RENDICONTI
del
SEMINARIO MATEMATICO
della
UNIVERSITÀ DI PADOVA

G. C. ROTA

On the representation of averaging operators

Rendiconti del Seminario Matematico della Università di Padova, tome 30 (1960), p. 52-64

<http://www.numdam.org/item?id=RSMUP_1960__30__52_0>

ON THE REPRESENTATION OF AVERAGING OPERATORS

Nota (*) *di* G. C. ROTA (*a Cambridge, Mass.*) (1)

1. Introduction. - Recent developments in the theory of turbulence (see (7), (8) and especially (6), where further references will be found) have led to a systematic study of some concepts which were formerly only implicitly used. One of these is that of a Reynolds operator, studied by Kampé de Fériet, G. Birkhoff and M.me Dubreil-Jacotin; a related notion is that of an *averaging operator,* whose importance was first pointed our in (6) and (7). The history of the subject is excellently explained in (6) and (7).

The present work intends to fill a gap in the theory of averaging operators. Our aim is to give a complete representation theory of averaging operators in some of the function spaces of most frequent occurrence. This is done by starting from a « minimal » system of postulates for averaging operators which are independent from any order structure of the function space. The thesis of Mrs. Moy (12) gives a result related to ours, starting however from a less stringent system of postulates than those set forth below.

The exposition is self-contained. Lemma 1 is stated in slightly different form by Mrs. Moy.

(*) Pervenuta in Redazione il 17 ottobre 1959.

Indirizzo dell:A.: Massachussetts Institute of Technology, Cambridge Mass. (U.S.A.)

1) Research supported by contract 7667 with the Office of Naval Research.

2. Averaging Operators. - In this and the following section we shall only consider probability spaces (S, Σ, μ), that is, measure spaces [2]) for which $\mu(S) = 1$. By « functions » we mean as usual equivalence classes modulo sets of measure zero. If Σ' is a subfield of the σ-field Σ, then we denote by (S, Σ', μ) the measure space obtained by restricting the measure μ to Σ'. All function spaces are tacitly taken over the real field endowed with its natural Borel structure.

DEFINITION. An *averaging operator* (or *average*) A in $L_p(S, \Sigma, \mu)$ (where p is a *fixed* real number, $1 \leq p \leq \infty$) is a linear operator in $L_p(S, \Sigma, \mu)$ with the following three properties:

(1) A is a contraction operator:

$$\int_S |(Af)(s)|^p \mu(ds) \leq \int_S |f(s)|^p \mu(ds),$$

for f in $L_p(S, \Sigma, \mu)$.

(2) If f is of class $L_p(S, \Sigma, \mu)$ and g is an essentially bounded function on (S, Σ, μ), then the function $(Af)(s)(Ag)(s)$ is of class $L_p(S, \Sigma, \mu)$ and

$$A(g\,Af) = (Ag)(Af).$$

(3) If $I(s)$ is the function identically equal to one on S, then $AI = I$ [3]).

The following examples may be found enlightening by the reader who is puzzled by this definition. Consider first the case when the range of the operator A is a one-dimensional subspace. Then our main theorem below will show that A is the operator

$$Af = \left(\int_S f(s)\mu(ds) \right) I,$$

[2]) We follow the notation in Dunford and Schwartz, (1).

[3]) It follows from the main theorem proved below that every averaging operator in this sense is also an averaging operator in the sense of Birkhoff (6) and Sopka (10).

that is, the ordinary « average » or mean of the function f. A slightly more general example is obtained when the range of A is a finite-dimensional subspace. Again, it is a consequence of the main theorem that the operator A can be described as follows. There is a partition of S into a finite number of disjoint measurable sets $E_1, E_2, \ldots, E_n$ of positive measure, and for every function f in the given space,

$$(3) \qquad (Af)(s) = [\mu(E_k)]^{-1} \int_S f(s)\mu(ds) \qquad \text{for } s \text{ in } E_k.$$

The second example is immediately extended to a countable partition of S. However, any further non-trivial example of averaging operator lies beyond this « naive » approach, and requires the following notion from measure theory.

Let Σ' be a σ-subfield of Σ, and let f be of class $L_p(S, \Sigma, \mu)$ (again for fixed p). Then the countably additive set function on Σ'

$$\mu_f(E) = \int_E f(s)\mu(dsZ), \qquad E \in \Sigma'$$

is μ-continuous. By the Radon-Nikodym Theorem there exists a *unique* Σ'*-measurable* function f' such that

$$\int_E f'(s)\mu(ds) = \int_E f(s)\mu(ds), \qquad E \in \Sigma'.$$

The function f' is the Radon-Nikodym derivative of f relative to the σ-subfield Σ'. It is evident that the operator $Af = f'$ is linear and bounded; in fact, it is easily verified that it is an averaging operator. In probability theory f' is the conditional expectation of f relative to the σ'-field Σ'.

The first of the examples given above is obtained by taking the Radon-Nikodym derivative of f relative to the minimal subfield consisting of the two sets ($\emptyset$ and S); the second is obtained by considering the σ-subfield Σ' generated by a partition, and again taking the Radon-Nikodym derivative. Thus the operator $f \rightarrow f'$ is a natural generalization of the elementary mean taken over a partition. Our main result states that every averaging operator in $L_p(S, \Sigma, \mu)$ is obtained by taking a Radon-Nikodym derivative relative to a suitably defined subfield.

Before embarking upon a proof of this result, consider the intuitive meaning of the crucial property (2) of averages. This property was obtained inductively from the study of certain operators introduced by Osborne Reynolds in turbulence theory; it also appears in Tarski's quantification axiom (see (6)). Algebraically it can be considered as a strenghthening of the property of linearity of the operator A, for it states that A remains linear when the vector space is endowed with the richer structure of module over a certain subring of the ring of essentially bounded functions (cf. Lemma 2 below).

3. Main Representation Theorem. - We begin by establihing some unavoidable technical lemmas, some of them of independent interest. It is always tacitly assumed that p is a real number greater than one.

LEMMA 1. Let g be an essentially bounded function on (S, Σ, μ), and let $\mathcal{A}$ be the collection of all functions of the form $p(g)$, where p is a real polynomial. Let $\Sigma(g)$ be the σ-field generated by sets of the form $\{s \mid g(s) \in F\}$, where F is a Borel set on the line. Then the closure of $\mathcal{A}$ in $L_p(S, \Sigma, \mu)$ is the subspace $L_p(S, \Sigma(g), \mu)$.

Proof. A measure μ_g on the line (the « distribution function » of g) is defined by the formula

$$\mu_g(F) = (\{s \mid g(s) \in F\})$$

where F is any Borel set. Since g is bounded, the measure μ_g has compact carrier, say the interval $[r, s]$. For any real-valued Borel function h on the line we have the identity

$$\int_S h(g(s))\mu(ds) = \int_r h(t)\mu_g(dt).$$

In particular, if E is a set in $\Sigma(g)$, and $g(E) = F$, then F

is a Borel set, and

$$\int_S \chi_E(s)\mu(ds) = \int_S \chi_F(g(s)\mu(ds) =$$

$$= \int_S \chi_F(t)\mu_g(dt).$$

Since the characteristic function χ_F is a pointwise limit everywhere on $[r, s]$ of a sequence of polynomials in the real variable t, it follows from the bounded convergence theorem that χ_F is a limit of polynomials in $L_p([r, s], \Sigma, \mu)$ and hence, that χ_E is a limit in norm of a sequence of polynomials in the function g in $L_p(S, \Sigma, (g), \mu)$. Thus the characte ristic function χ_E lies in the closure of $\mathcal{A}$. The statement then follows from the fact that simple functions are dense in $L_p(S, \Sigma, \mu)$.

PROPOSITION 1. (Birkhoff). Every average is a projection (that is, $A^2 = A$).

Proof. By (2) and (3) we have

$$A^2 f = A(IAf) = (AI)(Af) = Af.$$

LEMMA 2. Let g be an essentially bounded function for which $Ag = g$, and let $\Sigma(g)$ be the σ-field generated by g, as in Lemma 1. Then for every measurable set E in $\Sigma(g)$ we have

$$A\chi_E = \chi_E.$$

Proof. If g is fixed under A and bounded, then g^2 is also bounded and fixed under A, for by (2) and Prop. 1 we have

$$A(g^2) = A(gA(g)) = (Ag)^2 = g^2.$$

It follows by induction that $A(g^n) = g^n$ for any positive integer n, and hence that $A(p(g)) = p(g)$ for any real polynomial p. Lemma 1 and the uniform continuity of A entail that $Af = f$

for any f in $L_p(S, \Sigma(g), \mu)$. The assertion of the Lemma is a special case of this conclusion.

Proposition 2. Every averaging operator in $L_p(S, \Sigma, \mu)$ maps essentially bounded functions into essentially bounded functions.

Proof. Let f be essentially bounded. From the knowledge that A is bounded in L_p one can only infer that the function $g = Af$ is of class L_p. However, the « averaging identity » (2) shows that $A(fg) = g^2$ is also of class L_p, and by induction that all powers $g^n = A(fg^{n-1})$ are of class L_p for any positive integer n. Applying (1) and Hölder's inequality we find that

$$\int_S |g(s)|^{np} \mu(ds) \leq \int_S |(Afg^{n-1})(s)|^p \mu(ds) \leq$$

$$\leq \int_S |f(s)|^p |g(s)|^{(n-1)p} \mu(ds) \leq$$

$$\leq \left(\int_S |f(s)|^{np} \mu(ds)\right)^{1/n} \left(\int_S |g(s)|^{np} \mu(ds)\right)^{1-(1/n)}$$

and hence that

$$\left(\int_S |g(s)|^{np} \mu(ds)\right)^{1/n} \leq \left(\int_S |f(s)|^{np} \mu(ds)\right)^{1/n} \leq |f|_\infty^p$$

for all positive n. Letting n tend to infinity we find (cf. Loomis [3], p. 39, Thm. 14F) that the left hand side tends to $|g|_\infty^p$, which is thus seen to be finite, q. e. d.

Lemma 3. The equality

$$\int_S (Af)(s)\mu(ds) = \int_S f(s)\mu(ds)$$

holds for every function f in $L_p(S, \Sigma, \mu)$.

Proof. Let A^* be the adjoint operator of A in $L_q(S, \Sigma, \mu)$ $(p^{-1} + q^{-1} = 1)$. We shall first prove that $A^*I = I$. Note that for $g(s) = (A^*I)(s)$ we have by (3)

$$\int_S g(s)\mu(ds) = \int_S I(s)(AI)(s)\mu(ds) = 1.$$

Secondly, from

$$| g |_q = \sup_{|f|_p \leq 1} | \int_S g(s)f(s)\mu(ds) | =$$
$$= \sup_{|f|_p \leq 1} | \int_S (Af)(s)\mu(ds) | = 1$$

we infer that

$$1 = \int_S g(s)\mu(ds) \leq | g |_q | I |_p = 1.$$

Since equality obtains in Hölder's inequality, we conclude ((1), III.9.42) that $g = I$ almost everywhere.

Now to the proof of the Lemma. For f of class L_p we obtain

$$\int_S (Af)(s)\mu(ds) = \int_S g(s)f(s)\mu(ds) = \int_S f(s)\mu(ds),$$

q.e.d. [4]).

The proof of the main Theorem is now almost concluded.

THEOREM 1. Let A be an averaging operator in $L_p(S, \Sigma, \mu)$ (for fixed p). Then there exists a unique σ-subfield Σ' of Σ such that $Af = f'$ for f in L_p, where f' is the Radon-Nikodym derivative of f relative to the σ-field Σ'.

Proof. Let Σ' be the smallest σ-field containing all σ-fields $\Sigma(g)$, where g is any bounded function fixed under A. It is immediate from Lemma 2 and assumption (2) that for every E in Σ', $A\chi_E = \chi_E$. Therefore, by Lemma 3 and (2), for all E in Σ',

$$\int_S (Af)(s)\mu(ds) = \int_S \chi_E(s)(Af)(s)\mu(ds) =$$
$$= \int_S (A(\chi_E f))(s)\mu(ds) = \int_E f(s)\mu(ds).$$

[4]) The preceding argument does not include the case $p = 1$, but this case is even simpler.

If the function f is bounded, then Af is Σ'-measurable. Therefore the first and last members in the above equalities now show that Af is the Radon-Nikodym derivative f' of f relative to the σ-field Σ'. Thus $Af = f'$ whenever f is essentially bounded. Since A has a unique extension to the entire space, the proof is complete.

4. Averages of Essentially Bounded Functions. - Two properties of averaging operators in $L_p(S, \Sigma, \mu)$ for finite p which follow at once from the preceding Theorem are the following:

(A) An averaging operator is uniquely determined by its range.

(B) Every averaging operator is selfadjoint. That is, if the operator A, as well as its adjoint A^*, are restricted to the dense subspace of essentially bounded functions, then $A = A^*$. This follows from the identity for Radon-Nikodym derivatives, for bounded f and g,

$$\int fg' = \int f'g' = \int f'g.$$

Neither (A) nor (B) hold for averaging operators in $L_\infty(S, \Sigma, \mu)$. This happens because there is a far wider variety of averages of essentially bounded functions. We derive below the structure of such operators, as far as it can be described in full generality.

PROPOSITION 3. Every averaging operator in $L_\infty(S, \Sigma, \mu)$ is an order-preserving transformation.

Proof. We can find an order-preserving isometric isomorphism F of the algebra $L_\infty(S, \Sigma, \mu)$ onto $C(S_1)$, the B-algebra of all continuous functions on a compact Hausdorff space S_1 (for a proof, see (1), V.8.11). It follows that if A is an averaging operator in L_∞, then the operator $A' = FAF^{-1}$ is a bounded operator in the B-space $C(S^1)$ with the following three properties:

(1′) A' is a contraction,

(2′) $A'(fAg) = A'f\ A'g$

for any tho continuous functions f and g,

(3′) $A'I = I$.

Since the isomorphism F is order-preserving, the assertion will be established if it is shown that the operator A' preserves order. If this were not the case, then we could find a real function f in $C(S_1)$, which without loss of generality we can assume to take values between 0 and 1, such that $(A'f)(s_0) = -\varepsilon < 0$ at some point s_0 in S_1. Therefore $(A'(I - f))(s_0) = 1 + \varepsilon > 1$. Since the function $1 - f$, lies between 0 and 1, this contradicts the assumption that A′ is a contraction, q.e.d.

From the preceding proof it is seen that the family of averaging operators in $L_\infty(S, \Sigma, \mu)$ is «isomorphic» to the family of all averaging operators in $C(S_1)$, as defined by (1′) — (3′). The latter operators have been studied by G. Birkhoff in (7), J. Sopka in (9) and J. L. Kelley in (10). Using the results of these authors together with the representation of L_∞ as a C-space, one can obtain some information about general averaging operators in L_∞. The following discussion is therefore limited to those results which do not follow from such a representation technique.

Let $\Sigma(A)$ be the field of all measurable sets whose characteristic functions are invariant under A. We call $\Sigma(A)$ the field *associated with* the averaging operator A. In general, $\Sigma(A)$ is not a σ-field. One can characterize the range of the operator A as follows.

PROPOSITION 4. Let A be an averaging operator in $L_\infty(S, \Sigma, \mu)$, and let $\Sigma(A)$ be the associated field. Then the range of A is the subspace of all essentially bounded functions f such that, for any real a and b (possibly infinite) the set

$$[4] \qquad [s \mid a < f(s) \leq b]$$

belongs to the field $\Sigma(A)$.

Proof. By Prop. 1 the range of A consists of all functions invariant under A. Let f be one such function, and let E be the set [4]. Since the product of two invariant characteristic functions is again invariant, we can assume without loss of generality that $a = 0$ and b is infinite. From the identity $A(\chi_E f) = fA\chi_E$ we infer that the function $fA\chi_E$ is non-negative; therefore $A\chi_E$ is supported on E, that is, $A\chi_E = \chi_E A\chi_E$. Let $E' = S - E$; then $I = A\chi_E + A\chi_{E'} = \chi_E A\chi_E + \chi_{E'} A\chi_{E'}$ by (3). Multiplying both sides by χ_E we obtain $\chi_E = A\chi_E$, so that χ_E is invariant, as we wanted to show.

Conversely, suppose that f safisfies the condition of the Proposition. We show that f is invariant under A. If $Af \neq f$ on a set of positive measure, then there exist real numbers a and b such that the sets in [4] for f and Af do not coincide almost everywhere. We can assume without loss of generality that $a = 0$ and b is infinite, and that the set where Af is non-negative does not contain the set E in [4], where f is non-negative (otherwise take $-f$). Thus χ_E is invariant under f, by hypothesis, but $\chi_E Af$ is negative on a set of positive measure. This contradicts Prop. 3.

In virtue of Prop. 4, every function f invariant under A is measurable relative to the σ-field $\Sigma'(A)$ generated by the associated field $\Sigma(A)$. The converse of this statement is false in general; and this indicates that the introduction of purely finitely additive measures is essential to the problem.

Let ν be a finitely additive real measure on the σ-field Σ and vanishing on all μ-null sets. It is well-known that the indefinite integral

$$\nu(f;\ E) = \int_E f(s)\nu(ds), \qquad E \text{ in } \Sigma$$

is well-defined for all essentially bounded functions f. Now suppose that the restriction of the measure ν to a given sub-σ-field Σ' is countably additive. Then the set function $\nu\ (f;\ E)$ restricted to Σ' is absolutely continuous relative to μ, and the Radon-Nikodym derivative of $\nu(f;\ E)$ is well-defined; we call it the Radon-Nikodym *derivative of the pair* $(f,\ \nu)$ relative to the σ-field Σ'.

THEOREM 2. Let A be an averaging operator in $L_\infty(S, \Sigma, \mu)$ with associated field $\Sigma(A)$, and let $\Sigma'(A)$ be the σ-field generated by $\Sigma(A)$. Then there exists a unique positive finitely additive measure ν_A defined on Σ and vanishing on μ-null sets, such that:

(a) for every E in $\Sigma'(A)$, $\nu_A(E) = \mu(E)$,

(b) for every essentially bounded function f, the function Af is the Radon-Nikodym derivative of the pair (f, ν_A) relative to the σ-field $\Sigma'(A)$.

Proof. We recall that the conjugate space of $L_\infty(S, \Sigma, \mu)$ is the B-space $ba(S, \Sigma, \mu)$ of all finitely additive set functions which vanish on μ-mull sets (cf. (1), IV.8.16). Denoting again by A^* the adjoint operator of A. let $\nu_A = A^*\mu$. We have then the identity

$$\int_E (Af)(s)\mu(ds) = \int_S A(\chi_E f)(s)\mu(ds) = \tag{5}$$

$$= \int_S \chi_E(s)f(s)\nu_A(ds) = \int_E f(s)\nu_A(ds)$$

for E in $\Sigma(A)$. In particular for $f = I$ we obtain $\nu_A(E) = \mu(E)$; therefore ν_A is countably additive on $\Sigma(A)$, and by a well-known theorem of Carathéodory also countably additive on the σ-field $\Sigma'(A)$. Thus the Radon-Nikodym derivative of the pair (f, ν_A) is well-defined. and [5] holds for all E in $\Sigma'(A)$. The first and last members of [5] show that Af is the Radon-Nikodym derivative of the pair (f, ν_A) relative to $\Sigma'(A)$. The uniqueness and positivity of the measure ν_A are also immediately inferred from [5].

The converse of Thm. 2 is false, as may be expected from its counterpart in $C(S)$, as shown in (7). This raises the question of whether there is a natural characterization among all averaging operators with the same range of the one which is obtained by taking the Radon-Nikodym derivative relative to the σ-field $\Sigma'(A)$, as in Thm. 1. Such a characterization is derived below. We say that a bounded operator A in L_∞ is *weakly continuous* when it is continuous in the weak L_1-topology of L_∞.

PROPOSITION 5. Let A be an averaging operator in $L_\infty(S, \Sigma, \mu)$. The associated field $\Sigma(A)$ is a σ-field and $Af = f'$ (the Radon-Nikodym derivative of f relative to the σ-field $\Sigma(A)$) if and only if $A^*\mu = \mu$.

Proof. The necessity of the condition is clear, in virtue of Remark (B). To establish the sufficiency, we remark that the assumption gives the identity

$$\int_S f(s)\mu(ds) = \int_S (Af)(s)\mu(ds)$$

for all essentially bounded f. Writing $f = f^+ - f^-$, where f^+ and f^- are non-negative, we get by Prop. 3 that $(Af^+)(s) + (Af^-)(s) \geq |(Af)(s)|$, and hence

$$\int_S |f(s)|\mu(ds) = \int_S [f^+(s) + f^-(s)]\mu(ds) \geq \int_S |(Af)|(s)\mu(ds)$$

Therefore A can be extended to an averaging operator in $L_1(S, \Sigma, \mu)$, and the result follows from Thm. 1 (*).

Remark. The assumption of weak continuity of A is not strong enough to give the conclusion of Prop. 4. To see this, it suffices to consider the case when S is a finite atomic measure space, and apply the Theorem of Birkhoff (6) for continuous functions, noting that in this simple case the conditions given there for an averaging operator in $C(S)$ are also sufficient. Such an example also shows that in general the adjoint operator of an averaging operator is not an averaging operator.

(*) Dr. P. C. Shields of M. I. T. was kind enough to point out an oversight in the first draft of this proof, and to suggest an improvement.

BIBLIOGRAPHY

(1) DUNFORD and SCHWARTZ: *Linear Operators*, Vol. I, New York, 1959.

(2) LOÈVE: *Probability Theory*, New York, 1955.

(3) LOOMIS: *Abstract Harmonic Analysis*, New York, 1953.

(4) Mme. M.-L. DUBREIL - JACOTIN: *Étude Algébrique des Trasformations de Reynolds*, Colloque d'Algèbre superieure, CBRS, Louvain, 1957.

(5) J. KAMPÉ DE FÉRIET: *La notion de moyenne dans la théorie de la turbulence*, Rendiconti Seminario Matematico e Fisico di Milano (27), (1955-56), pp. 1-43.

(6) G. BIRKHOFF: *Moyennes des Fonctions Bornées*, Colloque Internationale No. 24 du CNRS, Paris 1950.

(7) G. BIRKHOFF: *Averaging Operators*, Symposium in Lattice Theory, AMS, 1960.

(8) SHU-TEH-CHEN-MOY: Pacific Journal of Mathematics, 4 (1954), pp. 47-63.

(9) J. SOPKA: Harvard Doctoral Dissertation, 1950.

(10) J. L. KELLEY: *Averaging Operators in $C_\infty(X)$*, Illinois J. M. 2 (1958), pp. 214-223.