Orthogonal Transition Kernels

R. Daniel Mauldin; David Preiss; Heinrich v. Weizsacker

The Annals of Probability, Vol. 11, No. 4 (Nov., 1983), 970-988.

The Annals of Probability
1983, Vol 11, No. 4, 970–988

ORTHOGONAL TRANSITION KERNELS

BY R. DANIEL MAULDIN[1], DAVID PREISS AND HEINRICH V. WEIZSÄCKER

North Texas State University, Charles University, and Universität Kaiserslautern

A transition kernel $(\mu_x)_{x\in X}$ between Polish spaces X and Y is called completely orthogonal if the μ_x are separated by the fibers of a Borel map φ: $Y \to X$. It is orthogonality preserving if orthogonal measures on X induce orthogonal mixtures on Y. We give a von Neumann "type" isomorphism theorem for atomless completely orthogonal kernels, and a theorem and some counterexamples concerning the separation of two orthogonal measure convex sets of probability measures by a measurable set. These techniques yield three results on orthogonality preserving kernels: (1) They need not be completely orthogonal but (2) are uniformly orthogonal (in the sense of D. Maharam) and (3) if X is σ-compact, $Y = \lim_{\leftarrow} Y_n$ and (μ_x) is orthogonality preserving and continuous in x then there is even a strongly consistent sequence of statistics φ_n: $Y_n \to X$ for (μ_x).

0. Introduction. Let X and Y be Polish spaces. We will address ourselves to the study of transition kernels of orthogonal probability measures from X to Y and their classification. To this end we first (Section 1) present a hierarchy of notions of orthogonality beyond the simple property of being pairwise orthogonal.

These include "complete orthogonality" and the property "orthogonality preserving". Complete orthogonality which in statistical language means existence of a perfect statistic φ: $Y \to X$, is known to occur in many contexts (ergodic decompositions, Gibbs states, extremal models, etc. ([6], [8], [9], [19])). Preservation of orthogonality is an "algebraic" condition which is e.g. equivalent to p-almost sure consistency of the a posteriori distributions for every a priori measure p on X. This condition naturally arises when looking at orthogonal kernels from a functional analytic point of view (cf. the comments to Theorem 4.1).

Section 2 contains a uniform von Neumann isomorphism theorem for atomless completely orthogonal kernels. It implies a complete classification of completely orthogonal kernels. It improves a parametrization theorem of Mauldin [21] and a classical result on decompositions of Lebesgue space of Rokhlin and Maharam.

Section 3 is devoted to the following problem. Suppose M and N are measurable measure convex sets of probability measures on a completely regular space Y such that every element of one is orthogonal to any element of the other. Does there exist a Baire subset B of Y which is a nullset for all measures in M and a one-set for those in N? (To see the role of measure convexity look at the trivial counterexample M = {Lebesgue measure} and $N = \{\varepsilon_y: y \in [0,1]\}$.) *If* M and N are σ-compact, the answer is known to be yes [11]. We prove the same if one of the sets is a singleton and the other arbitrary. Our method is based on the concept of a filter of countable type (essentially due to Grimeisen [13] and Katetov [16]) which allow us to formulate a weakened minimax theorem for certain convex measurable sets. The proof also yields the result of Mokobodzki [31] that under Martin's Axiom there is for every pair M, N, a universally measurable separating set. However, in Section 5, it is shown that there need not exist a Borel separating set even if Y is Polish and both M and N are narrowly closed or one of them is narrowly compact. This improves an example found independently by Blackwell [2] and Preiss [25].

Received November 1981; revised September 1982.

[1] Research supported by NSF: Grant MCS 81-01581.

AMS 1970 *subject classifications.* Primary 60A10; secondary, 28A75.

Key words and phrases. Orthogonal measures, perfect statistics, filters of countable type, classification of kernels, simultaneous Lebesgue decompositions.

Section 4 gives a number of reformulations of the condition of being orthogonality preserving. A new point is that it implies under some additional assumptions the existence of a strongly consistent sequence of statistics. In particular, an orthogonality preserving kernel is completely orthogonal if the set $\{\mu_x\}_{x\in X}$ is narrowly σ-compact. Also we prove using the results of Section 3 that orthogonality preserving kernels are always uniformly orthogonal in the sense of D. Maharam [20], i.e. every measure μ_x of the kernel is supported by a Borel set B_x which is a nullset for the other measures in the kernel. However, we then show in the last section by similar techniques as used for the example mentioned above that preservation of orthogonality does not imply complete orthogonality in general, i.e. sometimes the sets B_x must overlap.

A list of problems related to and left open in this paper appears in [22]. Two others are mentioned at the end of the paper.

Besides its intrinsic interest, our work may also more specifically be interpreted as an effort to understand the general structural conditions on a measure theoretic model $(Y, \mathscr{B}, \{\mu_x: x \in X\})$ of a statistical experiment which ensure that according to this model the parameter x can (perhaps asymptotically) be completely identified from the idealized observations in Y.

1. Notions of orthogonality. Let $(X, \mathscr{B}(X))$ and $(Y, \mathscr{B}(Y))$ be measurable spaces. We shall be mainly interested in the case where both are uncountable standard Borel spaces, i.e. Borel isomorphic to the unit interval I with its Borel σ-algebra. However the following definitions also make sense in more general settings. Let $\mathscr{P}(X)$ and $\mathscr{P}(Y)$ be the set of all probability measures on $\mathscr{B}(X)$ and $\mathscr{B}(Y)$, respectively. A transition kernel (μ_x) from X to Y is a family $(\mu_x)_{x\in X}$ of elements of $\mathscr{P}(\mathrm{Y})$ such that $x \to \mu_x(E)$ is Borel measurable for all Borel subsets E of Y.

In our opinion the following concept of "complete orthogonality" is the most important of the notions discussed in this paper.

DEFINITION 1.1. A transition kernel $(\mu_x)_{x\in X}$ from X to Y is said to be *completely orthogonal* provided there is a set $B \in \mathscr{B}(X) \otimes \mathscr{B}(Y)$ such for each x, $\mu_x(B_x) = 1$ and if $x \neq x'$, then $B_x \cap B_{x'} = \emptyset$. We say that the set B *completely separates* (μ_x).

Letting $\varphi(y) = x$ if $(x, y) \in B$ and $\varphi(y) = x_0(x_0 \in X$ arbitrary) if $y \notin \pi_Y(B)$ we see that B is the (anti)-graph of a map φ: $Y \to X$ satisfying $\mu_x(\varphi^{-1}\{x\}) = 1$ for all x, i.e. the measures μ_x are separated by a map. If X and Y are standard Borel the map is Borel measurable by Kuratowski's theorem. This justifies the definition of complete orthogonality given in our summary. Statisticians call such a φ an "exact estimator of the parameter x" or a "perfect statistic". In [6], the associated kernel Q: $y \mapsto \mu_{\varphi(y)}$ is called a *H-sufficient* statistic for the set $\{\mu_x\}_{x\in X}$. Kernels (Q_y) from Y to Y arising in this way from completely orthogonal kernels are characterized by the property $Q_y(\{y'\colon Q_y = Q_{y'}\}) = 1$ for all $y \in Y$ (decomposing kernels [17]).

Our next proposition illustrates one manner in which completely orthogonal transition kernels arise.

PROPOSITION 1.2. *Let Y be a Polish space and let $\mathscr{N}$ be a set of kernels (ν_y) from Y to Y such that for all but at most countably many elements of $\mathscr{N}$ the corresponding map $\tilde{\nu}$: $\rho \to \int_Y \nu_y(\cdot)\,d\rho(y)$ of $\mathscr{P}(Y)$ to $\mathscr{P}(Y)$ is continuous in the narrow topology. Let $\mathscr{P}_{\mathscr{N}}$ be the set $\{\rho \in \mathscr{P}(Y): \tilde{\nu}\rho = \rho$ for all $(\nu_y) \in \mathscr{N}\}$. If $\mathscr{P}_{\mathscr{N}} \neq 0$ then there is a completely orthogonal kernel (μ_x) from some standard X to Y such that $\{\mu_x\}_{x\in X}$ is exactly the set of extreme points of $\mathscr{P}_{\mathscr{N}}$ (i.e. the $\mathscr{N}$-ergodic measures are parametrized by (μ_x)).*

PROOF. Write $\mathscr{N} = \mathscr{N}_1 \cup \mathscr{N}_2$ where $\mathscr{N}_1$ is countable and $\tilde{\nu}$ is continuous for all (ν_y) in $\mathscr{N}_2$. Then $\mathscr{P}_{\mathscr{N}_2}$ is the set of common fixpoints of the family $\{\tilde{\nu}: (\nu_y) \in \mathscr{N}_2\}$ of continuous transformations of $\mathscr{P}(Y)$. By the Lindelöf property of the separable metrizable space

$\mathscr{P}(Y)$, there is a countable subset $\mathcal{N}'_2$ of $\mathcal{N}_2$ with the same invariant measures. Thus $\mathscr{P}_{\mathcal{N}} = \mathscr{P}_{\mathcal{N}_1} \cap \mathscr{P}_{\mathcal{N}_2} = \mathscr{P}_{\mathcal{N}_1} \cap \mathscr{P}_{\mathcal{N}'_2} = \mathscr{P}_{\mathcal{N}_1 \cup \mathcal{N}'_2}$ is the set of invariant probability measures of the countable set $\mathcal{N}_1 \cup \mathcal{N}'_2$. For countable sets $\mathcal{N}$ the proposition is proved in [17], Theorems 2.6–2.8.

COMMENT. If the elements of $\mathcal{N}$ are induced by point maps, the corresponding results are already contained in [8]. As pointed out in the introduction of [17], the present more general assumption also includes Gibbs measures and the other examples treated in [6].

A second concept of orthogonality is the following, which—except for the measurability requirement for (μ_x) and B—was introduced in [20].

DEFINITION 1.3. A transition kernel $(\mu_x)_{x\in X}$ from X to Y is said to be *uniformly orthogonal* provided there is a set $B \in \mathscr{B}(X) \otimes \mathscr{B}(Y)$ such that for each x, $\mu_x(B_x) = 1$ and if $x \neq x'$, then $\mu_{x'}(B_x) = 0$.

That this notion is weaker than that of complete orthogonality is demonstrated by the following example which we first learned from R.J. Gardner.

EXAMPLE 1.4. Take $X = I$ and $Y = I \times I$. For $x \in I$, let μ_x be the probability measure $\frac{1}{2}\varepsilon_x \otimes \lambda + \frac{1}{2}\lambda \otimes \varepsilon_x$ where λ is Lebesgue measure. Then the set

$$B = \{(x, (y_1, y_2)): y_1 = x \text{ or } y_2 = x\} \subset X \times Y$$

has the properties in the definition of uniform orthogonality. But it is not difficult to verify that the restriction of this kernel to any compact subset of X with positive Lebesgue measure fails to be completely orthogonal.

COMMENT. In [10] a transition kernel $(\mu_x)_{x\in I}$ of pairwise orthogonal measures is constructed such that the restriction of the kernel to any subset of I with positive Lebesgue measure fails to be uniformly orthogonal. From the "Cantorian theorem" in [3] it follows however that for a transition kernel (μ_x) of pairwise orthogonal measures there is always a nonempty compact perfect subset X_0 of X such that $(\mu_x)_{x\in X_0}$ is completely orthogonal. Even the analogue of this for non-measurable families of pairwise orthogonal measure fails, as was proved in the paper [20] which motivated the research of [3] and [10].

REMARK. One could consider *modified* notions of uniform and complete orthogonality by deleting the requirement that the set B be a Borel measurable subset of $X \times Y$. Let μ_0 be Lebesgue measure on [0,1] and μ_x be the point mass at x for each x in (0,1]. Then $(\mu_x)_{x\in[0,1]}$ is a transition kernel consisting of pairwise orthogonal measures which is not uniformly orthogonal even in the modified sense. With this modification, uniform orthogonality implies complete orthogonality if card $X = \aleph_1$. One simply sets $D_x = B_x - \bigcup \{B_{x'}: x' < x\}$ where $<$ is a well ordering of X into type $\aleph_1$. On the other hand the situation may change for Polish X if one drops CH:

EXAMPLE 1.5 (Fremlin). Take a model of set theory in which there is a set $A \subset [0,1]$ such that (a) A has positive outer measure and (b) if $\{N_\alpha: \alpha \in A\}$ is a family of sets with measure zero, then $\bigcup\{N_\alpha: \alpha \in A\} \neq [0,1]$. Then there is a uniformly orthogonal kernel which is not (modified) completely orthogonal: Let $X = ((0, 1] \times \{0\}) \cup (\{0\} \times (0,1])$ and $Y = (0,1] \times (0,1]$. Let $\mu_{(0,t)}$ be $\lambda \otimes \varepsilon_t$ and let $\mu_{(t,0)}$ be $\varepsilon_t \otimes \lambda$. As in the first example this kernel is uniformly orthogonal. Suppose now that there is a subset B of $X \times Y$ such that for each x, $\mu_x(B_x) = 1$ and if $x \neq x'$, then $B_x \cap B_{x'} = \varnothing$. According to property (b), there is some t' in $\bigcap_{t\in A}\{s:(s, t) \in B_{(0,t)}\}$. According to property (a), there is some t'' in A such that $(t', t'') \in B_{(t',0)}$. Then $(t', t'') \in B_{(t',0)} \cap B_{(0,t'')}$, contradiction.

The following condition depends on a (usually metric) topology on X having $\mathscr{B}(X)$ as Borel-σ-algebra and on an increasing sequence $(\mathscr{B}_n)$ of sub-σ-algebras of $\mathscr{B}(Y)$ generating $\mathscr{B}(Y)$. It can be verified in many probabilistic situations.

DEFINITION 1.6. The kernel (μ_x) is said to admit a *strongly consistent sequence of statistics* if for each n, there is a $\mathscr{B}_n$-measurable map φ_n of Y into X such that for each x, $\mu_x\{y: \lim_{n\to\infty} \varphi_n(y) = x\} = 1$.

Finally we turn to a lattice condition which is near to complete orthogonality. The name is taken from [12]. A number of equivalent reformulations of this property will be given in Section 4.

DEFINITION 1.7. A transition kernel $(\mu_x)_{x\in X}$ is called *orthogonality preserving*, if for any pair (p, q) of orthogonal elements of $\mathscr{P}(X)$ the corresponding mixtures $\mu^p = \int_X \mu_x p(dx)$ and $\mu^q = \int_X \mu_x q(dx)$ are orthogonal.

Now we summarize the relations between the notions of definitions 1.1, 1.3, 1.6, and 1.7.

THEOREM 1.8. *Let (μ_x) be a transition kernel between standard measurable spaces X and Y.*

a) *If (μ_x) admits a strongly consistent sequence of statistics then it is completely orthogonal.*

b) *Every completely orthogonal kernel is orthogonality preserving.*

c) *Every orthogonality preserving kernel is uniformly orthogonal.*

d) *The converse implications in* a), b) *and* c) *are false in general.*

e) (*Partial converse of* a) *and* b)). *Assume that X and Y have metric topologies so that X is σ-compact, $x \mapsto \mu_x$ is continuous for the narrow topology on $\mathscr{P}(Y)$ and $Y = \lim_n Y_n$ for some spaces Y_n with projections $\pi_n: Y \to Y_n$. If (μ_x) is orthogonality preserving then it admits a strongly consistent sequence of $\pi_n^{-1}(\mathscr{B}(Y_n))$-measurable statistics.*

PROOF. The proofs of c) and e) will be postponed to Section 4 and the counterexample to the converse of b) to Section 5.

a) is obvious: Consider $B = \{(x, y): \lim_{n\to\infty}\varphi_n(y) = x\}$ where (φ_n) has the properties of Definition 1.6.

b) Suppose (μ_x) is completely orthogonal. Let φ be a Borel map from Y to X satisfying $\mu_x(\varphi^{-1}\{x\}) = 1$ for all x. Let p and q be two orthogonal measures on X. Choose A in $\mathscr{B}(X)$ such that $p(A) = 1 = q(X\backslash A)$. Then $\mu^p(\varphi^{-1}(A)) = \int_X \mu_x(\varphi^{-1}(A))p(dx) = \int_A 1p(dx) = 1$ and similarly $\mu^q(Y\backslash\varphi^{-1}(A)) = 1$. So μ^p and μ^q are orthogonal. This proves b).

d) Example 1.4 also shows that a uniformly orthogonal kernel need not be orthogonality preserving, so the converse of c) fails.

Finally we give an example showing that the converse of a) fails: Let Y be the Cantor space $\{0,1\}^N$ and let $\psi: X \to Y$ be a Borel isomorphism such that ψ^{-1} is not of first Baire class. If $\mu_x = \varepsilon_{\psi(x)}$ then there is no sequence (φ_n) of continuous mappings of Y into X such that for each x, $\mu_x\{y: \varphi_n(y) \to x\} = 1$. Now take $\mathscr{B}_n$ to be the σ-algebra generated by the first n coordinates so that every $\mathscr{B}_n$-measurable map is continuous. For this choice of $(\mathscr{B}_n)$ there is no strongly consistent sequence of statistics. However, (μ_x) clearly is completely orthogonal.

2. Classification of completely orthogonal kernels. The following may be called the "canonical example of a completely orthogonal atomless transition kernel".

EXAMPLE 2.1. Let I be the unit interval (0, 1] and let $\nu: I \times \mathscr{B}(I \times I) \to [0, 1]$ be defined by $\nu_t(E) = (\varepsilon_t \otimes \lambda)(E) = \lambda(E_t)$. Certainly (ν_t) is a transition kernel. Let B be the set $\{(t, (t, s)) : (t, s) \in I \times I\}$. Then (1) B is a Borel subset of $I \times (I \times I)$. (2) $B_t \cap B_{t'} = \varnothing$ if $t \neq t'$ and (3) $\nu_t(B_t) = 1$ for all t. So (ν_t) is completely separated by B.

Indeed, we have the following isomorphism theorems which can be viewed as parametrized versions of von Neumann's theorem that if μ is an atomless probability measure on a Borel set B in a Polish space X, then there is a Borel isomorphism φ of B onto I such that $\mu\circ\varphi^{-1} = \lambda$.

THEOREM 2.2. *Let X and Y be uncountable standard Borel spaces. Let (μ_x) be an atomless completely orthogonal transition kernel from X to Y. Let B be a Borel subset of $X \times Y$ which completely separates (μ_x). Then for each Borel isomorphism ψ of X onto I there is a Borel isomorphism φ of $\pi_Y(B)$ onto $I \times I$ such that for each x in X the set B_x is mapped under φ onto $\{\psi(x)\} \times I$ and $\varepsilon_{\psi(x)} \otimes \lambda$ is the image measure of μ_x under φ.*

PROOF. If B completely separates (μ_x) then the projection $\pi_Y: B \to \pi_Y(B)$ is a Borel isomorphism, mapping the set $\{x\} \times B_x$ onto B_x and $\varepsilon_x \otimes \mu_x$ to μ_x for each x. Therefore, this theorem is implied by the following which improves Theorem 2.3 in [21]. (In fact our two results are essentially the same, because for every not necessarily orthogonal kernel (μ_x) from X to Y one may consider the completely orthogonal kernel $(\varepsilon_x \otimes \mu_x)_x$ from X to $X \times Y$ applied to which the first result easily yields the second.)

THEOREM 2.3. *Let X and Y be uncountable standard Borel spaces and (μ_x) be an atomless transition kernel from X to Y. Let B be a Borel subset of $X \times Y$ with $\mu_x(B_x) = 1$ for all x. Then to every Borel isomorphism ψ of X onto I there is a Borel isomorphism φ of B onto $I \times I$ such that under φ for each x in X*
(i) *the set $\{x\} \times B_x$ is mapped onto $\{\psi(x)\} \times I$*
(ii) *the measure $\varepsilon_x \otimes \mu_x$ has $\varepsilon_{\psi(x)} \otimes \lambda$ as its image.*

PROOF. According to Theorem 2.3 of [21] there are Borel isomorphisms ψ_0 and φ_0 of X onto I and of B onto $I \times I$ respectively which fulfill the analogue of property (i). Since an automorphism of the first factor (component) of $I \times I$ does not change this statement we may assume $\psi_0 = \psi$. In order to show that a suitable modification φ of φ_0 even satisfies (ii), apply the following assertion to the kernel (ρ_t) which is defined by $\varepsilon_{\psi(x)} \otimes \rho_{\psi(x)} = (\varepsilon_x \otimes \mu_x) \circ \varphi_0^{-1}$.

LEMMA 2.4. *Let (ρ_t) be an atomless transition kernel from I to I. Then there is a Borel isomorphism τ of $I \times I$ onto $I \times I$ which for each t maps the set $\{t\} \times I$ onto itself and transforms $\varepsilon_t \otimes \rho_t$ into $\varepsilon_t \otimes \lambda$.*

PROOF OF THE LEMMA. For each t let R_t be the measure defining function associated with ρ_t, $R_t(s) = \rho_t((0, s])$. The maps R_t are continuous, nondecreasing with 0 and 1 as fixpoints. In particular $R_t(I) = I$. Consider the set

$$L = \cup_{r \text{ rational}} \{(t, s) : r < s,\ R_t(r) = R_t(s)\}.$$

Clearly $L \in \mathscr{B}(I \times I)$. For each t the section L_t is the union of the intervals of constancy of R_t including the right and excluding the left endpoints. In particular $\lambda(L_t) > 0$ if $L_t \neq \varnothing$.

Fix an uncountable Borel set D in I of zero Lebesgue measure. Choose a Lebesgue measure preserving Borel isomorphism δ of I onto $I - D$. Since $\pi_1(L) = \{t : \lambda(L_t) > 0\}$ and L is a Borel set it follows that $\pi_1(L)$ is a Borel set. According to Theorem 2.3 of [21] (or by an elementary direct construction) there is a Borel isomorphism ω of L onto $\pi_1(L) \times D$ which maps $\{t\} \times L_t$ onto $\{t\} \times D$ for each t. Now define $\tau : I \times I \to I \times I$ by

$$\tau(t, s) = \begin{cases} (t, R_t(s)) & \text{if } t \notin \pi_1(L) \\ (t, \delta(R_t(s))) & \text{if } t \in \pi_1(L) \text{ but } (t, s) \notin L \\ \omega(t, s) & \text{if } (t, s) \in L. \end{cases}$$

As is easily checked, τ is bijective Borel measurable and it keeps the first coordinate fixed. It remains to be shown that for each t, λ is the image of ρ_t under the map $\tilde{R}_t : s \to \pi_2(\tau(t, s))$. For R_t instead of $\tilde{R}_t$ this is simple and well known. If $t \notin \pi_1(L)$ then $\tilde{R}_t = R_t$ and for $t \in \pi_1(L)$ the map $\tilde{R}_t$ is obtained from R_t by just sending the ρ_t-nullset L_t to D and on $I \backslash L_t$ composing R_t with the λ-preserving transformation δ. This completes the proof of the lemma and of the two theorems.

We would like to note the following application of our methods. For each x, one can choose in the second theorem a nonempty compact perfect μ_x-nullset $K_x \subseteq B_x$. The problem is, can one make this selection in a measurable fashion. In [21] it was demonstrated that this is so if we were selecting compact perfect sets with positive μ_x measure. Our corollary gives a positive answer for the case of nullsets.

COROLLARY 2.5. ("*Null set theorem*"). *In Theorem* 2.3 *let Y be equipped with a Polish topology compatible with the given Borel structure. Then the set B contains a Borel set K such that for each x the section K_x is a nonempty compact perfect μ_x-nullset.*

PROOF. Choose any Borel isomorphism ψ of X onto I. Let φ be as in the theorem. Let ρ be an atomless probability measure on I which is concentrated on a Lebesgue nullset D. Let M be the inverse image of $I \times D$ under φ. Then M is a Borel subset of B such that $\mu_x(M_x) = \lambda(D) = 0$ for all x. On the other hand $\kappa_x(M_x) = \rho(D) = 1$ if we define the atomless kernel (κ_x) by $(\varepsilon_x \otimes \kappa_x) \circ \varphi^{-1} = \varepsilon_{\psi(x)} \otimes \rho$. Thus, according to Theorem 2.4 of [16] the set M has a Borel subset K with nonempty compact perfect sections. Because of $K_x \subseteq M_x$ we have also $\mu_x(K_x) = 0$.

For the sake of completeness we now consider also measures with a discrete part. Let $z = (z_i)_{i \geq 0}$ be a nonincreasing sequence of nonnegative real numbers satisfying $\sum_0^\infty z_i = 1$. A probability measure μ is said to be of *type* z if it is of the form $\mu = \mu^c + \sum_1^\infty z_i \varepsilon_{y_i}$ where μ^c is an atomless measure and $y_i \neq y_j$ whenever $z_i > 0$, $z_j > 0$ and $i \neq j$. The type of μ is unique and will be denoted by $z(\mu)$ or $(z_i(\mu))_{i \geq 0}$. The proof of the following lemma is left to the reader. It is a consequence of the fact that $\mathscr{B}(Y)$ is countably generated.

LEMMA 2.6. *Let (μ_x) be a transition kernel from X to Y. Then the mappings $x \mapsto z_i(\mu_x)$ are measurable and there is a sequence $(f_i)_{i \geq 1}$ of Borel functions of X to $Y \cup \{y_\infty\}$ (where $y_\infty \notin Y$) such that for each x the discrete part μ_x^d of μ_x equals $\sum_1^\infty z_i(\mu_x)\varepsilon_{f_i(x)}$ and $f_i(x) = y_\infty$ if $z_i(\mu_x) = 0$.*

Let now X, X', Y, Y' be standard Borel spaces. Let (μ_x) and $(\nu_{x'})$ be completely orthogonal transition kernels from X and X' to Y and Y' respectively. Let B and B' be completely separating sets for (μ_x) and $(\nu_{x'})$. Denote by $\bar{Y}$ the set

$$(\cup_{\mu_x^c \neq 0} B_x) \cup (\cup_{\mu_x^c = 0} \{y : \mu_x\{y\} > 0\}).$$

Let $\overline{Y'}$ be defined similarly using B' and $(\nu_{x'})$.

THEOREM 2.7. *If there is a Borel isomorphism ψ of X onto X' such that $z(\nu_{\psi(x)}) = z(\mu_x)$ for all x then there is a Borel isomorphism φ of the Borel set $\bar{Y}$ onto $\overline{Y'}$ such that for every x*

(i) $\varphi(B_x) = B'_{\psi(x)}$ *if* $\mu_x^c \neq 0$,

(ii) $\mu_x \circ \varphi^{-1} = \nu_{\psi(x)}$.

PROOF. Let X_0 be the set of all x in X for which $\mu_x^c \neq 0$, or equivalently for which $z_0(\mu_x) > 0$. According to the lemma X_0 is a Borel subset of X. The restriction $\psi|_{X_0}$ is a Borel isomorphism of X_0 onto $X_0' = \{x' \in X' : z_0(\nu_{x'}) > 0\}$.

Choose $f_i : x \to Y \cup \{y_\infty\}$ according to the lemma for the kernel (μ_x), similarly $f_i' : X' \to Y' \cup \{y_\infty'\}$ for $(\nu_{x'})$. The set $B_0 = (B \backslash \cup_{i=1}^\infty \text{Graph}\,(f_i)) \cap (X_0 \times Y)$ completely separates the atomless kernel $((1/z_0(\mu_x))\mu_x^c)_{x \in X_0}$. Similarly $((1/z_0(\nu_{x'}))\nu_{x'}^c)_{x' \in X_0'}$ is separated by $B_0' = (B' \backslash \cup_{i=1}^\infty \text{Graph}\,(f_i')) \cap (X_0' \times Y')$. It is easy to deduce from the first theorem in this section the existence of a Borel isomorphism φ_0 of $\pi_Y(B_0)$ onto $\pi_{Y'}(B_0')$ such that $\varphi_0(B_{0x}) = B'_{0\psi(x)}$ and $\mu_x^c \circ \varphi^{-1} = \nu_{\psi(x)}^c$ for each $x \in X_0$.

Each f_i maps the set $X_i = \{x : z_i(\mu_x) > 0\}$ injectively onto its image Y_i, since the measures μ_x are pairwise orthogonal. For the same reason $\bar{Y} = \pi_Y(B_0) \cup \cup_{i=1}^\infty Y_i$, the union being disjoint. Similarly $\overline{Y'} = \pi_{Y'}(B_0') \cup \cup_{i=1}^\infty Y_i'$. So $\bar{Y}$ and $\overline{Y'}$ are Borel sets and we can define the map $\varphi: \bar{Y} \to \overline{Y'}$ by

$$\varphi(y) = \begin{cases} f_i' \circ \psi \circ f_i^{-1} & \text{if } \ y \in Y_i,\ i \geq 1 \\ \varphi_0(y) & \text{if } \ y \in \pi_Y(B_0). \end{cases}$$

It is easily verified that φ has all required properties.

COMMENTS. 1) Compare the "classification of measurable decompositions of a Lebesgue space" in [26], Section 4. Our results in this section are essentially the strengthened form of Rokhlin's which are obtained by dropping in his conclusions everywhere the qualification "mod 0". Rosenblatt ([27], Lemma 2) and Ershov [7] independently gave a simplified proof of Rokhlin's "theorem on independent complements," [26], Section 4, number 3, which is a reformulation of the atomless case. Their proof uses an argument similar to our Lemma 2.4.

2) The condition that card $\{x: z(\mu_x) = z\} = \text{card}\{x': z(\nu_{x'}) = z\}$ for every z does not imply the existence of a Borel isomorphism ψ of X onto X' satisfying $z(\mu_x) = z(\nu_{\psi(x)})$. This is why ψ is always given in advance. Indeed, let Z be the set of all possible types with its Borel structure induced from the Polish space ℓ_1. Let X be the product $Z \times I$ and let X' be a subset of $Z \times I$ such that for each $z \in Z$ the section X'_z is uncountable but X' does not admit a Borel isomorphism $\psi: X \to X'$ leaving the first coordinate fixed. Then for any completely orthogonal kernel $(\mu_x)_{x \in X}$ satisfying $z(\mu_{(c,t)}) = c$ for all $c \in Z$, the kernel (μ_x) and its restriction $(\nu_{x'})$ to X' may serve as a counterexample.

3. Separation of two sets of measures.

In this section we will work in the following setting unless otherwise stated.

GENERAL HYPOTHESIS. *Let $(Y, \mathscr{B})$ be any measurable space, $\mathscr{P}$ be a set of probability measures on $\mathscr{B}$ and let G be a linear subspace of $\mathscr{L}^\infty(Y, \mathscr{B})$ such that*

$$\|\kappa\|_1 = \sup\left\{\int g\, dx : g \in G,\ \|g\|_\infty \leq 1\right\}$$

for all κ in the linear hull of $\mathscr{P}$. The set $\mathscr{P}$ will always be equipped with the topology of pointwise convergence on G.

The most interesting case occurs when Y is a topological space, G is the set of all bounded continuous functions, $\mathscr{B}$ the Baire σ-algebra (generated by G) and $\mathscr{P}$ the set of all Baire probability measures $\mathscr{P}(Y)$ with the "narrow" topology (the topology induced by $\sigma(M(Y), C_b(Y))$).

We want to modify the following result of Goullet de Rugy [11]. If M and N are countable unions of compact convex sets and $\mu \perp \nu$ for all $\mu \in M$ and $\nu \in N$ and Y is compact, then there are pairwise disjoint $K_{\sigma\delta}$ subsets F and H of Y such that $\mu(F) = 1 = \nu(H)$, for all $\mu \in M$, $\nu \in N$. (Recently, S. Graf and G. Mägerl gave a new proof of this using capacities, [12].) We will improve this theorem by demonstrating that the sets F and H can be taken to be K_σ sets. However, it is not this result itself which we need. Rather, it is the quantitative version of Goullet de Rugy's theorem given in the following proposition and the minimax lemma in its proof which we will employ later on. Finally, let us note that if one substitutes for compactness, the existence of a σ-finite measure with respect to which the elements of K and L are absolutely continuous the proposition is (with a similar proof) already essentially contained in [18] where it is attributed to Le Cam. (This was pointed out to us by H. Strasser.)

PROPOSITION 3.1. *Let K and L be two compact convex subsets of $\mathscr{P}$. Suppose for some $\varepsilon > 0$ and each μ in K and each ν in L there is a set $S_{\mu\nu}$ in $\mathscr{B}$ such that $\mu(S_{\mu\nu}) < \varepsilon$ and $\nu(S_{\mu\nu}) > 1 - \varepsilon$. Then there is a function $g \in G$, $\|g\| \leq 1$ and a $\delta > 0$ such that $\mu\{g < \delta\} < 2\varepsilon$ and $\nu\{g > -\delta\} < 2\varepsilon$, for all $\mu \in K$, $\nu \in L$.*

PROOF. Computing integrals of the functions $1_{S_{\mu\nu}} - 1_{Y\setminus S_{\mu\nu}}$ we see that $\|\mu - \nu\|_1 > 2 - 2\varepsilon$ for each μ in K and each ν in L. Since $\| \|_1$ is lower semicontinuous with respect to our topology and the sets K and L are compact we infer the strict inequality $\mathrm{dist}_{\| \|_1}(K, L) > 2 - 2\varepsilon$. According to the minimax lemma given below, there is a function $g \in G$, $\|g\|_\infty \leq 1$ and a number $d > 2 - 2\varepsilon$ such that $\mu(g) - \nu(g) \geq d$ wherever $\mu \in K$, $\nu \in L$. For each $\delta > 0$, define $g_\delta := g \cdot 1_{\{|g| \geq \delta\}}$. Then $\|g - g_\delta\|_\infty \leq \delta$. Therefore, there is some $\delta_0 > 0$ such that $\mu(h) - \nu(h) > 2 - 2\varepsilon$, for all $\mu \in K$, $\nu \in L$, where $h = g_{\delta_0}$. Notice

$$1 - \mu\{h \leq 0\} = \mu\{h > 0\} \geq \int_{\{h>0\}} h\, d\mu \geq \mu(h) > 2 - 2\varepsilon + \nu(h) \geq 1 - 2\varepsilon$$

i.e. $\mu\{h \leq 0\} < 2\varepsilon$ for all μ in K. Similarly $\nu\{h \geq 0\} < 2\varepsilon$ if $\nu \in L$. Since $\{h \leq 0\} = \{g < \delta_0\}$ and $\{h \geq 0\} = \{g > -\delta_0\}$, the proof is complete.

MINIMAX LEMMA 3.2. *Let $(G, \| \|)$ be a normed space. If K and L are convex and $\sigma(G^*, G)$-compact subsets of G^* then*

$$\mathrm{dist}_{\| \|}(K, L) = \sup\{\inf_{e \in K}\langle e, g\rangle - \sup_{e' \in L}\langle e', g\rangle : g \in G, \|g\| \leq 1\}.$$

PROOF. The lemma is a special case of classical Minimax theorems (e.g. [30], Theorem 4.2′) applied to the map $\langle , \rangle : (K - L) \times B \to \mathbb{R}$ where B is the unit ball in G.

COROLLARY 3.3. *Let Y be a completely regular space and let $M = \cup_{n=1}^\infty K_n$ and $N = \cup_{n=1}^\infty L_n$ be subsets of $\mathscr{P}(Y)$ such that K_n and L_n are narrowly compact convex sets and $\mu \perp \nu$ whenever $\mu \in M$, $\nu \in N$. Then there are pairwise disjoint Baire-measurable F_σ-sets F and H such that $\mu(F) = 1 = \nu(H)$ for all $\mu \in M$, $\nu \in N$. If all measures in $M \cup N$ are tight and each of the sets K_n, L_n is uniformly tight, then F and H can be chosen to be K_σ-sets.*

PROOF. Assume $K_n \uparrow M$, $L_n \uparrow N$ for compact convex set K_n, L_n. Choose $\varepsilon_n > 0$ such that $\sum_{n=1}^\infty \varepsilon_n < \infty$. For every n there is a continuous function g_n and a number $\delta_n > 0$ such that $\mu\{g_n < \delta_n\} < 2\varepsilon_n$ and $\nu\{g_n > -\delta_n\} < 2\varepsilon_n$ if $\mu \in K_n$ and $\nu \in L_n$. Put $F_k = \cap_{n=k}^\infty \{g_n \geq \delta_n\}$ and $H_k = \cap_{n=k}^\infty \{g_n \leq -\delta_n\}$. Then $F_k \cap H_k = \phi$ and $\mu(F_k) \geq 1 - \sum_{n=k}^\infty \mu\{g_n < \delta_n\} > 1 - 2\sum_{n=k}^\infty \varepsilon_n$ for $\mu \in K_k$. Similarly $\nu(H_k) \geq 1 - 2\sum_{k=1}^\infty \varepsilon_n$ for $\mu \in L_k$. The sets $F = \cup_{k=1}^\infty F_k$, $H = \cup_{k=1}^\infty H_k$ are disjoint F_σ-sets such that $\mu(F) = 1 = \nu(H)$ for all $\mu \in M$, $\nu \in N$. Now assume that for each n there is a compact subset C_n of Y such that $\mu_n(B) < \varepsilon_n$ for all Baire sets with $B \cap C_n = \phi$ and all $\mu \in K_n$. Then we may in the definition of F_k and F replace $\{g_n \geq \delta_n\}$ by the compact set $\{g_n \geq \delta_n\} \cap C_n$ to get a K_σ-set $\tilde{F} \subset F$ with $\mu(\tilde{F}) = 1$ for all $\mu \in M$. Similarly uniform tightness of the L_n leads to a K_σ-set $\tilde{H} \subset H$ with $\nu(\tilde{H}) = 1$ for all $\nu \in N$.

COMMENT. The reader may notice a certain ambiguity in the last statement of the corollary: If the sets K_n, L_n are uniformly tight with respect to Baire measurable compact sets, then the K_σ-sets F and H can be made Baire measurable. Otherwise this need not be true and the equality $\mu(F) = 1 = \nu(H)$ refers to the Radon measure extension of the Baire measures μ, ν.

It is quite natural to ask the same questions for measurable measure convex sets K, L. The main difficulty is that in this situation only a weak minimax result is available. For its formulation we need the concept of a filter of countable type. The theory of these or rather closely related filters has been developed to some extent in [1], [13], [16].

DEFINITION 3.4. a) If $\mathscr{F}_1, \mathscr{F}_2, \cdots$ is a sequence of filters on a set I, the filter $\{\cup_{k=n}^\infty F_k : n \in \mathbb{N}, F_k \in \mathscr{F}_k\}$ is called the product of $\mathscr{F}_1, \mathscr{F}_2, \cdots$.

b) Let I be any (nonempty) set. A filter $\mathscr{F}$ on I is said to be of *countable type* if it belongs to the smallest family $\mathscr{I}$ of filters on I such that

(i) for each $i \in I$ the filter $\{F \subset I : i \in F\}$ belongs to $\mathscr{I}$, and
(ii) the product of each sequence of filters in $\mathscr{I}$ belongs to $\mathscr{I}$.

Recall that, if $\mathscr{F}$ is a filter on I and $(x_i)_{i\in I}$ is a family of real numbers, one defines $\limsup_{\mathscr{F}} x_i$ as $\inf_{F\in\mathscr{F}}\sup_{i\in F}x_i$. Similarly $\liminf_{\mathscr{F}}x_i$ and $\lim_{\mathscr{F}}x_i$ are defined. It is easily seen that $\limsup_{\mathscr{F}}x_i = \limsup_{n\to\infty}(\limsup_{\mathscr{F}_n} x_i)$ whenever $\mathscr{F}$ is the product of $\mathscr{F}_1, \mathscr{F}_2, \cdots$. Note also that each filter of countable type contains a countable set, hence instead of considering arbitrary index sets I, one often may restrict one's attention to the case $I = \mathbb{N}$. The importance of filters of countable type can be seen from the following remark.

REMARK. a) Let $\mathscr{C}$ be a set of functions on a set X, $\mathscr{B}$ the smallest family of functions containing $\mathscr{C}$ and closed under limits of sequences, $\mathscr{B}_c$ the family of all limits $\lim_{\mathscr{F}}f_i$, where $\mathscr{F}$ is a filter of countable type and $f_i \in \mathscr{C}$. Then $\mathscr{B} \subset \mathscr{B}_c$, and $\mathscr{B} = \mathscr{B}_c$ provided that $f, g \in \mathscr{C}$, $r \in \mathbb{R}$ implies that $\max(f, g)$, $\min(f, g)$, $\max(f, r)$, $\min(f, r) \in \mathscr{C}$.

b) Whenever $\mathscr{F}$ is a filter of countable type on a set I, $(\Omega, \mathscr{A}, P)$ is a probability space and $f_i(i \in I)$ are nonnegative measurable functions then $\liminf_{\mathscr{F}}f_i$ is measurable and $\int \liminf_{\mathscr{F}}f_i\, dP \le \liminf_{\mathscr{F}} \int f_i\, dP$.

PROOF. a) The inclusion $\mathscr{B} \subset \mathscr{B}_c$ is obvious since $\mathscr{B}_c$ is closed under limits of sequences. To prove $\mathscr{B}_c \subset \mathscr{B}$ under the conditions stated in a) first note that $\mathscr{B}$ is closed under lim sup and lim inf for uniformly bounded sequences in $\mathscr{B}$. This implies that the same property extends from sequences to all filters of countable type. If now $\mathscr{B}_c \ni g = \lim_{\mathscr{F}}f_i$ we have $g = \lim_{n\to\infty}\lim_{\mathscr{F}}\max(-n, \min(f_i, n)) \in \mathscr{B}$.

b) The set of all filters on I for which b) is true fulfills the conditions (i), (ii) by Fatou's lemma.

DEFINITION 3.5 (cf. [15]). A set M in a locally convex space is called *convexly analytic* if there is a map φ from a Polish space T into the set of nonempty compact subsets of M such that

i) (φ is onto) Every point of M is contained in some $\varphi(t)$, $t \in T$
ii) φ is upper semicontinuous, i.e. $\{t : \varphi(t) \subset V\}$ is open for every relatively open subset V of M,
iii) for every compact subset C of T there is a compact convex subset K of M such that $\varphi(t) \subset K$ for all t in C.

COMMENT. In this definition one may also take either $\mathbb{N}^{\mathbb{N}}$ or analytic spaces instead of the Polish space T. An analytic measure convex set is obviously convexly analytic but the converse is not true even for bounded sets.

For example, let M be the set of all probability measures μ on $\mathbb{R}$ for which $\int_{-\infty}^{+\infty} |x|\, d\mu(x) < \infty$ with the narrow topology. Then M is bounded but not measure convex. However M is convexly analytic: Let τ be the coarsest topology on M finer than the narrow one such that $\mu \mapsto \int |x|\, d\mu(x)$ is τ-continuous. Then for every τ-compact set C it follows $\sup_{\mu\in C} \int |x|\, d\mu(x) < \infty$ and hence every measure on C has its barycenter in M. From this one deduces that the map $\varphi : \mu \mapsto \{\mu\}$ defined on the Polish space $T = (M, \tau)$ has properties i), ii), iii), so M is convexly analytic.

For analytic finite dimensional sets, the three notions *convex, convexly analytic, measure convex* coincide since very finite dimensional convex set is automatically measure convex. On the other hand every compact convex metrizable infinite dimensional set contains a convex G_δ which is not measure convex [33]. The set constructed in that paper even fails to be convexly analytic.

We now give a weakened generalized minimax principle suitable for our purposes.

PROPOSITION 3.6. *Let $(G, \|\ \|)$ be a Banach space. For every pair of convexly analytic subsets M and N of $(G^*, \sigma(G^*, G))$ there are a family $(g_i)_{i\in I}$ of elements of the unit ball of G and a filter $\mathscr{F}$ of countable type on I such that*

$$\operatorname{dist}_{\|\ \|}(M, N) = \inf_{e\in N}\liminf_{\mathscr{F}}\langle e, g_i\rangle - \sup_{e'\in M}\limsup_{\mathscr{F}}\langle e', g_i\rangle.$$

PROOF. In fact we shall prove something more: Let φ and ψ be correspondences of $\mathbb{N}^{\mathbb{N}}$ onto N and M respectively with the properties described in the above definition. Then there is a filter $\mathscr{F}$ of countable type on the unit ball of G such that

$$\mathrm{dist}_{\| \|}(M, N) = \inf_{C\in\mathscr{C}}\lim\inf_{\mathscr{F}}\inf_{e\in\varphi(C)}\langle e, g\rangle - \sup_{C\in\mathscr{C}}\lim\sup_{\mathscr{F}}\sup_{e'\in\psi(C)}\langle e', g\rangle$$

where $\varphi(C) = \cup_{t\in C}\varphi(t)$ and $\mathscr{C}$ denotes the set of all compact subsets of $\mathbb{N}^{\mathbb{N}}$. (Here the index set I is the unit ball of G itself. This simplifies the proof.)

Let τ_ε for $\varepsilon \geq 0$ denote the set of all open subsets A of $\mathbb{N}^{\mathbb{N}}$ such that

$$\inf_{C\in\mathscr{C}(A)}\lim\inf_{\mathscr{F}}\inf_{e\in\varphi(C)}\langle e, g\rangle \geq \sup_{C\in\mathscr{C}(A)}\lim\sup_{\mathscr{F}}\sup_{e'\in\psi(C)}\langle e' g\rangle + \gamma_\varepsilon$$

for some filter $\mathscr{F}$ of countable type on the unit ball of G where $\mathscr{C}(A)$ is the set of compact subsets of A and $\gamma_\varepsilon = \mathrm{dist}_{\| \|}(M, N) - \varepsilon$.

We claim that $A \in \tau_\varepsilon$ whenever $\varepsilon_n \downarrow \varepsilon$, $A_n \uparrow A$ and $A_n \in \tau_{\varepsilon_n}$ for all $n = 1, 2, \cdots$. For each n choose a filter $\mathscr{F}_n$ associated with A_n in the described way. Passing to a suitable subsequence we may assume that $\lim_{n\to\infty}(\inf_{C\in\mathscr{C}(A_n)}\lim\inf_{\mathscr{F}_n}\inf_{e\in\varphi(C)}\langle e, g\rangle)$ exists. Let $\mathscr{F}$ be the product of the filters $\mathscr{F}_n$. Fix a compact subset C_0 of A. Since the A_n are open we have $C_0 \in \mathscr{C}(A_n)$ for all sufficiently large n and hence

$$\begin{aligned}\lim\inf_{\mathscr{F}}\inf_{e\in\varphi(C_0)}\langle e, g\rangle &= \lim\inf_{n\to\infty}\lim\inf_{\mathscr{F}_n}\inf_{e\in\varphi(C_0)}\langle e, g\rangle\\ &\geq \lim_{n\to\infty}\inf_{C\in\mathscr{C}(A_n)}\lim\inf_{\mathscr{F}_n}\inf_{e\in\varphi(C)}\langle e, g\rangle\\ &\geq \lim\sup_{n\to\infty}(\sup_{C\in\mathscr{C}(A_n)}\lim\sup_{\mathscr{F}_n}\sup_{e'\in\psi(C)}\langle e', g\rangle + \gamma_{\varepsilon_n})\\ &\geq \lim\sup_{n\to\infty}\lim\sup_{\mathscr{F}_n}\sup_{e'\in\psi(C_0)}\langle e', g\rangle + \gamma_\varepsilon\\ &= \lim\sup_{\mathscr{F}}\sup_{e'\in\psi(C_0)}\langle e', g\rangle + \gamma_\varepsilon.\end{aligned}$$

This proves $A \in \tau_\varepsilon$.

Assume now $\mathbb{N}^{\mathbb{N}} \notin \tau_\varepsilon$ for some $\varepsilon > 0$. Using what we just proved, a standard inductive procedure shows that there is a $\mathbf{n} = (n_1, n_2, \cdots)$ in $\mathbb{N}^{\mathbb{N}}$ such that none of the sets $A_k (k \in \mathbb{N})$ is in τ_ε where A_k is the open set $\{z \in \mathbb{N}^{\mathbb{N}}: z_i \leq n_i$ for all $i \leq k\}$. Let C be the compact set $\cap_{k=1}^{\infty} A_k$. Choose compact convex sets K, L such that $\varphi(C) \subseteq K \subseteq N$ and $\psi(C) \subseteq L \subseteq M$. According to the minimax lemma there are a $g \in G$ with $\|g\| = 1$ and real numbers c, d such that $c - d > \mathrm{dist}_{\| \|}(K, L) - \varepsilon \geq \gamma_\varepsilon$ and

$$\inf_{e\in K}\langle e, g\rangle > c > d > \sup_{e'\in L}\langle e', g\rangle.$$

The set $\{z \in \mathbb{N}^{\mathbb{N}}: \varphi(z) \subset \{\langle\cdot, g\rangle > c\}$ and $\psi(z) \subset \{\langle\cdot, g\rangle < d\}\}$ contains C and is open since φ and ψ are upper semicontinuous. Thus it contains A_k for some k in contradiction to $A_k \notin \tau_\varepsilon$.

So $\mathbb{N}^{\mathbb{N}} \in \tau_\varepsilon$ for every $\varepsilon > 0$. Letting ε tend to 0 we conclude $\mathbb{N}^{\mathbb{N}} \in \tau_0$, which completes the proof since the inequality "$\geq$" in the proposition is obvious for any filter $\mathscr{F}$ on $\{\|g\| \leq 1\}$.

COMMENTS. 1) In a similar way as we have passed from the minimax lemma to the above proposition, also other separation theorems can be shown to have noncompact analoga involving filters of countable type.

2) Part b) of the following theorem is essentially (i.e. for analytic measure convex sets) due to Mokobodzki (cf. [31]). We shall see in Section 5 that generally the set S in b) cannot be chosen to have the property of Baire even if Y is Polish. So the dependence of μ in part a) is not superfluous.

3) We recall that a map χ: $[0, 1]^{\mathbb{N}} \to [0, 1]$ is called a *medial limit for a filter* $\mathscr{F}$ on $\mathbb{N}$ if it is measure affine, i.e. universally measurable and $\chi(r(\mu)) = \int \chi(x)\mu(dz)$ for every Borel probability measure μ on $[0, 1]^{\mathbb{N}}$ where $r(\mu)$ is the barycenter of μ, and if $\lim\inf_{\mathscr{F}} z_\iota \leq \chi(z) \leq \lim\sup_{\mathscr{F}} z_i$ for all $z = (z_i) \in [0, 1]^{\mathbb{N}}$. If $\mathscr{F}$ is the Fréchet filter, one calls χ simply a *medial limit.* Medial limits are known to exist under CH [23] and even under Martin's Axiom. (An

introduction to these questions is given in [14].) One easily sees that, if a medial limit exists (for the Fréchet filter), then it exists for each filter of countable type.

THEOREM 3.7. *Under the general hypothesis of this section, let M and N be two convexly analytic subsets of $\mathscr{P}$ such that $\mu \perp \nu$ whenever $\mu \in M$, $\nu \in N$.*

a) *Let $\mathscr{G}^*$ be the σ-algebra over $\mathscr{P}$ generated by the evaluations $\mu \to \mu(g)$, $g \in G$. There is a set B in $\mathscr{G}^* \otimes \mathscr{B}$ such that $\mu(B_\mu) = 1$ and $\nu(B_\mu) = 0$ for all $\mu \in M$, $\nu \in N$.*

b) *If a medial limit χ: $[0, 1]^{\mathbb{N}} \to [0, 1]$ exists then there is a set S in the universal completion of $\mathscr{B}$ such that $\mu(S) = 1$ for all $\mu \in M$ and $\nu(S) = 0$ for all $\nu \in N$.*

PROOF. From the preceding proposition one deduces that there is a family $(g_i)_{i \in I}$ of measurable maps of Y into $[0, 1]$ and a filter $\mathscr{F}$ of countable type on I such that $\lim_{\mathscr{F}} \int g_i \, d\mu = 0$ for $\mu \in M$ and $\lim_{\mathscr{F}} \int g_i \, d\nu = 1$ for $\nu \in N$. (The function g_i is obtained from the one given by the proposition by first multiplying by ½ and then adding ½.) We may assume $I = \mathbb{N}$.

For the proof of b) it is sufficient to take $S = \{y \in Y: \chi((g_i(y))_{i \in \mathbb{N}}) = 0\}$ where χ is a medial limit for $\mathscr{F}$.

For the proof of a) put $B = \{(\mu, y): g(\mu, y) = 0\}$ where g is the function obtained from the following lemma applied to the space $(X, \mathscr{A}) = (\mathscr{P}, \mathscr{G}^*)$ and the identity on $\mathscr{P}$ considered as a transition kernel from $\mathscr{P}$ to Y.

LEMMA 3.8. *Let (μ_x) be a transition kernel between two measurable spaces $(X, \mathscr{A})$ and $(Y, \mathscr{B})$. Let $(g_i)_{i \in I}$ be a family of $\mathscr{A} \otimes \mathscr{B}$-measurable real functions such that for each x the family $(g_i(x, \cdot))_{i \in I}$ is uniformly bounded. Let $\mathscr{F}$ be a filter of countable type on I. Then there is a $\mathscr{A} \otimes \mathscr{B}$-measurable function g such that* $\inf g_i \leq g \leq \sup g_i$ *and for all x, $x' \in X$ we have*

$$\int g(x, y) \, d\mu_x(y) \leq \limsup_{\mathscr{F}} \int g_i(x, y) \, d\mu_x(y),$$

$$\int g(x, y) \, d\mu_{x'}(y) \geq \liminf_{\mathscr{F}} \int g_i(x, y) \, d\mu_{x'}(y).$$

PROOF. We may assume $I = \mathbb{N}$. First let us consider the case of sequences, i.e. let $\mathscr{F}$ be the Fréchet filter. The idea is just to take suitable convex combinations depending on x which converge μ_x – a.e. Since this choice has to be measurable in x we give a detailed proof.

Note that if C is a convex subset of a (pre-) Hilbert space H and $a = \inf\{\|z\|^2: z \in C\}$, we have for every pair (z, z') of C the estimate $\|z - z'\|^2 \leq 4 \max(\|z\|^2 - a, \|z'\|^2 - a)$. We are going to apply this to the bounded sets $C_n(x) = \operatorname{conv}\{g_i(x, \cdot): i \geq n\}$ in $\mathscr{L}^2(\mu_x)$. The numbers

$$a_n(x) = \inf\left\{\int h^2(y) \, d\mu_x(y): h \in C_n(x)\right\}$$

are increasing, uniformly bounded and measurable in x since it is sufficient to consider rational convex coefficients. Choose an increasing sequence of measurable functions n_r: $X \to \mathbb{N}$ such that $a_{n_{r'}}(x) - a_{n_r}(x) \leq \varepsilon_r$ for $r' > r$ where (ε_r) is some decreasing sequence of positive numbers with $\sum_{r=1}^{\infty} \varepsilon_r < \infty$. For fixed r choose (measurably) some $k_r(x) \geq n_r(x)$ and (rational) convex coefficients $(\alpha_r^k(x))_{n_r \leq k \leq k_r}$ such that $\|h_r(x, \cdot)\|_{2,\mu_x}^2 < a_{n_r}(x) + \varepsilon_r$ where $h_r(x, y) = \sum_{k=n_r(x)}^{k_r(x)} \alpha_r^k(x) g_k(x, y)$. Then from the geometric estimate given above we infer for $r' > r$

$$\|h_r(x, \cdot) - h_{r'}(x, \cdot)\|_{2,\mu_x}^2 \leq 4 \max(\varepsilon_r, a_{n_{r'}}(x) + \varepsilon_{r'} - a_{n_r}(x)) \leq 8 \, \varepsilon_r.$$

Thus $h_r(x, \cdot)$ converges μ_x – a.s. Let $g(x, y) = \limsup_{r\to\infty} h_r(x, y)$. Then by dominated convergence

$$\int g(x, y)\, d\mu_x(y) = \lim_{r\to\infty} \int h_r(x, y)\, d\mu_x(y) \le \limsup_{i\to\infty} \int g_i(x, y)\, d\mu_x(y)$$

and by Fatou's lemma

$$\int g(x, y)\, d\mu_{x'}(y) \ge \limsup_{r\to\infty} \int h_r(x, y)\, d\mu_{x'}(y) \ge \liminf_{i\to\infty} \int g_i(x, y)\, d\mu_{x'}(y).$$

Now the lemma has been proved for sequences. From this fact, it is easily seen that the set of all filters for which the lemma holds is closed under products of sequences. Hence it is proved for all filters of countable type.

We now want to establish the topological properties of the separating set.

THEOREM 3.9. *Let $(Y, \mathscr{B})$ be a completely regular space with the Baire σ-algebra and let G be the space $C_b(Y)$. Let M, N be as in Theorem 3.7. Then for each $\varepsilon > 0$ there is a set $F \in \mathscr{G}^* \otimes \mathscr{B}$ such that $\mu(F_\mu) > 1 - \varepsilon$, $\nu(F_\mu) = 0$ and F_μ is closed for every $\mu \in M$, $\nu \in N$. If Y is homeomorphic to a Baire subset of a compact space $\hat{Y}$ and $\mathscr{B}$ is the trace of the Baire σ-algebra of $\hat{Y}$, then the F_μ can be made compact.*

PROOF. We know that there is a $\mathscr{G}^* \otimes \mathscr{B}$ – measurable set B such that $\mu(B_\mu) = 1$ and $\nu(B_\mu) = 0$ for all $\mu \in M$, $\nu \in N$. Choose $F \in \mathscr{G}^* \otimes \mathscr{B}$ according to the following lemma, applied to $(X, \mathscr{A}) = (\mathscr{P}(Y), \mathscr{G}^*)$ and the "identity kernel." Then F has the required properties.

LEMMA 3.10. *Let $(\mu_x)_{x\in X}$ be a transition kernel from the measurable space $(X, \mathscr{A})$ to the completely regular space $(Y, \mathscr{B})$ with its Baire-σ-algebra. Let $B \in \mathscr{A} \otimes \mathscr{B}$ be such that $\mu_x(B_x) = 1$ for all x. Then for each $\varepsilon > 0$ there is a set $F \in \mathscr{A} \otimes \mathscr{B}$ such that F_x is a closed subset of B_x and $\mu_x(F_x) > 1 - \varepsilon$ for all $x \in X$. If Y is homeomorphic to a Baire subset of a compact space $\hat{Y}$ and $\mathscr{B}$ is the trace of the Baire σ-algebra of $\hat{Y}$, then the F_x can be made compact.*

PROOF. This lemma is a straightforward extension of Theorem 2.2 in [21].

4. Orthogonality preserving kernels. The following result collects a number of properties, most of which are known to be shared by all completely orthogonal kernels and which in fact are equivalent to the condition of being orthogonality preserving. For a kernel (μ_x) from X to Y and every $p \in \mathscr{P}(X)$ we denote by $\bar{\mu}^p$ the mixture $\int_X \varepsilon_x \otimes \mu_x\, dp$ on $X \times Y$ (generalized Fubini measure) and as before by μ^p the mixture $\int_X \mu_x\, dp$ on Y.

THEOREM 4.1. *Let X and Y be Borel subsets of Polish spaces. Let (μ_x) be a transition kernel from X to Y. Then the following conditions are equivalent:*

(i) *(μ_x) is orthogonality preserving.*

(ii) *The map r: $m \to \int_X \mu_x\, dm$ is a vector lattice isomorphism from the space $m(X)$ of signed measures on X to a linear subspace V of $m(Y)$ which is closed under the lattice operations in $m(Y)$.*

(iii) *For every $p \in \mathscr{P}(X)$ and every $B \in \mathscr{B}(X \times Y)$ there is some $C \in \mathscr{B}(Y)$ such that $\int_X \mu_x(B_x \Delta C) p(dx) = \bar{\mu}^p(B \,\Delta\, (X \times C)) = 0$.*

(iv) *For every $p \in \mathscr{P}(X)$ there is a Borel map $\varphi: Y \to X$ such that $p\{x: \mu_x(\varphi^{-1}\{x\}) = 1\} = 1$.*

(v) *For every $p \in \mathscr{P}(X)$ and every increasing sequence $(\mathscr{B}_n)$ of σ-algebras generating $\mathscr{B}(Y)$, $p\{x: \mu_x\{y: \phi_n^p(y, \cdot) \to_{n\to\infty} \varepsilon_x\} = 1\} = 1$ where $(\phi_n^p(y, \cdot))_{y\in Y}$ is a disintegration of $\bar{\mu}^p | \mathscr{B}(X) \otimes \mathscr{B}_n$ for the projection onto Y and the convergence is narrow convergence in $\mathscr{P}(X)$.*

(vi) *For every subset* X_0 *of* X *such that* $\{\mu_x\}_{x\in X_0}$ *is narrowly* σ*-compact, the kernel* $(\mu_x)_{x\in X_0}$ *is completely orthogonal.*
(vii) *Let* $X_0 \subseteq X$ *be a countable union of compact sets on each of which the map* $x \mapsto \mu_x$ *is continuous. Let* (Y_n) *be a sequence of spaces such that* $Y = \varprojlim_n Y_n$ *with corresponding projections* $\pi_n: Y \to Y_n$. *Then there are Borel maps* $\varphi_n: Y_n \to X_0$ *such that* $\mu_x\{y:\varphi_n(\pi_n(y)) \to_{n\to\infty} x\} = 1$ *for all* $x \in X_0$.
(viii) *There is a family* $(\varphi_i)_{i\in\mathbb{N}}$ *of Borel maps from* Y *to* X *and a filter* $\mathscr{F}$ *of countable type on* $\mathbb{N}$ *such that* $\lim_{\mathscr{F}}\mu_x\{y:\varphi_i(y) \notin U\} = 0$ *for all open sets* U *in* X *and all* $x \in U$.

Let us make some comments on these conditions. (The label of a comment coincides with the label of the corresponding condition in the theorem.)

(ii) Let e.g. $(\mu_x)_{x\in X}$ be the family of ergodic measures for a set $\mathscr{N}$ of kernels as in Proposition 1.2. Let V be the space of all $\mathscr{N}$-invariant signed measures. Then V is a vector lattice and—this is essential in (ii)—the lattice operations in V coincide with those induced from $\mathscr{m}(Y)$([4]). A Choquet type existence and uniqueness theorem (for a suitable non-compact version see e.g. [34]) then shows that the barycentric map r defines as isomorphism from $\mathscr{m}(X)$ to V. So here condition (ii) is easily verified by Choquet theory, whereas in this context even complete orthogonality is known by different arguments (see Proposition 1.2). We mention in [22], Problem 4 some kernels of *quasi*-ergodic measures which preserve orthogonality but for which complete orthogonality is still open.

(iii) A vague but perhaps more suggestive reformulation of (iii) is the formula $L^2(\mu^p) \cong \int_X^\oplus L^2(\mu_x)\,dp$. In fact the assignment $B \mapsto C$ induces canonically an isomorphism between $L^2(\bar{\mu}^p)$ and $L^2(\mu^p)$. On the other hand for any kernel (μ_x) and any $p \in \mathscr{P}(X)$ the space $L^2(\bar{\mu}^p)$ may be viewed as a concrete realization of the direct integral $\int_X^\oplus L^2(\mu_x)\,dp$ of the Hilbert spaces $L^2(\mu_x)$, $x \in X$. By the way, the equivalence of (i), (ii) and (iii) holds for all measurable spaces X and Y. (For (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) use the proofs given below whereas a direct proof of (i) $\Rightarrow$ (iii) is easy).

(iv) is just a nonsequential form of (v).

(v) states in statistical language that for any a priori distribution p on the parameter space X any sequence of a posteriori distribution concentrates for p-almost every parameter almost surely around this parameter. The implication (iv) $\Rightarrow$ (v) was an early success of martingale convergence (see [5], or, more accessible, [29], Section 3; consider the martingales $(x, y) \mapsto \phi_n^p(y, V)$ on $(X \times Y, \bar{\mu}^p)$ where V runs through a countable base of the topology of X).

(vi) The implication (i) $\Rightarrow$ (vi) had been proved independently in [8] and [32].

(vii) Condition (vii) does not imply $\mu_x\{y: Q_n(y)(f) \to \mu_x(f)\} = 1$ for all bounded *Borel* functions f on Y where $Q_n(y) = \mu_{\varphi_n(\pi_n(y))}$. Therefore, the sequence (Q_n) need not be an asymptotically H-sufficient statistic in the sense of [6], page 714, even if $X_0 = X$.

Proof. We prove (i) $\Rightarrow$ (vii) $\Rightarrow$ (vi) $\Rightarrow$ (iv) $\Rightarrow$ (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) and (i) $\Leftrightarrow$ (viii). The equivalence (iv) $\Leftrightarrow$ (v) has already been mentioned.

(i) $\Rightarrow$ (vii). Let d be a metric for the topology on X. As is easily seen, condition (vii) follows from the following statement which we are going to prove: If $x \mapsto \mu_x$ is continuous on some compact subset Z of X then for every $\varepsilon > 0$ there is some n and a Borel map φ_n: $Y_n \to Z$ such that $\mu_x\{y: d(x, \varphi_n(\pi_n(y))) \ge \varepsilon\} \le \varepsilon$ for all $x \in Z$.

For the proof of this, choose $z_1, \cdots, z_r$ in Z such that $Z = \bigcup_{i=1}^r S(z_i, \varepsilon/4)$ where $S(z, \delta)$ (resp. $\bar{S}(z, \delta)$) denotes the open (resp. closed) δ-neighbourhood of z in Z. For every i the sets $\{\mu^p: \operatorname{supp} p \subseteq \bar{S}(z_i, \varepsilon/4)\}$ and $\{\mu^q: \operatorname{supp} q \subseteq Z\backslash S(z_i, \varepsilon/2)\}$ are compact convex subsets of $\mathscr{P}(Y)$. Every element of the first is orthogonal to every element of the second since (μ_x) is orthogonality preserving. The subspace G of $C_b(Y)$ of all functions of the form $h \circ \pi_n (n \in \mathbb{N}, h \in C_b(Y_n))$ is norm-defining for $\mathscr{m}(Y)$. Our two sets of measures are also $\sigma(\mathscr{m}(Y), G)$-compact since this topology is Hausdorff and coarser than the narrow topology. Hence by Proposition 3.1, there is a subset V_i of Y of the form $\{h \circ \pi_{n_i} \ge 0\}$ such that $\mu_x(V_i) > 1 - \varepsilon/r$ for all x in $\bar{S}(z_i, \varepsilon/4)$ and $\mu_x(V_i) < \varepsilon/r$ for all x in $Z\backslash S(z_i, \varepsilon/2)$. Now define

ψ_i: $Y \to \{z_1, \cdots, z_r\}$ inductively by $\psi_1(y) = z_1$ and for $i > 1$, $\psi_i(y) = z_i$ on V_i and $\psi_i = \psi_{i-1}$ on $Y \setminus V_i$. Choose n sufficiently large and φ_n: $Y_n \to \{z_1, \cdots, z_r\}$ such that $\psi_r = \varphi_n \circ \pi_n$.

In order to show that φ_n does the job, fix x in Z. There is some m such that $x \in S(z_m, \varepsilon/4)$. If $d(z_m, \psi_r(y)) \geq (3/4)\varepsilon$ then either $\psi_m(y) \neq z_m$ and hence $y \notin V_m$, or $\psi_r(y) = z_i$ and hence $y \in V_i$ for some $i > m$ with $d(z_m, z_i) \geq (3/4)\varepsilon$. Therefore

$$\{y: d(x, \varphi_n(\pi_n(y))) \geq \varepsilon\} \subseteq \{y: d(z_m, \psi_r(y)) \geq (3/4)\varepsilon\}$$
$$\subseteq (Y \setminus V_m) \cup \cup \{V_i: i > m, d(z_m, z_i) \geq (3/4)\varepsilon\}.$$

But $d(z_m, z_i) \geq (3/4)\varepsilon$ implies $x \notin S(z_i, \varepsilon/2)$ and hence $\mu_x(V_i) < \varepsilon/r$. On the other hand $\mu_x(V_m) > 1 - \varepsilon/r$ and thus altogether $\mu_x\{y: d(x, \varphi_n(\pi_n(y))) \geq \varepsilon\} \leq \sum_{i=m}^{r} \varepsilon/r \leq \varepsilon$.

(vii) $\Rightarrow$ (vi). By identification of x with μ_x one gets a σ-compact topology on X_0 to which (vii) is applicable. The conclusion of (vii) implies in particular that $(\mu_x)_{x \in X_0}$ is completely orthogonal.

(vi) $\Rightarrow$ (iv) is clear by Lusin's theorem applied to the map $x \mapsto \mu_x$.

(iv) $\Rightarrow$ (iii). Let B in $\mathscr{B}(X \times Y)$ and the Borel map φ: $Y \to X$ be given such that $\mu_x(\varphi^{-1}\{x\}) = 1$ for all $x \in X_0$ where $p(X_0) = 1$. Let D be the set $\{(x, y): x \in X_0, x = \varphi(y)\}$ and define C to be $\pi_Y(D \cap B)$. The projection π_Y is injective on D, hence $C \in \mathscr{B}(Y)$. Also $(C \Delta B_x) \cap D_x = \varnothing$ and therefore $\mu_x(C \Delta B_x) = 0$ for all $x \in X_0$. This implies (iii).

(iii) $\Rightarrow$ (ii). Let V be the image of $\mathscr{m}(X)$ under the linear map r: $m \to \mu^m (= \int_X \mu_x dm)$. We shall see that $|\mu^m| = \mu^{|m|}$ for every $m \in \mathscr{m}(X)$. It follows then that r is a linear order isomorphism between $\mathscr{m}(X)$ and V and that V is closed under the lattice operations in $\mathscr{m}(Y)$, i.e. (ii) holds. Fix now m in $\mathscr{m}(X)$. Let A in $\mathscr{B}(X)$ satisfy $m^+(A) = 0$ and $m^-(X \setminus A) = 0$. According to (iii) there is a Borel subset C of Y such that $\mu_x(C \Delta (A \times Y)_x) = 0$ for all $x \in X_0$ where $|m|(X \setminus X_0) = 0$. Hence $\mu_x(C) = \mu_x((A \times Y)_x) = 1$ for $x \in X_0 \cap A$ and $\mu_x(C) = \mu_x((A \times Y)_x) = \mu_x(\varnothing) = 0$ for all $x \in X_0 \setminus A$. This implies $\mu^{m+}(C) = \int_{X_0 \setminus A} \mu_x(C)\, dm = 0$ and $\mu^{m-}(Y \setminus C) = \int_{A \cap X_0} \mu_x(Y \setminus C)\, dm = 0$. So μ^{m-} and μ^{m+} are orthogonal and hence $|\mu^m| = |\mu^{m+} - \mu^{m-}| = \mu^{m+} + \mu^{m-} = \mu^{|m|}$.

(ii) $\Rightarrow$ (i). Condition (ii) implies in particular $r(p) \wedge r(q) = 0$ (infimum in V and hence in $\mathscr{m}(Y)$!) whenever $p, q \in \mathscr{P}(X)$ and $p \wedge q = 0$. So the kernel is orthogonality preserving.

(i) $\Rightarrow$ (viii). Since there is a continuous surjection g: $N^N \to X$, we may assume for the proof of (viii) that $X = N^N$. For $k, l \in N$ let U_{kl} be the set $\{(x_n) \in N^N: x_k = l\}$. According to Proposition 3.6, condition (i) implies that there is a family $(g_i^{kl})_{i \in I^{kl}}$ and a filter $\mathscr{F}^{kl}$ of countable type on I^{kl} such that $\|g_i^{kl}\|_\infty \leq 1$ and $\lim_{\mathscr{F}^{kl}} \int_Y g_i^{kl}\, d\mu_x = 1$ for all $x \in U_{kl}$ and $\lim_{\mathscr{F}^{kl}} \int_Y g_i^{kl}\, d\mu_x = -1$ for all $x \notin U_{kl}$.

According to the lemma below we may assume $I^{kl} = N$ and $\mathscr{F}^{kl} = \mathscr{F}$ for all k, l where $\mathscr{F}$ is some filter of countable type on N. Let A_i^{kl} be the set $\{y: g_i^{kl}(y) > 0\}$. Then $\lim_{\mathscr{F}} \mu_x(A_i^{kl}) = \delta_{x_k l}$ (Kronecker symbol) for each $x = (x_n) \in N^N$. If we define φ_i: $Y \to N^N$ for each $i \in N$ by

$$(\varphi_i(y))_k = \begin{cases} \min\{l: y \in A_i^{kl}\} & \text{if } y \in \cup_{l \in N} A_i^{kl} \\ 1 & \text{otherwise} \end{cases}$$

it follows that $\lim_{\mathscr{F}} \mu_x\{y: (\varphi_i(y))_k = x_k\} = 1$ for all $k \in N$ and all $x \in N^N$. But this is just (viii) due to the definition of the topology in N^N.

(viii) $\Rightarrow$ (i) Let p and q be orthogonal measures on X. For every $\varepsilon > 0$ we may find a continuous function φ: $X \to [0, 1]$ such that $\int \varphi(x) p(dx) < \varepsilon^2$ and $\int (1 - \varphi(x)) q(dx) < \varepsilon^2$. Let $U = \{x \in X: \varphi(x) < \varepsilon\}$. Then $\limsup_{\mathscr{F}} \int_X \int_Y \varphi \circ \varphi_i(y) \mu_x(dy) p(dx) \leq p(X - U) + \varepsilon + \limsup_{\mathscr{F}} \int_U \mu_x\{y: \varphi_i(y) \notin U\} p(dx) < 2\varepsilon$. Similarly we see that $\limsup_{\mathscr{F}} \int_X \int_Y (1 - \varphi) \circ \varphi_i(y) \mu_x(dy) q(dx) \leq 2\varepsilon$, hence for some j we have $\int \varphi \circ \varphi_j(y) \mu^p(dy) < 2\varepsilon$ and $\int \varphi \circ \varphi_j(y) \mu^q(dy) > 1 - 2\varepsilon$. This shows that μ^p and μ^q are orthogonal.

LEMMA 4.2. *Let $(\mathscr{F}_m)_{m \in N}$ be a countable family of filters of countable type on corresponding sets I_m. Then there is a filter $\mathscr{F}$ of countable type on N and for each $m \in N$ a map ξ_m: $N \to I_m$ such $\{\xi_m(F): F \in \mathscr{F}\}$ is a base for $\mathscr{F}_m$* (i.e. *$\mathscr{F}_m$ is the image of $\mathscr{F}$ under ξ_m*).

PROOF. If $\mathscr{F}$ and $\mathscr{G}$ are filters of countable type on sets I and J respectively, we define $\mathscr{F} \times \mathscr{G}$ to be the filter on $I \times J$ with base $\{\bigcup_{k\in F}\{k\} \times G_k : F \in \mathscr{F}, G_k \in \mathscr{G}\}$. Then $\mathscr{F} \times \mathscr{G}$ is a filter of countable type since for fixed $\mathscr{G}$ the set $\{\mathscr{F} : \mathscr{F} \times \mathscr{G}$ is of countable type$\}$ is closed under products of sequences and contains the filters generated by singletons. Also, the images of $\mathscr{F} \times \mathscr{G}$ under the projections are $\mathscr{F}$ and $\mathscr{G}$ respectively.

Using this in the situation of the lemma we may define $\mathscr{G}_m$ on $I_1 \times \cdots \times I_m$ inductively by $\mathscr{G}_1 = \mathscr{F}_1$ and $\mathscr{G}_m = \mathscr{G}_{m-1} \times \mathscr{F}_m$. On $I = \bigcup_{m=1}^{\infty} I_1 \times \cdots \times I_m$ let $\mathscr{F}$ be the product of the sequence $(\{G : G \cap (I_1 \times \cdots \times I_m) \in \mathscr{G}_m\})_{m\in N}$ of filters of countable type. For each m define $\xi_m : I \to I_m$ by

$$\xi_m(i_1, \cdots, i_n) = \begin{cases} i_m & \text{if } n \geq m \\ i_m^* & \text{if } n < m \end{cases}$$

where i_m^* is some fixed element of I_m. Then one easily verifies that $\mathscr{F}_m$ is the image of $\mathscr{F}$ under ξ. Since we may assume I is countable, the same construction may be done on N.

The following proves part c) of Theorem 1.8.

THEOREM 4.3. *Let X and Y be standard measurable spaces. Any orthogonality preserving kernel (μ_x) from X to Y is uniformly orthogonal. If a medial limit exists then (μ_x) is also "universally measurably" completely orthogonal, i.e. the measures μ_x are separated by the fibers of a universally measurable map $\varphi : Y \to X$.*

PROOF. Let $(V_n)_{n\in N}$ be a sequence of subsets of X which generates $\mathscr{B}(X)$. Let M_n and N_n be defined as $\{\mu^p : p(V_n) = 1\}$ and $\{\mu^p : p(V_n) = 0\}$ respectively. Then M_n and N_n are measure convex analytic subsets of $\mathscr{P}(Y)$ so that every element of one is orthogonal to any element of the other. According to part a) of Theorem 3.7 there is a set B_n in $\mathscr{B}^*(M_n) \otimes \mathscr{B}(Y)$ such that $\mu_x(B_{n\mu_x}) = 1$ and $\mu_{x'}(B_{n\mu_x}) = 0$ whenever $x \in V_n$, $x' \notin V_n$. Let B be the Borel set $\bigcap_{n=1}^{\infty} \{(x, y) : x \notin V_n$ or $(\mu_x, y) \in B_n\}$. For every x we have $B_x = \bigcap \{B_{n\mu_x} : x \in V_n, n \geq 1\}$ and hence $\mu_x(B_x) = 1$. If $x \neq x'$ then there is an n such that $x \in V_n$ and $x' \notin V_n$ and hence $\mu_{x'}(B_x) \leq \mu_{x'}(B_{n\mu_x}) = 0$. This proves the first assertion. For the second assertion a similar argument applies using part b) of Theorem 3.7.

5. Counterexamples. Finally we show that the various (σ–) compactness conditions in the previous results cannot completely be omitted.

Recall that a set $A \subset Y$ is said to have the property of Baire if $(A - G) \cup (G - A)$ is a first category set for some open set $G \subset Y$. These sets form a σ-field, hence they contain the Borel sets; moreover they are closed under the Souslin operation [24] and hence they contain all sets belonging to the smallest σ-field containing Borel sets and closed under the Souslin operation.

The following lemma is the key to the counterexamples given in this section. The first part of the lemma is of a purely descriptive set theoretic nature and the second part is the measure theoretic tool which we will use. This lemma is closely related to one given by Blackwell [2] and was independently rediscovered by M. Talagrand and D. Preiss. It replaces our original arguments which concerned analytic sets [25].

LEMMA 5.1. *Let A be a subset of the topological product $Y = \prod_{i=1}^{\infty} Y_i$ of a sequence of finite spaces. Then*

a) *A is residual if and only if there is a $z \in Y$ and a sequence $1 = m(1) < m(2) < \cdots$ of integers such that $y \in A$ whenever the set*

$$\{k \in N : y_j = z_j \text{ for } j \in \{m(k), \cdots, m(k + 1) - 1\}\}$$

is infinite and contains 1.

b) *If A is residual then there is a real sequence (t_i) with $\lim_{i\to\infty} t_i = 0$ and a compact subset C of A such that for every $y^0 = (y_i^0) \in Y$ there is a probability measure μ satisfying $\mu(C) = 1$ and $\mu\{y : y_i \neq y_i^0\} \leq t_i$.*

PROOF. a) Let G be open such that A is comeager in G. Choose m and $z_1, \cdots z_m$ such that $G_1 = \{y \in Y: y_i = z_i \text{ for } i \le m\} \subset G$. Let $(G_k)_{k\ge1}$ be a decreasing sequence of dense open subsets of G_1 with $\cap_{k=1}^{\infty} G_k \subset A$. Define H_k for $k > m$ to be the set $\{y \in G_k$: for all $u_{m+1}, \cdots, u_k$ with $u_i \in Y_i$ the sequence $(z_1, \cdots, z_m, u_{m+1}, \cdots, u_k, y_{k+1}, y_{k+2}, \cdots)$ belongs to $G_k\}$. Then $G_k \supset H_k \supset H_{k+1}$ and each H_k is (a finite intersection of sets which are) open and dense in G_1. Thus there is a point z in $\cap_{k=1}^{\infty} H_k$. For each $r > m$ choose $p(r) \in N$ such that $p(r) > r$ and $\{y \in Y: y_i = z_i \text{ for } i < p(r)\} \subset H_r$.

We now define $m(n)$ by $m(1) = 1$, $m(2) = m + 1$ and $m(n+1) = p(m(n))$. Then z and the sequence $m(1), m(2), \cdots$ have the desired properties. In fact fix n for the moment and let y be such that $y_j = z_j$ for $j \le m$ and $m(n) \le j < m(n+1)$. Consider y' with $y'_j = y_j$ for $j \ge m(n)$ and $y'_j = z_j$ for $j < m(n)$. Then $y'_j = z_j$ even for all $j < p(m(n)) = m(n+1)$ and hence $y' \in H_{m(n)}$. Since y is of the form $(z_1, \cdots, z_m, u_{m+1}, \cdots, u_{m(n)}, y'_{m(n)+1}, y'_{m(n)+2}, \cdots)$ this implies $y \in G_{m(n)}$. Now if this holds for infinitely many n we conclude $y \in \cap_{k=1}^{\infty} G_k$ and hence $y \in A$.

Conversely, suppose the second condition in part a) holds. Let $G_k = \{y: y_j = z_j \text{ for } j = m(k), \cdots, m(k+1) - 1\}$. Then A contains the set $S = \cap_{n=1}^{\infty} (\cup_{k=n}^{\infty} G_k) \cap G_1$ which is a dense G_δ set in the clopen set G_1.

b) Let C be the set $\{y \in Y: y_j = z_j$ for $1 \le j < m(2) = m + 1$ and for each $r \in N$ there is some $i \in \{0, 1, \cdots, 2^r - 1\}$ such that $y_j = z_j$ for $m(2^r + i) \le j < m(2^r + i + 1)\}$. Then C is a compact subset of A according to part a). Let y^0 be any point in Y. The measure μ we are looking for is the image of Lebesgue measure under the map $g = (g_1, g_2, \cdots)$ of $[0, 1]$ into Y where $g_j(t) = z_j$ for $1 \le j \le m$, $g_j(t) = z_j$ for $m(2^r + i) \le j < m(2^r + i + 1)$ and $t \in [i2^{-r}, (i+1)2^{-r})(r \in N, i \in \{0, \cdots, 2^r - 1\})$ and $g_j(t) = y_j^0$ in all other cases. Then $g([0, 1]) \subset C$, hence $\mu(C) = 1$ and $\mu\{y: y_j \ne y_j^0\} \le \lambda[i2^{-r}, (i+1)2^{-r}] = t_j$ where $t_j = 2^{-r}$ for $m(2^r) \le j < m(2^{r+1})$. Clearly $t_j \to_{j\to\infty} 0$ and this sequence does not depend on y^0. This proves the lemma.

From Lemma 5.1, the following result can be obtained.

COROLLARY 5.2 (*Blackwell* [2]). *The two sets*

$$M = \{\mu \in \mathscr{P}(\{0, 1\}^N): \mu\{y_n = 0\} \to_{n\to\infty} 1\}, \quad N = \{\nu \in \mathscr{P}(\{0, 1\}^N): \nu\{y_n = 0\} \to_{n\to\infty} 0\}$$

cannot be separated by a set with the Baire property.

COMMENT. This corollary which gives two orthogonal measure convex Borel sets of probabilities which cannot be separated by a set with the property of Baire yields together with Mokobodzki's result (our Theorem 3.7(b)) the fact that in ZFC + MA there are universally measurable sets which do not have the property of Baire. In fact, we would know that some set theoretical assumption is needed to obtain Theorem 3.7 provided the answer to the following is "yes".

PROBLEM. Is there a model of ZFC in which every universally measurable set has the property of Baire?

THEOREM 5.3. *There is an atomless transition kernel* (μ_x) *from* $X = I$ *to a compact metric space* Y *such that*

a) *there is a sequence* (φ_n) *of continuous maps from* Y *to* X *such that* $\varphi_n \to x$ *in measure* μ_x *for each* x. *In particular* (μ_x) *is orthogonality preserving* (*compare* (viii) $\Rightarrow$ (i) *in Theorem* 4.1.)

b) (μ_x) *is not completely orthogonal*: *there is not even a map* $\varphi: Y \to X$ *such that* $\mu_x(\varphi^{-1}\{x\}) = 1$ *for all* $x \in X$ *and* $\varphi^{-1}(B)$ *has the property of Baire for all Borel sets* B *in* X.

PROOF. Let $X = [0, 1]$, $Y_l = \{i/2^l: i = 1, 2, \cdots, 2^l\}$ and $Y = \Pi_{l=1}^{\infty} Y_l$. To define the kernel, we introduce some other spaces. Let 2^Y be the space of all compact subsets of Y

with the Hausdorff topology and let S be the set of all positive sequences of real numbers with limit zero; S will have the topology of pointwise convergence. Denote by $\mathscr{C}$ the set of all pairs $(C, s) \in 2^Y \times S$ such that for each $x \in X$ there is some $\mu \in \mathscr{P}(Y)$ with $\mu(C) = 1$ and $\int |y_i - x| \, d\mu(y) \leq s_i$ for each $i \in \mathbb{N}$. Since $\mathscr{C}$ is a closed subset of $2^Y \times S$, there is a Borel measurable map γ of X onto $\mathscr{C}$. Finally, let $A = \{(x, C, s, \mu) \in X \times \mathscr{C} \times \mathscr{P}(Y): \mu(C) = 1$ and $\int |y_i - x| \, d\mu(y) \leq s_i$ for each $i \in \mathbb{N}\}$. The set A is closed in $X \times \mathscr{C} \times \mathscr{P}(Y)$ (recall that the set $\mathscr{P}(Y)$ is compact in the narrow topology) and the sections $A_{x,C,s}$ are compact, hence, according to a uniformization theorem [28], there is a Borel measurable map $(x, C, s) \to \nu_{x,C,s} \in A_{x,C,s}$. The kernel $(\mu_x)_{x \in X}$ will now be defined by $\mu_x = \nu_{x,\gamma(x)}$. This kernel satisfies a) since $y_i \to x$ in measure μ_x for each x.

Let $S^1, S^2 \subset S$ be disjoint Borel sets such that, for each $s \in S$, one can find $s^1 \in S^1$ and $s^2 \in S^2$ with $s_i \leq s_i^1$ and $s_i \leq s_i^2$ for each $i \in \mathbb{N}$. Put $B_k = \gamma^{-1}(2^Y \times S^k)$. We shall prove that the sets $\{\mu_x: x \in B_1\}$ and $\{\mu_x: x \in B_2\}$ are not separated by a set $A \subset Y$ with the property of Baire. To prove this, let $A \subset Y$ be a set with the property of Baire. Then, according to Lemma 5.1 b), there is a pair $(C, s) \in \mathscr{C}$ such that $C \subset A$ or $C \subset Y - A$. If we take $s^k \in S^k$ such that $s_i^k \geq s_i$, then $(C, s^k) \in \mathscr{C}(k = 1, 2)$. Hence $x^k = \gamma^{-1}(C, s^k) \in B_k$ and $\mu_{x^k}(C) = 1$. Consequently $\mu_{x^1}(A) = \mu_{x^2}(A)$, and A does not separate the sets $\{\mu_x: x \in B_1\}$ and $\{\mu_x: x \in B_2\}$. This proves b).

Starting from the example of the preceding theorem one may construct even topologically nicer examples.

THEOREM 5.4. *There is an atomless orthgonality preserving kernel* (μ_x) *between Polish spaces which satisfies Theorem* 5.3 b) *and moreover* c) *Y is locally compact and* d) *the map $p \mapsto \mu^p$ defines a topological isomorphism of $m^+(X)$ onto a closed subcone of $m^+(Y)$ in the narrow topology.*

PROOF. Let (ν_t) be any orthogonality preserving kernel between spaces X_0 and Y_0 which satisfies condition b) of Theorem 5.3. We may assume that Y_0 is a compact metric space. Find a closed subset X of $\mathbb{N}^{\mathbb{N}}$ and a one-to-one continuous map β of X onto $\{\nu_t: t \in X_0\}$. Let $\psi: \mathbb{N}^{\mathbb{N}} \mapsto [0, 1]$ be the homeomorphism (into) defined by continued fractions. Now take for Y the direct sum of Y_0 and $[0, +\infty)$ and define $\mu_x = \frac{1}{2}(\beta_x + \sum_{k=1}^{\infty} 2^{-k} \varepsilon_{\chi_k(x)})$ for $x \in X$ where $\chi_k(x) = \psi(x) + \sum_{i=1}^{k} x_i$.

Since $(\nu_t)_{t \in X_0}$ is orthogonality preserving and satisfies b) the same is true for $(\mu_x)_{x \in X}$. Further, c) is obvious. The maps χ_k are continuous. Thus, for every $f \in C_b(Y)$ the map $x \mapsto \int f \, d\mu_x = \frac{1}{2}(\int_{Y_0} f \, d\beta_x + \sum_{k=1}^{\infty} 2^{-k} f(\chi_k(x)))$ is continous, being the uniform limit of continous functions. Hence $x \mapsto \mu_x$ and $p \mapsto \mu^p$ are continuous for the narrow topology.

Finally let us show that $\{\mu^p: p \in m^+(X)\}$ is closed in $m^+(Y)$ and that $\mu^p \mapsto p$ is continuous. Suppose (p_n) is a sequence in $m^+(X)$ such that $\mu^{p_n} \to \mu$ for some $\mu \in m^+(Y)$. Then $\sup_n p_n(X) < \infty$ and the sequence (μ^{p_n}) is uniformly tight. Therefore for every $\varepsilon > 0$, $k \in \mathbb{N}$ there is some $M < \infty$ such that

$$p_n\{x: \textstyle\sum_{i=1}^{k} x_i > M\} \leq p_n\{x: \mu_x((M, +\infty)) > 2^{-k}\} \leq 2^k \mu^{p_n}((M, +\infty)) < \varepsilon.$$

This means that also the set $\{p_n\}$ is uniformly tight and hence relatively compact in $m^+(\mathbb{N}^{\mathbb{N}})$ since X is closed in $m^+(X)$. Now it is easy to see that $\mu = \mu^p$ for some $p \in m^+(X)$ with $p_n \to p$.

COROLLARY 5.5. *There is a locally compact second countable space Y and a pair of orthogonal closed convex subsets of $\mathscr{P}(Y)$ which cannot be separated by a subset S of Y with the Baire property.*

PROOF. Suppose in the preceding theorem that for every pair (X', X'') of disjoint closed subsets of X the two orthogonal closed sets $\{\mu^p: p(X') = 1\}$ and $\{\mu^q: q(X'') = 1\}$ could be separated by a subset of Y with the property of Baire. Then it would be easy to construct a map $\varphi: Y \mapsto X$ such that $\varphi^{-1}(B)$ has the property of Baire for all Borel sets B

and $\mu_x(\varphi^{-1}x) = 1$ for all $x \in X$ in contrast to condition b). Thus a pair of sets as in the corollary must exist.

COROLLARY 5.6. *Let (μ_x) be a completely orthogonal atomless transition kernel between uncountable Polish spaces X and Y. Then there is a Borel measurable measure convex subset N of $\mathscr{P}(Y)$ such that $\mu^p \perp \nu$, for all $p \in \mathscr{P}(X)$ and $\nu \in N$ but $\{\mu_x\}$ and N cannot be separated by a Borel subset of Y.*

PROOF. By the isomorphism theorem of Section 2 it is sufficient to exhibit one kernel (μ_x) with this property. One such construction follows.

Let $X = [0, 1]$ and $Y = [0, 1] \times \{0, 1\}^N$, $N = \{\nu \in \mathscr{P}(Y)\colon \int y_i\, d\nu(y) \to 1\}$, $M = \{\eta \in \mathscr{P}(Y)\colon \eta$ is atomless and $\int y_i\, d\eta(y) \to 0\}$. Let $x \to \eta_x$ be a Borel measurable map of [0, 1] onto M. For $x \in [0, 1]$ and $y = (y_0, y_1, \cdots) \in Y$ put $q_x(y) = (x, y_1, y_2, \cdots)$ and let $\mu_x = q_x\eta_x$. Assume that E is a Borel subset of Y such that $\nu(E) = 1$ for every $\nu \in N$. Let $E_1 = \{y \in Y\colon q_x y \in E$ for every $x \in [0, 1]\}$. Then E_1 is co-analytic, hence it is universally measurable and possesses the property of Baire. Moreover if $\nu \in N$ is such that $\nu(E - E_1) = 1$ then there is a ν-measurable map $T\colon E - E_1 \to Y - E$ such that $(Ty)_j = y_j$ for $j \geq 1$. Hence $T\nu \in N$ and $T\nu(T - E) = 1$, which is impossible. Hence $\nu E_1 = 1$ for every $\nu \in N$. Since each residual set supports some measure from N and since E_1 has the property of Baire, the set E_1 is residual. This implies that $\eta_x(E_1) = 1$ for some $x \in [0, 1]$, hence $\mu_x(E) = q_x\eta_x(E) = 1$.

COROLLARY 5.7. *There are a compact metric space Y and two Borel measure convex subsets M, N of $\mathscr{P}(Y)$ such that M is compact and $\mu \perp \nu$ for all $\mu \in M$, $\nu \in N$, but M and N cannot be separated by a Borel subset of Y.*

PROOF. Let X, Y, $(\mu_x)_{x\in X}$ and N be as constructed in Theorem 5.6. Our assumptions are invariant under the change of the topology on Y as long as we do not change the Borel structure. Hence by the isomorphism theorem in Section 2 we may even assume that $X = I$, $Y = I \times I$ and $\mu_x = \varepsilon_x \otimes \lambda$ for all x. The set $M = \{\int_X \mu_x\, dp\colon p \in \mathscr{P}(X)\} = \{p \otimes \lambda\colon p \in \mathscr{P}(I)\}$ is narrowly compact and convex. Thus the assertion follows from the preceding theorem.

The following is another example of the situation in the preceding corollary.

EXAMPLE 5.8. Let $Y = \{0, 1\}^{N^2}$, $M = \{\mu \in \mathscr{P}(Y)\colon \mu(y_{(n,k)} = 0) \leq 2^{-n}$, for all $n\}$ and $N = \{\nu \in \mathscr{P}(Y)\colon \lim_{n\to\infty} \liminf_{k\to\infty} \nu(y_{(n,k)} = 0) = 1\}$.
We omit the proof which is a modification of the proof of the last theorem.

We do not know the answer to the following questions:

1) Suppose that $(\mu_x)_{x\in X}$ and $(\nu_{x'})_{x'\in X'}$ are two completely orthogonal kernels such that $\mu^p \perp \nu^q$ for all mixing measures p, q. Can the two sets $\{\mu_x\}$ and $\{\nu_{x'}\}$ be separated by a Borel subset of Y?

2) Suppose $M \subset \mathscr{P}(Y)$ is compact and $N \subset \mathscr{P}(Y)$ closed such that both are convex and $\mu \perp \nu$ for all $\mu \in M$, $\nu \in N$. Can M and N be separated by a Borel subset of Y?

Acknowledgment. We are indebted to L. Louveau who pointed out an error in the first version of the paper.

REFERENCES

[1] ANISKOVIC, E. M. (1981). On subspaces of sequential spaces. *Sov. Math. Dokl.* **24** 202–205.
[2] BLACKWELL, D. (1980). There are no Borel SPLIFs. *Ann. Probab.* **8** 1189–1190.
[3] BURGESS, J. P. and MAULDIN, R. D. (1981). Conditional distribution and orthogonal measures. *Ann. Probab.* **9** 902–906.
[4] CHOQUET, G. (1969). *Lectures on Analysis, vol. II.* Benjamin, New York.
[5] DOOB, J. L. (1949). Application of the theory of martingales. *Coll. Int. du CNRS*, Paris 22–28.
[6] DYNKIN, E. B. (1978). Sufficient statistics and extreme points. *Ann. Probab.* **6** 705–730.

[7] Ershov, M. P. (1977). On the Rokhlin theorem on the independent complement (in Russian). *Uspekhi Mat. Nauk* **32** 1, 187–188.
[8] Farell, R. H. (1962). Representation of invariant measures. *Ill. J. Math.* **6** 447–467.
[9] Föllmer, H. (1975). Phase transition and Martin boundary. (Sém. de Prob. IX), *Springer Lecture Notes in Math.* **465** 305–317.
[10] Gardner, R. J. (1982). A note on conditional distributions and orthogonal measures. *Ann. Probab.* **10** 877–878.
[11] Goullet de Rugy (1971). Sur les mesures étrangères. *C. R. Acad. Sci. Paris* **272** 123–126.
[12] Graf, S. and Mägerl, G. (1981). Families of pairwise orthogonal measures. 9th Winter School in Abstract Analysis. Math. Inst. of the Cz. Acad. of Sciences, Praha 1981.
[13] Grimeisen, G. (1962). Ein Approximationsatz für Bairesche Funktionen. *Math. Ann.* **146** 189–194.
[14] Hoffman-Jørgensen, J. (1978). How to make a divergent series convergent by Martin's axiom. Aarhus Math. Inst. preprint series. Nr. 21.
[15] Holický, P. (1974). The convex generation of convex Borel sets in locally convex spaces. *Mathematika* **21** 207–215.
[16] Katetov, M. (1972). On descriptive classes of functions. *Theory of Sets and Topology*, edited by G. Asser, J. Flachsmeyer and W. Rinow, Deutscher Verlag der Wissenschaften, Berlin 265–278.
[17] Kerstan, J. and Wakolbinger, A. (1981). Ergodic decomposition of probability laws. *Z. Wahrsch. verw. Gebiete* **56** 399–414.
[18] Kraft, Ch. (1955). Some conditions for consistency and uniform consistency of statistical procedures. *Univ. of Calif., Publ. in Statistics* **1** 125–142.
[19] Lauritzen, S. (1974). Sufficiency prediction and extreme models. *Scand. J. Statist.* **1** 128–134.
[20] Maharam, D. (1982). Orthogonal measures: an example. *Ann. Probab.* **10** 879–880.
[21] Mauldin, R. D. (1979). Borel parametrizations. *Trans. Amer. Math. Soc.* **250** 223–234.
[22] Mauldin, R. D., Preiss, D. and v. Weizsäcker, H. (1981). A survey of problems and results concerning orthogonal transition kernels. Measure Theory Oberwolfach, pages 419–426, *Springer Lecture Notes in Math.* **945**.
[23] Meyer, P. A. (1973). Limites mediales, d'après Mokobodzki. (Sém de Prob. VII). *Springer Lecture Notes in Math* **321** 198–214.
[24] Nikodym, O. (1925). Sur une propiété de l'operation (A). *Fund. Math.* **7** 149–154.
[25] Preiss, D. (1981). Non separated sets of singular measures. 9th Winter School in Abstract Analysis. Math. Inst. of the Cz. Acad. of Sciences, Praha 1981.
[26] Rokhlin, V. A. (1949). On the fundamental ideas of measure theory. *Mat. Sbornik* **25** 107–150. (*Amer. Math. Soc. Trans.* **71** (1952)).
[27] Rosenblatt, M. (1959). Stationary processes as shifts of functions of independent random variables. *J. Math. Mech.* **8** 665–681.
[28] Saint-Raymond, J. (1976). Boreliens à coupes K_σ. *Bull. Soc. Math. France* **104** 389–400.
[29] Schwartz, L. (1965). On Bayes procedures. *Z. Wahrsch. verw. Gebiete.* **4** 10–26.
[30] Sion, M. (1958). On general minimax theorems. *Pacific J. Math.* **8** 171–176.
[31] Talagrand, M. (1981). Separation of orthogonal sets of measures (result of Mokobodzki). 9th Winter School in Abstract Analysis. Math. Inst. of the Cz. Acad. of Sciences, Praha 1981.
[32] v. Weizsäcker, H. (1980). Streng orthogonale Familien von W-Massen. Unpublished manuscript.
[33] v. Weizsäcker, H. (1976). A note on infinite dimensional convex sets. *Math. Scand.* **38** 321–324.
[34] v. Weizsäcker, H. and Winkler, G. (1979). Integral representations in the set of solutions of a generalized moment problem. *Math. Ann.* **246** 23–32.

R. D. Mauldin
Dept. of Mathematics
North Texas State University
Denton, Texas 76203

D. Preiss
Chair of Math. Analysis
Charles University
Praha 8
CSSR

H. v. Weizsäcker
Fachbereich Mathematik
Universität Kaiserslautern
6750 Kaiserslautern
West Germany