Riv. Mat. Univ. Parma (4) 3 (1977), 141-149

G. TEPPATI (*)

A set-theoretical approach to Lebesgue - Stieltjes measure. (**)

1. - Introduction.

It is often natural to study integrals from the point of view of considering finite-valued continuous monotone functions on the real line, particularly in probability and classical applied mathematics. This leads to a wide class of definitions of integrals, and specifically to the well-known Lebesgue-Stieltjes integral. In the section 2 of this Note we analize an approach to this kind of integral from a different point of view, which may be considered as a constructive one, and which only uses sets and mapping between them. In the section 3 we examine a suitable definition of measurable mappings, and we discuss an analogue of the Lebesgue-Stieltjes integral, which in this context can be understood as a real number which relates each other two mappings. Some elementary properties of this kind of integration are sketched. In the section 4 some applications to probability theory are briefly discussed.

2. - A Lebesgue-Stieltjes measure.

For our purposes we need a well-known general method for constructing measures [1]. Thus we recall the following definitions (we suppose from now on that the set X, and later Y, always are non-void sets).

(*) Indirizzo: Istituto di Matematica, Università, 43100 Parma, Italy.
(**) Work supported by C.N.R. (Italy) (under the contract n. 73.00283.02). — Ricevuto: 20-XII-1976.

Definition 1. Let X be a set, and let $\mathscr{C}$ be a class of subsets of X, containing $\emptyset$. A mapping $\tilde{\mu}: \mathscr{C} \to R$ (R the extended real line) will be called a *pre-measure* on X if and only if: (i) $\forall C \in \mathscr{C}$, $0 \leqslant \tilde{\mu}(C) \leqslant +\infty$; (ii) $\tilde{\mu}(\emptyset) = 0$. Let now $P(X)$ be the set of all subsets of X. A mapping $\mu: P(X) \to R$, will be called a *measure* (¹) on X if and only if: (a) $\forall S \in P(X)$, $0 \leqslant \mu(S) \leqslant +\infty$, $\mu(\emptyset) = 0$; (b) $\forall S_1, S_2 \in P(X)$, $S_1 \subset S_2 \Rightarrow \mu(S_1) \leqslant \mu(S_2)$; ($c$) if $\{S_i\}_{i \in I}$ (I the set of the positive integers) is any family of sets of X, then $\mu(\bigcup_{i=1}^{\infty} S_i) \leqslant \sum_{i=1}^{\infty} \mu(S_i)$.

Definition 2. Let X be a set and let $\mathscr{C}$ be a class of sets which satisfies to: (i) $\emptyset \in \mathscr{C}$; (ii) $\forall S \in P(X)$ a family $\{C_i\}_{i \in I}$ of sets of $\mathscr{C}$ exists such that $\bigcup_{i=1}^{\infty} C_i \supset S$ (we conventionally exclude the finite case). We call $\mathscr{C}$ a *sequential covering class* in X, and for any S we call the family $\{C_i\}_{i \in I}$ the *sequential covering* of S.

The link between pre-measures on X, defined possibly on a smaller class than $P(X)$, and measures on X, defined on the entire $P(X)$, is given by the following well-known result [**1**].

Theorem 1. *Let X be a set and let $\tilde{\mu}$ be a pre-measure defined on a class $\mathscr{C}$ of sets of X, which is a sequential covering class in X. Then the set function*

$$\mu(S) = \inf_{C_i \in \mathscr{C} | \bigcup C_i \supset S} \sum_{i=1}^{\infty} \tilde{\mu}(C_i)$$

is a measure on X.

This result together with the definition above is the motivation for the next concepts we want to specify. At this stage the following definition of pre-metrizability will be useful.

Definition 3. Let X be a set. Let $\varrho: X \times X \to R$ a mapping of the Cartesian product $X \times X$ into R such that $\forall x, x' \in X$, $\varrho(x, x') \geqslant 0$ (symmetry *not* required). A couple (X, ϱ) will be called a *pre-metrizable* set. Let now (X, ϱ) be a pre-metrizable set. Let we define the following non-negative number d, by means of

(i) $\quad \forall S \in P(X)$, $d(S) = \sup_{x, x' \in S} \varrho(x, x')$,

(ii) $\quad d(\emptyset) = 0$.

(¹) Also called *outer measure*.

We call d a *pre-diameter* (2) of S.

We are now in position to give in the next definition our set-theoretical analogue of a Lebesgue-Stieltjes pre-measure.

Definition 4. Let X, Y be two sets. Let $f: D_f \to Y$ be a mapping, D_f being a non-void subset of X. Let (Y, ϱ) be a pre-metrizable set, and let d be a pre-diameter. Let we call $f(S)$ the set $f(S) = \{y \in Y \mid \forall x \in S, f(x) = y\}$. Let we define a set function $\tilde{\mu}_f$ by means of (α) $\forall S \subset D_f$, $\tilde{\mu}_f(S) = d(f(S))$; ($\beta$) $\tilde{\mu}_f(\emptyset) = 0$.

We call μ_f a *set-theoretical analogue of the Lebesgue-Stieltjes pre-measure,* and f the *distribution mapping* for the pre-measure $\tilde{\mu}_f$.

We report here the following result, which is an immediate consequence of Theorem 1.

Theorem 2. *Let X, Y, f and $\tilde{\mu}_f$ be as in Def. 4. Let D_f coincide with the entire X. Then the set function*

$$\forall S \in P(X)\,, \qquad \mu_f(S) = \inf_{C_i \in \mathscr{C} \mid \bigcup C_i \supset S} \sum_{i=1}^{\infty} \tilde{\mu}_f(C_i)$$

is a measure on X.

The proof follows the standard one of Theorem 1, and for this reason will not be reproduced here.

In the last definition of this section we collect some standard notions concerning measurability of sets.

Definition 5. Let X, Y, f and μ_f be as above. We call μ_f a *set-theoretical analogue of the Lebesgue-Stieltjes measure,* and f the *distribution mapping* for the measure μ_f. Moreover, as usual, a subset S of X is said to be *μ_f-measurable* if for any set $T \subset X$, $\mu_f(T) > \mu_f(T \cap S) + \mu_f(T - S)$ ($T - S$ is the complement of S with respect to T). We call M_f the class of all μ_f-measurable subsets of X.

Thus in particular if $\mu_f(S) = 0$, it follows that S is μ_f-measurable.

3. - Measurable mappings and integrals.

In this section we briefly discuss the concept of measurability of a mapping and, consequently, its integration on subsets.

---

(2) We observe that the pre-diameter d is also a pre-measure, and satisfies to the requirements (a), (b) for a measure in Def. 1. Moreover we explicitely point out the fact that not necessarily $\varrho(x, x) = 0$.

It is usual to characterize measurable mappings in the following way. Let X be any set with a measure μ, let Y be another set, and let g be a mapping $g: D_g \to R_g$ (3) $(R_g \subset Y)$, D_g being a measurable non-void subset of X. Let $\mathcal{G}$ be a sequential covering class in Y; then the mapping $g: D_g \to R_g$ is said to be a measurable mapping if for any set $D \in \mathcal{G}$, $D \subset R_g$, $g^{-1}(D)$ is a μ-measurable set.

However we will proceed in a slightly more constructive way, which is a direct set-theoretical translation of the elementary procedure for measurability of ordinary functions.

Definition 6. Let X, Y be two sets. Let (Y, ϱ) be a pre-metrizable set together with a pre-diameter d. Moreover let Y be such that a non-void set, e, exists, which separates Y in the weak sense that two distinct-each-other (and from e) non-void subsets Y_1, Y_2 exist satisfying to $Y = Y_1 \cup Y_2 \cup e$. Let g be any mapping $g: D_g \to Y$ (D_g a non-void subset of X). We will call *height of g over a non-void subset $S \subset D_g$* the mapping $H: Y \to R$ defined by the following relations

(i) $H(g(S)) = d(g(S) \cup e)$ whenever $g(S) \subset Y_1$ and $g(S) \cap Y_2 = \emptyset$;

(ii) $H(g(S)) = - d(g(S) \cup e)$ whenever $g(S) \subset Y_2$ and $g(S) \cap Y_1 = \emptyset$;

(iii) $H(g(S)) = 0$ whenever $g(S) \subset e$ or $S = \emptyset$;

(iv) $H(g(S)) = d((g(S) \cap Y_1) \cup e) - d((g(S) \cap Y_2) \cup e)$ whenever $g(S)$ is not as above.

We call *null function* the mapping g such that $H(g(S)) = 0$ for any subset S of D_g.

Thus we can give our definition for a measurable mapping.

Definition 7. Let X, Y be two sets, and $f: X \to Y$ be a mapping. Let we suppose that the requirements from Def. 1 to Def. 5 of the preceding section are satisfied, including the ones of Theorem 2, and let μ_f be the set-theoretical analogue of the Lebesgue-Stieltjes measure, as previously constructed. Let g be any mapping $g: D_g \to Y$ and let the requirements of Def. 6 be satisfied. Let D_g be a μ_f-measurable subset of X. Let S be a non-void subset of D_g. We will say that *g is μ_f-measurable on S* if an only if for any sequential covering $\{C_i\}_{i \in I}$ of S in D_g, for any $i \in I$, for any $\alpha_i \in R$ such that $H(g(C_i)) = \alpha_i$, the sets C_i belong to M_f.

(3) Where R_g is the range of g.

At this point we can proceed to our definition of the analogue of the Lebesgue-Stieltjes integration.

Definition 8. Let X, Y be two sets, and f, g be two mappings submitted to the requirements from Def. 1 to Def. 6 (including the ones of Theorem 2). Let S be a non-void subset of D_g, and let g be μ_f-measurable over S, in the sense of Def. 7. Let any sequential covering $\{C_i\}_{i\in I}$ of S be made of subsets of D_g. We will call *analogue of the Lebesgue-Stieltjes integral of g over S* the real number $\int_S g \, \mathrm{d}\mu_f$ so defined

$$\int_S g \, \mathrm{d}\mu_f = \inf_{C_i \in \mathscr{C} / \bigcup_{i=1}^{\infty} C_i \supset S} \sum_{i=1}^{\infty} H(g(C_i))\, \mu_f(C_i) \, .$$

A μ_f-measurable mapping g such that all the above series are absolutely convergent (4) $\left(\int_S g\, d\, \mathrm{d}\mu_f < +\infty\right)$ will be said μ_f-integrable.

In this latter definition we have adopted the convention $0. \;\infty = 0$ which is all special to measure theory. We also observe that the null mapping is μ_f-measurable and μ_f-integrable on any subset.

We collect in the following theorem some elementary properties.

Theorem 3. *Let X, Y, f, g be as above. Let g be μ_f-measurable on two non-void subsets S_1, S_2 such that $S_1 \subset S_2$. Then*: (a) $\int_{S_1} g \, \mathrm{d}\mu_f \leqslant \int_{S_2} g \, \mathrm{d}\mu_f$. *Let moreover S be such that $\mu_f(S) = 0$. Then*: (b) *for any g, μ_f-measurable on S, we have* $\int_S g \, \mathrm{d}\mu_f = 0$. *Let $S_1, \ldots, S_n$ be disjoint μ_f-measurable subsets, let $S = \bigcup_{k=1}^{n} S_k$, and let g be a μ_f-integrable mapping on each set S_k. Then*: (c) *g is μ_f-integrable over S and* $\int_S g \, \mathrm{d}\mu_f \leqslant \sum_{k=1}^{n} \int_{S_k} g \, \mathrm{d}\mu_f$. *Moreover let g_1 and g_2 be two μ_f-integrable mappings over a non-void μ_f-measurabke set $S \subset D_{g_1} \cap D_{g_2}$. Let $\{C_i\}_{i\in I}$ be a sequential covering of S. Let the heights of g_1 and g_2 be such that $|H(g_1(C_i))| = |H(g_2(C_i))|$ for any $i \in I$ except on a subclass whose members all have μ_f-measure 0. Then*: (d) $\int_S g_1 \, \mathrm{d}\mu_f = \int_S g_2 \, \mathrm{d}\mu_f$. *At last, let $\mathscr{G}$ be a set of μ_f-integrable mappings over a same set S, with the following properties*: (i) *for any two mappings $g_1, g_2 \in \mathscr{G}$ a mapping $g \in \mathscr{G}$ exists such that* $\int_S g \, \mathrm{d}\mu_f = \int_S g_1 \, \mathrm{d}\mu_f + \int_S g_2 \, \mathrm{d}\mu_f$; (ii) *for any $\lambda \in R$ and $g \in \mathscr{G}$ a mapping $g_1 \in \mathscr{G}$ exists such that* $\int_S g_1 \, \mathrm{d}\mu_f = \lambda \int_S g \, \mathrm{d}\mu_f$; (iii) *the null mapping belongs to $\mathscr{G}$. Then*: (e) *$\mathscr{G}$ is a vector-space over R.*

---

(4) Otherwise, the number $\int_S g \, \mathrm{d}\mu_f$ is $-\infty$, for the Mertens theorem.

Proof. (a) follows directly by observing that a sequential covering of S_2 exists which contains a sequential covering of S_1. (b) comes out from the fact that a sequential covering of S exists such that for any member of it, $\mu_f(C_i) \leqslant \leqslant \mu_f(S) = 0$. (c) is proved by observing that for any C_i, $C_i = \bigcup_k C_{ki}$, where C_{ki} is an element of a sequential covering of S_k. (d) is proved by direct inspection. (e) follows when we define an abelian group law $+$ (with the null mapping as 0 element) and an external law as: $\int_S (g_1 + g_2)\, d\mu_f = \int_S g_1\, d\mu_f + \int_S g_2\, d\mu_f$ and $\int_S \lambda \cdot g\, d\mu_f = \lambda \int_S g\, d\mu_f$.

The proof is thus completed.

In the last theorem of this section we obtain, under a somewhat weak assumption, an analogous of a Radon-Nikodym type theorem. All the sets are supposed to be different from the void set.

Theorem 4. *Let X, Y, f and μ_f as above. Let σ be an everywhere finite set function such that*

(α) $\forall S \in M_f$, $\mu_f(S) \leqslant \sigma(S)$;

(β) *for any sequential covering of disjoint subsets of any set*

$$S \in M_f\colon \sum_{i=1}^{\infty} \sigma(C_i) = \sigma(S) .$$

Then a μ_f-integrable non-null mapping $g\colon X \to Y$ exists such that: $\int_S g\, d\mu_f \leqslant \leqslant \sigma(S)$.

Proof. Let Λ be any sequence of positive terms, all of them $\leqslant 1$, and let $\{C'_i\}_{i\in I}$ be a sequential covering of S. We define a function g on the whole X by posing

$$\forall i , \qquad H(g(C'_i)) = \Lambda_i \qquad \text{and} \qquad H(g(X - C'_i)) = 0 .$$

Thus, for any measurable set S

$$\int_S g\, d\mu_f = \inf \sum_{i=1}^{\infty} H(g(C_i))\, \mu_f(C_i) \leqslant \sum_{i=1}^{\infty} H(g(C'_i))\, \mu_f(C'_i)$$

$$\leqslant \sum_{i=1}^{\infty} \lambda_i \mu_f(C'_i) \leqslant \sum_{i=1}^{\infty} \mu(C'_i) \leqslant \sum_{i=1}^{\infty} \sigma(C'_i) = \sigma(S) .$$

The proof is then concluded.

We consider now a specific example, which will be useful for the applications we have in mind.

Let X be a set, and let Y be a Boolean algebra of projections of an abstract Hilbert space H, whose elements we denote with h, and whose scalar product and norm we denote with $\{\,,\}$ and $\|\;\|$ respectively. Let $\mathscr{C}$ be a sequential covering class of X. Let P be a mapping $P\colon X \to Y$ such that: (i) for any family C_i of mutually disjoint sets of $\mathscr{C}$, $\sum_{i=1}^{\infty} P(C_i) = P(\bigcup_{i=1}^{\infty} C_i)$; (ii) for any $C_1, C_2 \in \mathscr{C} P(C_1 \cap C_2) = P(C_1)P(C_2)$ and $P(C_1 \cup C_2) = P(C_1) + P(C_2) - P(C_1) \cdot \cdot P(C_2)$; $P(\emptyset) = 0_H$ and $P(X) = 1_H$ (respectively the null and the identity mapping of H). The mapping P is usually called a spectral measure. Let we define $\varrho(P(C_1), P(C_2)) = \{P(C_1)h, P(C_2)h\}$ for a fixed h in H; then for any $C \in \mathscr{C}$, $d(P(C)) = \{h, P(C)h\}$, from which we have the pre-measure $\tilde{\mu}_P^h(C) = \{h, P(C)h\}$, and the measure, defined on all subsets of X, $\mu_P^h(S) = \inf_{C_i \in \mathscr{C} | \cup C_i \supset S} \cdot \sum_{i=1}^{\infty}\{h, P(C_i)h\}$. Thus P is the distribution mapping for this measure, and $h \in H$ labels the so obtaine measures. We observe that, whenever we define $\varrho(P(C_1), P(C_2)) = \sup_{\|h\|=1}\{P(C_1)h, P(C_2)h\}$, we have the analogous result $\mu_P(S) = \inf_{C_i \in \mathscr{C} | \cup C_i \supset S} \cdot \sum_{i=1}^{\infty} N(P(C_i))$, where $N(P(C_i)) = \sup_{\|h\|=1} \|P(C_i)h\|$: in this latter case $\mu_P(X) = 1$.

4. - An application to probability.

Let X be a set and let μ be a measure on its subsets: μ is called a probability measure whenever $\mu(X) = 1$. We observe that a probability measure can be easily obtained from a given measure: for instance the set functions $\mu/(1+\mu)$ (whenever $\mu(X) = +\infty$), $\inf(1, \mu)$, $\mu/\mu(X)$ (whenever $\mu(X) < +\infty$) are all set functions induced by μ which satisfy requirements (a) to (c) of Def. 1, together with the above, so all they are probability measures A probability measure will be denoted by p.

We remember now that a general description of classical probability consists of a set X and of a probability measure p on X. An event is an element of $P(X)$ which is a measurable set and the probability of the event is the measure on it.

We outline here another description, which makes use of the notions discussed in the preceding section.

Definition 9. Let X, Y be two sets, f a mapping between them, and let X, Y satisfy to the requirements from Def. 1 to Def. 6 of the preceding sections. Let p_f be a probability measure induced by a measure μ_f described above. Let P_f be the class of the p_f-measurable subsets of X. We call a member of P_f an *event*, the p_f-measure of it the *probability*, and the probability measure p_f a *state* with respect to the distribution mapping f of X, which, in this context, will be called the *density state function*. Let now g be a p_f-measurable mapping between X and Y: we call g an *observable*. Then $\forall S \subset R_g \subset Y$ we call the number $p_f(g^{-1}(S))$ the probability, with respect to the state p_f, that a measure on the observable g will be in S. At last, let g be a p_f-integrable observable. We call *expectation value* of g in S with respect to the density state function f the integral $\int_S g\,\mathrm{d}p_f$ and we write: $e_f^S(g) = \int_S g\,\mathrm{d}p_f$.

The scheme outlined in this latter definition is general enough to allow both classical and quantum like probability to enter as particular cases: at this purpose, the remark at the end of the preceding section shows that a quantum scheme-follows whenever Y is an abstract Boolean algebra of projections of an abstract Hilbert space, and the observable g is a spectral measure [2] (5). Moreover, as a trivial consequence of the Theorem 4 of the preceding section, we have the result that, given a set function σ, $0 \leqslant \sigma \leqslant 1$, which satisfies its assumptions, an observable g exists for any S such that $e_f^S(g) \leqslant \sigma(S)$.

We introduce now a structure on the set of the density state function, in the following way.

Definition 10. Let P_F be the set $P_F = \{p_f \,|\, f \text{ is a density state function}\}$. We call P_F a *state space* if $\forall$ real $\lambda_1, \lambda_2, \ldots, > 0$, such that $\sum_{k=1}^{\infty} \lambda_k = 1$ and for any sequence $p_{f_1}, p_{f_2}, \ldots$ the probability measure $\sum_{k=1}^{\infty} \lambda_k p_k$ belongs to P_F. Let now we call F the set $F = \{f \,|\, f \text{ is a density state function for } P_F\}$. Let g be be a given observable. We define in F two composition laws, $+'$ and, respectively interval and external, by the following setting

$$\forall \lambda_1, \lambda_2 \in R\,, \qquad \lambda_1, \lambda_2 > 0\,, \qquad \lambda_1 + \lambda_2 = 1\,,$$

$$\forall f_1, f_2 \in F\colon\ \lambda_1 e_{f_1}(g) + \lambda_2 e_{f_2}(g) = e_{\lambda_1 \cdot f_1 +' \lambda_2 \cdot f_2}(g)\,.$$

At least, we call $v_f(g)$ (the *variance* of g in the state f) the following number: $\forall g\colon\ v_f(g) = e_f(gg) - e_f^2(g)$ (we omit, here and above, the subset S as an index).

From this latter definition we have the following result.

(5) In this way, f can be thought to be roughly a squared modulus of a wave function.

Lemma. *Let X, Y be as above. Let F be the convex set of the preceding Def. 10. Then, for any* $f_1, f_2 \in F$, *for any* $\lambda_1, \lambda_2 > 0$ *and such that* $\lambda_1 + \lambda_2 = 1$, *we have*

$$v_{\lambda_1 \cdot f_1 +' \lambda_2 \cdot f_2}(g) = \lambda_1 v_{f_1}(g) + \lambda_2 v_{f_2}(g) + \lambda_1 \lambda_2 (e_{f_1} - e_{f_2})^2(g) .$$

Proof. We observe that, e.g. $e_{f_1}(gg) = e_{f_1}^2(g) + v_{f_1}(g)$: thus, multiplying respectively by λ_1 and λ_2 the two relations so obtained and subtracting

$$e^2_{\lambda_1 \cdot f_1 +' \lambda_2 \cdot f_2}(g) = \lambda_1^2 e_{f_1}^2(g) + 2\lambda_1 \lambda_2 e_{f_1}(g)\, e_{f_2}(g) + \lambda_2^2 e_{f_2}^2(g)$$

we have the result.

In conclusion a scheme is sketched which tries to fit both the classical and the quantum case, and which seems at the same time flexible enough to include more general theories. We remark that further possibilities of extending our formalism are given by extending the number field to an ordered field, even not commutative, as far as possible. Results on this direction will published elsewhere.

References.

[1] M. E. Munroe, *Measure and Integration*, Addison-Wesley, Reading, Massachusetts 1971.

[2] J. M. Jauch, *Foundations of Quantum Mechanics*, Addison-Wesley, Reading, Massachusetts 1973.

Summary.

A set-theoretical analogue of the Lebesgue-Stieltjes measure is examined. Some consequences are briefly considered, particularly the measurability and the integrability of some classes of mappings. An application to the probability theory is discussed.

* * *