

Pacific Journal of Mathematics

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 \mathcal{A} 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 $\{T^i, i = 0, \pm 1, \dots\}$ defined in the usual way. If $A \in \mathcal{A}$, $T^i A$ is defined to be the set of images of the elements of A under the transformation T^i .

Let \mathcal{P} be the class of probability measures defined on \mathcal{A} for which T is invariant, i.e. if P is a probability measure defined on \mathcal{A} then $P \in \mathcal{P}$ if and only if $PA = PTA$ for every $A \in \mathcal{A}$. Let \mathcal{A}_1 be the subclass of \mathcal{A} which is invariant under T ; a set $A \in \mathcal{A}$ belongs to \mathcal{A}_1 if and only if $A = TA$. It is trivial to verify that \mathcal{A}_1 is sub- σ -algebra of \mathcal{A} . Finally let \mathcal{P}_1 be the subclass of \mathcal{P} for which T is ergodic, i.e. if $P \in \mathcal{P}$ then $P \in \mathcal{P}_1$ if and only if $PA = 0$ or $PA = 1$ for every $A \in \mathcal{A}_1$.

In § 2. several results are proved, concerning the structure of the class \mathcal{P} . 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 \mathcal{P} can be represented as a convex combination of the extreme points of \mathcal{P} . Several consequences of this theorem are pointed out.

2. Some properties of the class \mathcal{P} .

THEOREM 1. *Let P and Q be elements of \mathcal{P} . Suppose $PA = QA$ for $A \in \mathcal{A}_1$. Then $P \equiv Q$.*

Proof. Let $\mu = P - Q$. Then μ is a completely additive set function defined on \mathcal{A} . If μ is not identically zero, there exists $A \in \mathcal{A}$ such $\mu(A) > 0$ and $\mu(A) \geq \mu(B)$ for all $B \in \mathcal{A}$. This follows from the Hahn decomposition theorem. Write $\mu(A) = \alpha + \beta$, where $\alpha = \mu(A - A \cap TA)$ and $\beta = \mu(A \cap TA)$. Since $\mu(A - A \cap TA) = \mu(TA - A \cap TA)$ we have $\mu(A \cup TA) = 2\alpha + \beta$. Now if $\alpha < 0$, then $\mu(A \cap TA) > \mu(A)$ and A is not maximal, and if $\beta < 0$ then $\mu(A - A \cap TA) > \mu(A)$ and A is not maximal. Consequently $\alpha \geq 0$ and $\beta \geq 0$. But if A is maximal then $\alpha + \beta \geq 2\alpha + \beta$. Hence $\alpha = 0$ and $\mu(A \cup TA) = \mu(A)$. By the same argument we show that $\mu(T^{-1}A \cup A \cup TA) = \mu(A)$ and it follows by in-

Received December 9, 1959.

¹ Supported in part by the Office of Ordnance Research, U. S. Army, under Contract No. DA-33-008-ORD-965.

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 $B_n = \bigcup_{i=-n}^n T^i A$. Now B_n is an increasing sequence of sets. Let $B = \lim_{n \rightarrow \infty} B_n$. Then $\mu(B) = \mu(A) > 0$. But clearly $B = \bigcup_{i=-\infty}^{\infty} T^i A \in \mathcal{A}_1$ and μ is zero on \mathcal{A}_1 . Consequently we have a contradiction and the theorem is proved.

Suppose now that $P \in \mathcal{P}_1$ and $Q \in \mathcal{P}$ and suppose also that Q is absolutely continuous with respect to P . Then if $A \in \mathcal{A}_1$ we have $PA = 0$ or $PA = 1$ and hence Q agrees with P on \mathcal{A}_1 . Thus the theorem applies and we have

COROLLARY 1. *If $P \in \mathcal{P}_1, Q \in \mathcal{P}$, 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. $P \in \mathcal{P}_1$ and $Q \in \mathcal{P}_1$. 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 $A \in \mathcal{A}_1$ and $PA = 1$. Then if Q is not orthogonal to P and since $Q \in \mathcal{P}_1$ we must have $Q(A) = 1$ and it follows that $P = Q$ on \mathcal{A}_1 . We have

COROLLARY 2. *If $P \in \mathcal{P}_1, Q \in \mathcal{P}_1$, 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. $P \in \mathcal{P}_1$ if and only if P is an extreme point of \mathcal{P} .*

Proof. The first statement is obvious. Suppose $P \in \mathcal{P}_1$ and suppose we may represent P in the form $P \equiv \alpha P_1 + (1 - \alpha)P_2$ 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 $P \in \mathcal{P}_1$ it is an extreme point of \mathcal{P} . Conversely if $P \notin \mathcal{P}_1$ there exists a set $B \in \mathcal{A}_1$ with $0 < PB < 1$. Then we may write $P \equiv \alpha P_1 + (1 - \alpha)P_2$ 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}_1 . 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 $A \in \mathcal{A}_1$ and if $PA = 0$ for every $P \in \mathcal{P}_1$ 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 $P \in \mathcal{P}_1$. 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 α with $0 \leq \alpha \leq 1$ 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 $f_s(x)$ be the set characteristic function of the set S . Let $A \in \mathcal{A}$, and α be given. For every positive integer n define $g_{n,A}(x) = 1/n \sum_{i=1}^{n-1} f_A(T^i x)$, and define $A'_\alpha = \{x \mid \lim_{n \rightarrow \infty} g_{n,A}(x) = \alpha\}$. 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 α with $0 \leq \alpha \leq 1$ 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 $A \in \mathcal{A}_1$. Define π_A by $\pi_A = \{P \in \mathcal{P}_1 \mid PA = 1\}$. Let Π be the collection of all such sets π_A i.e. $\Pi = \{\pi_A \mid A \in \mathcal{A}_1\}$. 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}_1 . Since \mathcal{A}_1 is a σ -algebra it follows that Π is a σ -algebra. Now let $Q \in \mathcal{P}$. We define a set function μ_Q on Π by $\mu_Q(\pi_A) = Q(A)$.

We shall show that under the assumption at the beginning of this section μ_Q is in fact a probability measure defined on Π . Clearly $\mu_Q(\pi_A) \geq 0$ for each π_A , and $\mu_Q(\mathcal{P}_1) = \mu_Q(\pi_\Omega) = Q(\Omega) = 1$. Now suppose $\{\pi_{A_n}\}$ is a sequence of disjoint elements of π . It is easily verified that this is the case if and only if $PA_n \cap PA_m = 0$ 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 μ_Q is a probability measure defined on Π . We summarize in

THEOREM 5. *If Π and μ_Q are defined as above then Π is a σ -algebra of subsets of \mathcal{P}_1 . Under the assumption at the beginning of this section μ_Q is a probability measure defined on Π .*

THEOREM 6. *Let $A \in \mathcal{A}$. Consider the function $f_A(P)$ defined on \mathcal{S}_1 and with values $f_A(P) = PA$. Then $f_A(P)$ is measurable with respect to Π .*

Proof. We must show that for every α with $0 \leq \alpha \leq 1$ we have $\{P \in \mathcal{S}_1 | f_A(P) \leq \alpha\} = \{P \in \mathcal{S}_1 | PA \leq \alpha\} \in \Pi$. But it follows from Theorem 4 that $\{P \in \mathcal{S}_1 | PA \leq \alpha\} = \pi_{A \leftarrow A_\alpha}$ where $A_\alpha \in \mathcal{A}$ 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 Π . Now let $Q \in \mathcal{S}$ and μ_Q be the corresponding probability measure defined on Π . For each $A \in \mathcal{A}$ define $Q'(A)$ by

$$Q'(A) = \int_{\mathcal{S}_1} f_A(P) d\mu_Q = \int_{\mathcal{S}_1} PAd\mu_Q.$$

It follows immediately from this definition that Q' is an invariant probability measure defined on \mathcal{A} . But if $A \in \mathcal{A}_1$ we have $Q'(A) = \mu_Q\{\pi_A\} = Q(A)$. Hence $Q' = Q$ on \mathcal{A}_1 and it follows from Theorem 1 that $Q' \equiv Q$. Furthermore suppose we know that $Q(A) = \int_{\mathcal{S}_1} PAd\mu$, where μ is some probability measure defined on Π . Then if $A \in \mathcal{A}_1$ we have $Q(A) = \int_{\mathcal{S}_1} PAd\mu = \mu\{\pi_A\} = \mu_Q\{\pi_A\}$, 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 $Q \in \mathcal{S}$ there exists a unique probability measure μ_Q defined on Π such that*

$$Q(A) = \int_{\mathcal{S}_1} P(A) d\mu_Q \text{ for every } A \in \mathcal{A}.$$

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 $Q_i \in \mathcal{S}$, $i = 1, 2$. Then Q_1 and Q_2 are orthogonal if and only if the corresponding measures μ_{Q_1} and μ_{Q_2} are orthogonal.*

Proof. Suppose Q_1 and Q_2 are orthogonal. Let B be a set such that $Q_1(B) = 1 = Q_2(B^c)$ and let $A = \bigcup_{i=-\infty}^{\infty} T^i B$. Then $A \in \mathcal{A}_1$ and $Q_1(A) = 1 = Q_2(A^c)$ and we obtain $1 = \mu_{Q_1}\{\pi_A\} = \mu_{Q_2}\{(\pi_A)^c\}$. Thus μ_{Q_1} and μ_{Q_2} are orthogonal. Conversely if μ_{Q_1} and μ_{Q_2} are orthogonal there is a set $A \in \mathcal{A}_1$ such that $1 = \mu_{Q_1}\{\pi_A\} = Q_1(A)$ and $0 = \mu_{Q_2}\{\pi_A\} = Q_2(A)$ 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 $P \in \mathcal{P}_1$ 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 Ω and measurable with respect to \mathcal{A} . Let $f_n(x) = 1/n \sum_{i=0}^{n-1} f(T^i x)$. Then if $f \in L_1(Q)$ the ergodic theorem states that $\lim_{n \rightarrow \infty} f_n(x) = f^*(x)$ 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 $P \in \mathcal{P}_1$ then f^* is constant on a set of P -measure one, and we have

$$Q\{x | f^*(x) \leq u\} = \int_{\mathcal{P}_1} P\{x | f^*(x) \leq u\} d\mu_Q = \mu_Q\{P \in \mathcal{P}_1 | f^* \leq u\},$$

In particular we conclude f^* is a constant, say c , on a set of Q -measure one if and only if $\mu_P\{P \in \mathcal{P}_1 | P\{x | f^*(x) = c\}\} = 1$.

Finally, suppose f is again measurable with respect to \mathcal{A} . Let $Q \in \mathcal{P}$ and suppose $\mu_Q P \left\{ \in \mathcal{P}_1 \mid \int_{\Omega} |f| dP < \infty \right\} = 1$. Then we can easily prove

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

$$\int_{\Omega} f dQ = \int_{\mathcal{P}_1} \left[\int_{\Omega} f dP \right] d\mu_Q.$$

REFERENCES

1. T. E. Harris, *On chains of infinite order*, Pacific J. Math., **5** (1955), 707-724.
2. E. Hewitt and L. J. Savage, *Symmetric measures on Cartesian products*, TAMS, **80** (1955), 470-501.
3. J. Lamperti, *Stationary measures for certain stochastic processes*, Pacific J. Math., **8** (1958), 127-132.

SANDIA CORPORATION AND
INTERNATIONAL BUSINESS MACHINES CORPORATION

PACIFIC JOURNAL OF MATHEMATICS

EDITORS

DAVID GILBARG

Stanford University
Stanford, California

F. H. BROWNELL

University of Washington
Seattle 5, Washington

A. L. WHITEMAN

University of Southern California
Los Angeles 7, California

L. J. PAIGE

University of California
Los Angeles 24, California

ASSOCIATE EDITORS

E. F. BECKENBACH

T. M. CHERRY

D. DERRY

E. HEWITT

A. HORN

L. NACHBIN

M. OHTSUKA

H. L. ROYDEN

M. M. SCHIFFER

E. SPANIER

E. G. STRAUS

F. WOLF

SUPPORTING INSTITUTIONS

UNIVERSITY OF BRITISH COLUMBIA
CALIFORNIA INSTITUTE OF TECHNOLOGY
UNIVERSITY OF CALIFORNIA
MONTANA STATE UNIVERSITY
UNIVERSITY OF NEVADA
NEW MEXICO STATE UNIVERSITY
OREGON STATE COLLEGE
UNIVERSITY OF OREGON
OSAKA UNIVERSITY
UNIVERSITY OF SOUTHERN CALIFORNIA

STANFORD UNIVERSITY
UNIVERSITY OF TOKYO
UNIVERSITY OF UTAH
WASHINGTON STATE COLLEGE
UNIVERSITY OF WASHINGTON

* * *

AMERICAN MATHEMATICAL SOCIETY
CALIFORNIA RESEARCH CORPORATION
HUGHES AIRCRAFT COMPANY
SPACE TECHNOLOGY LABORATORIES
NAVAL ORDNANCE TEST STATION

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.

Pacific Journal of Mathematics

Vol. 10, No. 4

December, 1960

M. Altman, <i>An optimum cubically convergent iterative method of inverting a linear bounded operator in Hilbert space</i>	1107
Nesmith Cornett Ankeny, <i>Criterion for rth power residuacity</i>	1115
Julius Rubin Blum and David Lee Hanson, <i>On invariant probability measures I</i>	1125
Frank Featherstone Bonsall, <i>Positive operators compact in an auxiliary topology</i>	1131
Billy Joe Boyer, <i>Summability of derived conjugate series</i>	1139
Delmar L. Boyer, <i>A note on a problem of Fuchs</i>	1147
Hans-Joachim Bremermann, <i>The envelopes of holomorphy of tube domains in infinite dimensional Banach spaces</i>	1149
Andrew Michael Bruckner, <i>Minimal superadditive extensions of superadditive functions</i>	1155
Billy Finney Bryant, <i>On expansive homeomorphisms</i>	1163
Jean W. Butler, <i>On complete and independent sets of operations in finite algebras</i>	1169
Lucien Le Cam, <i>An approximation theorem for the Poisson binomial distribution</i>	1181
Paul Civin, <i>Involutions on locally compact rings</i>	1199
Earl A. Coddington, <i>Normal extensions of formally normal operators</i>	1203
Jacob Feldman, <i>Some classes of equivalent Gaussian processes on an interval</i>	1211
Shaul Foguel, <i>Weak and strong convergence for Markov processes</i>	1221
Martin Fox, <i>Some zero sum two-person games with moves in the unit interval</i>	1235
Robert Pertsch Gilbert, <i>Singularities of three-dimensional harmonic functions</i>	1243
Branko Grünbaum, <i>Partitions of mass-distributions and of convex bodies by hyperplanes</i>	1257
Sidney Morris Harmon, <i>Regular covering surfaces of Riemann surfaces</i>	1263
Edwin Hewitt and Herbert S. Zuckerman, <i>The multiplicative semigroup of integers modulo m</i>	1291
Paul Daniel Hill, <i>Relation of a direct limit group to associated vector groups</i>	1309
Calvin Virgil Holmes, <i>Commutator groups of monomial groups</i>	1313
James Fredrik Jakobsen and W. R. Utz, <i>The non-existence of expansive homeomorphisms on a closed 2-cell</i>	1319
John William Jewett, <i>Multiplication on classes of pseudo-analytic functions</i>	1323
Helmut Klingen, <i>Analytic automorphisms of bounded symmetric complex domains</i>	1327
Robert Jacob Koch, <i>Ordered semigroups in partially ordered semigroups</i>	1333
Marvin David Marcus and N. A. Khan, <i>On a commutator result of Tausky and Zassenhaus</i>	1337
John Glen Marica and Steve Jerome Bryant, <i>Unary algebras</i>	1347
Edward Peter Merkes and W. T. Scott, <i>On univalence of a continued fraction</i>	1361
Shu-Teh Chen Moy, <i>Asymptotic properties of derivatives of stationary measures</i>	1371
John William Neuberger, <i>Concerning boundary value problems</i>	1385
Edward C. Posner, <i>Integral closure of differential rings</i>	1393
Marian Reichaw-Reichbach, <i>Some theorems on mappings onto</i>	1397
Marvin Rosenblum and Harold Widom, <i>Two extremal problems</i>	1409
Morton Lincoln Slater and Herbert S. Wilf, <i>A class of linear differential-difference equations</i>	1419
Charles Robson Storey, Jr., <i>The structure of threads</i>	1429
J. François Treves, <i>An estimate for differential polynomials in $\partial/\partial z_1, \dots, \partial/\partial z_n$</i>	1447
J. D. Weston, <i>On the representation of operators by convolutions integrals</i>	1453
James Victor Whittaker, <i>Normal subgroups of some homeomorphism groups</i>	1469