ON INVARIANT PROBABILITY MEASURES I

JULIUS RUBIN BLUM AND DAVID LEE HANSON
ON INVARIANT PROBABILITY MEASURES I

J. R. BLUM¹ AND D. L. HANSON²

1. Introduction. Let Ω be a set and let Χ be a σ-algebra of subsets of Ω. Let T be a one-to-one bimeasurable transformation mapping Ω onto itself. T then induces the group of transformations \{T^i, i = 0, ±1, \cdots\} defined in the usual way. If A ∈ Χ, T^iA is defined to be the set of images of the elements of A under the transformation T^i.

Let Ρ be the class of probability measures defined on Χ for which T is invariant, i.e. if P is a probability measure defined on Χ then P ∈ Ρ if and only if PA = PTA for every A ∈ Χ. Let Χ₁ be the subclass of Χ which is invariant under T; a set A ∈ Χ belongs to Χ₁ if and only if A = TA. It is trivial to verify that Χ₁ is sub-σ-algebra of Χ. Finally let Ρ₁ be the subclass of Ρ for which T is ergodic, i.e. if P ∈ Ρ then P ∈ Ρ₁ if and only if PA = 0 or PA = 1 for every A ∈ Χ₁.

In § 2, several results are proved, concerning the structure of the class Ρ. These are not new, although several of them do not seem to have appeared in the literature. The main theorem of this paper is in § 3 where it is shown that each element of Ρ can be represented as a convex combination of the extreme points of Ρ. Several consequences of this theorem are pointed out.

2. Some properties of the class Ρ.

THEOREM 1. Let P and Q be elements of Ρ. Suppose PA = QA for A ∈ Χ. Then P ≡ Q.

Proof. Let μ = P − Q. Then μ is a completely additive set function defined on Χ. If μ is not identically zero, there exists A ∈ Χ such μ(A) > 0 and μ(B) ≥ μ(B) for all B ∈ Χ. This follows from the Hahn decomposition theorem. Write μ(A) = α + β, where α = μ(A ∩ TA) and β = μ(A Π TA). Since μ(A ∩ TA) = μ(TA ∩ TA) we have μ(A Π TA) = 2α + β. Now if α < 0, then μ(A ∩ TA) > μ(A) and A is not maximal, and if β < 0 then μ(A ∩ TA) > μ(A) and A is not maximal. Consequently α ≥ 0 and β ≥ 0. If A is maximal then α + β ≥ 2α + β. Hence α = 0 and μ(A Π TA) = μ(A). By the same argument we show that μ(T⁻¹A Π A Π TA) = μ(A) and it follows by in-

Received December 9, 1959.
² Work done while the author was a Standard Oil Co. Fellow at Indiana University.
duction that $\mu(B_n) = \mu(A)$ for every positive integer $n$, where $B_n = \bigcup_{i=-n}^{n} T_iA$. Now $B_n$ is an increasing sequence of sets. Let $B = \lim_{n \to \infty} B_n$. Then $\mu(B) = \mu(A) > 0$. But clearly $B = \bigcup_{i=-\infty}^{\infty} T_iA \in \mathcal{A}$ and $\mu$ is zero on $\mathcal{A}$. Consequently we have a contradiction and the theorem is proved.

Suppose now that $P \in \mathcal{P}$ and $Q \in \mathcal{P}$ and suppose also that $Q$ is absolutely continuous with respect to $P$. Then if $A \in \mathcal{A}$ we have $PA = 0$ or $PA = 1$ and hence $Q$ agrees with $P$ on $\mathcal{A}$. Thus the theorem applies and we have

**Corollary 1.** If $P \in \mathcal{P}$, $Q \in \mathcal{P}$, and $Q$ is absolutely continuous with respect to $P$ then $Q \equiv P$.

Theorem 1 also furnishes an elegant proof of a result which was proved by Lamperti [3], and in a special situation by Harris [1]. Suppose $P$ and $Q$ are both ergodic, i.e. $P \in \mathcal{P}$ and $Q \in \mathcal{P}$. Then either $P$ and $Q$ are orthogonal or for each $A \in \mathcal{A}$ for which $PA = 1$ we have $Q(A) > 0$. Now suppose $A \in \mathcal{A}$ and $PA = 1$. Then if $Q$ is not orthogonal to $P$ and since $Q \in \mathcal{P}$ we must have $Q(A) = 1$ and it follows that $P = Q$ on $\mathcal{A}$. We have

**Corollary 2.** If $P \in \mathcal{P}$, $Q \in \mathcal{P}$, then either $P \equiv Q$ or $P$ is orthogonal to $Q$.

In § 3, we shall show that this result can be considerably generalized.

**Theorem 2.** $\mathcal{P}$ is a convex set. $P \in \mathcal{P}$ if and only if $P$ is an extreme point of $\mathcal{P}$.

**Proof.** The first statement is obvious. Suppose $P \in \mathcal{P}$ and suppose we may represent $P$ in the form $P = \alpha P_1 + (1 - \alpha) P_2$ where $0 < \alpha < 1$ and $P_i \in \mathcal{P}$, $i = 1, 2$. Then clearly $P_1$ and $P_2$ are absolutely continuous with respect to $P$ and it follows from Corollary 1 that $P_1 \equiv P_2 \equiv P$. Thus if $P \in \mathcal{P}$ it is an extreme point of $\mathcal{P}$. Conversely if $P \not\in \mathcal{P}$ there exists a set $B \in \mathcal{A}$ with $0 < PB < 1$. Then we may write $P = \alpha P_1 + (1 - \alpha) P_2$ where $\alpha = PB$, and for $A \in \mathcal{A}$ we have $P_1(A) = P(A \cap B)/P(B)$ and $P_2(A) = P(A \cap B^c)/P(B^c)$. It is easily verified that $P_1$ and $P_2$ are invariant probability measures and it follows that $P$ is not an extreme point of $\mathcal{P}$.

Theorem 2 strongly suggests that it may be possible to obtain the elements of $\mathcal{P}$ as convex combinations of the extreme points of $\mathcal{P}$. Under a rather mild assumption this is in fact true, as will be shown in the next section. Examples of the kind of theorem we have in mind were proved by Hewitt and Savage [2].

3. The representation theorem. Throughout part of this section we shall assume that if $A \in \mathcal{A}$ and if $PA = 0$ for every $P \in \mathcal{P}$ then
PA = 0 for every $P \in \mathcal{P}$. Clearly such a condition is necessary for a convex representation theorem and the condition can actually be verified in many examples of interest.

Suppose now that $P \in \mathcal{P}$. Theorem 1 tells us that $P$ has a unique invariant extension from $\mathcal{A}_1$ to $\mathcal{A}$. This suggests that if $A \in \mathcal{A}$ we should be able to determine $PA$ by knowing only the values of $P$ on $\mathcal{A}_1$. A proof of this statement follows from the individual ergodic theorem.

**Theorem 3.** Let $A \in \mathcal{A}$. For every $0 \leq \alpha \leq 1$ there exists a set $A'_\alpha \in \mathcal{A}_1$ such that if $P \in \mathcal{P}_1$ then $PA = \alpha$ if and only if $PA'_\alpha = 1$.

**Proof.** Let $f_\alpha(x)$ be the set characteristic function of the set $S$. Let $A \in \mathcal{A}$, and $\alpha$ be given. For every positive integer $n$ define $g_{n,\alpha}(x) = 1/n \sum_{i=1}^n f_\alpha(T^i x)$, and define $A'_\alpha = \{x \mid \lim_{n \to \infty} g_{n,\alpha}(x) = \alpha\}$. Clearly $A'_\alpha \in \mathcal{A}_1$ and the individual ergodic theorem implies that $PA = \alpha$ if and only if $PA'_\alpha = 1$, whenever $P \in \mathcal{P}_1$.

Using the same technique we can prove

**Theorem 4.** Let $A \in \mathcal{A}$. For every $0 \leq \alpha \leq 1$ there exists a set $A_\alpha \in \mathcal{A}_1$ such that if $P \in \mathcal{P}_1$ then $PA \leq \alpha$ if and only if $PA_\alpha = 1$.

Let $A \in \mathcal{A}$. Define $\pi_A$ by $\pi_A = \{P \in \mathcal{P} \mid PA = 1\}$. Let $\Pi$ be the collection of all such sets $\pi_A$ i.e. $\Pi = \{\pi_A \mid A \in \mathcal{A}\}$. The following facts are easily verified:

(i) $\pi_\Omega = \mathcal{P}_1$

(ii) $[\pi_A]^c = \pi^c$

(iii) $\pi \bigcup_n A_n = \bigcup_n \pi A_n$

where $A$ and each $A_n$ is an element of $\mathcal{A}$. Since $\mathcal{A}_1$ is a $\sigma$-algebra it follows that $\Pi$ is a $\sigma$-algebra. Now let $Q \in \mathcal{P}$. We define a set function $\mu_Q$ on $\Pi$ by $\mu_Q(\pi_A) = Q(A)$.

We shall show that under the assumption at the beginning of this section $\mu_Q$ is in fact a probability measure defined on $\Pi$. Clearly $\mu_Q(\pi_A) \geq 0$ for each $\pi_A$, and $\mu_Q(\mathcal{P}_1) = \mu_Q(\Omega) = Q(\Omega) = 1$. Now suppose $\{\pi_{A_n}\}$ is a sequence of disjoint elements of $\pi$. It is easily verified that this is the case if and only if $PA_n \cap A_m = 0$ for every pair of sets $A_n, A_m$ in $\mathcal{A}_1$ with $n \neq m$ and for every $P \in \mathcal{P}_1$. It follows from the assumption that $Q(A_n \cap A_m) = 0$ for $n \neq m$. Hence $\mu_Q(\bigcup_n \pi A_n) = Q(\bigcup_n A_n) = \sum_n Q(A_n) = \sum_n \mu_Q(\pi A_n)$ and we have shown that $\mu_Q$ is a probability measure defined on $\Pi$. We summarize in

**Theorem 5.** If $\Pi$ and $\mu_Q$ are defined as above then $\Pi$ is a $\sigma$-algebra of subsets of $\mathcal{P}_1$. Under the assumption at the beginning of this section $\mu_Q$ is a probability measure defined on $\Pi$. 
THEOREM 6. Let $A \in \mathcal{A}$. Consider the function $f_A(P)$ defined on $\mathcal{P}$ and with values $f_A(P) = PA$. Then $f_A(P)$ is measurable with respect to $\Pi$.

Proof. We must show that for every $\alpha$ with $0 \leq \alpha \leq 1$ we have $\{P \in \mathcal{P} \mid f_A(P) \leq \alpha\} = \{P \in \mathcal{P} \mid PA \leq \alpha\} \in \Pi$. But it follows from Theorem 4 that $\{P \in \mathcal{P} \mid PA \leq \alpha\} = \tau_{A - A_\alpha}$ where $A_\alpha \in \mathcal{A}_1$ is the set guaranteed by Theorem 4, and the theorem follows.

Since $f_A(P)$ is bounded and measurable it is clearly integrable with respect to any probability measure defined on $\Pi$. Now let $Q \in \mathcal{P}$ and $\mu_Q$ be the corresponding probability measure defined on $\Pi$. For each $A \in \mathcal{A}$ define $Q'(A)$ by

$$Q'(A) = \int f_A(P) d\mu_Q = \int PA d\mu_Q.$$  

It follows immediately from this definition that $Q'$ is an invariant probability measure defined on $\mathcal{A}$. But if $A \in \mathcal{A}$ we have $Q'(A) = \mu_Q[\tau_A] = Q(A)$. Hence $Q' = Q$ on $\mathcal{A}$ and it follows from Theorem 1 that $Q' \equiv Q$.

Furthermore suppose we know that $Q(A) = \int PA d\mu$, where $\mu$ is some probability measure defined on $\Pi$. Then if $A \in \mathcal{A}$ we have $Q(A) = \int PA d\mu = \mu[\tau_A] = \mu_Q[\tau_A]$, i.e. $\mu \equiv \mu_Q$. We state these results in

THEOREM 7. Suppose the assumption at the beginning of the section holds. Then for every $Q \in \mathcal{P}$ there exists a unique probability measure $\mu_Q$ defined on $\Pi$ such that

$$Q(A) = \int P(A) d\mu_Q \text{ for every } A \in \mathcal{A}.$$  

We shall refer to Theorem 7 as the representation theorem, and the rest of this section is devoted to exploring some consequences of this theorem. One immediate consequence is a generalization of Corollary 2 to Theorem 1.

THEOREM 8. Let $Q_i \in \mathcal{P}$, $i = 1, 2$. Then $Q_1$ and $Q_2$ are orthogonal if and only if the corresponding measured $\mu_{Q_1}$ and $\mu_{Q_2}$ are orthogonal.

Proof. Suppose $Q_1$ and $Q_2$ are orthogonal. Let $B$ be a set such that $Q_i(B) = 1 = Q_2(B^c)$ and let $A = \bigcup_{n=-\infty}^{\infty} T^nB$. Then $A \in \mathcal{A}$ and $Q_i(A) = 1 = Q_2(A^c)$ and we obtain $1 = \mu_{Q_1}[\tau_A] = \mu_{Q_2}[\tau_A]$. Thus $\mu_{Q_1}$ and $\mu_{Q_2}$ are orthogonal. Conversely if $\mu_{Q_1}$ and $\mu_{Q_2}$ are orthogonal there is a set $A \in \mathcal{A}$ such that $1 = \mu_{Q_1}[\tau_A] = Q_1(A)$ and $0 = \mu_{Q_2}[\tau_A] = Q_2(A)$ and the theorem is proved.
Another interesting consequence of the theorem is the obvious fact that if $A \in \mathcal{A}$ and if $PA = 1$ for each $P \in \mathcal{P}$, then $Q(A) = 1$ for each $Q \in \mathcal{P}$. Thus the individual ergodic theorem for arbitrary invariant measures is an immediate consequence of that theorem for ergodic measures. Furthermore Theorem 7 throws some light on the evaluation of the limiting function in the individual ergodic theorem. Let $Q \in \mathcal{P}$ and let $f(x)$ be defined on $\Omega$ and measurable with respect to $\mathcal{A}$. Let $f_n(x) = 1/n \sum_{i=1}^{n} f(T^i x)$. Then if $f \in L_1(Q)$ the ergodic theorem states that $\lim_{n \to \infty} f_n(x) = f^*(x)$ say, exists on a set of $Q$-measure one. It is clear that $f^*$ is invariant i.e. $f^*(Tx) = f^*(x)$ for all $x$ for which $f^*$ exists. If $f$ is also integrable with respect to $P \in \mathcal{P}$ then $f^*$ is constant on a set of $P$-measure one, and we have

$$Q\{x \mid f^*(x) \leq u\} = \int_{\mathcal{P}} P \{x \mid f^*(x) \leq u\} \, d\mu_Q = \mu_Q \{P \in \mathcal{P} \mid f^* \leq u\},$$

In particular we conclude $f^*$ is a constant, say $c$, on a set of $Q$-measure one if and only if $\mu_P \{P \in \mathcal{P} \mid P\{x \mid f^*(x) = c\}\} = 1$.

Finally, suppose $f$ is again measurable with respect to $\mathcal{A}$. Let $Q \in \mathcal{P}$ and suppose $\mu_Q \{P \in \mathcal{P} \mid \int_{\Omega} |f| \, dP < \infty\} = 1$. Then we can easily prove

**Theorem 8.** If $\int_{\Omega} |f| \, dP$ is an integrable function of $P$ (with respect to $\mu_Q$) then $f \in L_1(Q)$ and

$$\int_{\Omega} f \, dQ = \int_{\mathcal{P}} \left[ \int_{\Omega} f \, dP \right] \, d\mu_Q.$$

**References**

Mathematical papers intended for publication in the Pacific Journal of Mathematics should be typewritten (double spaced), and the author should keep a complete copy. Manuscripts may be sent to any one of the four editors. All other communications to the editors should be addressed to the managing editor, L. J. Paige at the University of California, Los Angeles 24, California.

50 reprints per author of each article are furnished free of charge; additional copies may be obtained at cost in multiples of 50.

The Pacific Journal of Mathematics is published quarterly, in March, June, September, and December. The price per volume (4 numbers) is $12.00; single issues, $3.50. Back numbers are available. Special price to individual faculty members of supporting institutions and to individual members of the American Mathematical Society: $4.00 per volume; single issues, $1.25.

Subscriptions, orders for back numbers, and changes of address should be sent to Pacific Journal of Mathematics, 2120 Oxford Street, Berkeley 4, California.

Printed at Kokusai Bunken Insatsusha (International Academic Printing Co., Ltd.), No. 6, 2-chome, Fujimi-cho, Chiyoda-ku, Tokyo, Japan.

PUBLISHED BY PACIFIC JOURNAL OF MATHEMATICS, A NON-PROFIT CORPORATION
The Supporting Institutions listed above contribute to the cost of publication of this Journal, but they are not owners or publishers and have no responsibility for its content or policies.
M. Altman, *An optimum cubically convergent iterative method of inverting a linear bounded operator in Hilbert space* ................................................................. 1107
Nesmith Cornett Ankeny, *Criterion for rth power residuacity* .................................................. 1115
Julius Rubin Blum and David Lee Hanson, *On invariant probability measures I* .................. 1125
Frank Featherstone Bonsall, *Positive operators compact in an auxiliary topology* ............... 1131
Billy Joe Boyer, *Summability of derived conjugate series* .................................................. 1139
Delmar L. Boyer, *A note on a problem of Fuchs* ................................................................. 1147
Hans-Joachim Bremermann, *The envelopes of holomorphy of tube domains in infinite dimensional Banach spaces* ............................................................... 1149
Andrew Michael Bruckner, *Minimal superadditive extensions of superadditive functions* ........................................................................................................................................ 1155
Billy Finney Bryant, *On expansive homeomorphisms* .......................................................... 1163
Jean W. Butler, *On complete and independent sets of operations in finite algebras* ............ 1169
Lucien Le Cam, *An approximation theorem for the Poisson binomial distribution* ............... 1181
Paul Civin, *Involutions on locally compact rings* ............................................................... 1199
Earl A. Coddington, *Normal extensions of formally normal operators* .................................. 1203
Jacob Feldman, *Some classes of equivalent Gaussian processes on an interval* .................. 1211
Shaull Foguel, *Weak and strong convergence for Markov processes* .................................... 1221
Martin Fox, *Some zero sum two-person games with moves in the unit interval* .................... 1235
Robert Pertsch Gilbert, *Singularities of three-dimensional harmonic functions* .................. 1243
Branko Grünbaum, *Partitions of mass-distributions and of convex bodies by hyperplanes* ...................................................................................................................................... 1257
Sidney Morris Harmon, *Regular covering surfaces of Riemann surfaces* ............................ 1263
Edwin Hewitt and Herbert S. Zuckerman, *The multiplicative semigroup of integers modulo m* .................................................................................................................... 1291
Paul Daniel Hill, *Relation of a direct limit group to associated vector groups* ...................... 1309
Calvin Virgil Holmes, *Commutator groups of monomial groups* ........................................ 1313
James Fredrik Jakobsen and W. R. Utz, *The non-existence of expansive homeomorphisms on a closed 2-cell* ................................................................................... 1319
John William Jewett, *Multiplication on classes of pseudo-analytic functions* ....................... 1323
Helmut Klingen, *Analytic automorphisms of bounded symmetric complex domains* .............. 1327
Robert Jacob Koch, *Ordered semigroups in partially ordered semigroups*.............................. 1333
Marvin David Marcus and N. A. Khan, *On a commutator result of Taussky and Zassenhaus* ....................................................................................................................................... 1337
John Glen Marica and Steve Jerome Bryant, *Unary algebras* .................................................. 1347
Edward Peter Merkes and W. T. Scott, *On univalence of a continued fraction* ....................... 1361
Shu-Teh Chen Moy, *Asymptotic properties of derivatives of stationary measures* .................. 1371
John William Neuberger, *Concerning boundary value problems* .......................................... 1385
Marian Reichaw-Reichbach, *Some theorems on mappings onto* ........................................... 1397
Marvin Rosenblum and Harold Widom, *Two extremal problems* ......................................... 1409
Morton Lincoln Slater and Herbert S. Wilf, *A class of linear differential-difference equations* ...................................................................................................................................... 1419
Charles Robson Storey, Jr., *The structure of threads* ............................................................ 1429
J. François Treves, *An estimate for differential polynomials in \( \partial / \partial z_1, \ldots, \partial / \partial z_n \)* 1447
J. D. Weston, *On the representation of operators by convolutions integrals* ......................... 1453
James Victor Whittaker, *Normal subgroups of some homeomorphism groups* ...................... 1469