Contemporary Mathematics
Volume 94, 1989

One-to-one Selections and Orthogonal Transition Kernels

R. DANIEL MAULDIN*

ABSTRACT. Some questions concerning the existence of one-to-one selections and Borel structures are discussed. It is shown that the translates of Wiener measure form a completely orthogonal transition kernel and new examples of orthogonality preserving kernels which are not completely orthogonal are given.

In this paper, I will briefly discuss two problems to which Dorothy Maharam has made contributions by a combination of her results, talks and, not least, her mischievous innocent sounding coffee table questions. I will also raise some questions related to these problems.

1. One-to-one selections. The first problem concerns one-to-one selections. To put the problems in context, recall some of the history of selections or "uniformizations." In 1904, Hadamard, in a letter discussing Zermelo's axiom of choice, raised the selection problem [2]. In essence, he asked what sort of describable selector must exist for a given describable set. Lusin expanded on this theme [3] and discussed this issue in his famous monograph [4]. One of the early results concerning this problem is due to Lusin and Sierpinski:

THEOREM. *Let B be a Borel subset of $X \times Y$, where X and Y are Polish spaces. Then B can be uniformized by a coanalytic graph.*

In other words, there is a function $f\colon \mathrm{proj}_X(B) \to Y$ such that $\mathrm{Gr}(f) \subset B$. Of course, Kondo proved the ultimate theorem in this direction:

THEOREM. *Let B be a coanalytic subset of $X \times Y$, where X and Y are Polish spaces. Then B can be uniformized by a coanalytic graph.*

On the other hand, Novikov [9] showed that there is a Borel set which cannot be uniformized by a Borel set. So, as our first side issue we address

1980 *Mathematics Subject Classification.* 28B20, 54H05; Secondary 60A10. This paper is in final form and no version of it will be submitted for publication elsewhere.
*Supported in part by National Science Foundation Grant DMS 85-05923.

ourselves to the question of finding sufficient conditions for the existence of a Borel uniformization or selector.

Novikov showed that if each fiber of B, $B_x = \{y | (x, y) \in B\}$, is compact, then B has a Borel selector. Also, Novikov and Lusin showed that if each fiber of B is countable, then B possesses a Borel selector. Later these theorems were extended by Arsenin, Kunugui, Scegol′kov and ultimately Saint-Raymond **[10]**. Saint-Raymond proved the following beautiful

THEOREM. *If each B_x is σ-compact, then $B = \bigcup_{n=1}^{\infty} B_n$, where each B_n is a Borel set and, for each x, B_{nx} is compact.*

Saint-Raymond's theorem was in turn extended by A. Louveau. Of course, it follows from Saint-Raymond's theorem that if B is a Borel set and each B_x is *σ-compact*, then B has a Borel selector. Now, we modify and consider a set mentioned by Hadamard [2]. Let $H = \{(x, y) \in \mathbb{R}^2 |$ neither x nor y is related to the other by a polynomial equation with integer coefficients$\}$; i.e., y cannot be expressed as $y = \sum_{p=1}^{n} a_p x^p$ with integers a_p and likewise for x.

Of course, H is a G_δ subset of $\mathbb{R}^2$ and each horizontal and vertical fiber of H, being co-countable, is a dense G_δ set.

In order to determine whether H has a Borel selector, let us take a general approach. First, some definitions.

Let X and Y be uncountable Polish spaces and B a Borel subset of $X \times Y$ such that for each x, B_x is uncountable. A Borel parametrization of B is a Borel isomorphism, g, of $X \times E$ onto B where E is a Borel subset of Y such that for each x, $g(x, \cdot)$ maps E onto $B_x = \{y | (x, y) \in B\}$. A transition kernel is a map $x \mapsto \mu_x$ from X into the probability measures on Y such that for each Borel subset A of Y $x \mapsto \mu_x(A)$ is Borel measurable. In **[1]**, I proved the following:

THEOREM. *Let X and Y be uncountable Polish spaces and let B be a Borel subset of $X \times Y$ such that for each x, B_x is uncountable. The following are equivalent*

1. *B has a Borel parametrization.*
2. *There is a transition kernel $\{\mu_x\}_{x \in X}$ such that for each x, $\mu(x, B_x) = 1$ and, for each x, $\mu(x, \cdot)$ is atomless.*
3. *B contains a Borel set M such that for each x, M_x is a nonempty compact perfect set.*

It is easy to see that this theorem implies that H can be filled up by disjoint Borel uniformizations.

If Hadamard's question is slightly changed, we find unsolved problems even after all these years. For example, using the axiom of choice, one can prove the following:

THEOREM. *There is a one-to-one map F of $\mathbb{R} \times \mathbb{R}$ onto H such that for each y, f_y defined by $f_y(x) = \pi_Y(F(x, y))$ is a one-to-one map of $\mathbb{R}$ into $\mathbb{R}$.*

In other words, H can be filled up by the graphs of pairwise disjoint injections of $\mathbb{R}$ into $\mathbb{R}$.

QUESTION. Can the map F of the preceding theorem be taken to be Borel measurable?

What we really want are some reasonable sufficient conditions under which a Borel set can be filled up by one-to-one, or even, one-to-one and onto, Borel selectors.

Dorothy Maharam and A. H. Stone **[6]** have made some progress on this type of problem. For example, the next theorem is a special case of their theorem concerning the approximation of measurable functions by those that are also one-to-one.

THEOREM. *Let $X = Y = [0,1]$ and let $B \subset X \times Y$ be a Borel set, such that B_x has non-empty interior for every $x \in X$. Then there exists a Borel measurable one-to-one map $f: X \to Y$ with $f(x) \in B_x$ for every $x \in X$.*

This theorem remains true if X and Y are dense-in-themselves Polish spaces.

Also, Siegfried Graf and I **[1]** proved the following.

THEOREM. *Let B be a Borel subset of $X \times Y$, μ a probability measure on X and v a probability measure on Y. Suppose that for μ-a.e. x, B_x is uncountable and, for v-a.e. y, $B^y = \{x | (x,y) \in B\}$ is uncountable. Then there are Borel sets D and R and a Borel measurable isomorphism, f, of D onto R such that $\mu(D) = v(R) = 1$ and $\mathrm{Gr}(f)$ is a subset of B.*

On the other hand, Graf and I give the following example in **[1]**.

EXAMPLE. There is a Borel set $B \subset [0,1] \times [0,1]$ such that all fibers have positive measure and yet B does not possess a one-to-one Borel selection.

It is still possible that the following category version holds.

CONJECTURE. Let B be a Borel subset of $[0,1] \times [0,1]$ such that each horizontal and each vertical fiber is comeager (= complement of a first category set). Then B contains the graph of a Borel isomorphism. Perhaps B can be parametrized by Borel isomorphisms?

NOTE. During the preparation of this article, G. Debs and R. Saint-Raymond have signaled that first part of this conjecture is true at least in case all the fibers are dense G_δ sets (to appear Amer. J. Math.).

2. Orthogonal Transition Kernels. The second problem I want to discuss concerns the classification of the isomorphism classes of conditional probability distributions. We will say that two conditional distributions $x \mapsto \mu_x \in \mathrm{Pr}(Y)$, $x' \mapsto \mu_{x'} \in \mathrm{Pr}(Y')$ are isomorphic provided there are Borel isomorphisms $\varphi: X \to X'$ and $\psi: Y \to Y'$ such that $\mu_x(E) = \mu_{\varphi(x)}(\psi(E))$. We will focus on orthogonal distributions. Maharam **[5]** pointed out one type of behavior a family of orthogonal or mutually singular measures can exhibit.

DEFINITION. A family $\{\mu_\alpha\}$ of measures on a σ-field $\mathscr{B}$ on a space X is "*uniformly orthogonal*" means that for each α, there is some $H_\alpha \in \mathscr{B}$ such that $\mu_\alpha(X \backslash H_\alpha) = \mu_\beta(H_\beta) = 0$, if $\alpha \neq \beta$.

Maharam gave the following example, and as usual began asking questions.

THEOREM. *Assuming the continuum hypothesis, there is an uncountable family $\mathscr{M}$ of pairwise orthogonal Borel probability measures on the unit square such that no uncountable subset of $\mathscr{M}$ is uniformly orthogonal.*

One obvious question is: Can this phenomenon occur in a transition kernel? Also, what do the pairwise orthogonal transition kernels look like?

J. Burgess and I gave a negative answer to the first question. The answer is related to the second question. First a definition.

DEFINITION. A transition kernel $x \mapsto \mu_x \in \Pr(Y)$ is *completely orthogonal* means there is a Borel subset B of $X \times Y$ such that for each x, $\mu_x(B_x) = 1$ and if $x' \neq x$, then $B_x \cap B_{x'} = \varnothing$.

In other words, a transition kernel is completely orthogonal provided there is a Borel measurable map $f\colon Y \to X$ such that for each x, $\mu_x(f^{-1}(x)) = 1$.

Burgess and I proved that if μ_x is a pairwise orthogonal transition kernel, then there is a compact perfect subset K of X such that $\{\mu_x\}_{x \in K}$ is completely orthogonal. A prime example of a completely orthogonal transition kernel consists of the ergodic measures.

EXAMPLE. Let T be a continuous map of a compact metric space Y into itself. Let $X = \mathrm{Erg}(T)$, the set of all T-invariant ergodic probability measures on Y. Then X is a Polish space (provided with the usual weak*-topology) and the transition kernel $x \mapsto \mu_x$ given by the identity map on X is completely orthogonal.

PROOF. Let $\{V_n\}_{n=1}^\infty$ be a base for the topology of Y and let $\{U_n\}_{n=1}^\infty$ be an enumeration of the finite unions of elements of this base.

Let

$$B_n = \left\{ (\mu, y) | \lim_{p \to \infty} \frac{1}{p} \sum_{i=0}^{p-1} 1_{V_n}(T^i(y)) = \mu(V_n) \right\}$$

Clearly, each B_n is a Borel measurable and

$$\mu(B_{n\mu}) = 1.$$

Set

$$B = \bigcap B_n.$$

The Borel set B completely separates the transition kernel: if $\mu \neq \mu'$, then $B_\mu \cap B_{\mu'} = \phi$.

Weizsäcker and I give a second example arising in Brownian motion.

EXAMPLE. Let Y be $C_0[0, \infty)$ and let W be Wiener measure on Y. Let H be the group of autohomeomorphisms of $[0, \infty)$. For each $h \in H$, let

$$\mu_h = W \circ L_h^{-1},$$

where L_h maps Y onto Y by

$$L_h(f) = f \circ h.$$

Clearly, $h \mapsto \mu_h$ is a transition kernel. In order to see that this kernel is completely orthogonal we will use the mean square variation.

Let $\{q_n\}_{n=1}^{\infty}$ enumerate the positive rational numbers. For each n, let

$$B_n = \Big\{ (h, f) | \text{ if } \Pi_k \text{ is the partition of } [0, q_n] \text{ into } k \text{ congruent subintervals,}$$

$$\lim_{k\to\infty} \sum_{\pi_k} (f(t_i) - f(t_{i-1}))^2 = h(q_n) \Big\}.$$

Then, $B = \bigcap B_n$ is a Borel subset of $X \times Y$ which completely separates μ_x.

In [**8**], Weizsäcker, Preiss and I gave a classification of completely orthogonal kernels. One of the main facts is that there is essentially only one such kernel consisting of atomless measures. In particular, the kernels given in the preceding two examples are isomorphic. We now turn to the possibility of classifying kernels in terms of their functional analytic behaviour. It is easy to see that if μ_x is a completely orthogonal kernel, then the mixture map T determined by the kernel defines an *orthogonality preserving* or *lattice isomorphism* of $M(X)$, the Borel measures on X into $M(Y)$. The map T is defined by

$$T\tau(E) = \int_X \mu_x(E)\, d\tau(x).$$

Clearly, if $\tau \perp \sigma$, then $T(\tau) \perp T(\sigma)$. It is an intriguing fact that the property of being orthogonality preserving is not equivalent to being completely orthogonal and only one example is known.

PROBLEM. How many isomorphism classes of atomless orthogonality preserving kernels are there? In view of some of the results of [**8**], it could be that the answer depends on the axioms of set theory.

EXAMPLE. Let $Y = \prod_{n=1}^{\infty} Y_n$, where $Y_n = \{i/2^n | i = 1, 2, \ldots, 2^n\}$ and consider the spaces $\mathscr{K}(Y)$, of compact subsets of Y, and $\mathscr{P}r(Y)$, of probability measures on $\mathscr{B}(Y)$. We will need a generalized result of Blackwell in our construction [**8**, Lemma 5.1].

THEOREM. *Let A be a non-meager Baire property subset of Y. Then there is a sequence (t_i) of positive numbers converging to* 0 *and a compact subset K of A such that for every $y^0 = (y_i^0) \in Y$, there is a probability measure μ satisfying $\mu(K) = 1$ and $\mu(\{y: y_i \neq y_i^0\}) < t_i$.*

Now, for each n, let φ_n be the nth projection map of Y onto Y_n. Set

$$P = \Big\{ (K, s) \in \mathscr{K}(Y) \times c_0 | \forall x \in [0,1] \exists \mu \Big[\mu(K) = 1$$

$$\text{and } \forall i \int |\varphi_i - x|\, d\mu \le s_i \Big] \Big\}.$$

Let $\gamma(x) \mapsto (K_x, s_x, i_x)$ be a Borel isomorphism of $[0,1]$ onto $P \times \{0,1\}$ and let $x \mapsto \mu_x$ be a Borel map of $[0,1]$ into $\mathscr{P}_r(Y)$ such that for each x,

$$\mu_x(K_x) = 1$$

and

$$\varphi_n \to x \text{ in } \mu_x\text{-measure.}$$

According to Theorem 4.1 of **[8]**, $x \mapsto \mu_x$ is an orthogonality preserving kernel.

To see that μ_x is not completely orthogonal, set

$$G = \{x | i_x = 0\}.$$

Claim. There does not exist a set A with the Baire property (much less a Borel set) such that

$$\mu_x(A) = \begin{cases} 0 & x \in G \\ 1 & x \notin G \end{cases}.$$

Let us assume A is not meager and A has the Baire property. Then according to the preceding theorem, there are numbers x, y in $[0,1]$ such that $\gamma(x) = (K, s, 0)$ and $\gamma(y) = (K, s, 1)$. Thus, $x \in G$ and $y \notin G$. But,

$$\mu_x(K) = 1 = \mu_y(K).$$

A wide class of examples of orthogonality preserving kernels which are not completely orthogonal can be generated by modifying this procedure.

References

1. S. Graf and R. D. Mauldin, "Measurable one-to-one selections and transition kernels," American Journal of Mathematics, (2) **107** (1985), 407–425.
2. J. Hadamard, "Cinq lettres sur la theorie des ensembles," Societe Math. de France, Bulletin **33** (1905), 261–273.
3. N. Lusin, "Lecons sur les ensembles analytiques et leurs applications," Gauthier-Villars (1930), reprinted Chelsea, New York (1972).
4. N. Lusin, "Sur les ensembles analytiques," Fund. Math. **10** (1927), 1–95.
5. D. Maharam, "Orthogonal measures: an example," Annals of Probability **10** (1982), 879–880.
6. D. Maharam and A. Stone, "Expressing measurable functions by one-one ones," Advances Math., **46** (1982), 151–161.
7. R. D. Mauldin, "Borel Parametrizations," Transactions of the American Mathematical Society, **250** (1979), 223–234.
8. R. D. Mauldin, D. Preiss and H. Weizsäcker, "Orthogonal Transition Kernels," Annals of Probability **11** (1983), 970–988.
9. P. S. Novikov, "Sur les fonctions implicites mesurables B," Fund. Math. **17** (1931), 8–25.
10. J. Saint-Raymond, "Boréliens à coupes K_σ," Bull. Soc. Math. France **104** (1976), 389–400.
11. A. H. Stone, "Measure theory," Lecture Notes in Math., no. 541, Springer-Verlag, Berlin and New York, 1976, 43–48.