uniqueness of p in ii) follows from this by the identity $\mu_{\delta_x} = \delta_x \times p = \delta_x \times p(x)$ $(x \in E_1)$. □

The main idea of Ionescu-Tulcea's theorem is that the transition probabilities $p_1, p_2, \ldots$ can be combined into a single transition probability p. Applying the above proposition to the sequence $(p_m, p_{m+1}, \ldots)$ and the probability $\nu := \delta_{(x_1, \ldots, x_m)}$ on $E_1 \times \ldots \times E_m$ $(x_n \in E_n,\ m \geq 1)$ we obtain the usual formulation of Ionescu-Tulcea's theorem, e.g. the one given in [Neveu].

CHAPTER III
DYNAMIC PROGRAMMING

In this final chapter we shall show how analytic measurable spaces can be used to solve measurability and selection problems occurring in dynamic programming. As an analytic measurable space is a generalization of an analytic topological space (see the remark at the end of section 8), our treatment of the subject has many points in common with the formalism of dynamic programming based an analytic topological spaces and many of our propositions and proofs are mere adaptations of those met with in the topological theory. Two results should, however, be considered as new, viz. our definition of decision models, which is a generalization of the usual one, and the existence of a, what may be called uniformly optimal, strategy (see Proposition 11.4).

§ 10. Decision models

For heuristics concerning decision models, which provide the mathematical structure dynamic programming is based on, we refer to [Hinderer] and [Bertsekas & Shreve].

DEFINITION. A *decision model* is a quintuple (S,A,G,p,u), the elements of which can be characterized and interpreted as follows.

i)   S is a sequence of countably generated analytic spaces $(S_n)_{n \in \mathbb{N}}$. S_n is called the *state space* at time n. Members of S_1 are called *initial states*.

ii)   A also is a sequence of countably generated analytic spaces $(A_n)_{n \in \mathbb{N}}$. A_n is called the *space of actions* that are available at time n.

For each $n \in \mathbb{N}$ we define $H_n := S_1 \times A_1 \times S_2 \times \ldots \times S_n$ and $H := \prod_{n=1}^{\infty} (S_n \times A_n)$. H is called the *space of realizations* and H_n the *space of histories* at time n.

iii) G is a sequence $(G_n)_{n \in \mathbb{N}}$, where G_n is a Souslin subset of $H_n \times \tilde{A}_n$. Moreover for each $n \in \mathbb{N}$ and $h \in H_n$ the set

$G_{n,h} := \{\lambda \in \tilde{A}_n \mid (h,\lambda) \in G_n\}$ is nonempty; $G_{n,h}$ is called the set of *admissible probabilities* at time n given history h.

iv) p is a sequence $(p_n)_{n \in \mathbb{N}}$, where for each $n \in \mathbb{N}$ p_n is a measurable transition probability from $H_n \times A_n$ to S_{n+1}. p is called the *transition law*.

v) u is a Souslin function on H. It is called the *utility*.

A *strategy* for the decision model defined above is a sequence $(q_n)_{n \in \mathbb{N}}$, where q_n is a universally measurable transition probability from H_n to A_n whose graph is contained in G_n. For every decision model at least one strategy exists, because for every $n \in \mathbb{N}$ it follows from the analyticity of $\tilde{A}_n$ that there exists a universally measurable mapping q_n of H_n into $\tilde{A}_n$ whose graph is contained in G_n.

By an *initial probability* we simply mean a probability on S_1. When for each $n \in \mathbb{N}$ and every $h \in H_n$ the set $G_{n,h}$ of admissible probabilities is σ-convex, then the decision model will be said to *allow combination of strategies*.

In the usual definition of a decision model, instead of G a sequence $(D_n)_{n \in \mathbb{N}}$ is introduced, where D_n is a Souslin subset of $H_n \times A_n$. For each $n \in \mathbb{N}$ and $h \in H_n$ the set $D_{n,h} := \{y \in A_n \mid (h,y) \in D_n\}$ is called the set of admissible actions at time n given history h. A strategy is then defined as a sequence $(q_n)_{n \in \mathbb{N}}$, where q_n is a universally measurable transition probability from H_n to A_n such that $\forall_{h \in H_n} \; q_n(h; D_{n,h}) = 1$.

It is easily deduced from Proposition 9.10 that our definition of a decision model is a generalization of the usual one. The generalized decision model has the advantage that one can prescribe not only which actions may be used at each step of a decision process, but also how they may be combined into a probability on the action space. For example, the property that a strategy q is deterministic (i.e. that each of the probabilities $q_n(h)$ is deterministic) is equivalent to the condition that for each $n \in \mathbb{N}$ the graph of q_n is contained in $H_n \times \Delta_n$, where Δ_n is the measurable subset of $\tilde{A}_n$ consisting of all deterministic probabilities on A_n (see Proposition 9.14).

To every initial probability of a decision model and to each strategy there corresponds, in a natural way, a probability on the space of

realizations. Moreover, the set of all such probabilities turns out to have useful algebraic and measure-theoretic properties.

DEFINITION. For every probability μ on the space of realizations of a decision model $\mu^{(2n-1)}$ and $\mu^{(2n)}$ denote its marginals on H_n and $H_n \times A_n$, respectively ($n \in \mathbb{N}$) (where H_n and A_n are the space of histories and the space of actions, respectively, both at time n).

PROPOSITION 10.1. *Let ν be an initial probability and q a strategy for a decision model with transition law p. Then there exists a unique probability μ on the space of realizations satisfying*

$$\mu^{(1)} = \nu \; , \quad \mu^{(2n)} = \mu^{(2n-1)} \times q_n \; , \; and \quad \mu^{(2n+1)} = \mu^{(2n)} \times p_n \quad (n \in \mathbb{N}) .$$

PROOF. This is a particular case of Proposition 9.18. □

Note that the above proposition is valid for a more general decision model, viz. one with arbitrary state and action spaces, and a transition law that is merely universally measurable.

DEFINITION. For any decision model the probability on the space of realizations that is generated by an initial probability ν and a strategy q in the sense of Proposition 10.1 will be denoted by $\mu(\nu,q)$. Probabilities of this kind will be called *strategic probabilities*.

A strategic probability depends universally measurably on the initial probability. This property, and a kind of reverse of it, are treated in the next three propositions. Also, the set of all strategic probabilities, as well as a set related to it, turn out to be Souslin sets; this is the subject of Propositions 10.5 and 10.6.

PROPOSITION 10.2. *For every strategy q the strategic probability $\mu(\nu,q)$ depends universally measurably on ν.*

PROOF. Due to Propositions 10.1 and 9.18 there exists a universally measurable transition probability p such that $\mu(\nu,q) = \nu \times p$ for each initial probability ν. The result now follows from Lemma 9.16. □

As a particular case of Proposition 10.2, for every strategy q the mapping $x \mapsto \mu(\delta_x, q)$ of the space of initial states into the space of strategic probabilities is universally measurable. The converse also holds:

PROPOSITION 10.3. *Let* $\varphi\colon S_1 \to \Sigma$ *be a universally measurable mapping of the space* S_1 *of initial states into the set* Σ *of strategic probabilities of a decision model, and let* $\varphi(x)^{(1)} = \delta_x$ *for every* $x \in S_1$. *Then there is a strategy* q *such that for every* $x \in S_1$ *one has* $\varphi(x) = \mu(\delta_x, q)$.

PROOF. Let (S, A, G, p', u) be the decision model and let $n \in \mathbb{N}$. When we apply Proposition 9.6 to the space of histories H_n and the action space A_n, both at time n, and to the probabilities $\varphi(x)^{(2n)}$ $(x \in S_1)$ on $H_n \times A_n$, we get a universally measurable mapping $p\colon (H_n \times A_n)^{\sim} \times H_n \to \tilde{A}_n$ such that

(1) $$\varphi(x)^{(2n)}(A \times B) = \int_A p(\varphi(x)^{(2n)}, h; B)\varphi(x)^{(2n-1)}(dh)$$

for each measurable rectangle $A \times B \subset H_n \times A_n$ and for each $x \in S_1$. Now for each $x \in S_1$ we have $(\varphi(x)^{(2n-1)})^{(1)} = \varphi(x)^{(1)} = \delta_x$ and therefore by Lemma 9.17 equality (1) is equivalent to

(2) $$\varphi(x)^{(2n)}(A \times B) = \int_A p(\varphi(h_1)^{(2n)}, h; B)\varphi(x)^{(2n-1)}(dh) .$$

Let $q_n\colon H_n \to \tilde{A}_n$ be defined by $q_n(h) = p(\varphi(h_1)^{(2n)}, h)$. Then q_n is universally measurable because p and φ are, and (2) can be written as

(3) $$\varphi(x)^{(2n)} = \varphi(x)^{(2n-1)} \times q_n \quad (x \in S_1) .$$

The set $\{h \in H_n \mid (h, q_n(h)) \in G_n\}$ is the inverse image of the Souslin subset G_n of $H_n \times \tilde{A}_n$ under the universally measurable mapping $h \mapsto (h, q_n(h))$, and as a consequence this set is a universally measurable subset of H_n. When therefore q' is an arbitrary strategy and q_n is redefined by

$$q_n(h) \quad \begin{aligned} &:= q_n(h) \quad \text{if} \quad (h, q_n(h)) \in G_n , \\ &:= q'_n(h) \quad \text{else,} \end{aligned}$$

then q_n becomes a universally measurable mapping whose graph is contained in G_n. Moreover we still have (3): Let $x \in S_1$. Since $\varphi(x)$ is strategic there exists a strategy, q^x say, such that $\varphi(x)^{(2n)} = \varphi(x)^{(2n-1)} \times q_n^x$. By (3) and Proposition 9.5 this implies that $q_n(h) = q_n^x(h)$, and hence, that $(h, q_n(h)) \in G_n$ for $\varphi(x)^{(2n-1)}$-almost all $h \in H_n$. So the redefinition of q_n does affect its values on a $\varphi(x)^{(2n-1)}$-null set only and as a consequence (3) remains valid.

Now letting n run through $\mathbb{N}$, we get a sequence (q_n) which obviously is a strategy, q say. Now for each $x \in S_1$ we have $\varphi(x)^{(1)} = \delta_x$ and $\varphi(x)^{(2n+1)} = \varphi(x)^{(2n)} \times p_n$ for each $n \in \mathbb{N}$ because $\varphi(x)$ is strategic. Together with (3) and Proposition 10.1 this implies that $\varphi(x) = \mu(\delta_x, q)$ for each $x \in S_1$. □

An equivalent formulation of the preceding proposition runs as follows:

COROLLARY 10.4. *For every initial state x of a decision model let a strategy q^x be given, and let $\mu(\delta_x, q^x)$ depend universally measurably on x. Then a strategy q exists such that $\mu(\delta_x, q^x) = \mu(\delta_x, q)$ for each initial state x.*

In general, the strategy q in Corollary 10.4 cannot be obtained by simply combining the strategies q^x in the following way:

$$q_n(h_1, h_2, \ldots) := q_n^{h_1}(h_1, h_2, \ldots) \qquad (n \in \mathbb{N}, \ h \in H_n) \ .$$

In fact the q_n's thus obtained may fail to be universally measurable, as the following example shows.

For all $n \in \mathbb{N}$ let $S_n = A_n = [0,1]$ and $p_n \equiv \lambda$, where λ is Lebesgue measure on $[0,1]$. Moreover let, for all $x \in S_1$, $q_n^x \equiv \lambda$ ($n \neq 2$) and

$$q_2^x(h_1, h_2, h_3) \ \begin{array}{l} = \delta_0 \quad \text{if} \quad h_2 = 0 \text{ and } x \in K \ , \\ = \lambda \quad \text{else,} \end{array}$$

where K is a subset of $[0,1]$ which is not universally measurable. A simple computation gives $\mu(\delta_x, q^x) = \delta_x \times \Pi_{n \geq 2} \lambda$, which depends measurably on x. On the other hand

$$q_2^{h_1}(h_1,h_2,h_3) \quad \begin{aligned} &= \delta_0 \quad \text{if } h_2 = 0 \text{ and } h_1 \in K , \\ &= \lambda \quad \text{else,} \end{aligned}$$

so $q_2^{h_1}(h_1,h_2,h_3)$ does not depend universally measurably on (h_1,h_2,h_3).

PROPOSITION 10.5. *The strategic probabilities on the space* H *of realizations of a decision model constitute a Souslin subset of* $\tilde{H}$. *When the decision model allows combination of strategies, then this subset is* σ-*convex.*

PROOF. Let the decision model be (S,A,G,p,u). For each $n \in \mathbb{N}$ let H_n be the space of histories at time n, M_n the set of probabilities on $H_n \times A_n$ that can be decomposed into a marginal on H_n and a universally measurable transition probability whose graph is contained in G_n, and M_n' the probabilities on $(H_n \times A_n) \times S_{n+1}$ that can be decomposed into a marginal on $H_n \times A_n$ and the transition probability p_n. Then by Proposition 10.1 the set Σ of strategic probabilities can be written as

$$\Sigma = \bigcap_{n \in \mathbb{N}} [\{\mu \in \tilde{H} \mid \mu^{(2n)} \in M_n\} \cap \{\mu \in \tilde{H} \mid \mu^{(2n+1)} \in M_n'\}] .$$

Let $n \in \mathbb{N}$. Proposition 9.9 implies that M_n is a Souslin subset of $(H_n \times A_n)^{\sim}$ and it follows from Proposition 9.11 that M_n is σ-convex when combination of strategies is allowed. Since $\mu \mapsto \mu^{(2n)}$ is a measurable linear mapping of $\tilde{H}$ into $(H_n \times A_n)^{\sim}$, the set $\{\mu \in \tilde{H} \mid \mu^{(2n)} \in M_n\}$ is a Souslin subset of $\tilde{H}$ which, in addition, is σ-convex provided that combination of strategies is allowed.

Next let G_n' be the graph of p_n. Then, by Proposition 7.3, G_n' is a measurable subset of $(H_n \times A_n) \times \tilde{S}_{n+1}$ and each of the sets $\{\lambda \in \tilde{S}_{n+1} \mid (x,\lambda) \in G_n'\}$ $(x \in H_n \times A_n)$ is a singleton and therefore is σ-convex. Moreover, M_n' can be described as the set of probabilities on $(H_n \times A_n) \times S_{n+1}$ that can be decomposed into a marginal on $H_n \times A_n$ and a transition probability whose graph is contained in G_n'. As a consequence, we can repeat the argument above and conclude that $\{\mu \in \tilde{H} \mid \mu^{(2n+1)} \in M_n'\}$ is a Souslin (in fact measurable) subset of $\tilde{H}$ that is σ-convex.

From the foregoing it is easily deduced that Σ has the desired properties. $\quad\square$

PROPOSITION 10.6. *Let* S_1 *be the space of initial states and* H *the space of realizations of a decision model. Let* Π *be the subset of* $\tilde{S}_1 \times \tilde{H}$ *consisting of the pairs* (ν,μ) *for which a strategy* q *exists such that* $\mu = \mu(\nu,q)$. *Then* Π *is a Souslin subset of* $\tilde{S}_1 \times \tilde{H}$.

PROOF. Let Σ be the set of strategic probabilities on H and let

$$\Delta := \{(\nu,\mu) \in \tilde{S}_1 \times \tilde{H} \mid \nu = \mu^{(1)}\} .$$

Then $\Pi = (\tilde{S}_1 \times \Sigma) \cap \Delta$. Now Σ is a Souslin subset of $\tilde{H}$ by Proposition 10.5 and Δ is a measurable subset of $\tilde{S}_1 \times \tilde{H}$, being the graph of the measurable mapping $\mu \mapsto \mu^{(1)}$ of $\tilde{H}$ into the countably separated space $\tilde{S}_1$ (see Propositions 2.3 and 7.3). The foregoing implies that Π is a Souslin subset of $\tilde{S}_1 \times \tilde{H}$. $\square$

§ 11. The expected utility and optimal strategies

The object of interest in dynamic programming is the expected utility. We first define this quantity and prove its measurability.

DEFINITION. Let a decision model be given with space of realizations H and utility u. For every initial probability ν and for each strategy q we define

$$w(\nu,q) := \int_{H} u \, d\mu(\nu,q) .$$

The function w is called the *expected utility*. Also, for every initial probability ν we define

$$v(\nu) := \sup_{q} w(\nu,q) ,$$

where the supremum is taken over all strategies. The function v is called the *maximal expected utility*.

PROPOSITION 11.1. *The maximal expected utility* v *is a Souslin function, and for each strategy* q *the expected utility* $w(\nu,q)$ *is a universally measurable function of* ν.

PROOF. Let the Souslin subset Π of $\tilde{S}_1 \times \tilde{H}$ be defined as in Proposition 10.6 and let $f\colon \Pi \to \overline{\mathbb{R}}$ be defined by $f(\nu,\mu) = \int u\,d\mu$, where u is the utility. Then, by Proposition 6.2, f is a Souslin function because u is, and for each $\nu \in \tilde{S}_1$ we have $v(\nu) = \sup \{f(\nu,\mu) \mid \mu\colon (\nu,\mu) \in \Pi\}$. Application of the exact selection theorem (Proposition 7.1) yields the desired result.

Also it follows from the definition of w and Proposition 6.1 that $w(\nu,q)$ is a universally measurable (in fact, Souslin) function of $\mu(\nu,q)$. By Proposition 10.2 this implies that $w(\nu,q)$ depends universally measurably on ν. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

In the rest of this section we consider optimal strategies, i.e. strategies for which the expected utility equals the maximal expected utility. In particular the question is raised whether optimality, or a prescribed optimality defect, can be realized simultaneously for all initial probabilities by one and the same strategy. A second topic will be the linear dependence of the (maximal) expected utility on the initial probability. We conclude with a derivation of the optimality equation.

DEFINITION. Let ν be an initial probability and q a strategy for a decision model. Then q is called ν-*optimal* if $w(\nu,q) = v(\nu)$.

It may happen that for an initial probability ν no ν-optimal strategies exist (see the example following Proposition 11.4).

PROPOSITION 11.2. *Let* h *be a universally measurable function on the space* S_1 *of initial states of a decision model satisfying*

$$h(x) \begin{array}{l} < v(\delta_x) \quad \textit{if} \quad -\infty < v(\delta_x) \;, \\ = -\infty \qquad \textit{else.} \end{array}$$

Then there exists a strategy q *such that*

i) $w(\delta_x,q) \geq h(x)$ *for each* $x \in S_1$;

ii) q *is* δ_x-*optimal for each* $x \in S_1$ *for which a* δ_x-*optimal strategy exists.*

PROOF. Let $\Pi \subset \tilde{S}_1 \times \tilde{H}$ be defined as in Proposition 10.6 and let $\Pi' \subset S_1 \times \tilde{H}$ be the inverse image of the Souslin set Π under the measurable mapping

$(x,\mu) \mapsto (\delta_x,\mu)$ of $S_1 \times \tilde{H}$ into $\tilde{S}_1 \times \tilde{H}$. Then Π' is a Souslin subset of $S_1 \times \tilde{H}$ and for all $(x,\mu) \in S_1 \times \tilde{H}$ we have $(x,\mu) \in \Pi'$ iff $\mu = \mu(\delta_x,q)$ for some strategy q.

Moreover, let $f: S_1 \times \tilde{H} \to \overline{\mathbb{R}}$ be defined by $f(x,\mu) = \int u \, d\mu$, where u is the utility. Then f is a Souslin function and for each $x \in S_1$ we have $v(\delta_x) = \sup \{f(x,\mu) \mid \mu: (x,\mu) \in \Pi'\}$. Now from the exact selection theorem (Proposition 7.1) the existence follows of a universally measurable mapping $\varphi: S_1 \to \tilde{H}$ the graph of which is contained in Π', and such that $f(x,\varphi(x)) \geq h(x)$ for each $x \in S_1$, and $f(x,\varphi(x)) = v(\delta_x)$ for each $x \in S_1$ for which a δ_x-optimal strategy exists.

Now by Proposition 10.3 a strategy q exists such that for each $x \in S_1$ $\varphi(x) = \mu(\delta_x,q)$, and consequently $f(x,\varphi(x)) = \int u \, d\varphi(x) = w(\delta_x,q)$. The definition of φ now implies that q has the properties stated in the proposition. $\square$

At every instant of time in a decision process one knows both the state the system is in, and (given the strategy) the probability distribution of the state the system will enter at the next instant of time. Due to this alternation of states and probabilities, either a quantity is most naturally expressed as a function of states or as a function of probabilities. The following proposition gives the connection between the two representations for the (maximal) expected utility. Note that representing a state x by the probability δ_x amounts merely to an identification of non-separated points.

PROPOSITION 11.3. *Let ν be an initial probability for a decision model.*

i) *When $-\infty < v(\nu)$, then $x \mapsto v(\delta_x)$ is quasi-integrable with respect to ν and $\int v(\delta_x)\nu(dx) = v(\nu)$.*

ii) *$\int v(\delta_x)\nu(dx) = \sup_{q'} \int w(\delta_x,q')\nu(dx)$, where the supremum is taken over all strategies q'.*

iii) *When q is a strategy such that the utility is quasi-integrable with respect to $\mu(\nu,q)$, in particular, when $w(\nu,q) > -\infty$, then $x \mapsto w(\delta_x,q)$ is quasi-integrable with respect to ν and $\int w(\delta_x,q)\nu(dx) = w(\nu,q)$.*

iv) *When the utility is quasi-integrable with respect to every strategic probability, then $\int v(\delta_x)\nu(dx) = v(\nu)$.*

PROOF. The order in which we shall prove the various parts of the proposition is: ii), iii), iv), and i).

ii)   Suppose that $\int\limits_{-} v(\delta_x)\nu(dx) > -\infty$. As by Proposition 11.1 $v(\delta_x)$ is a universally measurable function of x, it follows from Proposition 11.2 that a sequence $(q^m)_{m\in\mathbb{N}}$ of strategies exists such that for each initial state x and $m \in \mathbb{N}$

$$w(\delta_x, q^m) \quad \begin{array}{l} \geq v(\delta_x) - m^{-1} \quad \text{if} \quad v(\delta_x) < \infty, \\[2mm] \geq m \quad\quad\quad\quad\quad \text{else.} \end{array}$$

Then the sequence $(x \mapsto w(\delta_x, q^m))_{m\in\mathbb{N}}$ of functions converges to the function $x \mapsto v(\delta_x)$, while the integrable function $x \mapsto \min\{v(\delta_x) - 1, 0\}$ is a common lower bound. So by Fatou's lemma (see [Ash] 1.6.8)

$$\int\limits_{-} v(\delta_x)\nu(dx) \leq \liminf_{m} \int w(\delta_x, q^m)\nu(dx) \leq \sup_{q'} \int\limits_{-} w(\delta_x, q')\nu(dx) .$$

This inequality is satisfied also when the left hand member equals $-\infty$. The reverse inequality on the other hand is a direct consequence of $v(\delta_x) \geq w(\delta_x, q')$.

iii)   From the definition of $\mu(\nu, q)$ it follows that for every measurable cylinder C we have $\mu(\nu, q)(C) = \int \mu(\delta_x, q)(C)\nu(dx)$. This results extends to integrals of positive functions in the usual way. In particular, one has for the utility u

$$(\ast) \quad w(\nu, q) = \int\limits_{-} u\,d\mu(\nu, q) = \int u^+ d\mu(\nu, q) - \int u^- d\mu(\nu, q) =$$

$$= \iint u^+ d\mu(\delta_x, q)\nu(dx) - \iint u^- d\mu(\delta_x, q)\nu(dx) .$$

As u is quasi-integrable with respect to $\mu(\nu, q)$, at least one of the repeated integrals is finite, say the one of u^- for definiteness. Then $\int u^- d\mu(\delta_x, q)$ is finite for ν-almost all x, and, as a function of x, is integrable with respect to ν. Hence the last expression in $(\ast)$ can be written as

$$\int \left[\int u^+ d\mu(\delta_x, q) - \int u^- d\mu(\delta_x, q) \right] \nu(dx) \ ,$$

which in turn equals $\int w(\delta_x, q) \nu(dx)$.

iv)   It follows from ii) and iii) that

$$\underline{\int} v(\delta_x) \nu(dx) = \sup_{q'} \underline{\int} w(\delta_x, q') \nu(dx) = \sup_{q'} w(\nu, q') = v(\nu) \ .$$

i)   For each $n \in \mathbb{N}$ let the function u^n be defined on the space of reali-
zations by $u^n(h) = \min \{u(h), n\}$ and let v^n be the maximal expected
utility corresponding to the utility u^n. Then for every n the function
u^n is quasi-integrable with respect to every strategic probability
and it follows from iv) that $v^n(\nu) = \underline{\int} v^n(\delta_x) \nu(dx)$. Also for every
initial distribution λ we have

$$v(\lambda) = \sup_q w(\lambda, q) = \sup_q \underline{\int} u \, d\mu(\lambda, q) = \sup_q \sup_n \int u^n d\mu(\lambda, q) =$$

$$= \sup_n \sup_q \int u^n d\mu(\lambda, q) = \sup_n v^n(\lambda) \ .$$

So

$$-\infty < v(\nu) = \sup_n v^n(\nu) = \sup_n \underline{\int} v^n(\delta_x) \nu(dx) =$$

$$= \underline{\int} \sup_n v^n(\delta_x) \nu(dx) = \int v(\delta_x) \nu(dx) \ ,$$

where the third equality follows from the fact that $v^n(\delta_x)$ is in-
creasing in n and that $-\infty < \underline{\int} v^n(\delta_x) \nu(dx)$ for sufficiently large n.   □

The condition on the utility in iii) cannot be omitted. As an example
we consider a decision model for which $S_1 = (0,1)$, $A_1 = \mathbb{R}$ and
$u(x, y, \ldots) = y$. Take a strategy q for which $q_1(x)(\{x^{-1}\}) = \frac{1}{2}$ and
$q_1(x)(\{-x^{-1}\}) = \frac{1}{2}$ ($x \in S_1$) and for the initial probability ν take Lebesgue
measure on $(0,1)$. Then $w(\delta_x, q) = 0$ for all x and $w(\nu, q) = -\infty$, so
$w(\nu, q) \neq \int w(\delta_x, q) \nu(dx)$.

With the aid of Proposition 11.3 we can now extend the results in Proposition 11.2 to arbitrary initial probabilities.

PROPOSITION 11.4. *Let the utility of a decision model be quasi-integrable with respect to every strategic probability and let combination of strategies be allowed. Then the expected utility* w *and the maximal expected utility* v *are bounded on at least one side. When moreover* $\varepsilon > 0$ *and* $\eta\colon S_1 \to (0, \infty)$ *is a universally measurable function on the space* S_1 *of initial states, then there is a strategy* q *such that*

i) $w(\delta_x, q) \geq v(\delta_x) - \eta(x)$ *for each initial state x;*

ii) $w(\nu, q) \geq v(\nu) - \varepsilon$ *for each initial probability* ν*;*

iii) q *is* ν*-optimal for each initial probability* ν *for which a* ν*-optimal strategy exists.*

PROOF. Suppose that the function w is unbounded on both sides. Then for each $m \in \mathbb{N}$ there exists a strategic probability μ_m such that $\int u \, d\mu_m \geq 2^m$ and therefore such that $\int u^+ d\mu_m \geq 2^m$, where u is the utility. When $\mu := \sum_{m=1}^{\infty} 2^{-m} \mu_m$, then by Proposition 10.5 μ is a strategic probability and

$$\int u^+ d\mu \geq \sum_m 2^{-m} 2^m = \infty .$$

In the same way one can show that a strategic probability μ' exists such that $\int u^- d\mu' = \infty$. Consequently, the utility u is not quasi-integrable with respect to the strategic probability $\frac{1}{2}(\mu + \mu')$, which contradicts our assumptions. The function w therefore is bounded on at least one side. The definition of v now implies that v is bounded on the same side.

In the proof of the remaining part of the proposition we suppose, without loss of generality, that $\eta(x) < \varepsilon$ for every $x \in S_1$. Let $h\colon S_1 \to \overline{\mathbb{R}}$ be defined by

$$h(x) \quad \begin{aligned} &= v(\delta_x) - \eta(x) \quad \text{if } v(\delta_x) \text{ is finite,} \\ &= -\infty \quad\quad\quad\quad \text{else.} \end{aligned}$$

Application of Proposition 11.2 with this function h yields a strategy q such that $w(\delta_x, q) \geq h(x)$ for each $x \in S_1$ and such that q is δ_x-optimal for every $x \in S_1$ for which a δ_x-optimal strategy exists. We shall prove that q satisfies i), ii), and iii).

To prove i) we merely have to show that a δ_x-optimal strategy exists for every $x \in S_1$ satisfying $v(\delta_x) = \infty$. Therefore, let $x \in S_1$ and $v(\delta_x) = \infty$. Moreover, for each $m \in \mathbb{N}$ let q^m be a strategy such that $w(\delta_x, q^m) \geq 2^m$ and let $\mu := \Sigma_{m \in \mathbb{N}} \, 2^{-m} \mu(\delta_x, q^m)$. Then μ is a strategic probability and

$$\mu^{(1)} = \sum_m 2^{-m} \mu(\delta_x, q^m)^{(1)} = \sum_m 2^{-m} \delta_x = \delta_x \ ,$$

so $\mu = \mu(\delta_x, q')$ for some strategy q'. With utility u we now have

$$\int u^+ d\mu(\delta_x, q') = \sum_m 2^{-m} \int u^+ d\mu(\delta_x, q^m) \geq \sum_m 2^{-m} \int u \, d\mu(\delta_x, q^m) =$$

$$= \sum_m 2^{-m} w(\delta_x, q^m) \geq \sum_m 2^{-m} 2^m = \infty$$

and since u is quasi-integrable with respect to $\mu(\delta_x, q')$ this implies that $w(\delta_x, q') = \infty$. So q' is δ_x-optimal and thereby i) has been proved.

Next let ν be an initial probability. For every $x \in S_1$ we have $\eta(x) < \epsilon$ and therefore $w(\delta_x, q) \geq v(\delta_x) - \epsilon$ by i). From this inequality ii) can be obtained by integration with respect to ν and application of Proposition 11.3.

To prove iii) we first consider an initial probability ν such that $v(\nu)$ is infinite. From ii) it then follows that $w(\nu, q) = v(\nu)$ and hence that q is ν-optimal. Next let $v(\nu)$ be finite and suppose that a ν-optimal strategy q' exists. Then by Proposition 11.3

$$\int w(\delta_x, q') \nu(dx) = w(\nu, q') = v(\nu) = \int v(\delta_x) \nu(dx) \ .$$

From the finiteness of these integrals together with the inequalities $w(\delta_x, q') \leq v(\delta_x)$ $(x \in S_1)$ we conclude that $w(\delta_x, q') = v(\delta_x)$ for ν-almost all $x \in S_1$. So for ν-almost all $x \in S_1$ a δ_x-optimal strategy exists and hence the equality $w(\delta_x, q) = v(\delta_x)$ holds, as follows from the definition of q. This in turn implies by Proposition 11.3 that

$$w(\nu, q) = \int w(\delta_x, q) \nu(dx) = \int v(\delta_x) \nu(dx) = v(\nu) \ ,$$

so q is ν-optimal. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

The preceding proposition cannot be essentially improved by allowing ε to depend on ν, as the following example shows.

Consider a decision model for which $S_1 := \{0,1\}$, $A_1 := (0,1)$ and $u(x,y,z,\ldots) := y$. Then for every initial probability ν and for every strategy q we have

$$w(\nu,q) = \nu(\{0\}) \int y\, q_1(0)(dy) + \nu(\{1\}) \int y\, q_1(1)(dy) \leq$$

$$\leq \max \left\{ \int y\, q_1(0)(dy), \int y\, q_1(1)(dy) \right\} ,$$

and the last expression is smaller than 1 and independent of ν. Moreover, $v(\nu) = 1$. So, for every strategy q, $w(\nu,q)$ is bounded away from $v(\nu)$ uniformly in ν.

The next proposition concerns the optimality equation and a related result on optimal strategies. For its formulation we introduce a transformation of decision models.

DEFINITION. The *contraction* of a decision model (S,A,G,p,u) is the decision model $(\hat{S},\hat{A},\hat{G},\hat{p},\hat{u})$ defined by

i) $\quad \hat{S}_1 = S_1 \times A_1 \times S_2$;

ii) $\quad \hat{S}_{n+1} = S_{n+2}, \; \hat{A}_n = A_{n+1}, \; \hat{G}_n = G_{n+1}, \; \hat{p}_n = p_{n+1} \quad (n \in \mathbb{N})$;

iii) $\quad \hat{u} = u$.

For every strategy q for the given model the strategy $\hat{q}$ for the contracted model is defined by $\hat{q}_n = q_{n+1} \quad (n \in \mathbb{N})$.

Note that in the above definition of $\hat{p}$, $\hat{u}$ and $\hat{q}$ we have tacitly identified domains of definitions like $(S_1 \times A_1 \times S_2) \times A_2 \times S_3 \times \ldots$ and $S_1 \times A_1 \times S_2 \times A_2 \times S_3 \times \ldots$. Moreover, for every contraction of a decision model we will use the symbols μ, w and v for strategic probability, expected utility, and maximal expected utility, respectively; in each case the arguments of these functions will indicate to which contraction they apply.

The contraction operation can be repeated a number (say n) of times; this results in a decision model having as its initial states the histories up to time n of the original model. To all these contracted models the results derived so far apply.

PROPOSITION 11.5. *Let ν be an initial probability and let q be a strategy for a decision model with transition law p. Then*

i) $\quad w(\nu,q) = w(\nu \times q_1 \times p_1, \hat{q})$,

ii) $\quad$ *(Optimality equation)* $\quad v(\nu) = \sup_{q'} v(\nu \times q_1' \times p_1)$, *where the supremum is over all strategies q',*

iii) $\quad q$ *is ν-optimal iff both $\hat{q}$ is $\nu \times q_1 \times p_1$-optimal and $v(\nu) = v(\nu \times q_1 \times p_1)$.*

PROOF.

i) $\quad$ As a consequence of the definition of $\hat{q}$ and of Proposition 10.1 we have $\mu(\nu,q) = \mu(\nu \times q_1 \times p_1, \hat{q})$. Consequently, for utility u,

$$w(\nu,q) = \int \underline{u} \, d\mu(\nu,q) = \int \underline{\hat{u}} \, d\mu(\nu \times q_1 \times p_1, \hat{q}) = w(\nu \times q_1 \times p_1, \hat{q}) \ .$$

ii) $\quad$ From i) we deduce

$$v(\nu) = \sup_{q'} w(\nu,q') = \sup_{q'} w(\nu \times q_1' \times p_1, \hat{q}') =$$

$$= \sup_{q_1'} \sup_{\hat{q}'} w(\nu \times q_1' \times p_1, \hat{q}') = \sup_{q_1'} v(\nu \times q_1' \times p_1) \ .$$

iii) $\quad$ Let q be ν-optimal. Then from ii) and i) we conclude

$$v(\nu \times q_1 \times p_1) \le \sup_{q'} v(\nu \times q_1' \times p_1) = v(\nu) = w(\nu,q) =$$

$$= w(\nu \times q_1 \times p_1, \hat{q}) \le v(\nu \times q_1 \times p_1) \ .$$

So $w(\nu \times q_1 \times p_1, \hat{q}) = v(\nu \times q_1 \times p_1)$, i.e. $\hat{q}$ is $\nu \times q_1 \times p_1$-optimal, and $v(\nu) = v(\nu \times q_1 \times p_1)$.

Next, let $\hat{q}$ be $\nu \times q_1 \times p_1$-optimal and $v(\nu) = v(\nu \times q_1 \times p_1)$. Together with i) this implies

$$w(\nu,q) = w(\nu \times q_1 \times p_1, \hat{q}) = v(\nu \times q_1 \times p_1) = v(\nu) \ .$$

So q is ν-optimal. □

We conclude this section by giving two examples of the way in which Proposition 11.5 can be used. The results are well known, be it in a less general version.

DEFINITION. Let ν be an initial probability and let q be a strategy for some decision model. For each $k \in \mathbb{N}$ let q^k be the (k-1)-th contraction of q and let μ^k be the marginal of $\mu(\nu,q)$ on the space H_k of histories up to time k. Then the strategy q is called ν-*conserving* if

$$\forall_{k \in \mathbb{N}} \ [v(\mu^k) = v(\mu^{k+1})]$$

and it is called ν-*equalizing* if

$$\lim_{k \to \infty} [v(\mu^k) - w(\mu^k, q^k)] = 0 \ .$$

PROPOSITION 11.6. *Let* ν *be an initial probability and let* q *be a strategy for some decision model. Then* q *is* ν-*optimal iff* q *is both* ν-*conserving and* ν-*equalizing.*

PROOF. Let q be ν-optimal. Applying Proposition 11.5 iii) to all contractions of the decision model we get for each $k \in \mathbb{N}$ (using the notation of the preceding definition) that q^k is μ^k-optimal, i.e. $w(\mu^k, q^k) = v(\mu^k)$, and that $v(\mu^k) = v(\mu^{k+1})$. So q is ν-equalizing and ν-conserving.

Next, let q be ν-conserving and ν-equalizing. The conservingness of q implies that $v(\mu^k) = v(\mu^{k+1})$ for each $k \in \mathbb{N}$. From Proposition 11.5 i), applied to all contractions, we deduce for each $k \in \mathbb{N}$ that $w(\mu^k, q^k) = w(\mu^{k+1}, q^{k+1})$, hence

$$v(\mu^k) - w(\mu^k, q^k) = v(\mu^{k+1}) - w(\mu^{k+1}, q^{k+1}) \ .$$

As q is ν-equalizing, we must have $v(\mu^k) - w(\mu^k, q^k) = 0$ for each $k \in \mathbb{N}$ and, in particular, for k = 1. So q is ν-optimal. □

As in the case of Proposition 10.1 the Propositions 11.5 and 11.6 are easily seen to be valid for a more general decision model: arbitrary state and action spaces, and a transition law and utility that are merely universally measurable.

PROPOSITION 11.7. *Let there be given a decision model and, for each* $k \in \mathbb{N}$, *let* v^k *be the function defined on the space* H *of realizations by* $v^k(h) = v(\delta_{h_k})$, *where, for each* $h \in H$, h_k *is the projection of* h *on the space* H_k *of realizations up to time* k. *Moreover, for each* $k \in \mathbb{N}$, *let* H_k *be the* σ-*algebra of measurable subsets of* H *depending on the coordinates up to time* k *only.*

When ν *is an initial probability and* q *a strategy such that* $-\infty < w(\nu,q)$, *then the sequence* $(v^k, H_k)_{k \in \mathbb{N}}$ *is a supermartingale with respect to* $\mu(\nu,q)$. *When, in addition,* q *is* ν-*conserving and* $v(\nu) < \infty$, *then this supermartingale is even a martingale.*

PROOF. Let p be the transition law of the given model. By Proposition 11.3 iii) we have

$$-\infty < w(\nu,q) = \int w(\delta_x,q)\nu(dx) \ ,$$

so for ν-almost every initial state x we have

(1) $\qquad -\infty < w(\delta_x,q) = w(\delta_x \times q_1 \times p_1, \hat{q}) \le v(\delta_x \times q_1 \times p_1) =$

$$= \int v(\delta_{(x',y,z)}) \delta_x \times q_1 \times p_1 (d(x',y,z)) \ ,$$

where the first and second equality are consequences of the Propositions 11.5 i) and 11.3 i), respectively.

Also we have, in a similar way,

$$-\infty < w(\nu,q) = w(\nu \times q_1 \times p_1, \hat{q}) \le v(\nu \times q_1 \times p_1) =$$

$$= \int v(\delta_{(x,y,z)}) \nu \times q_1 \times p_1 (d(x,y,z)) \ .$$

So for every positive bounded measurable function f on H_1 we have

$$(2) \qquad -\infty < \int f(x)v(\delta_{(x,y,z)})\nu{\times}q_1{\times}p_1(d(x,y,z)) =$$

$$= \int f(x) \int v(\delta_{(x',y,z)})\delta_x{\times}q_1{\times}p_1(d(x',y,z))\nu(dx) =$$

$$= \int f(x)v(\delta_x{\times}q_1{\times}p_1)\nu(dx) \ ,$$

where the first equality is due to the definition of $\nu \times q_1 \times p_1$ and quasi-integrability, while the second one is a consequence of (1).

The arbitrariness of f now implies that the function $x \mapsto v(\delta_x{\times}q_1{\times}p_1)$ is a (universally measurable) version of the conditional expectation of the function $(x,y,z) \mapsto v(\delta_{(x,y,z)})$ with respect to the probability $\nu \times q_1 \times p_1$ and the σ-algebra of measurable subsets of H_2 depending on the initial state only. Using the definition of v^2, we conclude that $x \to v(\delta_x{\times}q_1{\times}p_1)$ is a version of the conditional expectation of v^2 with respect to $\mu(\nu,q)$ and H_1.

Now, for every initial state x, we have by Proposition 11.5 ii)

$$(3) \qquad v^1(x) = v(\delta_x) = \sup_{q'} v(\delta_x{\times}q_1'{\times}p_1) \ge v(\delta_x{\times}q_1{\times}p_1) \ .$$

When, in addition, q is ν-conserving and $v(\nu) < \infty$, then

$$-\infty < v(\nu) = v(\nu{\times}q_1{\times}p_1) < \infty \ ,$$

so by Proposition 11.3 i)

$$-\infty < \int v(\delta_x)\nu(dx) = \int v(\delta_{(x,y,z)})\nu{\times}q_1{\times}p_1(d(x,y,z)) < \infty \ ,$$

and together with (3) this implies

$$(4) \qquad v^1(x) = v(\delta_x) = v(\delta_x{\times}q_1{\times}p_1)$$

for ν-almost every initial state x.

From (4) and (3) it follows that v^1 equals or majorizes a version of the conditional expectation of v^2 with respect to $\mu(\nu,q)$ and H_1, according to whether q is ν-conserving and $v(\nu) < \infty$ or not.

The only thing that remains to be proved is that a similar relation holds for each pair of functions v^k and v^{k+1} ($k = 2,3,\ldots$). Now the validity of any of the three conditions $-\infty < w(\nu,q)$, q is ν-conserving, or $v(\nu) < \infty$

implies by Proposition 11.5 the validity of a similar condition for each contraction of the model. Therefore the reasoning leading to the relation between v^1 and v^2 applies equally well to each contraction. $\square$

REFERENCES

ASH, R.B., *Measure, Integration and Functional Analysis*, Academic Press, New York etc., 1972.

BERTSEKAS, D.P. and S.E. SHREVE, *Stochastic Optimal Control*, Academic Press, New York etc., 1978.

BLACKWELL, D., *Discounted dynamic programming*, Ann. Math. Statist. 36 (1965) 226-235.

BLACKWELL, D., D. FREEDMAN, and M. ORKIN, *The optimal reward operator in dynamic programming*, Ann. Prob. 2 (1974) 926-941.

CHOW, Y.S. and H. TEICHER, *Probability Theory*, Springer-Verlag, New York etc., 1978.

CHRISTENSEN, J.P.R., *Topology and Borel Structure*, North-Holland Mathematics Studies No. 10, North Holland Publishing Company, Amsterdam-London, 1974.

COHN, D.L., *Measure Theory*, Birkhäuser, Boston etc., 1980.

HINDERER, K., *Foundations of Non-stationary Dynamic Programming with Discrete Time Parameter*, Lecture Notes in Operations Research and Mathematical Systems No. 33, Springer-Verlag, Berlin etc., 1970.

HOFFMANN-JØRGENSEN, J., *The theory of Analytic Spaces*, Various Publication Series No. 10, Dept. of Mathematics, University of Aarhus, Denmark, 1970.

NEVEU, J., *Mathematical Foundations of the Calculus of Probability*, Holden-Day, San Francisco etc., 1965.

PFANZAGL, J. and W. PIERLO, *Compact Systems of Sets*, Lecture Notes in Mathematics No. 16, Springer-Verlag, Berlin etc., 1966.

SHREVE, S.E., *Dynamic Programming in Complete Separable Spaces*, University of Illinois, 1977.

TOPSØE, F., *Topology and Measure*, Lecture Notes in Mathematics No. 133, Springer-Verlag, Berlin etc., 1970.

INDEX

A^c , $A \setminus B$	1
$A_{c,d,s,\delta,\sigma}$, $A_{s\delta}$, σA	1
$A \mid E$	3
E_μ	10
(E,E)	3
$\tilde{E}$, $\tilde{E}$, $(E,E)^{\sim}$	7
$\tilde{E}(E)$	6
$E \times F$, $E \otimes F$, $A \times B$	4
$\Pi_{i \in I} F_i$, $\otimes_{i \in I} F_i$	3
F^I	4
L	23
S	14
U	10
δ_x	5

φ^{-1}	2
$\mu \circ \varphi^{-1}$	8
$\mu^{(n)}$, $\mu(\nu,q)$	77
π_i	3
$\nu \times p$	57
f^+ , f^-	31
$(n_1,n_2,\ldots)$	4
$p(x;B)$, $p(x)(B)$	57
v , w	81
$\mathbb{N}$, $\mathbb{R}$, $\overline{\mathbb{R}}$	5
$\mathbb{N}^{\mathbb{N}}$	4
$\mathbb{D}$	30
1_A	6
$\mapsto$	6
$\int$	31

MATHEMATICAL CENTRE TRACTS

1 T. van der Walt. *Fixed and almost fixed points.* 1963.

2 A.R. Bloemena. *Sampling from a graph.* 1964.

3 G. de Leve. *Generalized Markovian decision processes, part I: model and method.* 1964.

4 G. de Leve. *Generalized Markovian decision processes, part II: probabilistic background.* 1964.

5 G. de Leve, H.C. Tijms, P.J. Weeda. *Generalized Markovian decision processes, applications.* 1970.

6 M.A. Maurice. *Compact ordered spaces.* 1964.

7 W.R. van Zwet. *Convex transformations of random variables.* 1964.

8 J.A. Zonneveld. *Automatic numerical integration.* 1964.

9 P.C. Baayen. *Universal morphisms.* 1964.

10 E.M. de Jager. *Applications of distributions in mathematical physics.* 1964.

11 A.B. Paalman-de Miranda. *Topological semigroups.* 1964.

12 J.A.Th.M. van Berckel, H. Brandt Corstius, R.J. Mokken, A. van Wijngaarden. *Formal properties of newspaper Dutch.* 1965.

13 H.A. Lauwerier. *Asymptotic expansions.* 1966, out of print; replaced by MCT 54.

14 H.A. Lauwerier. *Calculus of variations in mathematical physics.* 1966.

15 R. Doornbos. *Slippage tests.* 1966.

16 J.W. de Bakker. *Formal definition of programming languages with an application to the definition of ALGOL 60.* 1967.

17 R.P. van de Riet. *Formula manipulation in ALGOL 60, part 1.* 1968.

18 R.P. van de Riet. *Formula manipulation in ALGOL 60, part 2.* 1968.

19 J. van der Slot. *Some properties related to compactness.* 1968.

20 P.J. van der Houwen. *Finite difference methods for solving partial differential equations.* 1968.

21 E. Wattel. *The compactness operator in set theory and topology.* 1968.

22 T.J. Dekker. *ALGOL 60 procedures in numerical algebra, part 1.* 1968.

23 T.J. Dekker, W. Hoffmann. *ALGOL 60 procedures in numerical algebra, part 2.* 1968.

24 J.W. de Bakker. *Recursive procedures.* 1971.

25 E.R. Paërl. *Representations of the Lorentz group and projective geometry.* 1969.

26 European Meeting 1968. *Selected statistical papers, part I.* 1968.

27 European Meeting 1968. *Selected statistical papers, part II.* 1968.

28 J. Oosterhoff. *Combination of one-sided statistical tests.* 1969.

29 J. Verhoeff. *Error detecting decimal codes.* 1969.

30 H. Brandt Corstius. *Exercises in computational linguistics.* 1970.

31 W. Molenaar. *Approximations to the Poisson, binomial and hypergeometric distribution functions.* 1970.

32 L. de Haan. *On regular variation and its application to the weak convergence of sample extremes.* 1970.

33 F.W. Steutel. *Preservation of infinite divisibility under mixing and related topics.* 1970.

34 I. Juhász, A. Verbeek, N.S. Kroonenberg. *Cardinal functions in topology.* 1971.

35 M.H. van Emden. *An analysis of complexity.* 1971.

36 J. Grasman. *On the birth of boundary layers.* 1971.

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.

38 W.A. Verloren van Themaat. *Automatic analysis of Dutch compound words.* 1972.

39 H. Bavinck. *Jacobi series and approximation.* 1972.

40 H.C. Tijms. *Analysis of (s,S) inventory models.* 1972.

41 A. Verbeek. *Superextensions of topological spaces.* 1972.

42 W. Vervaat. *Success epochs in Bernoulli trials (with applications in number theory).* 1972.

43 F.H. Ruymgaart. *Asymptotic theory of rank tests for independence.* 1973.

44 H. Bart. *Meromorphic operator valued functions.* 1973.

45 A.A. Balkema. *Monotone transformations and limit laws.* 1973.

46 R.P. van de Riet. *ABC ALGOL, a portable language for formula manipulation systems, part 1: the language.* 1973.

47 R.P. van de Riet. *ABC ALGOL, a portable language for formula manipulation systems, part 2: the compiler.* 1973.

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.

49 H. Kok. *Connected orderable spaces.* 1974.

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.

51 A. Hordijk. *Dynamic programming and Markov potential theory.* 1974.

52 P.C. Baayen (ed.). *Topological structures.* 1974.

53 M.J. Faber. *Metrizability in generalized ordered spaces.* 1974.

54 H.A. Lauwerier. *Asymptotic analysis, part 1.* 1974.

55 M. Hall, Jr., J.H. van Lint (eds.). *Combinatorics, part 1: theory of designs, finite geometry and coding theory.* 1974.

56 M. Hall, Jr., J.H. van Lint (eds.). *Combinatorics, part 2: graph theory, foundations, partitions and combinatorial geometry.* 1974.

57 M. Hall, Jr., J.H. van Lint (eds.). *Combinatorics, part 3: combinatorial group theory.* 1974.

58 W. Albers. *Asymptotic expansions and the deficiency concept in statistics.* 1975.

59 J.L. Mijnheer. *Sample path properties of stable processes.* 1975.

60 F. Göbel. *Queueing models involving buffers.* 1975.

63 J.W. de Bakker (ed.). *Foundations of computer science.* 1975.

64 W.J. de Schipper. *Symmetric closed categories.* 1975.

65 J. de Vries. *Topological transformation groups, 1: a categorical approach.* 1975.

66 H.G.J. Pijls. *Logically convex algebras in spectral theory and eigenfunction expansions.* 1976.

68 P.P.N. de Groen. *Singularly perturbed differential operators of second order.* 1976.

69 J.K. Lenstra. *Sequencing by enumerative methods.* 1977.

70 W.P. de Roever, Jr. *Recursive program schemes: semantics and proof theory.* 1976.

71 J.A.E.E. van Nunen. *Contracting Markov decision processes.* 1976.

72 J.K.M. Jansen. *Simple periodic and non-periodic Lamé functions and their applications in the theory of conical waveguides.* 1977.

73 D.M.R. Leivant. *Absoluteness of intuitionistic logic.* 1979.

74 H.J.J. te Riele. *A theoretical and computational study of generalized aliquot sequences.* 1976.

75 A.E. Brouwer. *Treelike spaces and related connected topological spaces.* 1977.

76 M. Rem. *Associons and the closure statement.* 1976.

77 W.C.M. Kallenberg. *Asymptotic optimality of likelihood ratio tests in exponential families.* 1978.

78 E. de Jonge, A.C.M. van Rooij. *Introduction to Riesz spaces.* 1977.

79 M.C.A. van Zuijlen. *Emperical distributions and rank statistics.* 1977.

80 P.W. Hemker. *A numerical study of stiff two-point boundary problems.* 1977.

81 K.R. Apt, J.W. de Bakker (eds.). *Foundations of computer science II, part 1.* 1976.

82 K.R. Apt, J.W. de Bakker (eds.). *Foundations of computer science II, part 2.* 1976.

83 L.S. van Benthem Jutting. *Checking Landau's "Grundlagen" in the AUTOMATH system.* 1979.

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.

85 J. van Mill. *Supercompactness and Wallman spaces.* 1977.

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.

88 A. Schrijver. *Matroids and linking systems.* 1977.

89 J.W. de Roever. *Complex Fourier transformation and analytic functionals with unbounded carriers.* 1978.

90 L.P.J. Groenewegen. *Characterization of optimal strategies in dynamic games.* 1981.

91 J.M. Geysel. *Transcendence in fields of positive characteristic.* 1979.

92 P.J. Weeda. *Finite generalized Markov programming.* 1979.

93 H.C. Tijms, J. Wessels (eds.). *Markov decision theory.* 1977.

94 A. Bijlsma. *Simultaneous approximations in transcendental number theory.* 1978.

95 K.M. van Hee. *Bayesian control of Markov chains.* 1978.

96 P.M.B. Vitányi. *Lindenmayer systems: structure, languages, and growth functions.* 1980.

97 A. Federgruen. *Markovian control problems; functional equations and algorithms.* 1984.

98 R. Geel. *Singular perturbations of hyperbolic type.* 1978.

99 J.K. Lenstra, A.H.G. Rinnooy Kan, P. van Emde Boas (eds.). *Interfaces between computer science and operations research.* 1978.

100 P.C. Baayen, D. van Dulst, J. Oosterhoff (eds.). *Proceedings bicentennial congress of the Wiskundig Genootschap, part 1.* 1979.

101 P.C. Baayen, D. van Dulst, J. Oosterhoff (eds.). *Proceedings bicentennial congress of the Wiskundig Genootschap, part 2.* 1979.

102 D. van Dulst. *Reflexive and superreflexive Banach spaces.* 1978.

103 K. van Harn. *Classifying infinitely divisible distributions by functional equations.* 1978.

104 J.M. van Wouwe. *Go-spaces and generalizations of metrizability.* 1979.

105 R. Helmers. *Edgeworth expansions for linear combinations of order statistics.* 1982.

106 A. Schrijver (ed.). *Packing and covering in combinatorics.* 1979.

107 C. den Heijer. *The numerical solution of nonlinear operator equations by imbedding methods.* 1979.

108 J.W. de Bakker, J. van Leeuwen (eds.). *Foundations of computer science III, part 1.* 1979.

109 J.W. de Bakker, J. van Leeuwen (eds.). *Foundations of computer science III, part 2.* 1979.

110 J.C. van Vliet. *ALGOL 68 transput, part I: historical review and discussion of the implementation model.* 1979.

111 J.C. van Vliet. *ALGOL 68 transput, part II: an implementation model.* 1979.

112 H.C.P. Berbee. *Random walks with stationary increments and renewal theory.* 1979.

113 T.A.B. Snijders. *Asymptotic optimality theory for testing problems with restricted alternatives.* 1979.

114 A.J.E.M. Janssen. *Application of the Wigner distribution to harmonic analysis of generalized stochastic processes.* 1979.

115 P.C. Baayen, J. van Mill (eds.). *Topological structures II, part 1.* 1979.

116 P.C. Baayen, J. van Mill (eds.). *Topological structures II, part 2.* 1979.

117 P.J.M. Kallenberg. *Branching processes with continuous state space.* 1979.

118 P. Groeneboom. *Large deviations and asymptotic efficiencies.* 1980.

119 F.J. Peters. *Sparse matrices and substructures, with a novel implementation of finite element algorithms.* 1980.

120 W.P.M. de Ruyter. *On the asymptotic analysis of large-scale ocean circulation.* 1980.

121 W.H. Haemers. *Eigenvalue techniques in design and graph theory.* 1980.

122 J.C.P. Bus. *Numerical solution of systems of nonlinear equations.* 1980.

123 I. Yuhász. *Cardinal functions in topology - ten years later.* 1980.

124 R.D. Gill. *Censoring and stochastic integrals.* 1980.

125 R. Eising. *2-D systems, an algebraic approach.* 1980.

126 G. van der Hoek. *Reduction methods in nonlinear programming.* 1980.

127 J.W. Klop. *Combinatory reduction systems.* 1980.

128 A.J.J. Talman. *Variable dimension fixed point algorithms and triangulations.* 1980.

129 G. van der Laan. *Simplicial fixed point algorithms.* 1980.

130 P.J.W. ten Hagen, T. Hagen, P. Klint, H. Noot, H.J. Sint, A.H. Veen. *ILP: intermediate language for pictures.* 1980.

131 R.J.R. Back. *Correctness preserving program refinements: proof theory and applications.* 1980.

132 H.M. Mulder. *The interval function of a graph.* 1980.

133 C.A.J. Klaassen. *Statistical performance of location estimators.* 1981.

134 J.C. van Vliet, H. Wupper (eds.). *Proceedings international conference on ALGOL 68.* 1981.

135 J.A.G. Groenendijk, T.M.V. Janssen, M.J.B. Stokhof (eds.). *Formal methods in the study of language, part I.* 1981.

136 J.A.G. Groenendijk, T.M.V. Janssen, M.J.B. Stokhof (eds.). *Formal methods in the study of language, part II.* 1981.

137 J. Telgen. *Redundancy and linear programs.* 1981.

138 H.A. Lauwerier. *Mathematical models of epidemics.* 1981.

139 J. van der Wal. *Stochastic dynamic programming, successive approximations and nearly optimal strategies for Markov decision processes and Markov games.* 1981.

140 J.H. van Geldrop. *A mathematical theory of pure exchange economies without the no-critical-point hypothesis.* 1981.

141 G.E. Welters. *Abel-Jacobi isogenies for certain types of Fano threefolds.* 1981.

142 H.R. Bennett, D.J. Lutzer (eds.). *Topology and order structures, part 1.* 1981.

143 J.M. Schumacher. *Dynamic feedback in finite- and infinite-dimensional linear systems.* 1981.

144 P. Eijgenraam. *The solution of initial value problems using interval arithmetic; formulation and analysis of an algorithm.* 1981.

145 A.J. Brentjes. *Multi-dimensional continued fraction algorithms.* 1981.

146 C.V.M. van der Mee. *Semigroup and factorization methods in transport theory.* 1981.

147 H.H. Tigelaar. *Identification and informative sample size.* 1982.

148 L.C.M. Kallenberg. *Linear programming and finite Markovian control problems.* 1983.

149 C.B. Huijsmans, M.A. Kaashoek, W.A.J. Luxemburg, W.K. Vietsch (eds.). *From A to Z, proceedings of a symposium in honour of A.C. Zaanen.* 1982.

150 M. Veldhorst. *An analysis of sparse matrix storage schemes.* 1982.

151 R.J.M.M. Does. *Higher order asymptotics for simple linear rank statistics.* 1982.

152 G.F. van der Hoeven. *Projections of lawless sequences.* 1982.

153 J.P.C. Blanc. *Application of the theory of boundary value problems in the analysis of a queueing model with paired services.* 1982.

154 H.W. Lenstra, Jr., R. Tijdeman (eds.). *Computational methods in number theory, part I.* 1982.

155 H.W. Lenstra, Jr., R. Tijdeman (eds.). *Computational methods in number theory, part II.* 1982.

156 P.M.G. Apers. *Query processing and data allocation in distributed database systems.* 1983.

157 H.A.W.M. Kneppers. *The covariant classification of two-dimensional smooth commutative formal groups over an algebraically closed field of positive characteristic.* 1983.

158 J.W. de Bakker, J. van Leeuwen (eds.). *Foundations of computer science IV, distributed systems, part 1.* 1983.

159 J.W. de Bakker, J. van Leeuwen (eds.). *Foundations of computer science IV, distributed systems, part 2.* 1983.

160 A. Rezus. *Abstract AUTOMATH.* 1983.

161 G.F. Helminck. *Eisenstein series on the metaplectic group, an algebraic approach.* 1983.

162 J.J. Dik. *Tests for preference.* 1983.

163 H. Schippers. *Multiple grid methods for equations of the second kind with applications in fluid mechanics.* 1983.

164 F.A. van der Duyn Schouten. *Markov decision processes with continuous time parameter.* 1983.

165 P.C.T. van der Hoeven. *On point processes.* 1983.

166 H.B.M. Jonkers. *Abstraction, specification and implementation techniques, with an application to garbage collection.* 1983.

167 W.H.M. Zijm. *Nonnegative matrices in dynamic programming.* 1983.

168 J.H. Evertse. *Upper bounds for the numbers of solutions of diophantine equations.* 1983.

169 H.R. Bennett, D.J. Lutzer (eds.). *Topology and order structures, part 2.* 1983.

CWI TRACTS

1 D.H.J. Epema. *Surfaces with canonical hyperplane sections.* 1984.

2 J.J. Dijkstra. *Fake topological Hilbert spaces and characterizations of dimension in terms of negligibility.* 1984.

3 A.J. van der Schaft. *System theoretic descriptions of physical systems.* 1984.

4 J. Koene. *Minimal cost flow in processing networks, a primal approach.* 1984.

5 B. Hoogenboom. *Intertwining functions on compact Lie groups.* 1984.

6 A.P.W. Böhm. *Dataflow computation.* 1984.

7 A. Blokhuis. *Few-distance sets.* 1984.

8 M.H. van Hoorn. *Algorithms and approximations for queueing systems.* 1984.

9 C.P.J. Koymans. *Models of the lambda calculus.* 1984.

10 C.G. van der Laan, N.M. Temme. *Calculation of special functions: the gamma function, the exponential integrals and error-like functions.* 1984.

11 N.M. van Dijk. *Controlled Markov processes; time-discretization.* 1984.

13 D. Grune. *On the design of ALEPH.* 1985.

14 J.G.F. Thiemann. *Analytic spaces and dynamic programming: a measure theoretic approach.* 1985.