On a class of measures connected with past–dependent probability evolutions

Luigi Accardi

Laboratorio di Cibernetica
del C.N.R.
Arco Felice
Napoli – Italia

Institut of Control
Sciences
Profsoyuznaya 81
Moscow – U.R.S.S.

Contents

Abstract. A class of probability measures on the space of the trajectories of a dynamical system with discrete time is introduced. The concepts of "homogeneity of a measures" and of "stationarity of a measure" are discussed for the elements of this class. And some elementary examples of such measures are studied in connection with the problem of describing those evolutions of probability which depend on the whole past distributions of the system.

1 Statement of the problem

In the study of those systems whose evolutions is ruled by probabilistic laws, the probability measures on the "space of trajectories" present particular importance. For instance, Bernouilli's measure, can be considered as a probability measure connected with a completely casual[1] evolution. Markov's measures, as W. Feller repeatedly stressed (cfr. [7]; pg. 420), may be considered as associated to the most direct probabilistic generalization of a deterministic evolution described by ordinary differential equations (i.e., a dynamical system in the usual sense of the world), in the sense that, in such an evolution, the *state* of the system at any moment is completely determined by initial state: while, in a Markov process, the *probability distribution* at any moment is completely determined by initial distribution.

In the natural sciences, there are many examples of "hereditary" deterministic phenomena, whose evolution is not determined by an initial "state", but from a whole segment (possibly infinite) of the past history of the system.

The present work is an attempt at describing the "probabilistic analog" of such systems. That is, the study of probability measures on the space of the trajectories of those systems whose probability distributions along an arbitrary segment of the past, possibly infinite.

If W_k denotes the probability distribution of the system at the k–th moment, the law of evolution of this will be of the type:

$$w_{k+1} = F_{(k)}(W_k) \tag{1}$$

where W_k is a functional dependeing on all the distributions of the system relative to moments $h \leq k$. We will require that the probability measure

[1] We use this expression in its "naive" meaning, and refer to the deep articles of A.N. Kolmogorov [1] – [2] for a rigorous analysis of this concept.

connected with an evolution of the type (1) be *uniquely* determined by the "initial functional" and by the evolution law; and that it depends *explicitly*[2] on the latter. A problem of this kind naturally arises in various fields. For example, in the theory of stochastic neural networks the use of a purely "Markovian" formalism does not allow to take into account of the single neuron (cfr. [8]; part I; 2).

The first paragraph of this work is dedicated to a quantitative discussion of the problem outlined above. In this we will sum up a classification of the probability measures on product spaces proposed in [3] (Part II) and on which the determination of the possible solutions to the above stated problem is based.

In 2) we study some simple example of "past–dependent" probability–evolutions, for which, using Kakutani–Yoshida's extension of Krylov–Bogoliubov's condition (k).

It is clear that the validity of the interpretation of such measures as probability mesures connected to past–dependent evolution laws, can be confirmed only from the applications of these techniques to the study of phenomena which effectively appear in practise. We propose ourselves to return to this point in a future work.

The author is very grateful to Professor L.I. Rozonoer for many deep discussions and for his interest in this work.

2 Some structural properties of product measures

Let us consider a system ξ whose states belong to a certain (phase) space Ω, which we will always suppose endowed with a measurable structure $\mathcal{B}$. A probability measure on the product $\prod_{\mathbb{N}}(\Omega;\mathcal{B})$, of $\mathbb{N}$–copies of the space $(\Omega;\mathcal{B})$ can be interpreted as a measure on the space of the "trajectories" of the system ξ, with respect to a time variable $t \in \mathbb{N}$. If Ψ is such a measure, the quantities $(E_i \in \mathcal{B}; 1 \leq i \leq k)$ $\Psi_k(E_1;\ldots;E_k) = \Psi\left(\prod_{i=1}^{k} E_i x \prod_{k+1}^{\infty} \Omega\right)$ can be

[2]For instance, to every evolution of type (1) Bernoulli's measure can be associated. This one depends only on the sequence (w_k) of the probability distribution, and not (explicitly) on the evolution law.

interpreted as "joint" probabilities that the system ξ finds itself in the i–th moment in a state belonging to the set E; $1 \leq i \leq k$. For every $k \in \mathbb{N}$; Ψ_k is a probability measure on $\prod_1^k(\Omega;\mathcal{B})$ enjoying the fundamental property:

$$\Psi_{k+1}(E_1;\ldots;E_k;\Omega) = \Psi_k(E_1 \ldots E_k) \tag{2}$$

When a family $(\Psi_k)_{k\in\mathbb{N}}$ of probability measures on $\prod_1^k(\Omega;\mathcal{B})$; $(k \in \mathbb{N})$ is given, such that it can be uniquely extended[3] to a measure Ψ on the product $\prod_{\mathbb{N}}(\Omega;\mathcal{B})$, we shall say that (Ψ_k) is a *cylindrical measure*. Equality (1.1), which is the "agreement condition" for the measures Ψ_k's, can be written:

$$\Psi_l(E_1;\ldots;E_{l-k};\Omega;\ldots;\Omega) = \Psi_{l-k}(E_1;\ldots;E_{l-k})$$

In [3] (Part I) we have proposed the symmetric of this property as definition of the concept of "stationarity" for a measure.

Definition 1 A cylindrical measure $(\Psi_k)_{k\in\mathbb{N}}$, on the product $\Pi_{\mathbb{N}}(\Omega;\mathcal{B})$, is said to be "stationary" if

$$\Psi_l(\Omega;\ldots;\Omega;E_{k+1};\ldots;E_l) = \Psi_{l-k}(E_{k+1};\ldots;E_l)$$

We refer to [3] (Part I; § 2) for the discussion of some properties of stationary measures. For the moment, we limit if and only if it is of the type $\Pi_{\mathbb{N}}\varphi$; (φ – a probability measure on $(\Omega;\mathcal{B})$). A measure induced by a strictly positive Markov chain is stationary distribution (in particular: an homogeneous ergodic chain); (cfr. [3]; § 2).

The distinction between "homogeneity" and "existence of a stationary distribution" cannot be carried out in a natural way in the case of arbitrary cylindrical measures.

With the aim of carrying out this distinction in the most general class of probability measures, in [3] (Part II) there has been introduced an essentially

[3] This always happens if the (Ψ_k) are product measures, or (because of Ionescu–Tulcea's theorem; cfr. [5]; pg. 162) if they are Markov measures; or, at last, if the space $(\Omega;\mathcal{B})$ satisfies the conditions of Kolmogorov's extension theorem (it must be metrizable, separable; cfr. [5]; pg. 83). This justifies the fact that, in the following, we shall indifferently consider the measure Ψ, or the family $(\Psi_k)_{k\in\mathbb{N}}$ of measures connected with it.

algebraic classification of the probability measures on products of measurable spaces.

In the remainder of the paragraph we will sum up brifly this classification. However, in order to avoid the technical complications arising with topological tensor products we shall limit ourselves, here, to discussing the case when the space contains only a *finite number of points* which allows to conduct the discussion on a purely algebraic ground. We refer to [3] (Part II) for the general (infinite–dimensional and non–commutative) case.

It is known[4] that to every measurable space $(\Omega;\mathcal{B})$ is connected the abelian (von Neumann) algebra $L^\infty(\Omega;\mathcal{B})$ of the bounded measurable functions on Ω. There exists, furthermore, a bijective correspondence between measures on $(\Omega;\mathcal{B})$ and (continuous) linear forms on $L^\infty(\Omega;\mathcal{B})$. Conversely, if A is an abelian (von Neumann) algebra, and φ a positive (continuous) linear form on A, then there exist a measurable space $(\Omega;\mathcal{B})$ and a measure μ on $(\Omega;\mathcal{B})$ such that A is isomorphic to $L^\infty(\Omega;\mathcal{B})$ and, in the isomorphism, φ corresponds to the linear form induced by μ on $L^\infty(\Omega;\mathcal{B})$.

Therefore, to assign a cylindrical measure $(\Psi_k)_{k\in\mathbb{N}}$ on the product $\Pi_{\mathbb{N}}(\Omega;\mathcal{B})$ is equivalent to assigning a family of multilinear forms[5] $\hat{\Psi}_k$ on the products $\Pi_1^k A$ (of K copies of A); satisfying the agreement condition:

$$\Psi_k(a_1;\ldots;a_{k-1};I) = \Psi_{k-1}(a_1;\ldots;a_{k-1}) \tag{3}$$

where $a_i \in A$; $(1 \le i \le k)$; and I is the identity in A.

Now it is well known (cfr. [6]) that every multilinear map $\hat{\Psi}_k : \Pi_1^k A \to \mathbb{R}$ factorizes itself according to the commutative diagram

$$\begin{array}{ccc} \prod_1^k A & \xrightarrow{\tau_k} & \otimes_1^k A \\ \hat{\Psi}_k \downarrow & & \\ R & \swarrow \psi_k & \end{array}$$

where τ_k is a multilinear mapping (the tensor product) and ψ_k is a linear form.

The extension of this property lies at the base of the following definition (cfr. [3]; Part (II)):

[4] At this point we use the hypothesis that Ω contains a finite number of points. In this case A is isomorphic to the algebra of (real), diagonal, matrices on a finite–dimensional linear space.

[5] Cfr. 10 to which we refer also for bibliography.

Definition 2 A cylindrical measure $(\hat{\Psi}_k)$ on A is said to be of type (M)–A if the following condition is satisfied:
[i1] there exist an algebra A (not necessarily commutative!) a linear form ω on A; a sequence of linear mappings $\alpha_i : A \to$ A; such that each $\hat{\Psi}_k$ factorizes through the commutative digram

$$\begin{array}{ccc} \prod_1^k A & \overset{\prod_{i=1}^k \alpha_i}{\longrightarrow} & \mathrm{A} \\ \hat{\Psi}_k \downarrow & & \\ R & \swarrow \omega & \end{array}$$

where

$$\prod_{i=1}^k \alpha_i : (a_1; \ldots ; a_k) \in \prod_1^k A \to \prod_{i=1}^k \alpha_i(a_i) \in \mathrm{A}$$

Therefore, having assigned the algebra A, a cylindrical measure of type (M)–A is completely determined by: 1.) the sequence of linear maps α_l : A 2.) the linear form ω, on A.

This justifies the fact that, in the following, we will indifferently speak of the cylindrical measure $(\hat{\Psi}_k)$ or of the cylindrical measure $\{(\alpha_k); \omega\}$.

For a detailed analysis of the implications of the definition **??** we refer to [3] (Part II), the aim of the present work being to show how the class of measures of type (M)–A, (for an arbitrary algebra A) can represent the natural context for the discussion of concepts like "homogeneity" and "existence of a stationary distribution", for a cylindrical measure.

Above all, let us examine the following problem: "how is condition (3) expressed in the case of cylindrical measures of type (M)–A?". If A is an algebra, it is known that the regular right (resp. left) representation of A is defined by

$$R : \mathcal{A} \in \mathrm{A} \to R_{\mathcal{A}} \ ; \quad (\mathrm{Resp.}\ L_{\mathcal{A}})$$

where

$$R_{\mathcal{A}} x = x \cdot \mathcal{A} \ ; \quad (\mathrm{Resp.}\ L_{\mathcal{A}} x = \mathcal{A} \cdot x) \ ; \quad \forall x \in \mathrm{A}$$

Let us denote A^* the dual of A–considered as a vector space – then each of the linear mappings $\alpha_i : A \to$ A induces a linear map $R^*_{\alpha_i} : A \to \mathcal{L}(\mathrm{A}^*)$ of A into the space of linear maps of A^* into itself by means of the formula

$$(R^*_{\alpha_i(a)} \varphi)(\mathcal{A}) = \varphi(R_{\alpha_i(a)} \mathcal{A}) = \varphi(\mathcal{A} \cdot \alpha_i(a))$$

where $a \in A$; $\mathcal{A} \in$ A; and $\varphi \in$ A*.

Analogously is defined $L^*\alpha_i\varphi = \varphi \circ L\alpha_i$.

Thus, taking into account the equality

$$\hat{\Psi}_k(a_1; \ldots; a_k) = \omega\left(\prod_{i=1}^{k} \alpha_i(a_i)\right) = (R^*_{\alpha_k(a_k)}\omega)\left(\prod_{i=1}^{k-1} \alpha_i(a_i)\right)$$

one sees that condition (3) for cylindrical measures of type (M)–A, can be expressed:

$$(R^*_{\alpha_k(I)}\omega)(x) = \omega(x) \ ; \quad \forall\, x \in [\sqcup_{i=1}^{x-1} \alpha_i(A)]$$

where $[\sqcup_{i=1}^{k-1} \alpha_i(A)]$ is the subalgebra of A spanned by the $\alpha_i(A)$ $(1 \leq i \leq k-1)$. In general the above written equality does not imply $R^*_{\alpha_k(I)}\omega = \omega$.

In order to consider a class of measures where this last equality is valid, let us introduce the following definition:

Definition 3 We shall say that the cylindrical measure $\{(\alpha_k)_{k\in\mathbb{N}}; \omega\}$, of (M)–A type is "regular" if there exist:
1.) an imersion $T : A \to$ A

2.) a convex cone Q in A*

satisfying the collowing conditions:
[i1] $T(A)$ separates the points of Q: (i.e, if ω', $\omega'' \in Q$ and $\omega'(Ta) = \omega''(Ta)$; for every $a \in A$, then $\omega' = \omega'$).
[i2] S being the subsets of A defined by the equality:

$$S = \{s \in \text{A} : R^*_s(Q) \subset Q\} \cap \{t \in \text{A} : L^*_t(Q) \subset Q\}$$

there then exists a sequence $(s_k)_{k\in\mathbb{N}}$ in S, such that $\alpha_1 = T$; $\alpha_k = L_{s_k} \circ T$; for $k \geq 2$.

Proposition 1 If $\{(s_k); T; \omega\}$ is a cylindrical measure of type (M)–A, regular, then condition (3) is equivalent to the following:

$$R^*_{s_k}\omega = \omega \qquad \text{for every} \quad K \in \mathbb{N} \tag{4}$$

(The proof is given in appendix I).

In the following we will restrict ourselves to the study of *regular*, type (M)–A measure. In effect, Proposition **??** holds under a considerably broader hypothesis (cfr. [3] (Part II)) but cylindrical measures are regular (cfr. the examples at the end of the paragraph) and from the particularly simple form of equation (4) which is the generalization of the well known property of stochastic operators: $P^*u = u$ (where u is the unit function).

We will assume the symmetrical of property (4) as analogous of the existence of an invariant distribution. More precisely:

Definition 4 Let $\{(s_k); T; \omega\}$ be a cylindrical measure of type (M)–A, regular. We will say that ω is a "stationary state" for $\{(s_k); T; \omega\}$ if:

$$L^*_{s_k}\omega = \omega$$

As we will prove at the end of the paragraph this property is the generalization of the property $P_{(K)}w = w$, where w is a probability distribution and $P_{(K)}$ a stochastic operator.

Definition 5 Let $\{(s_k); T; \omega\}$ be a cylindrical measure of type (M)–A, regular. We will say that $\{(s_k); T; \omega\}$ is "homogeneous" if there exists $s \in$ A, such that $s_k = s$; for every $k \in \mathbb{N}$. The connection between the notions of "homogeneity" and "having a stationary state" with the property of stationareity introduced in Definition 2, is given by the following theorem.

Theorem 1 Every cylindrical measure, regular, homogeneous and with a stationary state, is stationary in the sense of definition 2. Conversely, let $\{(s_k); T; \omega\}$ be a cylindrical measure of (M)–A type, stationary in the sense of Definition 2. Conversely, let $\{(s_k); T; \omega\}$ satisfies the following condition: [t1] for every couple of natural integers i, J, the equality

$$\omega(T(a) \cdot [s_1 - s_J] \cdot T(b)) = 0 \; ; \quad \forall \, a, b \in A$$

is possible if and only if $s_i = s_J$.

Then the measure $\{(s_k); T; \omega\}$ is homogeneous and ω is a stationary state for it.

The proof is in appendix II.

A class of cylindrical measures staisfying condition [t1] of Theorem **??** is given, for example, by the discrete Markov chains $\{(P_{(k)}); w\}$ such that $\gamma^{(k)}_{iJ} > 0$; $w_J > 0$ for $1 \leq i, J \leq n$ and for every $k \in \mathbb{N}$.

We want to illustrate the definitions given with some usual example.

If A= $\mathbb{R}$; $\omega = id$ (identity on $\mathbb{R}$) then the diagram of Definition **??** is reduced to ()

$$(e_1; \dots; a_k) \in \prod_1^k A \to \alpha_1(a_1) : \dots \cdot \alpha_k(a_k)$$

where the α_i are linear forms. In this case the measure $\hat{\Psi}_k$ is the product measure. Let A= $\mathcal{L}(\mathbb{R}^n)$ the algebra of linear operators on $\mathbb{R}^n$; T; $A \to$ A, the map which transforms the vector ${}^t a = (a_1; \dots; a_n)$ into the digonal $n \times n$ matrix, with a_i as its i–th diagonal entry.

(T is even an algebra omomorphism). Let Q then be the convex cone of the linear forms on A defined by $s \mapsto (s \cdot \lambda; u)$ where λ is a vector with positive components; u the unit function and $(\cdot, \cdot)$ denote the scalar product in $\mathbb{R}^n$.

Condition [i1] of Definition **??** is clearly satisfied. Let us now consider the set:

$$S = \{s \in \text{A} : R_s^*(Q) \subset Q\} \cap \{s \in \text{A} : L_s^*(Q) \subset Q\}$$

If λ is a positive vector, and $s \in S$, then $R_s^*(t \cdot \lambda; u) = (t \cdot s\lambda; u)$ for every $t \in$ A and so the condition $R_s^*(Q) \subset Q$ implies that s maps positive vectors into positive vectors.

Analogously: $L_s^*(t \cdot \lambda; u) = (t \cdot \lambda; s^* u)$, so condition $L_s^*(Q) \subset Q$ implies $s^* u = u$. It is immediate to verify that these two properties define a stochastic operator.

Therefore n–dimensional Markov measures can be *characterized* as those regular, $(M) - \mathcal{S}(\mathbb{R}^n)$ type measures, with respect to the cone Q of positive vectors in $\mathbb{R}^n$.

The factorization introduced in Definition **??** in this case is realized through the linear form $\omega(s) = (s \cdot w; u)$ where w is a stochasti vector; and $\prod_{i=1}^k \alpha_i = \prod_{i=1}^k L_{s_i} \circ T$ where T is the diagonal immersion defined before and s_i a stochastic matrix $(1 \leq i \leq n)$.

Let us consider, finally, that both the Markov and the product measures can be distinguished, within the class of all the (M)–A measure, as those decomposable into couples $\{(\otimes_{i=1}^k \alpha_i)_{k \in \mathbb{N}}; \omega\}$; where $\otimes_{i=1}^k \alpha_i$ denotes a *pure* tensor product. The property of commutativity (cfr. [4]; pg. 29)

$$\hat{\Psi}_k(a_1; \dots; a_k) = \hat{\Psi}_k(a_{i_1}; \dots; a_{i_k})$$

is verified for these measures every time that the algebra A is *commutative*; therefore it holds for product measure (A= $\mathbb{R}$) and not for Markov measures (A $= \mathcal{L}(\mathbb{R}^n)$). There exist many other types of $(M) -$A type measures. For a general analysis of these we refer to [3] (Part II). In the following paragraph, some examples will be studied.

3 Some example of "past–dependent evolutions"

From the preceding paragraph one can derive the following three features as essentiaal for the construction of Markov measures.

1.) One considers the convex set Q^+ of all the probability distributions on a phase space $(\Omega; \mathcal{B})$. The algebra A of linear mappings of the Banach space $\mathcal{M}(\Omega; \mathcal{B})$[6] in itself. And the convex monoid S of A of all the linear operators which map Q^+ in itself.

2.) To every $W \in Q^+$ one associates in a natural way (i.e., depending only on W) a (continuous) linear form ω_w on A which enjoys the properties

$$\omega_w(R_s F) = \omega_w(F) \ ; \quad \omega_w(L_s F) = \omega_{sw}(F)$$

for every $s \in S$; $F \in$ A.

3.) One defines a representation T of the commutative algebra $A = L^\infty(\Omega; \mathcal{B})$ in A by means of the formula

$$(T(a)F)_m = a(Fm) \ ; \qquad F \in \text{A}; \ a \in A; \ m \in \mathcal{M}(\Omega; \mathcal{B})$$

After operations 1.) – 2.) 3.) have been performed the cylindrical measure on $\prod_{\mathbb{N}}(\Omega; \mathcal{B})$ is defined by means of the family of multilinear mappings:

$$\hat{\Psi}_k : (a_1 \ldots a_k) \in \prod_1^k A \to \omega_w(T(a_k) \cdot R_{s_k} \cdot \ldots \cdot T(a_1) R_{s_1} I')$$

[6] $\mathcal{M}(\Omega; \mathcal{B})$ is the space of all signed measures on $(\Omega; \mathcal{B})$. It is a Banach space for the norm: $\|x\| =$ total variation of X on $(\Omega; \mathcal{B})$.

$a_i \in A$; $(1 \leq i \leq k)$; I' is the identity in A.

From point 3.) one immediately sees that the family $(\hat{\Psi}_k)$ enjoys the agreement property (4), and that ω_w is a stationary state for $(\hat{\Psi}_k)$ if and only if: $S_k w = w$; for every k. Furthermore, the $\hat{\Psi}_k$ are positive, in the sense that if all the a_i are positive the value of $\hat{\Psi}_k$ is positive and ≤ 1.

The forms $\hat{\Psi}_k$ thus define an evolution law for the probability distribution by means of the equation:

$$w_k(\chi) = \omega_w(T(\chi) \cdot s_k \ldots s_1) = \hat{\Psi}_k(I; \ldots; I; \chi) \tag{5}$$

where χ is a projection operator in A and I is the identity. In this paragraph we shall apply the procedure described in points 1.) 2.) 3.) to the case when, instead of a single probability distribution $w \in Q^+$, one considers a sequence $W = (w_n)_{n=-\infty}^{0}$ of these.

The sequence W is interpreted as the sequence of the probability distributions relative to the past history of the system. Formula (2.1), with w substituted by W, will express the evolution of the probability distribution of the system as a function of the distribution relative to the past history.

Let us introduce some notations: the symbol $\prod_{\mathbb{N}^-} Q^+$ denotes the set of sequence $W = (w_k)$ of elements in Q^+, indexed by the set $\mathbb{N}^-$ of negative integers.

This is a convex subset of the space $L^\infty(\mathbb{N}; \mathcal{M}(\Omega; \mathcal{B}))$ of the bounded sequence of signed measures on $(\Omega; \mathcal{B})$. E denotes the space of linear mappings from $L^\infty(\mathbb{N}^-; \mathcal{M}(\Omega; \mathcal{B}))$ into $\mathcal{M}(\Omega; \mathcal{B})$ $\hat{Q}^+$ is the subset of the elements in E which map $\prod_{\mathbb{N}^-} Q^+$ into Q^+. Thus an element in E correspond to a law of transition from the sequence of probability distribution relative to the past of the system to a probability distribution. The algebra of linear mappings of E into itself will be denoted A, and S is the convex subset of A defined by:

$$S = \{s \in \mathrm{A} : s\hat{Q}^+ \subset \hat{Q}^+\}$$

At last, let us denote Q_0^+ the set of positive measures m on $(\Omega; \mathcal{B})$ such that $m(\Omega) \leq 1$.

The hypothesis that the space $(\Omega; \mathcal{B})$ contains exactly n points will be kept from now on. The following theorem is useful to clarify the structure of the set $\hat{Q}^+$, and justifies the choice of the operators in S which will be made in the following.

Theorem 2 A linear operator B which maps the set $\prod_{\mathbb{N}^-} Q^+$ into itself, satisfies the condition:

(i) B maps $\prod_{\mathbb{N}^-} Q_0^+$ into itself if and only if it has the form:

$$(BW)(\sigma) = \sum_{\sigma=-\infty}^{0} P(\sigma;\tau)w(\tau)\varphi(\tau) \tag{6}$$

$$W = (w(\sigma))_{\sigma\in\mathbb{N}^-} \ ; \quad \varphi(\tau) \geq 0 \ ; \quad \sum_{\tau=-\infty}^{0} \varphi(\tau) = 1$$

and $P(\sigma;\tau)$ is a stochastic matric for every τ such that $\varphi(\tau) > 0$.

The proof of this theorem is worked out in Appendix IV where, furthermore a counterexample is given which proves that the thesis does not subsist without condition (i).

In particular, the element of $\hat{Q}^+$ which lie in the class identified by conditions of Theorem **??** are expressed as

$$BW = \sum_{\sigma=-\infty}^{0} p(\sigma)w(\sigma)\varphi(\sigma) \tag{7}$$

The action of A defined in point 3.) is easily extended on E through the formula $(T'(a)B)W = T(a)(BW)$, which, in the case where B has the form (7), is expressed:

$$(T'(a)B)W = \sum_{\sigma=-\infty}^{0} T(a)P(\sigma)W(\sigma)\varphi(\sigma)$$

Inside the set $S \subset$ A let us choose those operators $s \in S$ defined by the relation:

$$(SB)W = \sum_{\sigma=-\infty}^{0} Q_s(\sigma)P(\sigma)w(\sigma)\varphi(\sigma) \tag{8}$$

("Diagonal" operators). Thus if $W_0 \in \prod_{\mathbb{N}^-} Q^+$; $B \in \hat{Q}^+$ has the form (7); $(S_k)_k \in \mathbb{N}$ is a sequence of operators in A of the form (8) the equation

$$W_k = (S_k \cdot S_{k-1} \cdot \ldots \cdot S_1 \cdot B)W_0 \ ; \quad k = 1, 2, \tag{9}$$

defines a sequence of probability distributions $(W_k)_{k=1}^{\infty}$ which is completely determined by the sequence of distributions $W_0 = (W_k)_{k=-\infty}^{0}$; relative to the past history of the system.

Every evolution of type (9) can be interpreted as resulting from the convex mean of an infinite number of Markov chains, each of which admits an initial distribution w_k, $(k = 0, -1, -2, \ldots)$. The multilinear function (i.e., the cylindrical measure) connected with evolution of type (9) will be given by

$$\hat{\Psi}_k(a_1; \ldots; a_k) = \omega_{W_0;B}(T(a_k)R_{s_k} \cdot \ldots \cdot T(a_1)R_{s_1}I) \tag{10}$$

where

$$\omega_{W_0;B}(s) = ((sB)W_0; u)$$

If $s_k B = B$ for every K, then $w_{W_0;B}$ is a stationary state for the measure $(\hat{\Psi}_k)$. If $S_k = S$ for every K then the cylindrical measure $(\hat{\Psi}_k)$ is homogeneous in the sense of definition **??**. At last, if both relations hold, from the first part of Theorem **??** it follows that the measure $\hat{\Psi}_k$ is stationary.

4 Evolutions of Volterra's type

In this paragraph we shall examine the possibility of associating a cylindrical measure to a probability evolution of Volterra's type.

That is, maintaining the notations of the preceding paragraphs we consider the evolution equation

$$W(n+1) = \sum_{-\infty}^{n} {}_{\sigma} P(n-\sigma) w(\sigma) \tag{11}$$

where $(W(\sigma))$ is a sequence of stochastic vectors; $P(\sigma) = Q(\sigma)\varphi(\sigma)$; $Q(\sigma)$ is a stochasti matrix;

$$\varphi(\sigma) \geq 0 \; ; \quad \sum_{-\infty}^{0} {}_{\sigma} \varphi(-\sigma) = 1$$

The evolution (11) can also be written

$$w(n+1) = \sum_{-\infty}^{0} {}_{\sigma} P(-\sigma) w(n+\sigma) \tag{12}$$

Therefore, making the hypothesis that equation (12) holds also for those $w(m)$, with $m \leq 0$, one sees:

$$w(n+1) = \sum_{-\infty \tau_{n+1}}^{0} \dots \sum_{-\infty \tau_1}^{0} P(-\tau_{n+1}) \cdot \dots P(-\tau_1) w(\tau_1 + \dots + \tau_{n+1}) \quad (13)$$

Now, if $W = (\sigma(\sigma)) \in \prod_{\mathbb{N}^-} Q^+$, then we can define the linear operator $\tilde{P}$ by the equality

$$(\tilde{P}W)(\tau) = \sum_{-\infty}^{\tau}{}_{\sigma} P(\tau - \sigma) w(\sigma) \quad (14)$$

And, putting $W_1 = (w(\tau))$ a sequence (W_n) in $\prod_{\mathbb{N}} Q^+$ is defined inductively by:

$$W_{n+1} = \tilde{P} W_n \quad (15)$$

Let then $p : \prod_{\mathbb{N}^-} Q^+ \to Q^+$ be the linear map

$$pW = \sum_{-\infty}^{0}{}_{\tau} P(-\tau) w(\tau) \quad (16)$$

It is easy to verify that the sequences (w_n) defined by (11), and (W_n) defined by (15) are connected by the equation:

$$w_n = pW_n \quad (17)$$

Thus we can conclude: to assign an evolution of probability distributions (w_n) ruled by the law

$$w_{n+1} = \sum_{-\infty}^{n}{}_{\sigma} P(n - \sigma) w(\sigma)$$

is equivalent to giving an evolution of *sequences* of probabiity distributions (i.e., an evolution in $\prod_{\mathbb{N}^-} Q^+$) ruled by the equation:

$$W_{n+1} = \tilde{P}^n W_1$$

the connection between the two sequences being given by the equation

$$w_{n+1} = pW_{n+1}$$

Let us now denote, as usual, with $T : L^\infty(\Omega;\mathcal{B}) \to \mathcal{L}(\mathcal{M}(\Omega;\mathcal{B}))$ the diagonal action, on $\mathcal{M}(\Omega;\mathcal{B}) = \mathbb{R}^n$ of the algebra of bounded functions on $(\Omega;\mathcal{B})$. This action induces naturally an action T', on the set $\prod_{\mathbb{N}^-} Q^+$ by means of the formula

$$(T'(a)W)(\sigma) = T(a)W(\sigma) \; ; \qquad W = (w(\sigma))$$

Let us introduce the family of multilinear forms:

$$\hat{\Psi}(a_1;\ldots;a_n) = \langle T(a_n)p \cdot T'(a_{n-1})\tilde{P} \cdot \ldots \cdot T(a_1)\tilde{P}W; u\rangle \tag{18}$$

Then, in order for the family $(\hat{\Psi}_n)$ to define a cylindrical measure, it is necessary (and sufficient too, because the space Ω has a finite number of points) that the agreement condition:

$$\hat{\Psi}_n(a_1;\ldots;a_{n-1};I) = \hat{\Psi}_{n-1}(a_1;\ldots;a_{n-1}) \tag{19}$$

holds.

Writing down this one explicitly, one finds:

$$\sum_{-\infty}^{0}{}_{\tau_{n-1}} \ldots \sum_{-\infty}^{0}{}_{\tau_1} < P(-\tau_{n+1}) \cdot T(a_n)P(-\tau_n) \cdot \ldots \cdot T(a_1)P(-\tau_1)w(\tau_1 + \ldots + \tau_{n+1}); u\rangle =$$

$$= \sum_{-\infty\tau_{n+1}}^{0} \ldots \sum_{-\infty}^{0}{}_{\tau_1} \varphi(-\tau_{n+1})\langle T(a_n)P(-\tau_n) \cdot \ldots \cdot T(a_1)P(-\tau_1)w(\tau_1 + \ldots + \tau_n + \tau_{n+1}); u\rangle =$$

$$= \sum_{-\infty\tau_n}^{0} \ldots \sum_{-\infty\tau_1}^{0} \langle T(a_n)P(-\tau_n) \cdot \ldots \cdot T(a_1)P(-\tau_1)w(\tau_1 + \ldots + \tau_n); u\rangle$$

Having set $(\tilde{\varphi}W)(\sigma) = \sum_{-\infty\tau}^{0} \varphi(-\tau)w(\sigma+\tau)$, the above equality is equivalent to

$$\sum_{-\infty}^{0}{}_{\tau_n} \ldots \sum_{-\infty}^{0}{}_{\tau_1} \langle T(a_n)P(-\tau_n) \cdot \ldots \cdot T(a_1)P(-\tau_1)[(\tilde{\varphi}W)(\tau_1 + \ldots + \tau)n)]; u\rangle =$$

$$= \sum_{-\infty}^{0}{}_{\tau_n} \ldots \sum_{-\infty}^{0}{}_{\tau_1} \langle T(a_n)P(-\tau_n) \cdot \ldots \cdot T(a_1)P(-\tau_1)w(\tau_1 + \ldots + \tau_n); u\rangle \tag{20}$$

If we require that the agreement condition (19) takes place independently of the particular choice of the sequence $(Q(\sigma))$ the above written equality is equivalent to:

$$\langle T(a)[\tilde{\varphi}W(\sigma)]; u\rangle = \langle T(a)W(\sigma); u\rangle \ ; \quad \forall\, a \in A$$

that is

$$\tilde{\varphi}W = W \tag{21}$$

or, equivalently, for every $\sigma \leq 0$

$$\sum_{0\tau}^{+\infty} \varphi(\tau)w(\sigma - \tau) = w(\sigma) \tag{22}$$

Now, the functions φ and $w = (w_1; \ldots; w_n)$ are bounded, therefore they admit a (discrete) Laplace transform for every $\rho > 0$. Applying the Laplace transform to each side of equation (22) one finds

$$[\mathcal{L}(\varphi)(\rho) - 1]\mathcal{L}(w_i)(\rho) \equiv 0 \quad \rho \geq 0 \quad 1 \leq i \leq n \tag{23}$$

($\mathcal{L}(f)(t)$ denote the Laplace transform of f in the point t). Then $\mathcal{L}(w_i)(\rho) = 0$ if $\mathcal{L}(\varphi)(\rho) \neq 1$, for $1 \leq i \leq n$. Since $w_i(\rho) \geq 0$, this implies $\varphi(0) = 1$ $\varphi(\rho) = 0$ for $\rho > 0$. Hence, one deduces that the only non zero solutions of equation (22) with $w_i(\rho) \geq 0$, are those for which $w(0)$ is an arbitrary positive vector and $w(\rho) = 0$ for $\rho > 0$. These solutions correspond to the usual Markov processes. In particular, there are no solutions of equation (21) in the set $\prod_{\mathbb{N}^-} Q^+$, and so the agreement condition (19) in general is not satisfied. Therefore concerning the probability evolution described by the equation (11), we conclude that it is not possible to associate to it a cylindical measure which depends uniquely on the evolution law and not on the particular form of the sequence of stochastic operators $(Q(\sigma))$ and on the function φ[7].

[7]This does not mean that for a particular choice of φ and the sequences $(Q(\sigma))$ it could not be possible to find a cylindrical measure with the required properties. It is sufficient to this aim, for a given φ, to consider a subset ϑ of $\prod_{\mathbb{N}^-} Q^+$ transformed into itself by the operator $\tilde{\varphi}$ and then to limit oneself to considering those sequences of stochastic operators $(Q(\sigma))$ for which the equality $(3.9')$ are identically satisfied in ϑ. Nevertheless the determination of the explicit form of the operators $(Q(\sigma))$ as functions of φ and ϑ is very complicated.

5 Conclusion

Analyzing from the standpoint of the considerations in § 1) the probability evolution

$$W_{n+1} = \sum_{-\infty}^{n}{}_{\sigma} P(n-\sigma)w(\sigma) \tag{24}$$

one sees that such an evolution law is completely determined by the assignment of:

1.) a continuous linear map $\tilde{P}$ of the space $L^{\infty}(\mathbb{N}^{-};\mathbb{R}^{n})$ into itself.

2.) A continuous form $\omega_n; w$ defined on the algebra A of the linear transformations of $L^{\infty}(\mathbb{N}^{-};\mathbb{R}^{n})$ into itself, which is positive on the convex cone of the positive elements of A.

Once given $\tilde{P}$ and $\omega_p; w$ the probability evolution is determined by the equation

$$W_{n+1}(T(a)) = \omega_{p;w}(T^{*}(a)\tilde{P}^{n}) = \langle T(a)p\tilde{P}^{n}W; u\rangle \tag{25}$$

Furthermore, we have also seen (§ 3) that the agreement condition

$$P^{*}\omega_{p;w} = \omega_{p;w} \tag{26}$$

does not take place for the linear forms of type (25), with the exception of the trivial case: $\varpi(0) = 1$; $\varphi(\sigma) = 0$, $\sigma > 0$. Nevertheless the fact that, once known the operator $\tilde{P}$, the evolution (24) is equivalent (cfr. (16); (17)) to the evolution

$$W_{n+1} = \tilde{P}W_n \tag{27}$$

suggests not to limit oneself to the consideration of the linear forms on A, of the type $\omega_{p;w}$; but to state the problem in the larger class of all the linear forms on A which take positive values on the convex cone of the positive operators in A. This statement of the problem leaves unaltered the "Volterra type" character of the probability evolution (cfr. (27) and (14)). More precisely, one consider all the probability evolution determined by the assignment of:

1.) A continuous linear operator $\tilde{P}$ (cfr. (14)) in the space $L^{\infty}(\mathbb{N}^{-};\mathbb{R}^{-})$.

2.) A continuous linear form $\omega \in$ A*, positive on the convex cone of positive operators in A; through the equation:

$$W_{n+1}(T(a)) = \omega(T_0(a)\tilde{P}^{n}) \tag{28}$$

(To denote an action of A on A).**Appendix I**: Proof of Proposition **??**.

By hypothesis $s_1 = e$ therefore equality $R^{+}_{s_1}\omega = \omega$ is trivial. Let us suppose that $R^{*}_{S^{\omega}_{J}} = \omega$ for $1 \leq J \leq k$, then one has:

$$\hat{\Psi}_{k+1}(a_1; I; \ldots; I) = \omega\left(Ta_1 \cdot \prod_{i=2}^{k+1} L_{S_i} \cdot T(I)\right) =$$

$$= \omega\left(Ta_1 \cdot \prod_{i=2}^{k+1} s_i\right) = R^{*}_{S_{k+1}}\left(\prod_{J=0}^{k-2} R^{*}_{S_{k-J}}\omega\right)(T_{a_1}) =$$

$$= (R^{*}_{S_{k+1}}\omega)(T_{a_1})$$

because of the inductive hypothesis. On the other hand,

$$\hat{\Psi}_{k+1}(a_1; I; \ldots; I) = \hat{\Psi}_1(a_1) = \omega(T_{a_1})$$

for every $a_1 \in A$. Therefore from Definition **??** it follows $R^{*}_{S_{k+1}}\omega = \omega$.

Consequently $R^{*}_{S_J}\omega = \omega$; for every J; and this ends the proof.

Then a cylindrical measure "naturally" (in the sense specified in § 0.) associated with the evolution (28) exists if and only if the agreement condition

$$L^{*}_{\tilde{p}}\omega = \omega \tag{29}$$

is satisfied.

If such an ω exists, the associated cylindrical measure will be automatically homogeneous, (in the sense of Definition **??**; and it will be stationary, in the sense of Definition 2 if and only if

$$R^{*}_{\tilde{p}}\omega = \omega \tag{30}$$

The existence of linear forms $\omega \in \mathrm{A}^{*}$ satisfying equation (29) and the positivity condition of point 2.), will be discussed in a subsequent paper.

Appendix II: Proof of Theorem **??**.

Let $\{(p_k); T; \omega\}$ be a cylindrical measure, regular, stationary, satisfying $[pl]$. Then, having set $p = p_2$ then one has

$$\hat{\Psi}(I; a_2) = \omega(L_{p_2} \cdot T(a_2)) = (L^{*}_{p_2}\omega)(T(a_2)) = \hat{\Psi}_1(a_2) = \omega(T(a_2))$$

Therefore $L^{*}_{p_2}\omega = \omega$ because of regularity. Suppose now $p_i = p$ for $i \leq k$. Then:

$$\hat{\Psi}_{k+2}(I; \ldots; I; a_{k+1}; a_{k+2}) = \omega(L_{p_{k+1}} \cdot T(a_{k+1}) \cdot L_{p_{k+2}} \cdot T(a_{k+2}) =$$

$$= (L^*_{p_{k+1}}\omega)(T(a_{k+1}) \cdot L_{p_{k+2}} \cdot T(a_{k+2})) =$$

$$= \hat{\Psi}_2(a_{k+1}; a_{k+2}) = \omega(T(a_{k+1}); L_p \circ T(a_{k+2}))$$

Setting $a_{k+2} = I$ from the equality

$$\omega(T(a_{k+1})) = (L^*_{p_{k+1}} R^*_{p_{k+2}}\omega)(T(a_{k+1})) = (L^*_{p_{k+1}}\omega)(T(a_{k+1}))$$

from regularity it follows $L^*_{p_{k+1}}\omega = \omega$.

¿From the preceding equality one then deduces

$$\omega(T(a_{k+1})) \cdot [p_{k+2} - p] \cdot T(a_{k+2})) = 0$$

for every a_{k+1}, a_{k+2} in A. Therefore, from condition $[p1]$ it follows $p_{k+2} = p$, which is the thesis.

Appendix III

Proof of Theorem **??**. It is clear that the condition is sufficient. Conversely, suppose that

$$v(\tau) = \sum_{-\infty}^{0}{}_{\sigma} A(\tau; \sigma) w(\sigma) \in Q^+ \tag{31}$$

for every $W = (w(\sigma)) \in \prod_{\mathbb{N}^-} Q^+$. Then

$$\sum_{-\infty}^{0}{}_{\sigma} \langle A(\tau; \sigma) w(\sigma); u \rangle = \sum_{-\infty}^{0}{}_{\sigma} \langle w(\sigma); q_\tau(\sigma) \rangle = 1 \tag{32}$$

(having set $q_\tau(\sigma) = \ldots\ {}^tA(\tau; \sigma)u)$ [8]. Let now be $tq_\tau(\sigma) = (q_{1,\tau}(\sigma); \ldots; q_{n,\tau}(\sigma))$: and suppose that there exists $\sigma^* \in \mathbb{N}^-$ such that $q_{i,\tau}(\sigma^*) \neq q_{J,\tau}(\sigma^*)$ for some i and j; $(1 \leq i; j \leq n)$. From the positiveness of $w_k(\sigma)$, it follows:

$$\langle w(\sigma^*); q_\tau(\sigma^*) \rangle = \sum_{1}^{n}{}_{i} w_i(\sigma^*) q_{i,\tau}(\sigma^*) < \tag{33}$$

$$< \max_{1 \leq i \leq n} q_{i,\tau}(\sigma^*) = \langle e_{J^*}; q_\tau(\sigma^*) \rangle$$

[8]We denote ${}^tA$ the transpose of the matrix A.

where by definition; $q_{J^*;\tau}(\sigma^*) = \max_{1\leq i\leq n} q_{i;\tau}(\sigma^*)$ and e_{J^*} denotes the vector $e_{J^*,i} = \delta_{J^*,i}$ (the Kroneker δ). Since $e_{J^*} \in Q^+$, the sequence $(\overline{W}(\sigma))$ such that

$$\overline{W}(\sigma) = w(\sigma) \quad \text{for} \quad \sigma \neq \sigma^* \ ; \quad \overline{W}(\sigma^*) = e_{J^*}$$

lies in $\prod_{\mathbb{N}^-} Q^+$. So from (32) it follows

$$1 = \sum_{-\infty}^{0}{}_{\sigma} \langle \overline{W}(\sigma); q_\tau(\sigma)\rangle > \sum_{-\infty}^{0}{}_{\sigma} \langle w(\sigma); q_\tau(\sigma) >= 1 \tag{34}$$

which is aburd. Hence $q_{i,\tau}(\sigma) = q_{J,\tau}(\sigma)$ for every $\sigma, \tau \in \mathbb{N}^-$ and $1 \leq i, J \leq n$; we can write

$$q_\tau(\sigma) = \lambda_\tau(\sigma) \cdot u; \ \sum_{-\infty}^{0}{}_{\sigma} \lambda_\tau(\sigma) = 1 \tag{35}$$

Till now we did not use hypothesis (i). Suppose now thata $W = (w(\sigma)) \in \prod_{\mathbb{N}^-} Q_0^+$. Then, because of hypothesis (i), one has

$$0 \leq \sum_{-\infty}^{0}{}_{\sigma} \langle w(\sigma); q_\tau(\sigma)\rangle \leq 1 \tag{36}$$

thus, if there exists a σ^* such that $\lambda_\tau(\sigma^*) < 0$ choosing $w(\sigma) = 0$ for $\sigma \neq \sigma^*$, and $w(\sigma^*) \neq 0$ one has:

$$\langle w(\sigma^*); q_\tau(\sigma^*)\rangle = \lambda_T(\sigma^*) < 0 \tag{37}$$

which contradicts (36).

Thus, $\lambda_\tau(\sigma) \geq 0$; $\sum_{-\infty}^{0}{}_{\sigma} \lambda_\tau(\sigma) = 1$. Setting ${}^tA(\tau;\sigma) = \lambda_T(\sigma)\ {}^tQ(T;\sigma)$ it will be suficient, for our thesis to prove that ${}^tQ(\tau;\sigma)$ is a positive matrix. Suppose, again, the contrary. Then there exists a $\sigma^* \in \mathbb{N}^-$, such that $q_{i,\tau}(\sigma^*) < 0$ for some i; $(1 \leq i \leq n)$. Therefore for a vector $w_0 \in Q_0^+$, one has $\langle w(\sigma^*); q_\tau(\sigma^*)\rangle < 0$ contradicting (36). Hence, ${}^tQ(\sigma,\tau)$ is a positive matrix and ${}^tQ(\sigma;\tau)u = u$, for every $\sigma, \tau \in \mathbb{B}^-$ i.e., $Q(\sigma;\tau)$ is a stochasti matrix, and this conclude the proof.

Let us now give a counterexample which proves that, without assumption (i), Theorem **??** is false. Let $\varphi(\sigma) > 0$; $\sum_{i=0}^{N} \varphi(\sigma) = 1$ (N a fixed integer). Let

$(A(\sigma))_{0\leq\sigma\leq N}$ be stochastic matrices such that $a_{iJ}(\sigma)\geq\varepsilon>0$; $1\leq i,j\leq n$; $0\leq\sigma\leq N$ for some $\varepsilon>0$.

Choose a sequence of numbers $(\varphi(\tau))_{\tau>N}$ such that $\sum\limits_{\tau=N+1}^{\infty}\varphi(\tau)=0$; $\sum\limits_{\tau\in G^+}\varphi(\tau)=\frac{\varepsilon}{2M}$ where G^+ is the set of integers ($>N$) such that $\varphi(\tau)>0$.

Let $(A(\tau))_{\tau>N}$ now be a sequence of matrices satisfying the following conditions:

$$^tA(\tau)u=u\ ;\quad \|a_i(\tau)\|<M$$

where $a_i(\tau)$ denotes the i–th row of the matrix $A(\tau)$. Then if $W=(w(\sigma))\in\prod\limits_{\mathbb{N}^-}Q^+$ one sees

$$\sum_0^\infty{}_\sigma\langle\varphi(\sigma)A(\sigma)w(\sigma);u\rangle=$$

$$=\sum_0^N{}_\sigma\varphi(\sigma)<A(\sigma)w(\sigma);u\rangle+\sum_{N+1}^\infty{}_\sigma\varphi(\sigma)\langle A(\sigma)w(\sigma);u\rangle=1$$

And futhermore:

$$\sum_0^\infty{}_\sigma\varphi(\sigma)[A(\sigma)w(\sigma)]_i=$$

$$=\sum_0^N{}_\sigma\varphi(\sigma)\left[\sum_{J=1}^n a_{iJ}(\sigma)w_J\right]+\sum_{N+1}^\infty{}_\sigma\varphi(\sigma)\left[\sum_{J=1}^n a_{iJ}(\sigma)w_J\right]\geq$$

$$\geq\varepsilon-\sum_{\tau=N+1}^{\infty}|\varphi(\tau)|\cdot|\langle a_i(\tau);w(\tau)\rangle|\geq$$

$$\geq\varepsilon-\sum_{\tau=N+1}^{a}|\varphi(\tau)|M\geq 0$$

because of our hypothesis. Therefore we conclude that the operator:

$$w=(w(\sigma))\to\left(\sum_{-\infty}^0{}_\sigma\varphi(-\sigma)A(-\sigma)w(\sigma)\right)$$

maps $\prod_{\mathbb{N}^-}Q^+$ into itself, and it is not of the form required by Theorem **??**.

References

[1] A.N. Kolmogorov, "Sankhyà: The Indian Journal of Statistics", Series A, 25, (1963), 369–373.

[2] A.N. Kolmogorov, "Logical Basis for Information Theory", *IEEE Trans. on Inf. Theory*, vol. II–14: No. 5, September (1968).

[3] L. Accardi, *Measures on Product Spaces*, Part I, Part II (to appear).

[4] A.N. Kolmogorov, *Foundation of Probability*, Chelsea (1956).

[5] J. Neven, *Mathematical Foundations of the Calculus of Probability*, Holden Day Inc., (1965).

[6] C. Chevalley, "Fundamental Concepts of Algebra", *Academic Press*.

[7] W. Feller, *An Introduction to Probability Theory and its Applications*, vol. I, J. Wiley ed. (1968).

[8] L.I. Royonoer, *On random neural networks*, Automatica i Telemecanika.

[9] K. Yoshida, S. Kakutani, "Operator Theoretical Treatment of Markoff's Process and Mean Ergodic Theorem", *Ann. of Math.* **42**, (1941), 188–228.

[10] I.E. Segal, "A non commutative extension of abstract integration", *Ann. of Math.* **57** (1953), 401.