CHARACTERIZATIONS OF CONDITIONAL EXPECTATION AS A TRANSFORMATION ON FUNCTION SPACES

SHU-TEH CHEN MOY

Introduction. Let $(\Omega, \mathfrak{F}, \mu)$ be a probability space; that is, Ω is a collection of elements $u, v, w, \cdots$; $\mathfrak{F}$, a σ-algebra of subsets of Ω; and μ, a countably additive measure on $\mathfrak{F}$ with $\mu(\Omega) = 1$. Let x be a function defined on Ω, having values on the real line extended by the adjunction of $+\infty$ and $-\infty$. Let x be measurable with respect to $\mathfrak{F}$. We say that *the expectation of x exists* if one of the integrals

$$\int x^+ d\mu, \quad \int x^- d\mu$$

is finite where x^+, x^- are the positive part and the negative part of x, respectively. *The expectation of x*, $\mathrm{E}\{x\}$, is then defined to be equal to the integral $\int x \, d\mu$. Let $\mathfrak{F}_1$ be a σ-algebra of subsets of Ω with $\mathfrak{F}_1 \subset \mathfrak{F}$. For every $A \in \mathfrak{F}_1$, the equation

$$\phi(A) = \int_A x \, d\mu$$

defines a countably additive set function ϕ on $\mathfrak{F}_1$ which is absolutely continuous with respect to the contraction of μ to $\mathfrak{F}_1$. By a generalized form of the Radon-Nikodym theorem there is an extended real-valued function y defined on Ω which is measurable with respect to $\mathfrak{F}_1$ and satisfies the equation

$$\phi(A) = \int_A y \, d\mu$$

Received January 13, 1953. The work covered in this paper was done under a predoctoral fellowship tenure from the Atomic Energy Commission. The author feels greatly indebted to the encouraging inspiration and comments of Professor J. L. Doob under whose direction this paper was prepared. It constitutes part of a doctoral thesis submitted to the University of Michigan.

Pacific J. Math. 4 (1954), 47-64

for every $A \in \mathfrak{F}_1$ [3]. Such a function is unique within μ-measure 0 and is defined to be *the conditional expectation of x relative to* $\mathfrak{F}_1$, denoted by $E\{x \mid \mathfrak{F}_1\}$. The integral or expectation of x, $E\{x\}$, is then the special case of conditional expectation of x relative to $\mathfrak{F}_1$, where

$$\mathfrak{F}_1 = \{\Omega, \text{null set}\}.$$

In the following we shall list some properties of conditional expectation [**2**, Ch. 1; **6**, Ch. 5]: x, y, z, x_n, $\cdots$ are extended real-valued, measurable functions whose expectations exist.

CE 1. If α is a finite real number, then

$$E\{\alpha x \mid \mathfrak{F}_1\} = \alpha E\{x \mid \mathfrak{F}_1\} \tag{1}$$

almost everywhere. If x is nonnegative almost everywhere, then (1) is true for either finite or infinite α.

CE 2. If

$$E\{x\} > -\infty, \quad E\{y\} > -\infty,$$

then

$$E\{x + y \mid \mathfrak{F}_1\} = E\{x \mid \mathfrak{F}_1\} + E\{y \mid \mathfrak{F}_1\}$$

almost everywhere.

CE 3. If $x \geq y$ almost everywhere, then

$$E\{x \mid \mathfrak{F}_1\} \geq E\{y \mid \mathfrak{F}_1\}$$

almost everywhere.

CE 4. The relation

$$|E\{x \mid \mathfrak{F}_1\}| \leq E\{|x| \mid \mathfrak{F}_1\}$$

holds almost everywhere. Therefore, if x is equal to a bounded function almost everywhere then $E\{x \mid \mathfrak{F}_1\}$ is also.

CE 5. If x is measurable with respect to $\mathfrak{F}_1$ and x is equal to a bounded function almost everywhere, and y has finite expectation, then

$$\mathrm{E}\{xy \mid \mathfrak{Z}_1\} = x\mathrm{E}\{y \mid \mathfrak{Z}_1\} \tag{2}$$

almost everywhere. If x, y are nonnegative, then (2) is true for any x which is measurable with respect to $\mathfrak{Z}_1$.

CE 6. If

$$\mathrm{E}\{x_n\} > -\infty$$

for all n, and $x_1 \leq x_2 \leq \cdots$ almost everywhere, then $\mathrm{E}\{x_n \mid \mathfrak{Z}_1\}$ converges almost everywhere to $\mathrm{E}\{x \mid \mathfrak{Z}_1\}$, where $x = \lim_{n\to\infty} x_n$ almost everywhere.

CE 7. If $p \geq 1$, then

$$|\mathrm{E}\{x \mid \mathfrak{Z}_1\}|^p \leq \mathrm{E}\{|x|^p \mid \mathfrak{Z}_1\}$$

almost everywhere. Therefore, if

$$\mathrm{E}\{|x|^p\} < \infty,$$

then

$$\mathrm{E}\{|\mathrm{E}\{x \mid \mathfrak{Z}_1\}|^p\} < \infty;$$

and if

$$\mathrm{E}\{|x_n|^p\} < \infty$$

for every n and

$$\lim_{n\to\infty} \mathrm{E}\{|x_n - x|^p\} = 0,$$

then

$$\lim_{n\to\infty} \mathrm{E}\{|\mathrm{E}\{x_n \mid \mathfrak{Z}_1\} - \mathrm{E}\{x \mid \mathfrak{Z}_1\}|^p\} = 0.$$

In this paper we shall study the relation of the conditional expectation with a transformation T on some spaces of measurable functions into themselves satisfying the conditions:

$$T(x+y) = Tx + Ty,$$

$$T\alpha x = \alpha Tx, \qquad \text{where } \alpha \text{ is a constant},$$

$$T(x\, Ty) = (Tx)\cdot(Ty).$$

Under the restrictions that the transform of a function which is equal to a bounded function almost everywhere is also equal to a bounded function almost everywhere, and that T satisfies a certain continuity condition, we are able to identify such a transformation as the one which takes x to $\mathrm{E}\{xg \mid \mathfrak{F}_T\}$, where g is a measurable function and $\mathfrak{F}_T$ is a σ-algebra of subsets of Ω with $\mathfrak{F}_T \subset \mathfrak{F}$. This is an attempt to answer the search for an appropriate definition of average which would be desirable for the establishment of a mathematical theory of the dynamics of turbulence [5, 7]. In the past, J. Kampé De Fériet has studied the transformation on the collection of real functions which takes only a finite number of values [4]. Garret Birkhoff and John Sopka have studied the transformation on the space of continuous functions on a compact Hausdorff space [1, 8]. Garrett Birkhoff also treated the subject from an abstract algebraic point of view [1]. Since the modern treatment of fluid dynamics theory is based on probability theory, it seems to the author that a probability solution is the natural one.

In the first section of this paper we shall consider the transformation on the space of nonnegative measurable functions into itself. In some respect it is analogous to the integration theory of nonnegative functions. In the second section we consider T as a linear continuous transformation on L_p into L_p. In the case of L_1 it is also proved, if $T1 = 1$ and $\|Tx\|_1 \leq \|x\|_1$ where $\|\cdot\|_1$ denotes the L_1 norm, then Tx is the conditional expectation of x relative to a σ-algebra of subsets.

1. Transformation on the space of nonnegative measurable functions. Let $\mathcal{S}$ be the collection of all nonnegative, extended real-valued functions on Ω which are measurable with respect to $\mathfrak{F}$. Elements of $\mathcal{S}$ are denoted by x, y, z, $\cdots$. Two functions are considered equal if they are equal almost everywhere. By $x = y$ or $x > y$ we mean that $x = y$ almost everywhere or $x > y$ almost everywhere, respectively. When we say that x is bounded we mean that x is equal to a bounded function almost everywhere. We shall use the symbol $x_n \longrightarrow x$ to mean that

$$\lim_{n \to \infty} x_n(w) = x(w)$$

for almost all w in Ω. Addition and multiplication in $\mathcal{S}$ are ordinary pointwise addition and multiplication with the conventions that $\alpha + \infty = \infty$ for every nonnegative number α, and $\alpha \cdot \infty = \infty$ if $\alpha > 0$, $\alpha \cdot \infty = 0$ if $\alpha = 0$. Thus $\mathcal{S}$ is closed under addition and multiplication. We shall use symbols α, β, $\cdots$ to denote nonnegative real numbers. We shall use the same symbols to denote functions which take on a constant value α, β, $\cdots$ almost everywhere.

We are considering a transformation T on $\mathcal{S}$ into $\mathcal{S}$ satisfying the following conditions:

T1. a) *$T(x + y) = Tx + Ty$ for every pair of elements x, y in $\mathcal{S}$.*

b) *$T\alpha x = \alpha Tx$ for every nonnegative number α and every x in $\mathcal{S}$.*

T2. *If x is bounded then Tx is bounded.*

T3. *$T(x \cdot Ty) = (Tx) \cdot (Ty)$ for every pair of elements x, y in $\mathcal{S}$.*

T4. *If $\{x_n\}$ is a nondecreasing sequence of elements of $\mathcal{S}$ for which $x_n \longrightarrow x$, then $Tx_n \longrightarrow Tx$.*

If T_E is the transformation which takes x to $\mathrm{E}\{x \mid \mathfrak{F}_1\}$, then by CE 3 T_E is a transformation on $\mathcal{S}$ into $\mathcal{S}$. CE 1 and CE 2 imply that T_E satisfies the condition T1; CE 4 implies that T_E satisfies the condition T2; CE 5 implies that T_E satisfies the condition T3; and CE 6 implies that T_E satisfies the condition T4. Therefore the transformation which takes x to the conditional expectation of x relative to a σ-algebra of subsets $\mathfrak{F}_1 \subset \mathfrak{F}$ satisfies T1, T2, T3, and T4. It is easy to check that the transformation which takes x to $\mathrm{E}\{xg \mid \mathfrak{F}_1\}$, where g is a nonnegative measurable function with $\mathrm{E}\{g \mid \mathfrak{F}_1\}$ bounded, also satisfies T1, T2, T3, and T4. We shall prove that the last example is actually the most general form of a transformation satisfying T1, T2, T3, and T4.

LEMMA 1.1. *The inequality $x \geq y$ implies $Tx \geq Ty$.*

Proof. If y is finite valued almost everywhere, then

$$Tx = T(y + (x - y)) = Tx + T(x - y) \geq Ty.$$

If y is not finite valued, let

$$A = [w : y(w) = \infty];$$

then $x(w) = \infty$ for $w \in A$. Let x', y' be defined as $x'(w) = x(w)$ for $w \notin A$, $x'(w) = 0$ for $w \in A$; $y'(w) = y(w)$ for $w \notin A$, and $y'(w) = 0$ for $w \in A$. Then $Tx' \geq Ty'$. Let z be a function defined as $z(w) = \infty$ if $w \in A$, $z(w) = 0$ if $w \notin A$; then

$$x = x' + z, \quad y = y' + z,$$

and

$$Tx = Tx' + Tz \geq Ty' + Tz = Ty.$$

LEMMA 1.2. *If $x_n \longrightarrow x$, then $Tx \leq \liminf_n Tx_n$.*

Proof. Let

$$y_n = \inf [x_i : i \geq n];$$

then $y_n \leq x_n$ for every n and $\{y_n\}$ is a nondecreasing sequence of elements for which $y_n \longrightarrow x$. By T4, $Ty_n \longrightarrow Tx$. But by Lemma 1.1, $Ty_n \leq Tx_n$ for every n; hence

$$Tx \leq \liminf_n Tx_n.$$

LEMMA 1.3. *If $x_n \longrightarrow x$ and $x_n \leq y$ for every n where y and Ty are finite valued, then $Tx_n \longrightarrow Tx$.*

Proof. By Lemma 1.2,

$$Tx \leq \liminf_n Tx_n$$

and

$$T(y - x) \leq \liminf_n T(y - x_n).$$

By Lemma 1.1, $Tx_n \leq Ty$ for every n and $Tx \leq Ty$; hence the second inequality can be written as

$$Ty - Tx \leq Ty - \limsup_n Tx_n;$$

that is,

$$Tx \geq \limsup_n Tx_n.$$

Now we have

$$\limsup_n Tx_n \leq Tx \leq \liminf_n Tx_n;$$

hence

$$Tx_n \longrightarrow Tx.$$

LEMMA 1.4. *Let* $\mathcal{E}$ *be the totality of elements* y *of* $\mathcal{S}$ *for which* $T(xy) = y\,Tx$ *for every* $x \in \mathcal{S}$.

1. *If* $y_1, y_2 \in \mathcal{E}$, *then* $y_1 + y_2$ *and* $y_1 \cdot y_2 \in \mathcal{E}$; *and if* $y_1 \leq y_2$ *with* y_2 *bounded, then* $y_2 - y_1 \in \mathcal{E}$.

2. *If* $\alpha \geq 0$, $y \in \mathcal{E}$, *then* $\alpha y \in \mathcal{E}$.

3. *If* $\{y_n\}$ *is a nondecreasing sequence of elements with* $y_n \in \mathcal{E}$ *for every* n *and* $y_n \longrightarrow y$, *then* $y \in \mathcal{E}$.

4. *If* $\{y_n\}$ *is a sequence of elements in* $\mathcal{E}$ *with* $y_n \longrightarrow y$ *and there is a bounded function* z *for which* $y_n \leq z$ *for every* n, *then* $y \in \mathcal{E}$.

Proof. 1. We have

$$T[x(y_1 + y_2)] = T(xy_1) + T(xy_2) = y_1 Tx + y_2 Tx = (y_1 + y_2)\,Tx,$$

and

$$T(xy_1 y_2) = y_1\,T(xy_2) = y_1\,y_2\,Tx.$$

If y_2 is bounded and $y_1 \leq y_2$, then

$$y_2 Tx = T(xy_2) = T[x(y_1 + y_2 - y_1)] = T(xy_1) + T[x(y_2 - y_1)]$$

$$= y_1 Tx + T[x(y_2 - y_1)].$$

If x is bounded, then Tx is bounded. We have

$$y_2 Tx - y_1 Tx = (y_2 - y_1)\,Tx = T[x(y_2 - y_1)].$$

The last equality is true for all bounded x, and therefore is true for all x.

2. Clearly

$$T(x\alpha y) = \alpha\,T(xy) = \alpha y Tx.$$

3. We have $T(xy_n) = y_n Tx$ for every n. Further, $T(xy_n)$ and $y_n Tx$ are nondecreasing sequences with $T(xy_n) \longrightarrow T(xy)$, by T4; and $y_n Tx \longrightarrow y\,Tx$. Hence $T(xy) = y\,Tx$.

4. We have $T(xy_n) = y_n Tx$ for every n. If x is bounded, then xz is bounded and $xy_n \leq xz$ for each n. By Lemma 1.3, $T(xy_n) \longrightarrow T(xy)$. On the other hand, we have $y_n Tx \longrightarrow yTx$. Hence $T(xy) = yTx$. The equality is true for all bounded

x, and therefore is true for all x by T4.

We remark that 1. implies that any nonnegative polynomial $P(y)$, of a bounded $y \in \mathcal{E}$ (that is, $P(y(w)) \geq 0$ for all $w \in \Omega$) is also in $\mathcal{E}$.

The following lemma is obvious by T3.

LEMMA 1.5. *For each* $x \in \mathcal{B}$, $Tx \in \mathcal{E}$.

LEMMA 1.6. *Let* $\mathfrak{F}_T$ *be the collection of all sets* $E \in \mathfrak{F}$ *whose characteristic functions* χ_E *are in the* $\mathcal{E}$ *of Lemma* 1.4. *Then* $\mathfrak{F}_T$ *is a* σ*-algebra of subsets of* Ω.

Proof. We shall establish:

1. Clearly, $1 \in \mathcal{E}$ for $1Tx = T(x \cdot 1)$; hence $\Omega \in \mathfrak{F}_T$.

2. If $E_1, E_2 \in \mathfrak{F}_T$, then $E_1 \cap E_2 \in \mathfrak{F}_T$ by Lemma 1.4 and the equation

$$\chi_{E_1 \cap E_2} = \chi_{E_1} \cdot \chi_{E_2}.$$

3. If $E_1, E_2 \in \mathfrak{F}_T$, then $E_1 - E_2 \in \mathfrak{F}_T$ by the equality

$$\chi_{E_1 - E_2} = \chi_{E_1} - \chi_{E_1 \cap E_2}$$

and Lemma 1.4.

4. If $E_1, E_2 \in \mathfrak{F}_T$, then $E_1 \cup E_2 \in \mathfrak{F}_T$ by Lemma 1.4 and the equality

$$\chi_{E_1 \cup E_2} = \chi_{E_1} + \chi_{E_2} - \chi_{E_1 \cap E_2}.$$

By induction, for any finite number of sets $E_1, E_2, \cdots E_n \in \mathfrak{F}_T$,

$$\bigcup_{i=1}^{n} E_i \in \mathfrak{F}_T.$$

5. For a sequence of sets $E_1, E_2, \cdots, E_n, \cdots$ in $\mathfrak{F}_T$.

$$\bigcup_{n=1}^{\infty} E_n \in \mathfrak{F}_T.$$

For, by 4,

$$\chi_{\bigcup_{i=1}^{n} E_i} \in \mathcal{E}$$

and

$$\left\{\chi_{\bigcup_{i=1}^{n} E_i}\right\}$$

is a nondecreasing sequence for which

$$\chi_{\bigcup_{i=1}^{n} E_i} \longrightarrow \chi_{\bigcup_{i=1}^{\infty} E_i}.$$

LEMMA 1.7. *Let* $\mathfrak{M}$ *be the collection of all nonnegative functions which are measurable with respect to* $\mathfrak{F}_T$. *Then* $\mathfrak{M} \subset \mathcal{E}$.

Proof. Functions which are linear combinations with nonnegative coefficients of characteristic functions of sets in $\mathfrak{F}_T$ are in $\mathcal{E}$. For any element of $\mathfrak{M}$ there is a nondecreasing sequence of such functions converging to it; therefore, by Lemma 1.4, it is in $\mathcal{E}$.

LEMMA 1.8. *Let* $y \in \mathcal{E}$ *and* y *be bounded. Let* $\mathfrak{F}_y$ *be the least* σ*-algebra with respect to which* y *is measurable. Then* $\mathfrak{F}_y \subset \mathfrak{F}_T$, *or, equivalently,* $y \in \mathfrak{M}$.

Proof. Let Φ be a nonnegative continuous function on a finite interval containing the range of y. We want to prove $\Phi(y) \in \mathcal{E}$. We may assume that

$$0 < \alpha \leq \Phi(y) \leq \beta,$$

for $\Phi(y) + \alpha \in \mathcal{E}$ with $\alpha > 0$ implies that $\Phi(y) \in \mathcal{E}$, by Lemma 1.4.

Since y is bounded by the Weierstrass theorem there is a sequence $\{P_n(y)\}$ of polynomials such that $P_n(y)$ converges uniformly to $\Phi(y)$. We may assume $P_n(y(w)) \geq 0$ for all $w \in \Omega$; therefore, by Lemma 1.4, $P_n(y) \in \mathcal{E}$ for each n, and $\Phi(y) \in \mathcal{E}$.

For each $E \in \mathfrak{F}_y$, there is a sequence $\{\Phi_n(y)\}$ of continuous functions of y with $0 \leq \Phi_n(y(w)) \leq 1$ for each n and w for which

$$\Phi_n(y) \longrightarrow \chi_E.$$

Hence, again by Lemma 1.4, $\chi_E \in \mathcal{E}$; that is, $E \in \mathfrak{F}_T$.

LEMMA 1.9. *For each* $x \in \mathfrak{L}$, $Tx \in \mathfrak{M}$.

Proof. Let $\{x_n\}$ be a nondecreasing sequence of bounded functions for which $x_n \longrightarrow x$. Then $\{Tx_n\}$ is also a nondecreasing sequence of bounded functions, and $Tx_n \longrightarrow Tx$. By Lemma 1.5 and Lemma 1.8, $Tx_n \in \mathfrak{M}$ for every n; therefore $Tx \in \mathfrak{M}$.

THEOREM 1.1. *If T is a transformation on the collection $\mathfrak{S}$ of all nonnegative measurable functions on a probability space $(\Omega, \mathfrak{Z}, \mu)$ into itself satisfying* T1, T2, T3, T4, *then T is of the form*

$$Tx = \mathrm{E}\{xg/\mathfrak{Z}_T\},$$

where $\mathfrak{Z}_T$ is a σ-algebra of subsets of Ω with $\mathfrak{Z}_T \subset \mathfrak{Z}$ and g is a nonnegative measurable function for which $\mathrm{E}\{g \mid \mathfrak{Z}_T\}$ is bounded.

Proof. Consider the set function ν on $\mathfrak{Z}$ defined by

$$\nu(A) = \int T\chi_A \, d\mu,$$

where χ_A is the characteristic function of A; then T1, T2, T4 imply that ν is a finite measure on $\mathfrak{Z}$. For a linear combination of characteristic functions with nonnegative coefficients

$$\sum_{i=1}^{n} \alpha_i \chi_{A_i},$$

we have

$$\int T\left(\sum_{i=1}^{n} \alpha_i \chi_{A_i}\right) d\mu = \sum_{i=1}^{n} \alpha_i \int T\chi_{A_i} \, d\mu = \sum_{i=1}^{n} \alpha_i \, \nu(A_i).$$

Hence for each $x \in \mathfrak{S}$,

$$\int Tx d\mu = \int x \, d\nu.$$

Since ν is absolutely continuous with respect to μ, by the Radon-Nikodym theorem there is a nonnegative function g for which

$$\nu(A) = \int_A g \, d\mu$$

for every $A \in \mathfrak{Z}$. Therefore for each $x \in \mathfrak{S}$,

$$\int Tx\, d\mu = \int xg\, d\mu.$$

For every $E \in \mathfrak{Z}_T$,

$$\int_E Tx\, d\mu = \int \chi_E \ Tx\, d\mu = \int T(x\chi_E)\, d\mu = \int x\, \chi_E g\, d\mu = \int_E xg\, d\mu.$$

The previous equality and the fact that Tx is measurable with respect to $\mathfrak{Z}_T$ (Lemma 1.9) imply that

$$Tx = \mathrm{E}\{xg \mid \mathfrak{Z}_T\}.$$

In particular, $\mathrm{T}1 = \mathrm{E}\{g \mid \mathfrak{Z}_T\}$; hence $\mathrm{E}\{g \mid \mathfrak{Z}_T\}$ is bounded.

REMARK. The representation $Tx = \mathrm{E}\{gx \mid \mathfrak{Z}_T\}$ is not unique. For example, if $g(w) = 0$ for $w \in E$, where $E \in \mathfrak{Z}_T$ and $\mu(E) > 0$, and we let

$$\mathfrak{Z}' = [E \cup F : F \in \mathfrak{Z}_T, F \cap E = \phi] \cup [F : F \in \mathfrak{Z}_T, F \cap E = \phi],$$

then $\mathrm{E}\{xg \mid \mathfrak{Z}_T\} = \mathrm{E}\{xg \mid \mathfrak{Z}'\}$ for every x in $\mathfrak{S}$.

COROLLARY 1.1. *If the collection of nonnegative constant functions is invariant under T satisfying* T1, T2, T3, T4, *then, except for the trivial case $Tx = 0$ for all x, the range of T is $\mathfrak{M}$, and g of Theorem* 1.1 *satisfies*

$$\mathrm{E}\{g \mid \mathfrak{Z}_T\} = \alpha,$$

where $\alpha = \mathrm{T}1 \neq 0$. In particular, if $\mathrm{T}1 = 1$, then T is a projection of $\mathfrak{S}$ on $\mathfrak{M}$.

Proof. $\mathrm{T}1$ must not be 0. For if $\mathrm{T}1 = 0$ then

$$\int g\, d\mu = \int \mathrm{T}1\, d\mu = 0.$$

Therefore $g = 0$, and $Tx = 0$ for every x. For each $y \in \mathfrak{M}$,

$$Ty = \mathrm{E}\{yg \mid \mathfrak{Z}_T\} = y\ \mathrm{E}\{g \mid \mathfrak{Z}_T\} = \alpha y;$$

therefore

$$T\frac{1}{\alpha}y = y.$$

This fact together with Lemma 1.9 implies that $\mathfrak{M}$ is the range of T.

If $\mathrm{T}1 = 1$, then

$$T^2x = T(1 \cdot Tx) = Tx\,\mathrm{T}1 = Tx;$$

that is, $T^2 = T$. Hence T is a projection in this case.

2. Transformation on the space L_p. Let L_p, $p \geq 1$, be the usual space of all real pth power integrable functions on the probability space $(\Omega, \mathfrak{F}, \mu)$. If the transformation discussed in the previous section takes functions in L_p into functions in L_p, then it can be extended to be a linear transformation on L_p into L_p by defining $Tx = Tx^+ - Tx^-$, where x^+, x^- are the positive and negative parts of x, respectively; therefore we still have the same representation

$$Tx = \mathrm{E}\{xg \mid \mathfrak{F}_T\}$$

for every x in L_p with a nonnegative function g for which $\mathrm{E}\{g \mid \mathfrak{F}_T\}$ is bounded. The restriction that T transforms nonnegative functions into nonnegative functions is the same as that T be order preserving; that is, if $x \geq y$ then $Tx \geq Ty$. In this section we shall consider a transformation on L_p into L_p similar to the previous one, but with no restriction that the order is to be preserved. We shall employ the usual norm topology in L_p and assume the transformation to be continuous in norm topology. We are able to prove with essentially the same argument as in the previous section that the same representation

$$Tx = \mathrm{E}\{xg \mid \mathfrak{F}_T\}$$

is to be arrived at, but with a function g which is no longer nonnegative.

We shall state the assumptions precisely. In the following, elements of L_p are denoted by $x, y, z, \cdots$. As before, two functions are considered equal if they are equal almost everywhere. A function is said to be bounded if it is equal to a bounded function almost everywhere. The symbols $\alpha, \beta, \cdots$ are to denote both real constants and functions which take a constant value almost everywhere. For a sequence of functions $\{x_n\}$ we shall use the expression "$x_n \longrightarrow x$ in L_p" to denote

$$\lim_{n \to \infty} \int |x_n - x|^p d\mu = 0.$$

T is to be a transformation on L_p into L_p satisfying the following conditions:

T′1. *T is linear; that is,*

$$T(\alpha x + \beta y) = \alpha Tx + \beta Ty.$$

T′2. *If x is bounded, then Tx is bounded.*

T′3. *$T(x\,Ty) = (Tx)\cdot(Ty)$ for every pair x, y of bounded elements of L_p.*

T′4. *T is continuous; that is, if $x_n \longrightarrow x$ in L_p then $Tx_n \longrightarrow Tx$ in L_p.*

From the properties of the conditional expectation CE 1, CE 2, CE 5, CE 6, and CE 7, we see that the transformation which takes x to $E[x \mid \mathfrak{F}_1]$ satisfies the foregoing conditions.

LEMMA 2.1. *Let $\mathcal{E}$ be the totality of elements y of L_p for which $T(xy) = y\,Tx$ holds for every bounded x in L_p. Then $\mathcal{E}$ is a closed linear subspace of L_p. Moreover, if y_1, $y_2 \in \mathcal{E}$ and y_1, y_2 are bounded, then $y_1 \cdot y_2 \in \mathcal{E}$; therefore every polynomial of a bounded function in $\mathcal{E}$ is also in $\mathcal{E}$.*

Proof. The proof that $\mathcal{E}$ is a linear set and closed under multiplication of bounded functions is similar to that of Lemma 1.4. To prove that it is a closed subset of L_p, let $y_n \longrightarrow y$ in L_p and $y_n \in \mathcal{E}$ for every n; then $xy_n \longrightarrow xy$ in L_p for every bounded x, and hence $T(xy_n) \longrightarrow T(xy)$ in L_p, by T′4. On the other hand, if x is bounded, then Tx is also bounded by T′2; hence

$$y_n Tx \longrightarrow y\,Tx$$

in L_p. Since

$$y_n Tx = Txy_n$$

for every n, we have $yTx = Txy$; that is, $y \in \mathcal{E}$. Hence $\mathcal{E}$ is a closed subset of L_p.

LEMMA 2.2. *For each $x \in L_p$, $Tx \in \mathcal{E}$.*

Proof. By T′3, if x is bounded then $Tx \in \mathcal{E}$. If x is not bounded, there is a sequence $\{x_n\}$ of bounded functions for which $x_n \longrightarrow x$ in L_p; then $Tx_n \longrightarrow Tx$ in L_p, by T′4. Now $Tx_n \in \mathcal{E}$ for each n, and the fact that $\mathcal{E}$ is closed (Lemma 2.1), imply that $Tx \in \mathcal{E}$.

LEMMA 2.3. *Let $\mathfrak{F}_T$ be the collection of all sets $E \in \mathfrak{F}$ whose characteristic functions χ_E are in $\mathcal{E}$; then $\mathfrak{F}_T$ is a σ-algebra of subsets of Ω.*

The proof is the same as that of Lemma 1.6 except that we have to use the fact that $\mathcal{E}$ is a closed subset of L_p and

$$\chi_{\bigcup_{i=1}^{n} E_i} \longrightarrow \chi_{\bigcup_{i=1}^{\infty} E_i}$$

in L_p to prove that $\mathfrak{Z}_T$ is closed under countable unions.

LEMMA 2.4. *Let* $\mathfrak{M}$ *be the totality of elements of* L_p *which are measurable with respect to the* $\mathfrak{Z}_T$ *of Lemma* 2.3 , *then* $\mathfrak{M} \subset \mathcal{E}$.

Proof. Linear combinations of characteristic functions of sets of $\mathfrak{Z}_T$ are in $\mathcal{E}$. The facts that the totality of such functions is dense in $\mathfrak{M}$, and that $\mathcal{E}$ is a closed subset of L_p, imply that $\mathfrak{M} \subset \mathcal{E}$.

LEMMA 2.5. *Let* $y \in \mathcal{E}$ *and* y *be bounded. Let* $\mathfrak{Z}_y$ *be the smallest* σ*-algebra of subsets of* Ω *with respect to which* y *is measurable. Then* $\mathfrak{Z}_y \subset \mathfrak{Z}_T$ *or, equivalently,* $y \in \mathfrak{M}$.

Proof. Let Φ be a continuous real function on a finite interval containing the range of y. Since y is bounded, by the Weierstrass theorem there is a sequence $\{P_n(y)\}$ of polynomials of y for which $P_n[y(w)]$ converges to $\Phi[y(w)]$ uniformly in w. Hence $P_n(y) \longrightarrow \Phi(y)$ in L_p; therefore, by Lemma 2.1, $P_n(y) \in \mathcal{E}$ for every n, and $\Phi(y) \in \mathcal{E}$.

For each $E \in \mathfrak{Z}_y$, there is a sequence $\{\Phi_n(y)\}$ of continuous functions of y for which $\Phi_n(y) \longrightarrow \chi_E$ in L_p, where χ_E is the characteristic function of E. Hence, again by Lemma 2.1, $\chi_E \in \mathcal{E}$; that is, $E \in \mathfrak{Z}_T$.

LEMMA 2.6. *For each* $x \in L_p$, $Tx \in \mathfrak{M}$.

Proof. For each $x \in L_p$ there is a sequence $\{x_n\}$ of bounded functions in L_p for which $x_n \longrightarrow x$ in L_p. Hence $Tx_n \longrightarrow Tx$ in L_p by T′4. By T′2, Tx_n is bounded for every n. By Lemma 2.2 and Lemma 2.5, $Tx_n \in \mathfrak{M}$ for every n. Since $\mathfrak{M}$ is a closed subset of L_p, $Tx \in \mathfrak{M}$.

THEOREM 2.1. *If* T *is a transformation of* L_p *into* L_p *satisfying* T′1, T′2, T′3, T′4, *then* T *is of the form*

$$Tx = \mathrm{E}\{xg \mid \mathfrak{Z}_T\},$$

where $\mathfrak{Z}_T$ *is a* σ*-algebra of subsets of* Ω *with* $\mathfrak{Z}_T \subset \mathfrak{Z}$ *and* $g \in L_q$, *where* $1/p + 1/q = 1$, *for which* $\mathrm{E}\{g \mid \mathfrak{Z}_T\}$ *is bounded* (*in the case of* $p = 1$, *then* g *is a bounded function*).

Proof. We consider the function $\ell(x)$ defined on L_p by

$$\ell(x) = \int Tx \, d\mu.$$

$T'1$ implies that ℓ is linear. Also, ℓ is continuous, for if $x_n \longrightarrow x$ in L_p then $Tx_n \longrightarrow Tx$ in L_p, which implies $Tx_n \longrightarrow Tx$ in L_1; therefore

$$\ell(x_n) = \int Tx_n d\mu \longrightarrow \int Tx\, d\mu = \ell(x).$$

Now ℓ can be expressed as

$$\ell(x) = \int xg\, d\mu,$$

where $g \in L_q$ with $1/p + 1/q = 1$. (In the case $p = 1$, g is bounded almost everywhere.) Hence,

$$\int Tx\, d\mu = \int xg\, d\mu$$

for every $x \in L_p$. The same argument as in the proof of Theorem 1.1 shows that

$$Tx = \mathrm{E}\{xg \mid \mathfrak{F}_T\},$$

and that $\mathrm{E}\{g \mid \mathfrak{F}_T\}$ is bounded.

THEOREM 2.2. *If $p = 1$, and T satisfies the further conditions that*

$$T1 = 1 \quad \text{and} \quad \|Tx\|_1 \leq \|x\|_1$$

for every $x \in L_1$, where

$$\|x\|_1 = \int |x|\, d\mu,$$

then

$$Tx = \mathrm{E}\{x \mid \mathfrak{F}_T\}$$

for every x.

Proof. Consider the set function ν defined on $\mathfrak{F}$ by

$$\nu(E) = \ell(\chi_E) = \int T\chi_E\, d\mu,$$

where χ_E is the characteristic function of E, and ℓ is the linear functional defined in the proof of Theorem 2.1. Since ℓ is continuous, ν is a completely additive set function. $T1 = 1$ implies that $\nu(\Omega) = 1$. For any set $E \in \mathfrak{F}$,

$$|\nu(E)| = \left|\int T\chi_E \, d\mu\right| \leq \int |T\chi_E| \, d\mu = \|T\chi_E\|_1$$

$$\leq \|\chi_E\|_1 = \int \chi_E \, d\mu = \mu(E).$$

Now we want to show that $\nu(E) = \mu(E)$ for every $E \in \mathfrak{F}$. First we shall prove

$$|\nu(E)| = \mu(E).$$

Suppose

$$|\nu(E)| < \mu(E)$$

for a certain E; then

$$1 = \mu(\Omega) = \mu(E) + \mu(\Omega - E) > |\nu(E)| + |\nu(\Omega - E)| \geq \nu(\Omega) = 1.$$

This is a contradiction. Hence

$$|\nu(E)| = \mu(E).$$

Now for any E, we have either

$$\nu(E) = \mu(E) \quad \text{or} \quad \nu(E) = -\mu(E).$$

Suppose

$$\nu(E) = -\mu(E);$$

then

$$\nu(\Omega - E) = 1 + \mu(E).$$

This is possible only when $\mu(E) = 0$. Therefore

$$\nu(E) = \mu(E)$$

for every $E \in \mathfrak{F}$.

The fact that $\nu \equiv \mu$ implies that the g in Theorem 2.1 is equal to 1 almost everywhere, for

$$\nu(E) = \ell(\chi_E) = \int_E g \, d\mu.$$

Hence the theorem is proved.

REFERENCES

1. Garrett Birkhoff, *Moyennes des fonctions bornées*, Congrès d'Algèbre et Théorie des Nombres, Paris, 1949, pp. 143-154.

2. J. L. Doob, *Stochastic processes* (forthcoming book).

3. P. R. Halmos, *Measure theory*, New York, 1950.

4. J. Kampé De Fériet, *Sur un problems d'algèbre abstrait posé par la definition de la moyenne dans la theorie de la turbulence*. Ann. Soc. Sci. Bruxelles, 1, 63 (1949), 156-172.

5. ———, *Introduction to the statistical theory of turbulence, correlation and spectrum*. Lecture Series No. 8, prepared by S. I. Pai, The Institute of Fluid Dynamics and Applied Mathematics, University of Maryland, 1950-51.

6. A. N. Kolmogorov, *Foundations of the theory of probability*, translated by N. Morrison, 1950.

7. Osborne, Reynolds, *On the dynamical theory of incompressible viscous fluids*, Philos. Trans. Roy. Soc. A, 136 (1895), 123-164.

8. John Sopka, *On the characterization of Reynolds operators on the algebra of all continuous functions on a compact Hausdorff space*, Harvard thesis, June, 1950.

UNIVERSITY OF MICHIGAN
UNIVERSITY OF ILLINOIS