PACIFIC JOURNAL OF MATHEMATICS
Vol. 91, No. 1, 1980

DIRECT FACTORIZATIONS OF MEASURES

THEODOR EISELE

In this paper we want to investigate the question, to what extent can the disintegration of some measure on an arbitrary Suslin space with respect to some measurable function f be replaced by the image measure under some function g inverting f, such that the "outcome" of the situation under a function h is not changed. Such a direct factorization, as we call it, is modulo some conditions about atoms of the measures in general only possible, if the range of h is countable. But there are always solutions to the problem in a weak sense. The results have applications in game theory to the problem of "elimination of randomization".

Our staring point are some results about the compactness and convexity of the range of some measure operations. They are closely related to Lyapunov's theorem [10].

In § 2 we recall some known results about the disintegration of measures on Suslin spaces.

The problem of direct factorizations of measures is made precise in § 3 and solved there for the case where the "outcome"-set C is countable. Of course, some restrictions concerning the atoms of the measures are necessary. A counterexample shows that this result cannot be generalized to compact metrizable C. Thus we introduce in § 4 the notion of a weak direct factorization and show that such a weak direct factorization exists even if C is an arbitrary Suslin space.

It is quite obvious that the solutions to the direct factorization problem are extreme points in a certain convex space of measures on a Suslin set. In fact, we show in § 5 that if C is countable or if we regard the weak problem, the extreme points of this set are exactly the solutions to the corresponding factorization problem. Under somewhat different situations such characterizations have been found in [5].

As mentioned at the beginning, we shall apply the results to a question in game theory in § 6. The application shows, when random strategies (=behavior strategies) can be equivalently replaced by nonrandom ones. Such questions of "elimination of randomization" have been treated in [3] and [4] in the finite case and are here generalized to arbitrary Suslin spaces.

1. Convex ranges by nonatomic measures. Since in later sections we are interested when some integral operators have compact convex range, we shall give here some descriptions of such cases. Of course, our propositions lie in the vicinity of Lyapunov's theorem.

Let $(A, \mathfrak{A})$ be a fixed measurable space and $m, m_1, m_2, \cdots$ σ-additive, $\boldsymbol{R}^d$-valued measures on $(A, \mathfrak{A})$ with finite variation $|m|, |m_1|, |m_2|, \cdots$.

DEFINITION. $B \in \mathfrak{A}$ is called an m-atom if it is an atom with respect to $|m|$, i.e., $|m|(B') \in \{0, |m|(B) \neq 0\}$ for all $B' \subseteqq B$. m is called nonatomic if it has no atoms.

The following result is known as Lyapunov's theorem [10]:

THEOREM 1. *The range of a nonatomic, $\boldsymbol{R}^d$-valued measure of finite variation is compact and convex.*

Changing slightly the excellent proof of this theorem, due to Lindenstrauss [8], we get what we shall need in the sequel:

THEOREM 2. *If (m_i) is a sequence of nonatomic measures with $|m_i| \leqq \mu$ for some finite measure μ, then the set of all points in $(\boldsymbol{R}^d)^N$ of the form*

$$(m_1(B_1), m_2(B_2), \cdots)$$

with $(B_i)_{i \in N}$ ranging over the $\mathfrak{A}$-partitions[1] *of A, is convex and weakly compact.*

Proof. Set $\mathscr{L} = \{(g_i) \in (L^\infty(\mu))^N, 0 \leqq g_i \text{ and } \sum_{i \in N} g_i \leqq 1\}$. $\mathscr{L}$ is convex and compact in the product topology of the weak-$*$-topology. The linear mapping $M: \mathscr{L} \to (\boldsymbol{R}^d)^N$, $M(g_i) = \left(\int g_i dm_i\right)_{i \in N}$ is continuous in this topology, since $|m_i| \ll \mu$. Thus the proof is finished if

$$(1) \qquad M(\mathscr{L}) = M(\{(1_{B_i}) \in \mathscr{L}, (B_i) \text{ a partition of } A\}) ,$$

since the first set is weakly compact and convex. But "$\supseteqq$" is obvious; so let $(r_i) \in M(\mathscr{L})$. $M^{-1}\{(r_i)\}$ is a weakly compact convex subset of $\mathscr{L}$ and contains an extremal point (g_i) by the Krein-Milman theorem. If (g_i) were not in the right set of (1), then we could assume, without loss of generality, that there exist $\varepsilon > 0$ and

$$B \subseteqq \{\varepsilon \leqq g_1, g_2 \leqq 1 - \varepsilon\} \text{ with } |m_1(B)| > 0 .$$

Since m_1 and m_2 are nonatomic, there are disjoint $B_1, B_2 \subseteqq B$ with

[1] i.e. sequences of disjoint subsets from $\mathfrak{A}$.

$m_1(B_i) \neq 0$ $(i = 1, 2)$ and applying Theorem 1 to m_2/B_i, we find disjoint $B_3, B_4 \subseteqq B_1$ as well as disjoint $B_5, B_6 \subseteqq B_2$ with

$$m_2(B_{2i+j}) = \frac{1}{2} m_2(B_i) \qquad (i, j = 1, 2) .$$

Let α, β with $\varepsilon \geqq |\alpha| + |\beta| > 0$ satisfy

$$\alpha(m_1(B_3) - m_1(B_4)) = \beta(m_1(B_5) - m_1(B_6))$$

and

$$h = \alpha(1_{B_3} - 1_{B_4}) - \beta(1_{B_5} - 1_{B_6}) \neq 0 .$$

Then $m_1(h) = m_2(h) = 0$ and thus $(g_1 \pm h, g_2 \mp h, g_3, \cdots) \in M^{-1}\{(r_i)\}$ contradicting the extremality of (g_i).

Thus "$\subseteqq$" holds in (1) and the proof is complete.

COROLLARY. *Let μ be a nonatomic finite positive measure on $(A, \mathfrak{A})$ and $f_i: A \to \boldsymbol{R}^d$ uniformly bounded measurable functions. Then the set of all points in $\boldsymbol{R}^d$ of the form $\sum_{i \in N} \int f_i 1_{D_i} d\mu$ with $(D_i)_{i \in N}$ ranging over all $\mathfrak{A}$-partitions of A, is compact and convex.*

2. Some reminiscences. In this section let the sets A, B, C be Suslin spaces, i.e., continuous images of polish spaces, and $\mathfrak{B}(A)$, $\mathfrak{B}(B)$, $\mathfrak{B}(C)$ their Borel algebras. We remind the following, well known results about the factorization of measurable functions (cf. [7]).

(a) If $f: A \to B$ is a surjective Borel-measurable function, then there exists a universally measurable function $g: B \to A$ with $f \circ g = id_B$.

(b) If $f: A \to B$ is Borel-measurable and $h: A \to C$ is $f^{-1}(\mathfrak{B}(B)) - \mathfrak{B}(C)$-measurable, then there exists a universally measurable function $g: B \to A$ with $h \circ g \circ f = h$ and $f \circ g = id_B$.

The situation becomes more involved, if we regard measure spaces.

Let $R(A)$ denote the set of all positive finite Radon measures on A. If $\mu \in R(A)$ and $\nu \in R(B)$ we write

$$f: (A, \mu) \longrightarrow (B, \nu)$$

if f is a μ-measurable function and $\nu = Rf(\mu)$ the image of μ under f. μ is called a preimage measure of ν under f.

If $f: (A, \mu) \to (B, \nu)$ with f a Borel-measurable surjection and $g: B \to A$ as in (a), then in general we cannot expect $Rg(\nu) = \mu$, though both measures are preimages of ν.

(c) The preimage measure of ν is unique if and only if the universally measurable set $\{b \in B;\ \sharp f^{-1}(b) \geqq 2\}$ has ν-measure zero (see [6] and [8]). There is however a nice and rather deep representation theorem for all preimage measures of ν, known as disintegration of measures (cf. [11] and [12] 2, 21).

(d) Let $f\colon A \to B$ be a Borel-measurable mapping between the two Suslin spaces A, B and let $\mu \in R(A)$, $\nu \in R(B)$. Then are equivalent

(i) $Rf(\mu) = \nu$

(ii) there exists a family $(\mu_b)_{b \in B} \subseteqq W(A) \subseteqq R(A)$[2] with

(α) $b \mapsto \mu_b(A')$ is Borel-measurable for all $A' \in \mathfrak{B}(A)$

(β) $\mu = \int_B \mu_b \nu(db)$

(γ) for ν-almost all b, $|\mu_b|(A \backslash f^{-1}\{b\}) = 0$.

Moreover, if (μ_b) and (μ_b') are two families with $(\alpha) - (\gamma)$, then $\mu_b = \mu_b'$ for ν-almost all b. The uniqueness assertion in (d) shows especially, that

(e) if $\mu = Rg(\nu)$ for a universally measurable function $g\colon B \to A$ with $f \circ g = id_B$ (see (a)), then $\mu_b = \delta_{g(b)}$ is the ν-almost unique disintegration of $Rg(\nu)$. Conversely, if the disintegration of $\mu = \int \mu_b \nu(db)$ satisfying $(\alpha) - (\gamma)$ is of the form $\mu_b = \delta_{g(b)}$ for some $g\colon B \to A$, then g is Borel-measurable by (α) and by (γ)

$$f \circ g = id_B \ \nu\text{-almost surely .}$$

But in this paper we are more interested in situations as in (b) in the presence of measures.

We regard the situation

$$(*) \qquad \begin{array}{ccc} (A, \mu) & \xrightarrow{\ f\ } & (B, \nu) \\ \big\downarrow{\scriptstyle h} & & \\ (C, \pi) & & \end{array}$$

where f and h are Borel-measurable functions; μ, ν and π are positive finite Radon measures on A, B, C respectively and $\nu = Rf(\mu)$, $\pi = Rh(\mu)$.

If we were only interested in ν-measurable functions $g\colon B \to C$ with $\pi = Rg(\nu)$, then the following proposition would give a complete answer. First we need the following

DEFINITION. Let $T(\nu)$ denote the countable set of all atoms of ν. We say that ν is atomicly adapted to π if there exists a com-

[2] $W(A)$ denotes the space of all probability measures on A.

plete decomposition $(T_c(\nu))_{c \in T(\pi)}$ of $T(\nu)$ into disjoint subsets:

$$T(\nu) = \cup\{T_c(\nu),\ c \in T(\pi)\}$$

such that $\nu(T_c(\nu)) \leqq \pi\{c\}$ for all $c \in T(\pi)$.

Trivially, a nonatomic measure is atomicly adapted to all other measures.

PROPOSITION. *There exists a ν-measurable function $g: B \to C$ with $Rg(\nu) = \pi$ if and only if ν is atomicly adapted to π.*

Proof. If $Rg(\nu) = \pi$, then $T_c(\nu) = T(\nu) \cap g^{-1}\{c\}$ $(c \in T(\pi))$ shows, that ν is atomicly adapted to π.

Conversely, let $(T_c(\nu))_{c \in T(\pi)}$ be a decomposition of $T(\nu)$ as in the definition. Since $\nu(B) = \pi(C)$ and $\nu/B\backslash T(\nu)$ is nonatomic, we find a Borel set $B_0 \subseteqq B\backslash T(\nu)$ with $\nu(B_0) = \pi(T(\pi)) - \nu(T(\nu))$ and a complete disjoint decomposition $(B_c)_{c \in T(\pi)}$ of B_0 into Borel sets with

$$\nu(B_c) = \pi\{c\} - \nu(T_c(\nu)) \geqq 0 \qquad (c \in T(\pi))\ .$$

On $B_0 \cup T(\nu)$ define g by $g(b) = c$ if $b \in B_c \cup T_c(\nu)$. Now $B_1 = B\backslash(B_0 \cup T(\nu))$ and $C_1 = C\backslash T(\pi)$ are Suslin spaces with nonatomic measures $\nu \restriction B_1$ and $\pi \restriction C_1$ and $\nu(B_1) = \pi(C_1) = r$. But such spaces are Borel-isomorphic to a Borel subset of $[0, r]$ with Lebesgue measure.

This shows, that we can extend g to a Borel-measurable function from B_1 to C_1 with $Rg(\nu \restriction B_1) = \pi \restriction C_1$ and the proof is complete.

The last result is completely independent of A. But for later applications we are interested in factorization results which regard A.

3. **Direct factorizations.** Let again A, B, C be Suslin spaces with their Borel algebras $\mathfrak{B}(A)$, $\mathfrak{B}(B)$ and $\mathfrak{B}(C)$. In the situation

$$(**) \qquad \begin{array}{ccc} (A, \mu) & \xrightarrow{\ f\ } & (B, \nu) \\ \big\downarrow h & \nwarrow_{g(?)} & \\ (C, \pi) & & \end{array}$$

with the usual notation, we are now looking for ν-measurable functions $g: B \to A$ with $f \circ g = id_B$ ν-almost everywhere and $R(h \circ g)(\nu) = \pi$.

Let us call such a function g a direct factorization of $(**)$.

THEOREM 3. *Let (μ_b) be the unique disintegration of μ under f. If for ν-almost all b the function h is constant μ_b-almost every-*

where, then a direct factorization of $(**)$ *exists.*

Proof. We may assume f to be surjective.

We know that there exists a σ-compact subset D' of A with $\pi(C - D') = 0$. But D' is metrizable (cf. [11]). So we find a countable basis $(Q'_m)_m$ of relatively open subsets of D'. Set $D = h^{-1}(D')$ and $Q_m = h^{-1}(Q'_m)$. Furthermore we may assume $\mu_b(D) = 1$ for all b, neglecting a ν-nullset. Put

$$G = \{(a, f(a)); a \in D\} \cap \bigcap_m [D \times \{b, \mu_b(Q_m) > 0\} \cup (D \backslash Q_m) \times B] .$$

G is a Suslin set with $G|_b = \{a, (a, b) \in G\} \neq \phi$ and h is constant on $G|_b$ for all b. By a theorem of Blackwell ([2], III, 26) we see that[3] $h \circ \mathrm{pro}_A\colon G \to C$ is $(\mathrm{pro}_B)^{-1}(\mathfrak{B}(B)) - \mathfrak{B}(C)$-measurable. By (2, b) we get a ν-measurable function $\tilde{g}\colon B \to G$ with $h \circ \mathrm{pro}_A \circ \tilde{g} \circ \mathrm{pro}_B = h \circ \mathrm{pro}_A$ and $\mathrm{pro}_B \circ \tilde{g} = id_B$ ν-almost everywhere. Now $\mathrm{pro}_A \circ \tilde{g}$ is a direct factorization of $(**)$.

We remind that $T(\nu)$ denotes the at most countable set of all atoms of ν. The following theorem shows that we can restrict the hypothesis of Theorem 3 to the set $T(\nu)$, if C is countable.

THEOREM 4. *Let C be countable and suppose that h is constant μ_b-almost everywhere for all $b \in T(\nu)$, where (μ_b) is the disintegration of μ under f.*

Then $(**)$ *has a direct factorization.*

Proof. For $c \in C$ define the Suslin sets $A_c = h^{-1}\{c\} \backslash f^{-1}(T(\nu))$ and $B_c = f(A_c)$.

$\nu \upharpoonright B \backslash T(\nu)$ being nonatomic, we see by the proof of Theorem 2 of § 1, that the set

$$\Pi = \left\{\left(\int 1_{B_c \cap D_c} d\nu\right)_c ;\ (D_c) \text{ disjoint Borel subsets of } B\right\}$$

contains $\left(\int 1_{B_c} \mu_b(A_c) \nu(db)\right)_c$.

Hence $\pi = Rh(\mu \upharpoonright f^{-1}(T(\nu))) + \sum_{c \in C} \nu(D_c)\delta_c$ with $D_c \subseteqq B_c$ and (D_c) a Borel-measurable partition of $f(A) \backslash T(\nu)$.

By (2, a) we have universally measurable functions $g_c\colon D_c \to A_c$ with $f \circ g_c = id_{D_c} (c \in C)$. Further there exist measurable functions $g'\colon T(\nu) \to A$ with $g'(b) \in f^{-1}(b) \cap h^{-1}\{c_b\} \neq \phi$, where $\mu_b(h^{-1}\{c_b\}) = 1$ and $g''\colon B \backslash f(A) \to A$, $b \mapsto g''(b) \equiv a \in A$. Then $g = \bigcup_c g_c \cup g' \cup g''$ is a ν-measurable function: $B \to A$ with $f \circ g = 1_B$ ν-almost everywhere and

[3] pro_A denotes the projection to the space A.

$$R(h \circ g)(\nu) = Rh(\mu \upharpoonright f^{-1}(T(\nu))) + \sum_c \nu(D_c) \cdot \delta_c = \pi \ .$$

COROLLARY. *If ν is nonatomic and C countable, then there exists a direct factorization of* $(**)$.

The following example shows however, that the theorem cannot be generalized to arbitrary compact C.

Counterexample. Let $B = C = [0, 1]$, $A = \{(b, c) \in [0, 1]^2,\ c = (1/2)b$ or $c = 1/2(b + 1)\}$ and f, h the projections to the first, resp. second coordinate. Let μ be given by

$$\mu\left\{\left(b, \frac{1}{2} b\right);\ \alpha < b \leqq \beta\right\} = \mu\left\{\left(b, \frac{1}{2}(b + 1)\right);\ \alpha < b \leqq \beta\right\}$$
$$= \frac{1}{2}(\beta - \alpha) \text{ for } 0 \leqq \alpha < \beta \leqq 1 \ .$$

Then $\pi = \nu$ is the Lebesgue measure λ on $[0, 1]$. Assume, $g: B \to A$ would be a direct factorization and let $B_0 = g^{-1}(\{(b, (1/2)b);\ b \in [0, 1]\})$, a Lebesgue-measurable set. Then for all $0 \leqq \alpha < \beta \leqq 1$ we have $\lambda(B_0 \cap [\alpha, \beta]) = (\beta - \alpha)/2 = (1/2)\lambda([\alpha, \beta])$. Since $D \mapsto \lambda(B_0 \cap D)$ is a σ-additive measure on $\mathfrak{B}([0, 1])$ which equals $(1/2)\lambda$ on the intervals, we have $\lambda(B_0 \cap D) = (1/2)\lambda(D)$ for all Lebesgue-measurable D. Especially, $1/2 = \lambda(B_0) = \lambda(B_0 \cap B_0) = (1/2)\lambda(B_0) = 1/4$, contradiction!

Thus there does not exist a direct factorization, though ν is nonatomic.

4. Weak direct factorizations. Since our research for direct factorizations of measures has been knocked down by the above counterexample, we want to weaken the notion of a direct factorization to treat also noncountable C.

DEFINITION. Recall the situation $(**)$, where A, B, C are Suslin spaces, $f: A \to B$ and $h: A \to C$ Borel measurable functions and μ, ν, π positive finite Radon measures with $Rf(\mu) = \nu$ and $Rh(\mu) = \pi$. We say, $(**)$ is weakly directly factorizable, if for each bounded Borel measurable function $r: C \to \boldsymbol{R}^d$, there exists a ν-measurable function $g: B \to A$ with

$$f \circ g = id_B \ \nu\text{-almost everywhere and}$$
$$\int_C r d\pi = \int_B r \circ h \circ g d\nu \ .$$

THEOREM 5. *The situation* (**), *as described above, is weakly directly factorizable, if h is constant μ_b-almost everywhere for all $b \in T(\nu)$, (μ_b) being the disintegration of μ under f.*

Proof. We first define the Suslin space G' by $G' = (f, r \circ h)(A)$ with the measure $\rho = R(f, r \circ h)(\mu)$ which admits a unique disintegration $(\rho_b) \in R(\boldsymbol{R}^d)$ with $R\,\mathrm{pro}_{R^d}(\rho) = \int_B \rho_b \nu(db)$ and $G'|_b$ contains the support of ρ_b. Let (Q_m) be a countable open basis of $\boldsymbol{R}^d$ and set

$$G = G' \cap \bigcap_m [\{b, \rho_b(Q_m) > 0\} \times \boldsymbol{R}^d \cup B \times (\boldsymbol{R}^d \backslash Q_m)] .$$

$G|_b$ is contained in the support of ρ_b for ν-almost all b and $\sharp G|_b = 1$ if $b \in T(\nu)$. We regard the set $\mathfrak{G}$ of all ν-measurable maps $g\colon B \to A$ with $(f \circ g(b),\ r \circ h \circ g(b)) \in \{b\} \times G_{|b}$ for ν-almost all b. If $g_1, g_2, \cdots$ are from $\mathfrak{G}$ and (B_i) is a Borel partition of B, then $\bigcup_i (g_i \upharpoonright B_i) \in \mathfrak{G}$.

Since for all $g \in \mathfrak{G}$ $h \circ g$ are identical on $T(\nu)$ and $\nu \upharpoonright B \backslash T(\nu)$ is nonatomic, we see by the corollary of § 1 that

$$\Pi = \left\{\int_B r \circ h \circ g d\nu,\ g \in \mathfrak{G}\right\}$$

is convex. Put

$$\tilde{p} = \int_C r d\pi = \int_G \mathrm{pro}_{R^d} d\rho = \int_B r(b)\nu(db)$$

with

$$r(b) = \int_{R^d} x \rho_b(dx) ,$$

a Borel measurable function. We have to show $\tilde{p} \in \Pi$.

Otherwise, there would be a vector $v \in \boldsymbol{R}^d$ and a constant γ with

$$\langle v, p\rangle \leqq \gamma \leqq \langle v, \tilde{p}\rangle$$

for all $p \in \Pi$ and at least for one $p_0 \in \Pi$ we have $\langle v, p_0\rangle < \gamma$.

The sets

$$\underset{(\gtrless)}{G\geqq} = \{(b, x) \in G,\ \langle v, x\rangle \underset{(\gtrless)}{\geqq} \langle v, r(b)\rangle\}$$

are Suslin spaces. But for ν-almost all b $G_{\geqq|b} \neq \phi$, since otherwise

$$\langle v, x\rangle < \langle v, r(b)\rangle \quad \text{for all } x \in G_{|b}$$

which would yield a contradiction by

$$\langle v, r(b)\rangle = \int_{G_{|b}} \langle v, x\rangle \rho_b(dx) < \langle v, r(b)\rangle ,$$

the equality holding ν-almost everywhere.

If $p_0 = \int_B r \circ h \circ g_0 d\nu \in \Pi$ with $g_0 \in \mathfrak{G}$ we have by definition of p

$$\nu\{b;\ \langle v, r \circ h \circ g_0(b)\rangle < \langle v, r(b)\rangle\} > 0 ,$$

hence $\rho(G_<) > 0$. [Recall that $G_{|b}$ is in the support of ρ_b ν-a.s.] But now also $\rho(G_>) > 0$ to guarantee

$$\int_G \langle v, x\rangle \rho(d(b, x)) = \int_B \langle v, r(b)\rangle \nu(db) .$$

This gives us the existence of a ν-measurable function $\bar{g} \in \mathfrak{G}$ with

$$r \circ h \circ \bar{g}(b) \in G_{\geq|b} \quad \nu\text{-a.s.}$$

and

$$\nu(\{b,\ r \circ h \circ \bar{g}(b) \in G_{>|b}\}) > 0 .$$

With $\bar{p} = \int_B r \circ h \circ \bar{g} d\nu$ we get the contradiction

$$\begin{aligned} \gamma \leqq \langle v, \tilde{p}\rangle &= \int_B \langle v, r(b)\rangle \nu(db) \\ &< \int_B \langle v, r \circ h \circ \bar{g}(b)\rangle \nu(db) = \langle v, \bar{p}\rangle \leqq \gamma . \end{aligned}$$

Hence $\tilde{p} \in \Pi$ and the proof is complete.

COROLLARY. *If ν is nonatomic, then* (**) *always admits weak direct factorizations.*

REMARK. It is not difficult to generalize the above theorems to arbitrary Blackwell spaces if we add the usual "consistency assumptions on the atoms" (cf. [2]).

5. Extreme point problems. In this section we want to reinterpret Theorems 3, 4 and 5 as theorems about the existence of extreme points of certain convex sets. In fact, let us regard the following subsets of $R(A)$:

$$P = \{\lambda \in R(A),\ Rf(\lambda) = \nu \text{ and } Rh(\lambda) = \pi\}$$

$$P(r) = \left\{\lambda \in R(A),\ Rf(\lambda) = \nu \text{ and } \int r \circ h d\lambda = \int r \circ h d\mu\right\}$$

where $r: C \to \boldsymbol{R}^d$ is a bounded Borel measurable function. Then P and $P(r)$ are nonempty convex sets.

COROLLARY. *If the condition of Theorems* 3 *or* 4 [*resp. of*

Theorem 5] *is satisfied, then* P [*resp.* $P(r)$] *has extreme points.*

Proof. Let g be the solution to the [weak] direct factorization problem found in Theorems 3 or 4 [resp. Theorem 5]. Then $\lambda = Rg(\nu)$ is in P [resp. in $P(r)$], and since the disintegration of λ w.r.t. f is $(\delta_{g(b)})$ by (2, e), λ is an extreme point of P [resp. $P(r)$].

REMARK. The counterexample of § 3 shows however, that in some situations there are extreme points of P, which all are not of the form $Rg(\nu)$. That this can not happen in the situation of Theorems 4 or 5, show the following results:

THEOREM 6. *Let us assume in* (**) *that* h *is constant on* $f^{-1}(b)$ *for all* $b \in T(\nu)$ *and that* C *is countable. Then* λ *is extremal in* P *if and only if* $\lambda = Rg(\nu) \in P$ *for some direct factorization* g *of* (**).

Proof. We have only to show the necessity.

So let λ be extremal in P with the disintegration (λ_b) w.r.t. f. By (2.e) it suffices to show that ν-almost all λ_b are 0-1-measures. But if $0 < \lambda_b(Q) < 1$ for some $b \in T_\nu$ and $Q \in \mathfrak{B}(A)$ then with

$$\lambda^1(\cdot) = \nu\{b\}\lambda_b(\cdot \cap Q)/\lambda_b(Q) + \lambda(\cdot \backslash f^{-1}(b))$$

and

$$\lambda^2(\cdot) = \nu\{b\}\lambda_b(\cdot \backslash Q)/\lambda_b(A \backslash Q) + \lambda(\cdot \backslash f^{-1}(b))$$

we have

$$0 \leqq \lambda^2(Q) < \lambda(Q) < \lambda^1(Q),\ \lambda^1 \text{ and } \lambda^2 \in P$$

and

$$\lambda = \lambda_b(Q)\lambda^1 + \lambda_b(A \backslash Q)\lambda^2 ,$$

which contradicts the extremality of λ. Hence, for $b \in T(\nu)$ λ_b is a 0-1-measure.

If λ_b for $b \notin T(\nu)$ are not ν-almost everywhere 0-1-measures, we find (not necessarily different) $c_1, c_2 \in C$, $Q \in \mathfrak{B}(A)$ and $\varepsilon > 0$ such that with $E_1 = h^{-1}(c_1) \cap Q$, $E_2 = h^{-1}(c_2) \backslash Q$ the Borel set $D = \{b \in B \backslash T(\nu);\ \varepsilon \leqq \lambda_b(E_1),\ \varepsilon \leqq \lambda_b(E_2)\}$ has a positive ν-measure. Set $0 < \gamma \leqq \varepsilon/(1-\varepsilon)$ and

$$d_b^{\pm} = 1 \pm \gamma\lambda_b(E_1)/\lambda_b(E_2) \geqq 0 \qquad \text{for } b \in D .$$

Since ν is nonatomic on D we find a Borel set $D_1 \subseteqq D$ such that

$$\int_{D_1} \lambda_b(E_1)\nu(db) = \int_{D/D_1} \lambda_b(E_1)\nu(db) .$$

Define

$$\begin{aligned}\lambda^{\pm}(\cdot) = &\int_{D_1} (1 \pm \gamma)\lambda_b(\cdot \cap E_1) + d_b^{\mp}\lambda_b(\cdot \cap E_2) + \lambda_b(\cdot\backslash(E_1 \cup E_2))\nu(db) \\ &+ \int_{D/D_1} (1 \mp \gamma)\lambda_b(\cdot \cap E_1) + d_b^{\pm}\lambda_b(\cdot \cap E_2) + \lambda_b(\cdot\backslash(E_1 \cup E_2))\nu(db) \\ &+ \lambda(\cdot\backslash f^{-1}(D)) .\end{aligned}$$

It is easy to check that $\lambda^+, \lambda^- \in P$ and $\lambda = (\lambda^+ + \lambda^-)/2$, in contradiction to the extremality of λ. The proof is complete.

Similarly in the weak situation:

THEOREM 7. *Assume again in* $(**)$ *that* h *is constant on* $f^{-1}(b)$ *for all* $b \in T(\nu)$, *and let* $r: C \to \boldsymbol{R}^d$ *be a bounded Borel measurable function. Then* λ *is extremal in* $P(r)$ *if and only if* $\lambda = Rg(\nu) \in P(r)$ *for some with respect to* r *weak direct factorization* g *of* $(**)$.

Proof. Since again only the necessity has to been shown, we start with an extremal λ in $P(r)$ and its disintegration (λ_b) with respect to f. That for $b \in T(\nu)$ λ_b must be 0-1-measures, is shown as in the preceding proof. Assuming that λ has not the required representation, i.e., that for $b \notin T(\nu)$ λ_b are not ν-almost everywhere 0-1-measures, we find $Q \in \mathfrak{B}(A)$ and $\varepsilon > 0$ such that

$$D = \{b \in B\backslash T(\nu),\ \varepsilon \leqq \lambda_b(Q),\ \varepsilon \leqq \lambda_b(A\backslash Q)\} \in \mathfrak{B}(B)$$

has a positive ν-measure. Define

$$+Q = f^{-1}(D) \cap Q,\quad -Q = f^{-1}(D)\backslash Q$$

and for $b \in D$

$$p^{\pm}(b) = \int_{\pm Q} r \circ h d\lambda_b / \lambda_b(\pm Q) .$$

The set

$$\Pi = \left\{\int (1_E p^+ + 1_{D\backslash E} p^-) d\nu,\ E \in \mathfrak{B}(B),\ E \subseteqq D\right\}$$

is compact and convex by the corollary of § 1, since ν is nonatomic on D. We claim, that Π contains

$$\tilde{p} = \int_{f^{-1}(D)} r \circ h d\lambda = \int_D \tilde{p}(b)\nu(db)$$

where $\tilde{p}(b) = \int r \circ h d\lambda_b$. Otherwise, we find $v \in \boldsymbol{R}^d$ and a constant γ such that for all $p \in \Pi$

$$\langle v, p\rangle \leqq \gamma < \langle v, \tilde{p}\rangle .$$

Let E_0 be the set of all $b \in D$ with $\langle v, \tilde{p}(b)\rangle \leqq \langle v, p^+(b)\rangle$. Since $\langle v, \tilde{p}(b)\rangle = \lambda_b(+Q)\langle v, p^+(b)\rangle + \lambda_b(-Q)\langle v, p^-(b)\rangle$ we have $\langle v, \tilde{p}(b)\rangle \leqq \langle v, p^-(b)\rangle$ for $b \in D\backslash E_0$. With

$$p_0 = \int (1_{E_0}p^+ + 1_{D\backslash E_0}p^-)d\nu \in \Pi$$

we get the contradiction

$$\begin{aligned}\langle v, p_0\rangle \leqq \gamma < \langle v, \tilde{p}\rangle &= \int \langle v, \tilde{p}(b)\rangle \nu(db)\\ &\leqq \int (1_{E_0}(b)\langle v, p^+(b)\rangle + 1_{D\backslash E_0}(b)\langle v, p^-(b)\rangle)\nu(db)\\ &= \langle v, p_0\rangle .\end{aligned}$$

This shows $\tilde{p} = \int (1_E p^+ + 1_{D\backslash E}p^-)d\nu = \int_D\left(\int_A r \circ h d\tilde{\lambda}_b\right)\nu(db)$ with $\tilde{\lambda}_b = 1_E(b)\lambda_b(\cdot \cap + Q)/\lambda_b(+Q) + 1_{D\backslash E}(b)\lambda_b(\cdot \cap -Q)/\lambda_b(-Q)$ for some $E \in \mathfrak{B}(B)$, $E \subseteqq D$.

Let $0 < \gamma \leqq \varepsilon/(1-\varepsilon)$ and

$$\lambda^{\pm}(\cdot) = \int_D (1 \pm \gamma)\lambda_b(\cdot) + (\mp\gamma)\tilde{\lambda}_b(\cdot)\nu(db) + \lambda(\cdot\backslash f^{-1}(D)) .$$

$\lambda^{\pm}$ are positive measures, since

$$(1+\gamma)\lambda_b(\cdot) - \gamma\lambda_b(\cdot \cap \pm Q)/\lambda_b(\pm Q) \geqq \lambda_b(\cdot \cap \pm Q)(1 + \gamma(1 - 1/\varepsilon)) \geqq 0 .$$

Now it is easy to check that $\lambda^{\pm} \neq \lambda$, $\lambda^{\pm} \in P(r)$ and $\lambda = (\lambda^+ + \lambda^-)/2$. This contradiction to the extremality of λ completes the proof.

REMARK. Similar characterizations of extremal measures have been given for different situations in [5]. The results of [13] apply here, such that there are integral representations for P and $P(r)$ with respect to their sets of extreme points (see also [14]).

6. An application to game theory. There are several applications of the above theorems in game theory and statistics. Most of them can be subsumed under the notion of "elimination of randomization". So long, these applications were restricted to cases, where the parameter set or the set of strategies were finite (see [3], [4]). Here, we shall however confine us to the following.

Application. Elimination of behavior strategies in random games. Let

$$\mathfrak{G} = \langle \{1, \cdots, n\}, X_1, \cdots, X_n, E_1, \cdots, E_n, Y_1, \cdots, Y_n, \\ w_1, \cdots, w_n, u_1, \cdots u_n \rangle$$

be a n-person-game with in advance randomly chosen personal events, i.e., $X_i \neq \phi$ is the set of (a-posteriori) strategies of player i, $E_i \neq \phi$ is the set of possible personal events of player i, $Y_i \subseteqq E_i \times X_i$, where $Y_{i|e_i} \neq \phi$ is the subset of X_i of all strategies, which are still available for player i, after e_i has happened, $w_i \in W(E_i)$ is the probability, with which the personal events will occur, and

$u_i \colon \prod_i X_i \times \prod_i E_i \longrightarrow \boldsymbol{R}$ is a bounded function, the utility function of player i ($i = 1, \cdots, n$).

The game is played in such a manner, that first a personal event e_i, which can also consist of some information, will occur for player i with probability distribution w_i. Then each player i has to choose a strategy $x_i \in Y_{i|e_i}$. The outcome for player i, if all this has happened, will be

$$u_i(x_1, \cdots, x_n, e_1, \cdots, e_n) .$$

We assume that all X_i, E_i and Y_i are Suslin spaces and that the u_i are Borel measurable. $\mathfrak{B}(X_i)$ is the σ-algebra of Borel subsets of X_i. To get a more unified representation, it is convenient to introduce the w_i-measurable functions $\bar{x}_i \colon E_i \to X_i$ with $\bar{x}_i(e_i) \in Y_{i|e_i}$ as a-priori strategies, which can be chosen before the random event e_i takes place, and to regard then the expected outcome for player i:

$$N_i(\bar{x}_1, \cdots, \bar{x}_n) = \int_{E_1} \cdots \int_{E_n} u_i(\bar{x}_1(e_1), \cdots, \bar{x}_n(e_n), e_1, \cdots, e_n) , \\ w_1(de_1) \cdots w_n(de_n) .$$

Let $\bar{X}_i$ be the set of all a-priori strategies.

In some cases even a wider class of strategies is of interest, namely the behavior strategies k_i. These are Markov kernels from E_i to X_i; i.e.,

$k_i \colon E_i \times \mathfrak{B}(X_i) \to [0, 1]$ with $k_i(e_i, \cdot) \in W(Y_{i|e_i})$ and $k_i(\cdot, B_i)$ is Borel-measurable for all $B_i \in \mathfrak{B}(X_i)$. For these behavior strategies the expected outcome of player i is

$$N_i(k_1, \cdots, k_n) = \int_{E_1} \cdots {}_{E_n}\int_{X_1} \cdots \int_{X_n} u_i(x_1, \cdots, x_n, e_1, \cdots, e_n) \\ k_1(e_1, dx_1) \cdots k_n(e_n, dx_n) w_1(de_1) \cdots w_n(de_n) .$$

THEOREM 8. *Suppose that for any w_i-atom $\bar{e}_i \in T(w_i)$ the functions $u_j(x_1, \cdots, x_i, \cdots, x_n, e_1, \cdots, \bar{e}_i, \cdots, e_n)$ are independent of $x_i \in Y_{i|\bar{e}_i}$. Then to any tupel $(k_1, \cdots, k_n)$ of behavior strategies there exist a-priori strategies $\bar{x}_i \in \bar{X}_i$ $(i = 1, \cdots, n)$ such that*

$$(1) \qquad N_j(k_1, \cdots, k_n) = N_j(\bar{x}_1, \cdots, \bar{x}_n) \quad \text{for all } j = 1, \cdots, n .$$

Moreover, if there exist Borel sets $D_i \subseteqq Y_i$ such that

$$(2) \qquad k_i(e_i, D_{i|e_i}) = 1 \quad \text{for all } e_i \in E_i$$

then the a-priori strategies $\bar{x}_i$ can be chosen to satisfy $\bar{x}_i(s_i) \in D_{i|e_i}$ $(i = 1, \cdots, n)$.

Proof. Set $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$. Since any a-priori strategy $\bar{x}_i$ can be identified with the behavior strategy $k(e_i, B_i) = \delta_{\bar{x}_i(e_i)}(B_i)$, it suffices to show that we can replace any k_i by some $\bar{x}_i$ without changing the value of (1). Assume moreover, that $i = 1$ and $D \subseteqq Y_1$ satisfies (2). For $e \in T(w_1)$ let $\bar{x}(e) \in D_{|e} \neq \phi$ and set

$$A = D \cap \{(e, x),\ e \in T(w_1) \Longrightarrow x = \bar{x}(e)\} .$$

We apply Theorem 5 to A, $B = E_1$, $C = A$, $f = \text{pro}_B$, $h = id_A$

$$\mu(F) = \int_{E_1 \setminus T(w_1)} k_1(e, F_{|e}) w_1(de) + \sum_{e \in T(w_1)} \delta_{\bar{x}(e)}(F_{|e}) w_1\{e\}$$

$\nu = w_1$ and $\pi = \mu$. Let $r\colon A \rightarrow \boldsymbol{R}^n$ be

$$r(e, x) = \int_{E_2} \cdots \int_{E_n} \int_{x_2} \cdots \int_{x_n} u(e, e_2 \cdots, e_n, x, x_2, \cdots, x_n) k_2(e_2, dx_2) \cdots k_n(e_n, dx_n) w_2(de_2) \cdots w_n(de_n) .$$

Theorem 5 gives the required a-priori strategy $\bar{x}_1$, satisfying $\bar{x}_1(e_1) \in D_{|e_1}$ and

$$N_j(\bar{x}_1, k_2, \cdots, k_n) = N_j(k_1, \cdots, k_n) \quad (j = 1, \cdots, n) .$$

REMARK. Under the assuptions, that all X_i are finite and the w_i on E_i are nonatomic $(i = 1, \cdots, n)$, a result of the above kind has been shown in [4] and has also been used by W. Armbruster, Heidelberg, to obtain the existence of equilibrium points in a-priori strategies.

ACKNOWLEDGMENTS. I would like to thank the referee who pointed out to me the connection to the papers [5] and [13]. My thanks are also due to Mrs. Reinemuth for her fine typing.

References

1. R. J. Aumann and L. S. Shapley, *Values of Non-Atomic Games*, Princeton University Press, 1974.
2. C. Dellacherie, P. A. Meyer, *Probabilitês et Potentiel*, chapitre I á IV, Paris, Hermann, 1975.
3. A. Dvoretzky, A. Wald and J. Wolfowitz, *Elimination of Randomisation in certain Problems of Statistics and of the Theory of Games*, Proc. Nat. Acad. Sci. USA, **36** (1950), 256-259.
4. ———, *Relation among certain ranges of vector measures*, Pacific J. Math., **1** (1951), 59-74.
5. G. A. Edgar, *Measurable weak sections*, Illinois J. Math., **20** (1976), 630-646.
6. K.-Th. Eisele, *On the uniqueness of pre-images of measures*, In: Measure Theory, Proceedings of the Conference Held at Oberwolfach, Lecture Notes in Mathematics, **541**, 5-12, Berlin-Heidelberg-New York: Springer, 1976.
7. J. Hoffmann-Jørgensen, *The Theory of Analytic Spaces*, Various Publication Series 10, Aarhus Matematisk Institut 1970.
8. J. Lehn and G. Mägerl, *On the Uniqueness of Pre-image Measures*, Z. Wahrscheinlichkeitstheorie verw. Gebiete, **38** (1977), 333-337.
9. J. Lindenstrauss, *A short proof of Liapounoff's convexity theorem*, J. Math. Mech., **15** (1966), 971-972.
10. A. Lyapunov, *Sur les functions-vecteurs completement additives*, Bull. Acad. Sci. URSS Ser. Math., **4** (1940), 465-478.
11. L. Schwartz, *Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures*, Oxford University Press, 1973.
12. ———, *Surmartingales rêgulières á valeurs mesures et dêsintêgrations rêgulières d'une mesure*, J. Anal. Math., **26** (1973), 1-168.
13. H. V. Weizsäcker and G. Winkler, *Integral representation in the set of solutions of a generalized moment problem*, (to appear in Math. Ann.).
14. ———, *Non-compact Extremal Integral Representations: Some Probabilistic Aspects*, to appear in: Proc. of 2nd Paderborn Meeting of Functional Analysis, North-Holland Publ. Co., Amsterdam 1979.

Received July 31, 1979 and in revised form November 29, 1979.

INSTITUT FÜR ANGEWANDTE MATHEMATIK
UNIVERSITÄT HEIDELBERG
IM NEUENHEIMER FELD 294
D-6900 HEIDELBERG, GERMANY