Czechoslovak Mathematical Journal

Zbyněk Šidák

Integral representations for transition probabilities of Markov chains with a general state space

Czechoslovak Mathematical Journal, Vol. 12 (1962), No. 4, 492–522

Persistent URL: http://dml.cz/dmlcz/100535

Terms of use:

© Institute of Mathematics AS CR, 1962

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.

This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* http://dml.cz

Чехословацкий математический журнал т. 12 (87) 1962, Прага

INTEGRAL REPRESENTATIONS FOR TRANSITION PROBABILITIES OF MARKOV CHAINS WITH A GENERAL STATE SPACE

ZBYNĚK ŠIDÁK, Praha

(Received September 21, 1960)

The paper is devoted to the study of representations of types (1) and (2) (see Introduction) for transition probabilities of Markov chains with a general state space. Its sections deal with sub-invariant measures of a chain, representations of type (1) for general chains, continuity properties of representing functions v of bounded variation in (1), conditions for self-adjointness of a chain, representations of type (2) for self-adjoint chains. Attention is paid to finitely additive sub-invariant measures and representations derived on the basis of them.

1. INTRODUCTION

1.1. Introduction and summary. In the theory of homogeneous Markov chains the representation of transition probabilities appears to be an old problem. These probabilities can be derived from the one-step transition probabilities generally only by recurrent formulae, so that their dependence on the number of steps n is often too complicated. For this reason representations for transition probabilities have been sought in the form of closed formulae in which the dependence on n would be more simple, because such formulae are more useful for many purposes both theoretical and practical; let us recall here only the classical Perron's formula for powers of a matrix. This paper is a contribution to the solution of the problem; it deals with a certain very general solution of the representation problem. For individual chains (or classes of chains), of course, it would be necessary to specialize these general representations.

The paper appears to be a generalization of a paper by D. G. KENDALL [8] who studied the case of a denumerable state space. Our paper is based essentially on his ideas. Unfortunately, the representations derived by us are less satisfactory than those by Kendall; this is caused, on the one hand, by the basic fact that on a non-denumerable space there exists no σ-finite measure giving to each point a positive measure, on the other hand by the contemporaneous state of the Markov chain theory, since chains with a general state space are considerably less explored than those with a denumerable state space.

In the present paper some methods of functional analysis are used as fundamental ideas, a common practice now in probability theory. The well-known Markov operator T generated by the transition probabilities is considered in the space $\mathbf{L}_2$ and the spectral representations of a unitary, resp. self-adjoint operator in a Hilbert space are then used. In this way we get representations for the transition probabilities $p^{(n)}$ of the following types which are studied in the paper:

$$(1) \qquad p^{(n)}(.\,, A) = \int_0^{2\pi} e^{int}\mathbf{v}(dt)\,, \quad \int_B p^{(n)}(x, A)\,\mu(dx) = \int_0^{2\pi} e^{int}v(dt)\,,$$

$$(2) \qquad p^{(n)}(.\,, A) = \int_{-1}^{1} t^n\mathbf{v}(dt)\,, \quad \int_B p^{(n)}(x, A)\,\mu(dx) = \int_{-1}^{1} t^n v(dt)\,.$$

Of course, in order to apply the operator T in $\mathbf{L}_2$, we must have for the Markov chain a so called sub-invariant measure μ with respect to which integration is performed.

Following the introduction (part 1) the paper is concerned with representations for general Markov chains (part 2). First, some theorems on sub-invariant measures are proved. Here, let us draw attention to the theorem asserting the existence of an invariant finitely additive measure for an arbitrary Markov chain; by this theorem and its further applications, finitely additive measures are brought into consideration in Markov chain theory. Then theorems on representations of type (1) are presented, and finally the continuity properties of representing functions v are studied.

Part 3 is concerned with self-adjoint operators T. Firstly some conditions for self-adjointness and then theorems on representations of type (2) are given.

The concluding section (part 4) mentions a possibility of a further generalization and recapitulates the most interesting unsolved problems.

1.2. Notation and two known basic theorems. In the paper, we shall generally make use of the following system of notation.

Abstract spaces will be denoted by X, Z, their elements $x, z, \ldots$, their subsets $A, B, \ldots, E, F, G, \ldots$, σ-algebras in them Σ, Σ_X, Σ_Z. The set of all $x \in X$ such that the relation $R(x)$ holds will be denoted by $\{x; R(x)\}$ or shortly $\{R(x)\}$. The letter t will be reserved for a real variable.

Let us introduce the following convention: if Φ is a mapping then this mapping will be denoted either by the letter Φ only or when necessary by $\Phi(.)$. The symbol $\Phi(\omega)$ will always mean the value of the mapping Φ at the point ω. For numerical functions we shall use the letters $f, g, h, \ldots$, except that the characteristic function of the set E will be denoted by χ_E.

By the term "measure" we shall mean (somewhat differently from the usual meaning) a non-negative *finitely additive* set function which may take also infinite values, but which is *σ-finite*. An exception to σ-finiteness will be found in one place only, and we shall emphasize it explicitly there. If we have in mind σ-additive measures we shall always explicitly use this adjective. Measures in abstract spaces will be denoted by $\mu, \nu, \lambda, \ldots$, where, as a rule, μ is σ-additive, λ finitely additive. Complex (or real) numerical functions with bounded variation of a real variable will be denoted by $v, w, \ldots$. Further, for a mapping of a real interval into a Banach space we shall use the bold-type $\boldsymbol{v}$.

Abstract Lebesgue integrals with respect to the measure λ will be denoted by $\int_E f(x)\,\lambda(\mathrm{d}x)$, Riemann-Stieltjes integrals with respect to the function v of bounded variation by $\int_a^b f(t)\,v(\mathrm{d}t)$, and the symbol $\int_a^b f(t)\,\boldsymbol{v}(dt)$ is defined (in a manner similar to the definition used for the Riemann-Stieltjes integral) as the limit in norm (of the basic Banach space) of the Riemann sums; this integral, of course, is an element of that Banach space.

The letter i will always mean the imaginary unit, e the base of natural logarithms. For indices, i.e. natural numbers, we shall use the letters $j, k, l, m, n, \ldots$.

As in N. DUNFORD-J. T. SCHWARTZ [3] we will use the following notation of Banach spaces: $\mathbf{B}(X, \Sigma)$ is the space of all Σ-measurable bounded functions on X, $\mathbf{C}(X)$ is the space of all continuous bounded functions on the topological space X; the norm in both these spaces is given by $\|f\| = \sup_{x \in X} |f(x)|$. $\boldsymbol{ba}(X, \Sigma)$ is the space of all bounded finitely additive set functions on Σ with bounded variation, $\boldsymbol{ca}(X, \Sigma)$ is the space of all σ-additive set functions in $\boldsymbol{ba}(X, \Sigma)$, $\boldsymbol{rca}(X)$ is the space of all regular σ-additive set functions in $\boldsymbol{ba}(X, \Sigma)$ on the σ-algebra Σ of Borel sets in the topological space X; the norm in these three spaces is equal to the total variation of the set function. Complex (real) $\mathbf{L}_p(\mu)$ is the well-known space of complex-valued (real-valued) functions on X integrable in their p-th power with respect to measure μ, the norm in it being given by the usual formula. The norms of elements of Banach spaces will be denoted by $\|.\|$, the scalar product in a Hilbert space by (f, g). The letters $T, U, \ldots$ will denote operators in Banach spaces and by the term operator we shall always understand a linear continuous operator.

The transition probabilities of a Markov chain or a so called Markov system (see below) will be denoted by $p^{(n)}(x, E)$ or shortly $p^{(n)}$, their Radon-Nikodym derivatives or densities by $h^{(n)}$. For random variables we use the Greek letter ξ. The probability of the set E will be denoted by $\mathbf{P}\{E\}$, and if $E = \{x; R(x)\}$ then by $\mathbf{P}\{R(x)\}$. Similarly in another places we omit one bracket in fact and we write e.g. $\lambda\{R(x)\}$.

We must further clarify the development by defining the mathematical concepts of basic concern to this paper. Let X be an abstract space (of states of a Markov chain) with a σ-algebra Σ. By the term "*Markov chain*" we will mean a Markov chain in space X with a discrete parameter which is homogeneous in time (i.e. with stationary transition probabilities) acccording to the usual definition, i.e. the sequence of random variables ξ_n, $n = 0, 1, 2, \ldots$, with values in X, satisfying

$$\mathbf{P}\{\xi_{m+n} \in E | \xi_m, \ldots, \xi_0\} = \mathbf{P}\{\xi_{m+n} \in E | \xi_m\} = p^{(n)}(\xi_m, E)$$

with probability 1 for $m, n = 0, 1, 2, \ldots$, $E \in \Sigma$, where $p^{(n)}$, $n = 0, 1, 2, \ldots$, is a system of transition probabilities after n steps defined by this equation. On the other hand, we can choose a more general approach, from a functional point of view, as follows:

Definition 1.1. *A Markov system* (more completely *a Markov system of transition probabilities in the space* (X, Σ)) is a system of functions $p^{(n)}(.,.)$ of the variables $x \in X$, $E \in \Sigma$, for $n = 0, 1, 2, \ldots$ which has the following properties:

(a) $p^{(n)}(x, .)$ is a probability measure on Σ for each $x \in X$, $n = 0, 1, 2, \ldots$,
(b) $p^{(n)}(., E)$ is a real bounded Σ-measurable function on X for each $E \in \Sigma$, $n = 0, 1, 2, \ldots$,
(c) for $n, m = 0, 1, 2, \ldots$ so called transition relations

$$p^{(n+m)}(x, E) = \int_X p^{(n)}(x_1, E)\, p^{(m)}(x, \mathrm{d}x_1) \quad \text{for} \quad x \in X,\ E \in \Sigma$$

are satisfied.

We now have the following situation (under certain conditions which need not be discussed here): Every Markov chain defines a Markov system of transition probabilities; conversely, every Markov system defines a Markov chain supposing an initial probability p_0 is chosen according to which ξ_0 is distributed. (These facts are well-known and are exactly treated e.g. in J. L. Doob's book [2] pp. 86–88, 190.) Nevertheless, as P. Lévy and W. Feller [4] have noticed, there exist stochastic sequences ξ_n, $n = 0, 1, 2, \ldots$, which are not Markov chains although their transition probabilities

$$\mathbf{P}\{\xi_{m+n} \in E \mid \xi_m = x\} = p^{(n)}(x, E)$$

form a Markov system. Consequently, in this sense the concept of a Markov system is more general in its scope than a Markov chain.

Throughout the paper we shall always deal with Markov systems of transition probabilities in (X, Σ), we shall refer to them simply as Markov systems and denote them by $p^{(n)}$, and we shall also write $p^{(1)}(x, E) = p(x, E)$. Some common terms of Markov chain theory will be used also for Markov systems.

The letter T will always denote the mapping which a Σ-measurable function f on X (from a certain class of functions) maps onto a function Tf on X defined by the formula

$$Tf = \int_X f(x)\, p(., \mathrm{d}x)\,. \tag{3}$$

The mapping T is said to be generated by the Markov system $p^{(n)}$. The domain of definition of T will vary and will always be specified in the necessary place. Further, the symbol T^* will always denote the mapping which a set function $\lambda \in \boldsymbol{ba}(X, \Sigma)$ maps onto a set function $T^*\lambda$ on Σ defined by the formula

$$T^*\lambda = \int_X p(x, .)\, \lambda(\mathrm{d}x)\,. \tag{4}$$

The mappings T and T^* are well known, see e.g. [17], [13]. If $\boldsymbol{B}(X, \Sigma)$ is taken for the domain of T then T is an operator in $\boldsymbol{B}(X, \Sigma)$ the norm of which is 1. Similarly T^*

is an operator in $\boldsymbol{ba}(X, \Sigma)$ with the norm 1. Also it is well known that there is an isometric isomorphism between the space $\boldsymbol{ba}(X, \Sigma)$ and the space of functionals on $\boldsymbol{B}(X, \Sigma)$, see the theorem IV. 5. 1 in [3].

The following lemma is easily shown:

Lemma 1.1. *The operator T^* in $\boldsymbol{ba}(X, \Sigma)$ is the adjoint of the operator T in $\mathbf{B}(X, \Sigma)$.*

Proof. We have to prove that for arbitrary $f \in \mathbf{B}(X, \Sigma)$, $\lambda \in \boldsymbol{ba}(X, \Sigma)$

$$\lambda(Tf) = (T^*\lambda) f \tag{5}$$

or according to the theorem just quoted [3],

$$\int_X \left[\int_X f(x)\, p(x_1, \mathrm{d}x)\right] \lambda(\mathrm{d}x_1) = \int_X f(x) \left[\int_X p(x_1, \mathrm{d}x) . \lambda(\mathrm{d}x_1)\right]. \tag{6}$$

But if $f = \chi_E$, $E \in \Sigma$, the equality (6) obviously holds, and proceeding in the well-known way from the theory of integrals we arrive at a general $f \in \mathbf{B}(X, \Sigma)$ in (6). As to integration with respect to a finitely additive measure needed here see Dunford-Schwartz [3] III. 2. The analogue of the lemma is well known for $\mu \in \boldsymbol{ca}(X, \Sigma)$, and we give it here only for explicit elucidation of its more general validity.

For facilitation of reading we shall quote here two known theorems which are basic to our development using, of course, the present notation. First, however, it is necessary to introduce another basic definition.

Definition 1.2. We say that the Markov system $p^{(n)}$ has a *sub-invariant measure* λ, when λ is a measure on Σ for which the inequality

$$\int_X p(x, E)\, \lambda(\mathrm{d}x) \leqq \lambda(E) \tag{7}$$

holds for each $E \in \Sigma$. If in (7) the sign of equality holds for each $E \in \Sigma$ then we speak similarly of an *invariant measure.*

E. Nelson [13] has proved the following theorem:

Theorem I. *Let the Markov system $p^{(n)}$ have a sub-invariant σ-additive measure μ. If the mapping T is defined by the formula (3) for $f \in \mathbf{L}_p(\mu)$, where $1 \leqq p \leqq \infty$, then T is an operator in $\mathbf{L}_p(\mu)$ whose norm is $\|T\|_p \leqq 1$.*

Here $\mathbf{L}_p(\mu)$ may be complex or real, as it is easily seen from [13].

B. Sz.-Nagy [16] has proved the following theorem:

Theorem II. *Let V be an arbitrary contraction operator in a complex Hilbert space $\mathbf{H}_0$, i.e. $\|V\| \leqq 1$, and let V^+ be the adjoint of V. Then there exists a Hilbert space $\mathbf{H} \supset \mathbf{H}_0$ and a unitary operator U in $\mathbf{H}$ such that*

$$V^n = PU^n, \quad (V^+)^n = PU^{-n}, \quad n = 0, 1, 2, \ldots,$$

where P is an orthogonal projector of $\mathbf{H}$ onto $\mathbf{H}_0$.

2.1. Sub-invariant and invariant measures. As noted in introductory section 1.2, we shall suppose that we are given a Markov system $p^{(n)}$ in an abstract space X with a σ-algebra Σ. By Nelson's theorem (our Theorem I) if $p^{(n)}$ has a sub-invariant σ-additive measure μ then T generated by $p^{(n)}$ is an operator in $\mathbf{L}_p(\mu)$, particularly in complex $\mathbf{L}_2(\mu)$. This fact forms a basis for our representations of Markov systems which will be explained more thoroughly in the following section (2.2).

Therefore, the first question is the problem of existence of a sub-invariant measure. For a denumerable state space this problem was solved satisfactorily by D. G. Kendall [8] who showed that in this case every irreducible chain has a σ-additive sub-invariant measure. We shall derive now a somewhat similar result for a general state space, which is, however, considerably less satisfactory. We shall show that every irreducible Markov system (under rather general assumptions) has a non-trivial sub-invariant σ-additive measure which need not be, however, σ-finite. Here irreducibility is understood according to Nelson [13]:

Definition 2.1. A Markov system $p^{(n)}$ is said to be *irreducible* if the measures $\sum_{n=1}^{\infty} 2^{-n}p^{(n)}(x, .)$ are equivalent for each $x \in X$, i.e. they take the value 0 on the same sets.

Now let us use the Markov chain ξ_n generated by the Markov system $p^{(n)}$ in the following way: if x is a point from X we choose the initial probability p_0 for ξ_0 concentrating all its mass in the point x, and then we take $p^{(n)}$ as transition probabilities for ξ_n. (See also the text after definition 1.1 in section 1.2.) The resulting probability for the whole Markov chain ξ_n (i.e. on the space of sequences of elements from X) will be denoted by $\mathbf{P}_x$. Let us suppose that $p^{(n)}$ is irreducible and draw attention to the following two definitions of recurrence. Following Nelson [13], p. 683, $p^{(n)}$ is said to be recurrent if $\mathbf{P}_x\{\xi_n \in E$ for at least one $n\} = 1$ for each $x \in X$ and each $E \in \Sigma$ for which $\sum_{n=1}^{\infty} 2^{-n}p^{(n)}(x, E) > 0$; Harris [6] supposes that a σ-additive measure ν on Σ is given and then introduces condition C requiring $\mathbf{P}_x\{\xi_n \in E$ for infinitely many $n\} = 1$ for each $x \in X$ and each $E \in \Sigma$ for which $\nu(E) > 0$. Nelson's corollary 4.2 then shows that Nelson's condition of recurrence evidently implies Harris's condition C whichever of the measures $\sum_{n=1}^{\infty} 2^{-n}p^{(n)}(x, .)$ is taken for ν. Let us emphasize the important fact that the just mentioned probabilities in the conditions of recurrence depend on transition probabilities $p^{(n)}$ only, so that both kinds of recurrence are clearly properties of the Markov system $p^{(n)}$ in spite of their definition by means of the chain ξ_n.

Theorem 2.1. *Let $p^{(n)}$ be an irreducible Markov system in the space (X, Σ), where Σ is separable, i.e. Σ is a σ-algebra generated by a denumerable class of sets. If $p^{(n)}$ satisfies condition C for some σ-additive measure ν, then it has an* *invariant σ-additive σ-finite measure μ which is not identically zero. If* $p^{(n)}$ *does not satisfy condition* C *for* $v = \sum_{n=1}^{\infty} 2^{-n} p^{(n)}(x, .)$ *then it has at least a sub-invariant σ-additive measure μ which, however, need not be σ-finite, but which has the following property: there exists a set* $E_0 \in \Sigma$ *for which* $0 < \mu(E_0) < \infty$.

Proof. The first part of the theorem is merely the assertion proved by Harris [6] repeated here for completeness. In the second part, it can be seen from our preceding exposition that $p^{(n)}$ does not satisfy Nelson's condition. Nelson's theorem 4.1, however, gives the existence of $x_0 \in X$, $E_0 \in \Sigma$ such that $\sum_{n=1}^{\infty} 2^{-n} p^{(n)}(x_0, E_0) > 0$, $\sum_{n=1}^{\infty} p^{(n)}(x_0, E_0) < \infty$. Let us define now the measure $\mu(.) = \sum_{n=1}^{\infty} p^{(n)}(x_0, .)$. Clearly μ is a σ-additive measure (which, however, need not be σ-finite) and since

$$0 < \sum_{n=1}^{\infty} 2^{-n} p^{(n)}(x_0, E_0) \leqq \sum_{n=1}^{\infty} p^{(n)}(x_0, E_0) = \mu(E_0) < \infty ,$$

μ has the last property required in the theorem. Finally μ is seen to be sub-invariant because using the well-known theorem on interchangeability of summation and integration of a non-negative function we have

$$\int_X p(x, E)\, \mu(\mathrm{d}x) = \sum_{n=1}^{\infty} \int_X p(x, E)\, p^{(n)}(x_0, \mathrm{d}x) = \sum_{n=1}^{\infty} p^{(n+1)}(x_0, E) \leqq$$

$$\leqq \sum_{n=1}^{\infty} p^{(n)}(x_0, E) = \mu(E) .$$

Theorem 2.1 just proved is not, however, too satisfactory because the sub-invariant σ-additive measure μ constructed here can take the value ∞ for "too many" sets. It would be interesting to know under which conditions $p^{(n)}$ has a sub-invariant σ-additive σ-finite measure. Towards a solution of this problem, the already quoted result by Kendall [8] is known for a denumerable state space, further, a result by Harris [6] on invariant measures and some results by Nelson [13]. Nevertheless, none of the known results solves the problem satisfactorily for a general state space, and we also have not succeed in doing so.

Therefore, we turn to another direction: we shall abandon the requirement of σ-additivity and then we shall be able to prove a quite satisfactory result.

Theorem 2.2. *Every Markov system* $p^{(n)}$ *has an invariant measure* $\lambda \in \boldsymbol{ba}(X, \Sigma)$ *which is not identically zero.*

Proof. We shall here often use known facts from the book by Dunford-Schwartz [3]. There is an isometric isomorphism between $\boldsymbol{ba}(X, \Sigma)$ and the space of functionals on $\boldsymbol{B}(X, \Sigma)$, so we can introduce into $\boldsymbol{ba}(X, \Sigma)$ the well-known weak topology which will be called $\boldsymbol{B}$-topology. Explicitly written: a generalized sequence (in the sense of [3] I.7) λ_α of elements from $\boldsymbol{ba}(X, \Sigma)$ $\boldsymbol{B}$-converges to λ if and only if for each

$f \in \mathbf{B}(X, \Sigma)$ the equation $\lim_\alpha \int_X f(x)\, \lambda_\alpha(dx) = \int_X f(x)\, \lambda(dx)$ holds. The space $\mathbf{ba}(X, \Sigma)$ in its $\mathbf{B}$-topology is a locally convex linear topological space, see [3] V.3.

Let us now define $\mathbf{M} \subset \mathbf{ba}(X, \Sigma)$ as the set of all $\lambda \in \mathbf{ba}(X, \Sigma)$ which are non-negative and for which $\lambda(X) = 1$. Theorem V.4.2 in [3] states that the closed unit sphere in $\mathbf{ba}(X, \Sigma)$ is $\mathbf{B}$-compact. Let a generalized sequence λ_α of elements from $\mathbf{M}$ $\mathbf{B}$-converge to $\lambda \in \mathbf{ba}(X, \Sigma)$ If we take for f characteristic functions of the sets from Σ, we see that λ is non-negative and $\lambda(X) = 1$, which implies $\lambda \in \mathbf{M}$. Therefore $\mathbf{M}$ is $\mathbf{B}$-closed according to I.7.2 in [3], and being a subset of the $\mathbf{B}$-compact unit sphere $\mathbf{M}$ is $\mathbf{B}$-compact itself. Clearly $\mathbf{M}$ is a convex set.

Now let us use the operator T in $\mathbf{B}(X, \Sigma)$ generated by $p^{(n)}$ and its adjoint T^* in $\mathbf{ba}(X, \Sigma)$. Let λ_α $\mathbf{B}$-converge to λ. If $f \in \mathbf{B}(X, \Sigma)$ then also $Tf \in \mathbf{B}(X, \Sigma)$, and the definition of $\mathbf{B}$-convergence gives $\lim_\alpha \lambda_\alpha(Tf) = \lambda(Tf)$. However, using lemma 1.1 this equation can be rewritten in the form $\lim_\alpha (T^*\lambda_\alpha) f = (T^*\lambda) f$ which states that $T^*\lambda_\alpha$ $\mathbf{B}$-converges to $T^*\lambda$. Now by I.7.4 in [3] it is seen that T^* is a $\mathbf{B}$-continuous mapping. Finally it is evident that T^* maps $\mathbf{M}$ into $\mathbf{M}$.

The well-known Schauder-Tichonov fixed point theorem – see V.10.5 in [3] – now implies the existence of $\lambda \in \mathbf{M}$ such that $T^*\lambda = \lambda$, that is $\int_X p(x, .)\, \lambda(dx) = \lambda(.)$ which concludes the proof.

Of course, by this method a result on σ-additive measures can also be obtained.

Theorem 2.3. *Let $p^{(n)}$ be a Markov system in (X, Σ) where X is a compact Hausdorff space and Σ the σ-algebra of Borel sets in X. If the mapping T generated by $p^{(n)}$ maps $\mathbf{C}(X)$ into itself, then there exists $\mu \in \mathbf{rca}(X)$, not identically zero and satisfying $\int_X p(x, E)\, \mu(dx) = \mu(E)$ for all Baire sets E. If, moreover, T^* defined by (4) in 1.2 maps $\mathbf{rca}(X)$ into itself (e.g. this is the case for X second countable), then this measure μ is invariant (i.e. the preceding equation holds for all Borel sets E).*

The method of proof follows essentially that of theorem 2.2; we now need to change $\mathbf{B}(X, \Sigma)$ into $\mathbf{C}(X)$ and $\mathbf{ba}(X, \Sigma)$ into $\mathbf{rca}(X)$, and instead of simple integrals $\int_X \chi_E(x)\, \mu(dx) = \mu(E)$ we must make use of integrals $\int_X f(x)\, \mu(dx)$ with $f \in \mathbf{C}(X)$.

The following theorem is proved quite simply, but its result serves well for elucidation of the situation.

Theorem 2.4. *Let the Markov system $p^{(n)}$ have a sub-invariant finitely additive measure λ which is finite. Then λ is invariant.*

Proof. If $\int_X p(x, E)\, \lambda(dx) < \lambda(E)$ for some set $E \in \Sigma$, then

$$\int_X p(x, X \quad E)\, \lambda(dx) = \int_X [p(x, X) - p(x, E)]\, \lambda(dx) = \int_X p(x, X)\, \lambda(dx) -$$
$$- \int_X p(x, E)\, \lambda(dx) > \lambda(X) - \lambda(E) = \lambda(X - E)\,,$$

but this is a contradiction of the sub-invariance of the measure λ.

Let us recall the following definition and theorem – (see Dunford-Schwartz [3] III.7.7–8): A non negative $\lambda \in \mathbf{ba}(X, \Sigma)$ is said to be *purely finitely additive* in case $0 \leqq \mu \leqq \lambda$ and $\mu \in \mathbf{ca}(X, \Sigma)$ imply that $\mu = 0$. To every non-negative $\lambda \in \mathbf{ba}(X, \Sigma)$ there exists its unique decomposition into a σ-additive and purely finitely additive part expressed by $\lambda = \sigma + \kappa$, where $\sigma \geqq 0$ and $\sigma \in \mathbf{ca}(X, \Sigma)$, $\kappa \geqq 0$, $\kappa \in \mathbf{ba}(X, \Sigma)$ and κ is purely finitely additive.

Theorem 2.5. *Let the Markov system $p^{(n)}$ have an invariant measure $\lambda \in \mathbf{ba}(X, \Sigma)$ If $\lambda = \sigma + \kappa$ is its decomposition into a σ-additive and purely finitely additive part, then both λ and κ are invariant measures.*

Proof. If we use the operator T^* in $\mathbf{ba}(X, \Sigma)$ given by formula (4) in section 1.2, we have $T^*\lambda = \lambda$, that is $T^*(\sigma + \kappa) = T^*\sigma + T^*\kappa = \sigma + \kappa$. Because evidently $T^*\sigma$ is σ-additive, we see that the set function $\kappa - T^*\kappa = T^*\sigma - \sigma$ is σ-additive. Further also its upper variation $(\kappa - T^*\kappa)^+$ is σ-additive, and since $\kappa \geqq 0$, $T^*\kappa \geqq 0$, it is easily seen that $0 \leqq (\kappa - T^*\kappa)^+ \leqq \kappa$. But by assumption κ is purely finitely additive, therefore $(\kappa - T^*\kappa)^+ = 0$, that is $\kappa - T^*\kappa \leqq 0$. Thus we obtained $\kappa \leqq T^*\kappa$ from which $T^*\sigma \leqq \sigma$ follows. Now theorem 2.4 states that a finite sub-invariant measure is invariant; because $\sigma \in \mathbf{ca}(X, \Sigma)$, or σ is finite, we conclude $T^*\sigma = \sigma$ and also $T^*\kappa = \kappa$.

Corollary 2.1. *If $p^{(n)}$ has an invariant measure $\lambda \in \mathbf{ba}(X, \Sigma)$ which is not purely finitely additive, then it has an invariant σ-additive measure $\mu \in \mathbf{ca}(X, \Sigma)$ which is not identically zero.*

The proof is obvious from theorem 2.5.

2.2. Representations for transition probabilities. As usual we are given a Markov system $p^{(n)}$ and the mapping T generated by it. Let $p^{(n)}$ have a sub-invariant σ-additive σ-finite measure μ. If we take complex $\mathbf{L}_2(\mu)$ as the domain of definition of T, then Nelson's result [13] (see our theorem I) states that T is an operator in $\mathbf{L}_2(\mu)$. Let us point out that $\mathbf{L}_2(\mu)$ is a complex Hilbert space. We denote by T^+ the adjoint operator of T in $\mathbf{L}_2(\mu)$. Let us suppose that (X, Σ) has the same structure as a Borel set of the real line, i.e. there is a one-to-one mapping between X and a Borel set X' of the real line such that a subset of X is in Σ if and only if its image in X' is a Borel set. Under these assumptions Nelson's theorem 3.1 in [13] states that there always exists a function $p^+(.,.)$ which is a sub-stochastic (i.e. $p^+(x, X) \leqq 1$) transition probability such that

$$(1) \qquad T^+f = \int_X f(x)\, p^+(., \mathrm{d}x) \quad \text{for} \quad f \in \mathbf{L}_2(\mu)\,.$$

Let us define $p^{+(n)}(.,.)$ for $n = 1, 2, \ldots$ by a recurrent formula

$$(2) \qquad p^{+(n)}(x, E) = \int_X p^{+(n-1)}(z, E)\, p^+(x, \mathrm{d}z)\,, \quad x \in X,\; E \in \Sigma\,,$$

where we put $p^{+(1)} = p^+$.

Further we shall use the notation $h(x, .) = \mathrm{d}p(x, .)/\mathrm{d}\mu(.)$ for the Radon-Nikodym derivative provided it exists. In the case of an irreducible Markov system by Nelson's theorem 3.2 in [13] this derivative exists as well as the derivative $\mathrm{d}p^+(x, .)/\mathrm{d}\mu(.) = h^+(x, .)$ and it is $h^+(x, y) = h(y, x)$ for $(\mu \times \mu)$-almost all (x, y).

Theorem 2.6. *Let $p^{(n)}$ be a Markov system which has a sub-invariant σ-additive measure μ. Let T be the operator in complex $\mathbf{L}_2(\mu)$ generated by $p^{(n)}$, T^+ its adjoint, and let the functions $p^{+(n)}$ satisfy the formulae (1), (2). Then for each set $A \in \Sigma$ for which $\mu(A) < \infty$ there is a mapping $\mathfrak{v}(A; .)$ of the interval $[0, 2\pi]$ into complex $\mathbf{L}_2(\mu)$ such that*

$$(3) \qquad p^{(n)}(., A) = \int_0^{2\pi} e^{int}\mathfrak{v}(A; \mathrm{d}t) \quad \textit{for} \quad n = 0, 1, 2, \ldots ,$$

$$p^{+(n)}(., A) = \int_0^{2\pi} e^{-int}\mathfrak{v}(A; \mathrm{d}t) \quad \textit{for} \quad n = 0, 1, 2, \ldots ,$$

where equalities of functions in $\mathbf{L}_2(\mu)$ are meant, i.e. μ-almost everywhere.

Proof. By Nelson's theorem (our theorem I) the operator T in the space $\mathbf{L}_2(\mu)$ has the norm $\|T\| \leqq 1$, i.e. it is a contraction operator in the Hilbert space $\mathbf{L}_2(\mu)$. By the Sz.-Nagy's theorem [16], quoted in our paper as theorem II, there is a Hilbert space $\boldsymbol{H} \supset \mathbf{L}_2(\mu)$, a unitary operator U in $\boldsymbol{H}$ and a projection operator P of $\boldsymbol{H}$ onto $\mathbf{L}_2(\mu)$ such that

$$T^n f = PU^n f, \quad (T^+)^n f = PU^{-n} f$$

for $n = 0, 1, 2, \ldots$ and for arbitrary $f \in \mathbf{L}_2(\mu)$. If we make use of the spectral representation of a unitary operator and if we denote by E_t its corresponding spectral system of projections, we have

$$T^n f = P\int_0^{2\pi} e^{int}\, \mathrm{d}(E_t f) = \int_0^{2\pi} e^{int}\mathfrak{v}(f; \mathrm{d}t),$$

where $\mathfrak{v}(f; t) = PE_t f \in \mathbf{L}_2(\mu)$ for $t \in [0, 2\pi]$.

If now $\mu(A) < \infty$, then the characteristic function $\chi_A \in \mathbf{L}_2(\mu)$ and we obtain

$$T^n \chi_A = \int_X \chi_A(x)\, p^{(n)}(., \mathrm{d}x) = p^{(n)}(., A) = \int_0^{2\pi} e^{int}\mathfrak{v}(A; \mathrm{d}t)$$

for a certain mapping $\mathfrak{v}(A; .)$. It is easily seen that $(T^+)^n f = \int_X f(x)\, p^{+(n)}(., \mathrm{d}x)$ and therefore we obtain similarly

$$(T^+)^n \chi_A = \int_X \chi_A(x)\, p^{+(n)}(., \mathrm{d}x) = p^{+(n)}(., A) = \int_0^{2\pi} e^{-int}\mathfrak{v}(A; \mathrm{d}t).$$

Corollary 2.2. *For an irreducible Markov system $p^{(n)}$ equalities (3) can be written in the form*

$$\int_A h^{(n)}(., x)\, \mu(\mathrm{d}x) = \int_0^{2\pi} e^{int}\mathfrak{v}(A; \mathrm{d}t) \quad \textit{for} \quad n = 0, 1, 2, \ldots ,$$

$$\int_A h^{(n)}(x, .)\, \mu(\mathrm{d}x) = \int_0^{2\pi} e^{-int}\mathfrak{v}(A; \mathrm{d}t) \quad \textit{for} \quad n = 0, 1, 2, \ldots ,$$

where $h^{(n)}$ is the usual n-th iteration of the density h defined by the recurrent formula

$$h^{(n)}(x, y) = \int_X h^{(n-1)}(x, z)\, h(z, y)\, \mu(\mathrm{d}z)\,, \quad h^{(1)} = h\,.$$

Proof. Firstly $h^{(n)}(x, .) = \mathrm{d}p^{(n)}(x, .)/\mathrm{d}\mu(.)$ almost everywhere which implies immediately the first equality. If we define further $h^{+(n)}$ by a similar recurrent formula as $h^{(n)}$, then we see similarly $h^{+(n)}(x, .) = \mathrm{d}p^{+(n)}(x, .)/\mathrm{d}\mu(.)$ almost everywhere, and by an induction process it is easily shown $h^{+(n)}(x, y) = h^{(n)}(y, x)$ for $(\mu \times \mu)$-almost all (x, y). Therefore $p^{+(n)}(., A) = \int_A h^{+(n)}(., x)\, \mu(\mathrm{d}x) = \int_A h^{(n)}(x, .)\, \mu(\mathrm{d}x)$.

Theorem 2.7. *Let $p^{(n)}$ be a Markov system which has a sub-invariant σ-additive measure μ. Then for each two sets $A, B \in \Sigma$, for which $\mu(A) < \infty$, $\mu(B) < \infty$, there is a complex-valued function $v(A, B; .)$ of bounded variation in $[0, 2\pi]$ such that*

$$\text{(4)} \qquad \int_B p^{(n)}(x, A)\, \mu(\mathrm{d}x) = \int_0^{2\pi} e^{int} v(A, B; \mathrm{d}t) \quad \textit{for} \quad n = 0, 1, 2, \ldots\,,$$

$$\int_A p^{(n)}(x, B)\, \mu(\mathrm{d}x) = \int_0^{2\pi} e^{-int} v(A, B; \mathrm{d}t) \quad \textit{for} \quad n = 0, 1, 2, \ldots\,.$$

The equality $\overline{v(A, B; t)} = v(B, A; t)$ holds where the bar denotes the complex conjugate number provided the functions v are standardized in some appropriate way, e.g. they are right-continuous and $v(A, B; 0) = v(B, A; 0) = 0$.

The proof is similar to the proof used in theorem 2.6. For the unitary operator U we now use the representation of scalar products

$$(U^n f, g) = \int_0^{2\pi} e^{int} v(f, g; \mathrm{d}t) \quad \text{for} \quad f, g \in \mathbf{L}_2(\mu)\,,$$

$n = \ldots, -1, 0, 1, \ldots$, where $v(f, g; .)$ is a certain function of bounded variation in $[0, 2\pi]$. If we put the characteristic function χ_A in place of f and χ_B in place of g, we obtain

$$(T^n \chi_A, \chi_B) = \int_X \left[\int_X \chi_A(x)\, p^{(n)}(y, \mathrm{d}x)\right] \chi_B(y)\, \mu(\mathrm{d}y) = \int_B p^{(n)}(y, A)\, \mu(\mathrm{d}y)$$

for $n = 0, 1, 2, \ldots$. On the other hand

$$(T^n \chi_A, \chi_B) = (PU^n \chi_A, \chi_B) = (U^n \chi_A, P\chi_B) = (U^n \chi_A, \chi_B)\,,$$

from which the first formula in (4) follows. The second formula follows quite analogously, for

$$((T^+)^n \chi_A, \chi_B) = (\chi_A, T^n \chi_B) = \int_A p^{(n)}(x, B)\, \mu(\mathrm{d}x)\,, \quad (T^+)^n \chi_A = PU^{-n} \chi_A\,.$$

The last assertion of the theorem is immediately evident if we take in (4) complex conjugate numbers, having in mind a well-known fact that equality of all Fourier-

Stieltjes coefficients implies equality of standardized functions of bounded variation (see e.g. [18], p. 41).

Let us notice that the theorems 2.6 and 2.7 generalize similar representations for a denumerable state space given by D. G. Kendall [7], [8] in his theorem II.

We return now to our result in theorem 2.2 that every Markov system has an invariant measure $\lambda \in \boldsymbol{ba}(X, \Sigma)$ and, in order to make use of it, we shall be concerned with the representations on the basis of a finitely additive measure. We shall use here the theory of integration with respect to a finitely additive measure λ as it is developed in Dunford-Schwartz [3] III.1.−3. Of course, we work here with functions measurable with respect to a σ-algebra Σ and with a non-negative measure λ, so much is simplified in comparison with [3]. Let us recall that a measurable function f is called a null function if for every $\varepsilon > 0$ we have $\lambda\{|f(x)| > \varepsilon\} = 0$. We restrict ourselves further only to bounded measurable functions and to functions differing from these by a null function. As usual, two functions differing by a null function are treated as identical.

Notation. Let $1 \leqq p < \infty$. By the symbol $\boldsymbol{B}_p(\lambda)$ let us denote the space of all bounded measurable functions f (more exactly: of classes of equivalence) for which $[\int_X |f(x)|^p \, \lambda(\mathrm{d}x)]^{1/p} < \infty$ and this quantity will be denoted by $\|f\|_p$.

It is easily seen that $\|f\|_p$ has the properties of a norm. With this norm and with the usual summation of two functions and multiplication of a function by a real number, $\boldsymbol{B}_p(\lambda)$ is a normed linear space (which need not, however, be complete); it is a dense sub-space of the Lebesgue space $\boldsymbol{L}_p$ defined in III. 3 of the book [3].

Theorem 2.8. *Let the Markov system $p^{(n)}$ have a sub-invariant measure $\lambda \in \in \boldsymbol{ba}(X, \Sigma)$. If the mapping T is defined by formula (3) in section 1.2 for $f \in \boldsymbol{B}_p(\lambda)$, $1 \leqq p < \infty$, then T is an operator in $\boldsymbol{B}_p(\lambda)$, whose norm is $\|T\|_p = 1$.*

Remark. For the proof, sub-invariance of λ suffices formally, but by theorem 2.4 such a finite λ is invariant.

Proof of theorem 2.8. Firstly T has to be well defined on the equivalence classes, for which it is sufficient to show that $\|f\|_1 = 0$ implies $\|Tf\|_1 = 0$. But this fact is easily seen, because then f is a null function and therefore for every $\varepsilon > 0$ we have

$$\|Tf\|_1 = \int_X |(Tf)(x)| \, \lambda(\mathrm{d}x) = \int_X \left| \int_X f(y) \, p(x, \mathrm{d}y) \right| \lambda(\mathrm{d}x) \leqq$$

$$\leqq \int_X \int_{\{|f(y)| \leqq \varepsilon\}} |f(y)| \, p(x, \mathrm{d}y) \, \lambda(\mathrm{d}x) + \int_X \int_{\{|f(y)| > \varepsilon\}} |f(y)| \, p(x, \mathrm{d}y) \, \lambda(\mathrm{d}x) \leqq$$

$$\leqq \int_X \varepsilon \, . \, p(x, \{|f(y)| \leqq \varepsilon\}) \, \lambda(\mathrm{d}x) + \int_X \infty \, . \, p(x, \{|f(y)| > \varepsilon\}) \, \lambda(\mathrm{d}x) \leqq$$

$$\leqq \varepsilon \, . \, \lambda(X) + \infty \, . \, \lambda\{|f(y)| > \varepsilon\} = \varepsilon \, . \, \lambda(X) \, .$$

Now let f be a simple measurable function $f = \sum_{j=1}^{k} a_j \chi_{E_j}$. Then for $1 \leqq p < \infty$ we obtain

$$\|Tf\|_p^p = \int_X |(Tf)(x)|^p \lambda(dx) = \int_X \left| \int_X f(y)\, p(x, dy) \right|^p \lambda(dx) \leqq$$

$$\leqq \int_X \int_X |f(y)|^p p(x, dy)\, \lambda(dx) = \sum_{j=1}^{k} |a_j|^p \int_X p(x, E_j)\, \lambda(dx) \leqq \sum_{j=1}^{k} |a_j|^p \lambda(E_j) =$$

$$= \int_X |f(x)|^p \lambda(dx) = \|f\|_p^p .$$

Because the uniform convergence implies the convergence in norm of $\mathbf{B}_p(\lambda)$ by means of the well-known approximation by simple functions from the above inequalities we obtain $\|Tf\|_p \leqq \|f\|_p$ for each $f \in \mathbf{B}_p(\lambda)$. If we take for f the function identically equal to 1, it is $Tf = f$, $\|Tf\|_p = \|f\|_p$. As all of the other properties of T are clear, the proof is finished.

If a scalar product is introduced into $\mathbf{B}_2(\lambda)$ by the usual formula as into $\mathbf{L}_2$, the space $\mathbf{B}_2(\lambda)$ clearly satisfies the axioms of Hilbert space except it need not be complete. Nevertheless, we can use the results by Sz.-Nagy [16]; if we go through the proof of his theorem I in [16] (our theorem II in section 1.2) we see that he has proved, in fact, the following theorem:

Theorem II′. *If $\mathbf{B}$ is a space satisfying the axioms of Hilbert space except possibly completeness and if T is an arbitrary contraction operator in it, then for each $f, g \in \mathbf{B}$ there is a complex-valued function $v(f, g; .)$ of bounded variation on $[0, 2\pi]$ such that*

$$(T^n f, g) = \int_0^{2\pi} e^{int} v(f, g; dt) \quad \textit{for} \quad n = 0, 1, 2, \dots ,$$

$$(f, T^n g) = \int_0^{2\pi} e^{-int} v(f, g; dt) \quad \textit{for} \quad n = 1, 2, \dots .$$

Therefore, it is possible to prove an analogue of theorem 2.7. As we have already pointed out, the advantage of this modification lies in that every Markov system has an invariant finitely additive finite measure.

Theorem 2.9. *If the Markov system $p^{(n)}$ has a sub-invariant measure $\lambda \in \mathbf{ba}(X, \Sigma)$, then for each two sets $A, B \in \Sigma$ there is a complex-valued function $v(A, B; .)$ of bounded variation on $[0, 2\pi]$ such that*

$$(5) \qquad \int_B p^{(n)}(x, A)\, \lambda(dx) = \int_0^{2\pi} e^{int} v(A, B; dt) \quad \textit{for} \quad n = 0, 1, 2, \dots ,$$

$$\int_A p^{(n)}(x, B)\, \lambda(dx) = \int_0^{2\pi} e^{-int} v(A, B; dt) \quad \textit{for} \quad n = 1, 2, \dots .$$

The equality $\overline{v(A, B; t)} = v(B, A; t)$ holds provided the functions v are standardized as in theorem 2.7.

The proof is easy on the basis of Nagy's theorem II′ if we take $\mathbf{B}_2(\lambda)$ in place of $\mathbf{B}$ and characteristic functions χ_A, χ_B in place of the elements f, g.

Unfortunately, an analogue of theorem 2.6 cannot be proved, for here the completeness of the Hilbert space plays an essential role; by an analysis of proofs we can see that it is needed in order that the bilinear functional $v(f, g; t)$ may be generated by an operator V_t by the identity $v(f, g; t) = (V_t f, g)$.

But if we took the completed space $\overline{\mathbf{B}}$ of the space $\mathbf{B}_2(\lambda)$, evidently it would be possible to prove

$$p^{(n)}(., A) = \int_0^{2\pi} e^{int}\mathbf{v}(A; \mathrm{d}t) \quad \text{for} \quad n = 0, 1, 2, \dots ,$$

where $\mathbf{v}(A; t) \in \overline{\mathbf{B}}$ for each $t \in [0, 2\pi]$ and it is understood that the equalities are in the sense of space $\overline{\mathbf{B}}$.

According to the results by S. LEADER [12] it is possible to take as a completed space $\overline{\mathbf{B}}$ the space $\mathbf{V}^2$ of finitely additive set functions absolutely continuous with respect to λ for which a certain norm is finite. Leader has shown that $\mathbf{V}^2$ is complete and that simple functions are dense in $\mathbf{V}^2$; therefore, a fortiori, $\mathbf{B}_2(\lambda)$ is dense in $\mathbf{V}^2$. Nevertheless, we obtain nothing new: first of all, speaking now more exactly, $\mathbf{B}_2(\lambda)$ must be mapped onto a sub-space of $\mathbf{V}^2$ (as in Leader) and then the element $p^{(n)}(., A) \in \mathbf{B}_2(\lambda)$ is mapped onto the set function $v^{(n)}(A; .) \in \mathbf{V}^2$ defined by $v^{(n)}(A; B) = \int_B p^{(n)}(x, A)\, \lambda(\mathrm{d}x)$ for $B \in \Sigma$. Thus we should obtain a mapping $\mathbf{v}(A; .)$ of the interval $[0, 2\pi]$ into $\mathbf{V}^2$ such that

$$v^{(n)}(A; .) = \int_0^{2\pi} e^{int}\mathbf{v}(A; \mathrm{d}t) \quad \text{for} \quad n = 0, 1, 2, \dots ,$$

where it is understood that the equalities are in the space $\mathbf{V}^2$, but these equations express only a little more than the theorem 2.9.

At the end of this section let us point out that in the often treated K. YOSIDA-S. KAKUTANI [17] case of a quasi-compact operator T (we use here a more consistent terminology speaking of a "compact" operator as in [3] instead of a "completely continuous" operator) a better representation can be simply derived.

Theorem 2.10. *Let $p^{(n)}$ be a Markov system and let the mapping T generated by it be defined in $\mathbf{B}(X, \Sigma)$. Let there exist a compact operator V in $\mathbf{B}(X, \Sigma)$ and a natural number m such that $\|T^m - V\| < 1$. Then for each $x \in X$, $E \in \Sigma$ there exists a function $w(x, E; .)$ of bounded variation on $[0, 2\pi]$ such that*

$$p^{(n)}(x, E) = \int_0^{2\pi} e^{int} w(x, E; \mathrm{d}t) \quad \textit{for} \quad n = 1, 2, \dots .$$

The function w may have jumps only at the points $2\pi j/d$, $j = 0, 1, \dots, d - 1$, where d is a natural number, and the continuous part of w is absolutely continuous.

Proof. Yosida-Kakutani [17] under the said assumptions have shown that

$$p^{(n)}(x, E) = \sum_{j=1}^{d} r_j^n p_j(x, E) + s^{(n)}(x, E),$$

where r_j are d-th roots of 1, $\sup_{x,E} |s^{(n)}(x, E)| \leqq K(1 + \varepsilon)^{-n}$, K, ε being positive constants.

Since $\sum_{n=1}^{\infty} |s^{(n)}(x, E)|^2 \leqq \sum_{n=1}^{\infty} K^2(1 + \varepsilon)^{-2n} < \infty$, by the well-known Riesz-Fischer theorem – see e.g. A. ZYGMUND [18] IV. 1. 1 – we obtain

$$s^{(n)}(x, E) = \int_0^{2\pi} e^{int} g(t)\, dt, \quad n = 1, 2, \dots,$$

where g is square integrable on $[0, 2\pi]$. Integrating g and adding the jumps $p_j(x, E)$ at the d-th roots of unity r_j we get the function w and the assertion of the theorem.

2.3. Continuity properties of representing functions v of bounded variation. In the whole section we suppose $v(A, B; t) = \frac{1}{2}[v(A, B; t + 0) + v(A, B; t - 0)]$ or right-continuity or left-continuity of the functions $v(A, B; .)$ by which, of course, the values of integrals with respect to v are not changed.

Quite similarly as in Kendall's paper [8] the following theorem is proved.

Theorem 2.11. *Let $p^{(n)}$ be a Markov system. Let $A, B \in \Sigma$, λ be a measure on Σ, $\lambda(A) < \infty$, $\lambda(B) < \infty$, and let $v(A, B; .)$ be a function of bounded variation on $[0, 2\pi]$ such that formulae (5) hold. Let there exist a natural number d and natural numbers $r(A, B)$, $r(B, A)$ between 0 and $d - 1$ such that*

$$\lim_m \int_A p^{(md+r(A,B))}(x, B)\, \lambda(dx) = L(A, B), \quad \int_A p^{(n)}(x, B)\, \lambda(dx) = 0$$
$$\text{for} \quad n \neq md + r(A, B),$$

$$\lim_m \int_B p^{(md+r(B,A))}(x, A)\, \lambda(dx) = L(B, A), \quad \int_B p^{(n)}(x, A)\, \lambda(dx) = 0$$
$$\text{for} \quad n \neq md + r(B, A).$$

Then the function $v(A, B; .)$
(a) is continuous except at most the points $2\pi j/d$, $j = 0, 1, \dots, d - 1$;
(b) at the point $2\pi j/d$ has a jump of magnitude

$$(2d)^{-1}[L(A, B)\, e^{ir(A,B)2\pi j/d} + L(B, A)\, e^{-ir(B,A)2\pi j/d}].$$

Proof. Let us denote for brevity $q_m = \int_A p^{(md+r(A,B))}(x, B)\, \lambda(dx)$. For $j = 0, 1, \dots, \dots, d - 1$ we have

$$(6)\ \lim_n n^{-1} \sum_{k=1}^{n} \int_A p^{(k)}(x, B)\, \lambda(dx)\, e^{ik2\pi j/d} = \lim_n n^{-1} \sum_{m=1}^{[(n-r(A,B))/d]+1} q_m e^{i(md+r(A,B))2\pi j/d}.$$

If we let $N = [(n - r(A, B))/d] + 1$, then (6) becomes

$$\lim_N (Nd)^{-1} \sum_{m=1}^{N} q_m e^{ir(A,B)2\pi j/d} = d^{-1}L(A, B)\, e^{ir(A,B)2\pi j/d}, \tag{7}$$

for clearly the assumptions imply also $\lim_N N^{-1} \sum_{m=1}^{N} q_m = L(A, B)$. Similarly, it can be proved that

$$\lim_n n^{-1} \sum_{k=1}^{n} \int_B p^{(k)}(x, A)\, \lambda(dx)\, e^{-ik2\pi j/d} = d^{-1}L(B, A)\, e^{-ir(B,A)2\pi j/d}, \tag{8}$$

so that by theorem III. 9. 3 in Zygmund [18] formulae (7) and (8) imply the truth of assertion (b).

Further, let us suppose that $t \in [0, 2\pi]$ does not have the form $2\pi j/d$. We have

$$\lim_n n^{-1} \sum_{k=1}^{n} \int_A p^{(k)}(x, B)\, \lambda(dx)\, e^{ikt} = \lim_N (Nd)^{-1} \sum_{m=1}^{N} q_m e^{i(md+r(A,B))t} = \tag{9}$$

$$= \lim_N (Nd)^{-1} \Big[\sum_{m=1}^{N-1} S_m(q_m - q_{m+1}) + S_N q_N \Big],$$

where the last expression was obtained by means of Abel's partial summation, and where

$$S_m = \sum_{k=1}^{m} e^{it(kd+r(A,B))}. \tag{10}$$

For $t \neq 2\pi j/d$, $td \neq 2\pi j$, and $e^{itd} \neq 1$, and summing the geometrical series in (10) we have

$$|S_m| = \Big|\sum_{k=1}^{m} e^{it(kd+r(A,B))}\Big| = \left| e^{it(d+r(A,B))} \frac{e^{itmd} - 1}{e^{itd} - 1} \right| \leqq \frac{2}{|e^{itd} - 1|} = c < \infty. \tag{11}$$

By (9) and (11) follows that

$$\lim_n \left| n^{-1} \sum_{k=1}^{n} \int_A p^{(k)}(x, B)\, \lambda(dx)\, e^{ikt} \right| \leqq \lim_N (Nd)^{-1} \Big[\sum_{m=1}^{N-1} |S_m| \cdot |q_m - q_{m+1}| + \tag{12}$$

$$+ |S_N| \cdot |q_N| \Big] \leqq \lim_N c\,.\,(Nd)^{-1} \sum_{m=1}^{N-1} |q_m - q_{m+1}| + \lim_N c\,.\,(Nd)^{-1}\, |q_N|\,.$$

Since $\lim_m q_m = L(A, B)$, then $\lim_m |q_m - q_{m+1}| = 0$, and therefore $\lim_N N^{-1} \sum_{m=1}^{N-1} |q_m - q_{m+1}| = 0$; finally $|q_N| = |\int_A p^{(Nd+r(A,B))}(x, B)\, \lambda(dx)| \leqq \lambda(A) < \infty$. In summary, (12) is equal to 0 and therefore

$$\lim_n n^{-1} \sum_{k=1}^{n} \int_A p^{(k)}(x, B)\, \lambda(dx)\, e^{ikt} = 0\,.$$

In a quite similar manner it may be proved that

$$\lim_n n^{-1} \sum_{k=1}^{n} \int_B p^{(k)}(x, A)\, \lambda(dx)\, e^{-ikt} = 0\,,$$

and again by theorem III.9.3 in Zygmund [18] assertion (a) of our theorem is true.

The theorem just proved can be used for periodic chains, provided the transition probabilities converge in the usual sense:

Corollary 2.3. *Let $p^{(n)}$ be a Markov system which has a sub-invariant σ-additive measure μ. Let $p^{(n)}$ have the period d, i.e. there is a decomposition of the space X into disjoint sets $C_0, C_1, \ldots, C_{d-1}, D$ from Σ such that $\mu(D) = 0$ and for $x \in C_k$ we have $p(x, C_{k+1}) = 1$ (where we put $C_d = C_0$). Further let $r(k, l)$ denote the least non-negative number n such that for $x \in C_k$, $p^{(n)}(x, C_l) = 1$. Let $\lim\limits_m p^{(md+r(k,l))}(x, E) = p_0(x, E)$ for arbitrary $x \in X$, $E \in \Sigma$ with $x \in C_k$, $E \subset C_l$. If $A, B \in \Sigma$ are such that $\mu(A) < \infty$, $\mu(B) < \infty$, then theorem 2.7 states the validity of formulae (4) and we have further:*

(*a*) *the function $v(A, B; .)$ is continuous except at most the points $2\pi j/d$, $j = 0, 1, \ldots, d-1$;*

(*b*) *if $A \subset C_k$, $B \subset C_l$, then $v(A, B; .)$ has at the point $2\pi j/d$ a jump of magnitude*

$$(2d)^{-1} e^{ir(k,l)2\pi j/d} \left[\int_A p_0(x, B)\, \mu(dx) + \int_B p_0(x, A)\, \mu(dx) \right].$$

Proof. The assumptions of theorem 2.11 are evidently satisfied for the measure μ in place of λ, $A \subset C_k$, $B \subset C_l$, for 1 is an integrable dominating function on A and B and therefore $\lim\limits_m \int_A p^{(md+r(k,l))}(x, B)\, \mu(dx) = \int_A p_0(x, B)\, \mu(dx) = L(A, B)$ and analogously for A, B interchanged. We have obviously $r(k, l) + r(l, k) = d$ for $k \neq l$ and therefore

$$e^{ir(k,l)2\pi j/d} = e^{i(d-r(l,k))2\pi j/d} = e^{-ir(l,k)2\pi j/d};$$

for $k = l$ the equality of both extreme terms is evident. The assertion (b) is now easily seen, as

$$(2d)^{-1}\left[L(A, B)\, e^{ir(k,l)2\pi j/d} + L(B, A)\, e^{-ir(l,k)2\pi j/d}\right] =$$
$$= (2d)^{-1}\, e^{ir(k,l)2\pi j/d} \left[\int_A p_0(x, B)\, \mu(dx) + \int_B p_0(x, A)\, \mu(dx) \right].$$

Further, let us have arbitrary sets $A, B \in \Sigma$ of a finite measure. Since $\mu(D) = 0$ and for $x \in X - D$ we have $p^{(n)}(x, D) = 0$, we obtain $\int_B p^{(n)}(x, A)\, \mu(dx) = \int_{B-D} p^{(n)}(x, A - D)\, \mu(dx)$. Let us denote $A_k = (A - D) \cap C_k$, $B_l = (B - D) \cap C_l$ for $k, l = 0, 1, \ldots, d-1$. Each function $v(A_k, C_l; .)$ is by theorem 2.11 continuous at the points $t \neq 2\pi j/d$. But we have

$$\int_0^{2\pi} e^{int} v(A, B; dt) = \int_B p^{(n)}(x, A)\, \mu(dx) = \sum_{l=0}^{d-1} \int_{B_l} \sum_{k=0}^{d-1} p^{(n)}(x, A_k)\, \mu(dx) =$$
$$= \sum_{l=0}^{d-1} \sum_{k=0}^{d-1} \int_0^{2\pi} e^{int} v(A_k, B_l; dt)$$

and analogously for A, B interchanged. The equality of all Fourier-Stieltjes coefficients implies (e.g. by [18]) the equality $v(A, B; t) = \sum_{l=0}^{d-1} \sum_{k=0}^{d-1} v(A_k, B_l; t)$ and thus assertion (a) is proved.

It is easily seen that in the special case of chains with a denumerable state space our results agree with those by Kendall [7], [8].

Theorem 2.12. *Let $p^{(n)}$ be a Markov system and let all conditions of theorem 2.11 be satisfied. Let in addition*

$$(13) \qquad \sum_{m=1}^{\infty} \left| \int_A p^{(md+r(A,B))}(x, B)\, \lambda(dx) - \int_A p^{((m+1)d+r(A,B))}(x, B)\, \lambda(dx) \right| < \infty$$

and analogously for A, B interchanged. Then the continuous part of the function $v(A, B; .)$ is absolutely continuous.

Proof. We shall use the same notation q_m as in the proof of theorem 2.11. Condition (13) is the well-known condition of bounded variation of the sequence q_m (see e. g. Zygmund [18], p. 4).

Let a closed interval $[t_1, t_2] \subset [0, 2\pi]$ contain no point $2\pi j/d$, $j = 0, 1, \dots, d - 1$. We shall prove that for all $t \in [t_1, t_2]$ and for all n the Cesàro sums of the trigonometrical series corresponding to $v(A, B; .)$ are uniformly bounded. These Cesàro sums are equal to (omitting a constant factor)

$$(14) \qquad s_n(t) = \sum_{k=0}^{n} \left(1 - \frac{k}{n+1}\right) e^{ikt} \int_A p^{(k)}(x, B)\, \lambda(dx) + \sum_{k=1}^{n} \left(1 - \frac{k}{n+1}\right) e^{-ikt} \int_B p^{(k)}(x, A)\, \lambda(dx)\,.$$

As a consequence of symmetry we can restrict ourselves only to the first sum in (14). We divide all natural numbers into d groups (with respect to congruence): the j-th group ($j = 0, 1, \dots, d - 1$) contains the numbers n of the form $n = Nd + r(A, B) + j$. Clearly it is sufficient to prove the uniform boundedness of $s_n(t)$ for each of these groups separately, for we have a finite number d of them. But in the j-th group the first sum in (14) is by the assumptions of the theorem equal to

$$(15) \qquad \sum_{k=0}^{Nd+r(A,B)+j} \left(1 - \frac{k}{Nd + r(A, B) + j + 1}\right) e^{ikt} \int_A p^{(k)}(x, B)\, \lambda(dx) = \sum_{m=0}^{N} \left(1 - \frac{md + r(A, B)}{Nd + r(A, B) + j + 1}\right) e^{it(md + r(A,B))}\, q_m\,.$$

As in the proof of theorem 2.11 we shall again make use of Abel's partial summation, but now $S_m = \sum_{k=0}^{m} e^{it(kd + r(A,B))}$ and again as in (11) we obtain

$$|S_m| = \left| \sum_{k=0}^{m} e^{it(kd + r(A,B))} \right| \leq \frac{2}{|e^{itd} - 1|}$$

where the last expression is bounded by a finite constant c uniformly for all $t \in [t_1, t_2]$ and for all m. Analogously as in (9) and (12) we obtain

$$\left|\sum_{m=0}^{N}\left(1-\frac{md+r(A,B)}{Nd+r(A,B)+j+1}\right)e^{it(md+r(A,B))}\,q_m\right| \leqq$$
$$\leqq \sum_{m=0}^{N-1}|S_m|\cdot\left|\left(1-\frac{md+r(A,B)}{Nd+r(A,B)+j+1}\right)q_m-\right.$$
$$\left.-\left(1-\frac{(m+1)\,d+r(A,B)}{Nd+r(A,B)+j+1}\right)q_{m+1}\right| + |S_N|\cdot\left|\left(1-\frac{Nd+r(A,B)}{Nd+r(A,B)+j+1}\right)q_N\right| \leqq$$
$$\leqq c\,.\sum_{m=0}^{N-1}\left|q_m-q_{m+1}-\frac{md+r(A,B)}{Nd+r(A,B)+j+1}\,q_m+\frac{md+r(A,B)}{Nd+r(A,B)+j+1}\,q_{m+1}+\right.$$
$$\left.+\frac{d}{Nd+r(A,B)+j+1}\,q_{m+1}\right| + c\,.\,|q_N| \leqq c\,.\left[\sum_{m=0}^{N-1}|q_m-q_{m+1}|+\right.$$
$$+\sum_{m=0}^{N-1}\frac{md+r(A,B)}{Nd+r(A,B)+j+1}\,|q_m-q_{m+1}|+\frac{d}{Nd+r(A,B)+j+1}\sum_{m=0}^{N-1}|q_{m+1}|+$$
$$\left.+\,|q_N|\right] \leqq c\,.\left[\sum_{m=0}^{\infty}|q_m-q_{m+1}|+\sum_{m=0}^{N-1}|q_m-q_{m+1}|+\frac{d}{Nd}\sum_{m=0}^{N-1}|q_{m+1}|+|q_N|\right].$$

In the last expression the first two terms are bounded by assumption (13), and the last two terms are bounded since q_m and therefore $N^{-1}\sum_{m=0}^{N-1} q_{m+1}$ also converges.

Summing up we have proved that the Cesàro sums $s_n(t)$ in (14) are bounded by a finite constant uniformly in $[t_1, t_2]$. Since that constant is an integrable dominating function in $[t_1, t_2]$ we have

$$\lim_n \int_{t_1}^{t_2} |s_n(t)|\,\mathrm{d}t = \int_{t_1}^{t_2} |\lim_n s_n(t)|\,\mathrm{d}t$$

and by the theorem IV.4.20 in Zygmund [18] $v(A, B;\,.)$ is absolutely continuous in $[t_1, t_2]$, i.e. clearly everywhere in $[0, 2\pi]$ except the points $2\pi j/d$, $j = 0, 1, \ldots, \ldots, d-1$. Theorem 2.12 is proved.

Finishing part 2 let us give a recapitulation. A Markov system always has a finitely additive invariant measure by theorem 2.2 so that the representations of the theorem 2.9 always can be used. As for σ-additive sub-invariant measures, we have such a situation: If the corresponding chain is recurrent according to the definition of T. E. Harris [6] then there is a σ-finite σ-additive invariant measure (see Harris [6]) and representations of theorem 2.7 can be used. If this invariant measure is finite, then by the results by S. Orey [14] on convergence of transition probabilities we can use corollary 2.3, eventually also theorem 2.12. On the other hand if an irreducible chain is not recurrent according to Harris [6], it is not recurrent according to Nelson [13] as well (see the text after definition 2.1 and the proof of theorem 2.1). Then by the theorem 4.1 in [13] there are non-void sets $E \in \Sigma$ such that $\sum_{n=1}^{\infty} p^{(n)}(x, E) < \infty$. For such sets the following representation theorem is easily proved.

Theorem 2.13. *Let $p^{(n)}$ be a Markov system and let $x \in X$, $E \in \Sigma$ be such that $\sum_{n=1}^{\infty} p^{(n)}(x, E) < \infty$. Then there is a function $g(x, E; .)$ square integrable on $[0, 2\pi]$ such that*

$$p^{(n)}(x, E) = \int_0^{2\pi} e^{int} g(x, E; t)\,\mathrm{d}t \quad \textit{for} \quad n = 1, 2, \dots .$$

Proof. Since $\sum_{n=1}^{\infty} |p^{(n)}(x, E)|^2 \leqq \sum_{n=1}^{\infty} p^{(n)}(x, E) < \infty$, the well-known Riesz-Fischer theorem (see e.g. Zygmund [18] IV.1.1) gives the required formula immediately.

3. REPRESENTATIONS FOR SELF-ADJOINT MARKOV SYSTEMS

3.1. Conditions for self-adjointness. Let us begin with a definition.

Definition 3.1. A Markov system $p^{(n)}$ is called *μ-self-adjoint*, if it has a sub-invariant σ-additive measure μ and if for arbitrary $A, B \in \Sigma$ the equation

$$(1) \qquad \int_B p(x, A)\,\mu(\mathrm{d}x) = \int_A p(x, B)\,\mu(\mathrm{d}x)$$

holds. $p^{(n)}$ is called *self-adjoint*, if for some μ it is μ-self-adjoint.

Let us observe at once: If a Markov system $p^{(n)}$ is μ-self-adjoint, then μ is an invariant measure. This fact is clearly a consequence of (1) putting there X in place of B.

Some authors use the term "reversible chain" because self-adjointness in some special cases is equivalent to "reversibility in time" understood in a certain intuitive way; this situation appears particularly for irreducible chains which is seen from Nelson's results [13], in theorem 3.2, or for chains with a denumerable state space (see Kendall [7], [8]). In this paper we prefer the term "self-adjointness", for the intuitive content of the term "reversibility" disappears generally.

Self-adjoint or reversible chains were studied by A. N. Kolmogorov [10] as early as in 1936 (for a finite state space and only reversibility without relation to self-adjointness), D. G. Kendall [7], [8], [9] (for a denumerable state space), H. P. Kramer [11] (for processes in continuous time, but the results are quite analogous); further some results follow immediately from Nelson's paper [13].

Theorem 3.1. *A Markov system $p^{(n)}$ is μ-self-adjoint if and only if the mapping T defined by formula* (3) *in section* 1.2 *for $f \in \mathbf{L}_2(\mu)$ is a self-adjoint operator in $\mathbf{L}_2(\mu)$. If this is the case, then we have more generally*

$$(2) \qquad \int_B p^{(n)}(x, A)\,\mu(\mathrm{d}x) = \int_A p^{(n)}(x, B)\,\mu(\mathrm{d}x) \quad \text{for} \quad n = 1, 2, \dots .$$

Proof. First, let T be a self-adjoint operator in $\mathbf{L}_2(\mu)$, which means the equality of scalar products

$$(3) \qquad (Tf, g) = (f, Tg)$$

for arbitrary $f, g \in \mathbf{L}_2(\mu)$. If $A, B \in \Sigma$ are such that $\mu(A) < \infty$, $\mu(B) < \infty$, then (3) gives

$$(4) \qquad (T\chi_A, \chi_B) = (\chi_A, T\chi_B)$$

which is (1). If $A, B \in \Sigma$ are arbitrary, they can be written as disjoint unions $A = \bigcup_{j=1}^{\infty} A_j$, $B = \bigcup_{j=1}^{\infty} B_j$ where $A_j, B_j \in \Sigma$, $\mu(A_j) < \infty$, $\mu(B_j) < \infty$ for $j = 1, 2, \ldots$. Now (1) holds for A_j, B_j and interchanging the summation and integration, which is permitted for the non-negative function p (e.g. by theorem 2 § 27 in P. Halmos [5]), we obtain (1) for arbitrary A, B.

If, on the other hand, (1) holds, then by a common method of the theory of integrals we get (3).

If T is a self-adjoint operator, then also all powers T^n, $n = 1, 2, \ldots$, are self-adjoint, and thus (2) is proved.

In the following we shall use the Cartesian product of measurable spaces $(X \times X, \Sigma \times \Sigma, \nu \times \nu)$ defined in the usual way ν being some σ-additive measure, and in the usual corresponding way we shall understand the propositions concerning measurability, almost everywhere properties etc. Analogously we shall have k-fold Cartesian products. In the rest of the section 3.1 we suppose for simplicity that Σ is generated by a denumerable class of sets (although this supposition can be weakened at the costs of a certain complication of theorems). Let us introduce the following convention: if the Radon-Nikodym derivative $\mathrm{d}p^{(n)}(x, .)/\mathrm{d}\nu(.) = h^{(n)}(x, .)$ exists for ν-almost all x, in the rest of the section we shall always take for $h^{(n)}(., .)$ a non-negative $(\Sigma \times \Sigma)$-measurable function; this is possible by the assumed structure of Σ following Doob [1], p. 753, or [2], p. 616, similarly as in Nelson [13], lemma 3.3. We denote $h^{(1)}(., .) = h(., .)$.

Theorem 3.2. *Let $p^{(n)}$ be a Markov system. Let μ be a σ-additive measure on Σ, for which Radon-Nikodym derivatives $\mathrm{d}p(x, .)/\mathrm{d}\mu(.) = h(x, .)$ exist for μ-almost all x. Then $p^{(n)}$ is μ-self-adjoint if and only if*

$$(5) \qquad h(x, y) = h(y, x) \quad \textit{for } (\mu \times \mu)\textit{-almost all } (x, y)\,.$$

If this is the case, and if we define recurrently $h^{(n)}(x, y) = \int_X h^{(n-1)}(x, z)\, h(z, y)\, \mu(\mathrm{d}z)$, $h^{(1)} = h$, then $h^{(n)}(x, .) = \mathrm{d}p^{(n)}(x, .)/\mathrm{d}\mu(.)$ almost everywhere and

$$(6) \qquad h^{(n)}(x, y) = h^{(n)}(y, x) \quad \textit{for} \quad (\mu \times \mu)\textit{-almost all } (x, y)$$

and for all $n = 1, 2, \ldots$.

Remark. This theorem generalizes a little the characterization of self-adjoint chains, following easily in the irreducible case from Nelson's theorem 3.2 in [13].

Proof of the theorem 3.2. In the following we shall always have in mind that $h(x, y) \geqq 0$ so that the order of integration can be arbitrarily interchanged. If (5) is true, then for arbitrary sets $A, B \in \Sigma$ we have

$$\int_B p(x, A)\, \mu(\mathrm{d}x) = \int_B \left[\int_A h(x, y)\, \mu(\mathrm{d}y)\right] \mu(\mathrm{d}x) = \int_{B\times A} h(y, x)\, (\mu \times \mu)\, (\mathrm{d}(x, y)) =$$
$$= \int_A \left[\int_B h(y, x)\, \mu(\mathrm{d}x)\right] \mu(\mathrm{d}y) = \int_A p(y, B)\, \mu(\mathrm{d}y)$$

therefore (1) holds. Conversely, if (1) holds, then evidently by a simple change of notation we get

$$\int_B \left[\int_A h(x, y)\, \mu(\mathrm{d}y)\right] \mu(\mathrm{d}x) = \int_A \left[\int_B h(y, x)\, \mu(\mathrm{d}x)\right] \mu(\mathrm{d}y) =$$
$$= \int_B \left[\int_A h(y, x)\, \mu(\mathrm{d}y)\right] \mu(\mathrm{d}x)$$

or

$$(7) \qquad \int_{B\times A} h(x, y)\, (\mu \times \mu)\, (\mathrm{d}(x, y)) = \int_{B\times A} h(y, x)\, (\mu \times \mu)\, (\mathrm{d}(x, y))$$

for arbitrary $A, B \in \Sigma$. Each side of (7) represents a set function on the semi-ring of sets of the type $B \times A$, therefore it can be uniquely extended to $\Sigma \times \Sigma$ and equation (7) is preserved even when writing an arbitrary set from $\Sigma \times \Sigma$ in place of $A \times B$. But from this we see that (5) holds.

If the system is μ-self-adjoint then the first assertion about $h^{(n)}$ is easily proved by induction, and the last assertion is proved by the same method as (5) using the equations (2) of theorem 3.1.

The content of following theorems may become more comprehensive if the points of X are now called the *states* of the Markov system $p^{(n)}$.

Theorem 3.3. *For a Markov system* $p^{(n)}$ *to be self-adjoint, the set of following conditions is sufficient:*

For some σ-additive measure ν on Σ Radon-Nikodym derivatives $\mathrm{d}p^{(n)}(x, .)/\mathrm{d}\nu(.) = h^{(n)}(x, .)$ *exist for ν-almost all x. There exists a state $x_0 \in X$ having the following properties: for each state $x \in X$ there is a natural n such that*

$$(8) \qquad h^{(n)}(x_0, x) > 0\,, \quad h^{(n)}(x, x_0) > 0$$

and

$$(\mathrm{K}_1) \qquad h^{(k)}(x_0, x_1)\, h(x_1, x_2)\, h^{(l)}(x_2, x_0) = h^{(l)}(x_0, x_2)\, h(x_2, x_1)\, h^{(k)}(x_1, x_0)$$

for all natural k, l and $(\nu \times \nu)$-almost all (x_1, x_2).

Conversely: if the system is ν-self-adjoint and if $\mathrm{d}p^{(n)}(x, .)/\mathrm{d}\nu(.) = h^{(n)}(x, .)$ *exist for ν-almost all x, then for ν-almost all $x_0 \in X$ the condition* (K_1) *holds.*

Remark. Condition (8) means a certain kind of intercommunication between the state x_0 and an arbitrary state x (see e.g. T. A. SARYMSAKOV [15]), i.e. a condition similar to irreducibility (of course, here $h^{(n)}$ need not be continuous). If for each $x \in X$ the least natural n satisfying (8) is denoted by $n(x)$, then from the proof it is seen that it suffices to verify condition (K_1) only for $k = n(x_1)$, $l = n(x_2)$. Practically e.g. if $n(x)$ have an upper bound N, then it suffices to verify (K_1) for $k, l = 1, 2, \ldots, N$.

Proof of theorem 3.3. Let us denote $E_n = \{x; n(x) = n\}$; then $\bigcup_{n=1}^{\infty} E_n = X$. By our permanent assumption $h^{(n)}(.,.)$ are $(\Sigma \times \Sigma)$-measurable, so $h^{(n)}(x_0, .)$, $h^{(n)}(., x_0)$ are Σ-measurable. Therefore $F_n = \{x; h^{(n)}(x_0, x) > 0\} \cap \{x; h^{(n)}(x, x_0) > 0\} \in \Sigma$ for each $n = 1, 2, \ldots$. Clearly $E_n = F_n - \bigcup_{j=1}^{n-1} F_j$ which gives $E_n \in \Sigma$ for each $n = 1, 2, \ldots$.

Let us define a function g by the prescription that for $x \in E_n$ it takes the value $g(x) = h^{(n)}(x_0, x)/h^{(n)}(x, x_0)$. Since the definition of g is composed from denumerably many subsets $E_n \in \Sigma$ and on each E_n the function g is measurable, evidently g is measurable on X; furthermore g is clearly finite.

Let us define a measure μ on Σ by the formula $\mu(E) = \int_E g(x)\, \nu(\mathrm{d}x)$, $E \in \Sigma$. The measure ν is σ-finite, which means $X = \bigcup_{j=1}^{\infty} D_j$, $D_j \in \Sigma$, $\nu(D_j) < \infty$; we denote $D_{j,k} = D_j \cap \{x; g(x) \leqq k\}$ for k natural. There are denumerably many sets $D_{j,k}$, and we have $X = \bigcup_{j,k=1}^{\infty} D_{j,k}$ and $\mu(D_{j,k}) = \int_{D_{j,k}} g(x)\, \nu(\mathrm{d}x) \leqq k \,.\, \nu(D_{j,k}) \leqq k \,.\, \nu(D_j) < \infty$, so that the measure μ is σ-finite.

Let us have two sets from Σ, $A_k \subset E_k$, $B_l \subset E_l$. Then by condition (K_1) and by Fubini's theorem for non-negative functions we obtain

$$\int_{B_l} p(x_2, A_k)\, \mu(\mathrm{d}x_2) = \int_{B_l} p(x_2, A_k)\, g(x_2)\, \nu(\mathrm{d}x_2) = \int_{B_l} p(x_2, A_k) \frac{h^{(l)}(x_0, x_2)}{h^{(l)}(x_2, x_0)}\, \nu(\mathrm{d}x_2) =$$

$$= \int_{B_l} \left[\int_{A_k} h(x_2, x_1)\, \nu(\mathrm{d}x_1) \right] \frac{h^{(l)}(x_0, x_2)}{h^{(l)}(x_2, x_0)}\, \nu(\mathrm{d}x_2) =$$

$$= \int_{A_k} \left[\int_{B_l} \frac{h^{(k)}(x_0, x_1)}{h^{(k)}(x_1, x_0)} h(x_1, x_2)\, \nu(\mathrm{d}x_2) \right] \nu(\mathrm{d}x_1) = \int_{A_k} p(x_1, B_l) \frac{h^{(k)}(x_0, x_1)}{h^{(k)}(x_1, x_0)}\, \nu(\mathrm{d}x_1) =$$

$$= \int_{A_k} p(x_1, B_l)\, \mu(\mathrm{d}x_1)\,.$$

Thus we see that (1) is true for these particular $A = A_k$, $B = B_l$ and with the measure μ. Finally let $A, B \in \Sigma$ be arbitrary. We know already that (1) holds for $A_k = A \cap E_k$, $B_l = B \cap E_l$ and in general (1) is obtained by summation over all k, l. Therefore the system $p^{(n)}$ is μ-self-adjoint.

Conversely, let $p^{(n)}$ be ν-self-adjoint. Then by theorem 3.2 we have

$$h^{(k)}(x_0, x_1) = h^{(k)}(x_1, x_0) \tag{9}$$

for $(\nu \times \nu)$-almost all (x_0, x_1), $h(x_1, x_2) = h(x_2, x_1)$ for $(\nu \times \nu)$-almost all (x_1, x_2), $h^{(l)}(x_2, x_0) = h^{(l)}(x_0, x_2)$ for $(\nu \times \nu)$-almost all (x_0, x_2). By Fubini's theorem on null sets (9) is satisfied for ν-almost all x_0 except some x_1 from a ν-null set N_{x_0}. If we consider $h^{(k)}(x_0, x_1)$ as a function of the triple (x_0, x_1, x_2) (which remains constant for varying x_2) we see also that (9) is satisfied for ν-almost all x_0 except some (x_1, x_2) from a $(\nu \times \nu)$-null set $N_{x_0} \times X$. An analogous conclusion can be drawn concerning $h^{(l)}(x_0, x_2)$. Combining the above results and intersecting the sets on which all equalities for the functions h hold, we obtain the last assertion of theorem 3.3 concerning the truth of condition (K_1).

In the following we shall use this notation: If Z, X are two spaces and M some subset of $Z \times X$, then its projection into the space X (i.e. the set of all $x \in X$ for which there is a $z \in Z$ such that $(z, x) \in M$) will be denoted by M^X. Analogously M^Z denotes the projection of M into the space Z.

Lemma 3.1. *Let $(Z \times X, \Sigma_Z \times \Sigma_X)$ be the Cartesian product of two topological measurable spaces (Z, Σ_Z) and (X, Σ_X), let Σ_X be the σ-algebra generated by open sets and let Z be separable. If G is an open set in $Z \times X$, then on its projection G^X there exists a mapping $z(.)$ of G^X into Z with the following properties:*

(a) $(z(x), x) \in G$ for each $x \in G^X$,
(b) $z(.)$ is measurable.

Proof. Let $z_1, z_2, \ldots$ be a denumerable dense subset of Z. We denote by R_j, $j = 1, 2, \ldots$, the sections of the set G determined by the point z_j, i.e. R_j is the set of all $x \in X$ such that $(z_j, x) \in G$. Clearly R_j are open, and we shall show

$$G^X = \bigcup_{j=1}^{\infty} R_j\,. \tag{10}$$

To show it, let $x \in G^X$. Then there is a $z \in Z$ such that $(z, x) \in G$. But then G contains a "rectangular" neighbourhood N of this point, the projection N^Z into the space Z is open, and therefore there exists $z_j \in N^Z$. Clearly $(z_j, x) \in N \subset G$, $x \in R_j$, which gives $G^X \subset \bigcup_{j=1}^{\infty} R_j$. The converse inclusion is obvious, therefore really (10) is true. It also follows from this that G^X is open.

Now, for $x \in R_1$ we define $z(x) = z_1$, for $x \in R_j - \bigcup_{k=1}^{j-1} R_k$, $j = 2, 3, \ldots$, we define $z(x) = z_j$. Clearly by such a prescription the mapping $z(.)$ is defined on the whole G^X and it satisfies (a). If E is a set from Σ_Z, then $z(x) \in E$ is true for at most denumerable union of subsets of X having the form $R_j - \bigcup_{k=1}^{j-1} R_k$, i.e. for a set from Σ_X, since R_j being open belong to Σ_X. Thus (b) is proved.

Theorem 3.4. *For a Markov system $p^{(n)}$ to be self-adjoint, the set of following conditions is sufficient:*

X is a topological separable space, Σ_X is the σ-algebra generated by open sets of X. For some σ-additive measure ν on Σ_X Radon-Nikodym derivatives $\mathrm{d}p(x, .)/\mathrm{d}\nu(.) = h(x, .)$ exist for ν-almost all x, and $h(., .)$ is a continuous function of (x, y). There exists a state $x_0 \in X$ having the following properties: for each state $x \in X$ there are finite sequences of states $x'_1, x'_2, \ldots, x'_k$ and $x''_1, x''_2, \ldots, x''_l$ such that

$$(11) \quad h(x_0, x'_1)\, h(x'_1, x'_2) \ldots h(x'_k, x) > 0\,, \quad h(x, x''_l)\, h(x''_l, x''_{l-1}) \ldots h(x''_1, x_0) > 0$$

and

$$(\mathrm{K}_2) \qquad h(x_0, y_1)\, h(y_1, y_2) \ldots h(y_{m-1}, y_m)\, h(y_m, x_0) = \\ = h(x_0, y_m)\, h(y_m, y_{m-1}) \ldots h(y_2, y_1)\, h(y_1, x_0)$$

for all natural m and all $y_1, y_2, \ldots, y_m \in X$.

Conversely: if the system is ν-self-adjoint and if $\mathrm{d}p(x, .)/\mathrm{d}\nu(.) = h(x, .)$ exist for ν-almost all x, then (K_2) holds for $(\nu \times \nu \times \ldots \times \nu)$-almost all $(x_0, y_1, y_2, \ldots, y_m)$.

Proof. For brevity let us denote by $Z_k = X \times \ldots \times X$ the k-fold Cartesian product of the space X, k being a natural number; the elements of Z_k are $z_k = (x_1, \ldots, x_k)$. We see that Z_k is separable. By our assumption of continuity of h it follows immediately that the function $h_k(., .)$ defined by $h_k(z_k, x) = h(x_0, x_1)\, h(x_1, x_2) \ldots \ldots h(x_k, x)$ is continuous as a function of $(z_k, x) = (x_1, x_2, \ldots, x_k, x)$. Therefore, the set of all (z_k, x) for which $h_k(z_k, x) > 0$ is open, and we denote it by G_k. By the lemma 3.1 now there exists a measurable mapping $z_k(.) = (x_1(.), \ldots, x_k(.))$ defined on the projection G_k^X, for which

$$(12) \qquad h_k(z_k(x), x) > 0 \quad \text{for} \quad x \in G_k^X\,.$$

But by the assumption of the theorem for each $x \in X$ there exist $x''_1, \ldots, x''_l$ such that the second inequality of (11) holds. By (12), (11) and (K_2) we obtain

$$0 < h(x_0, x_1(x)) \ldots h(x_k(x), x)\, h(x, x''_l) \ldots h(x''_1, x_0) = \\ = h(x_0, x''_1) \ldots h(x''_l, x)\, h(x, x_k(x)) \ldots h(x_1(x), x_0)$$

for $x \in G_k^X$, which implies also $h(x, x_k(x)) \ldots h(x_1(x), x_0) > 0$. Thus for $x \in G_k^X$ it is possible to define a function g by

$$(13) \qquad g(x) = \frac{h(x_0, x_1(x)) \ldots h(x_k(x), x)}{h(x, x_k(x)) \ldots h(x_1(x), x_0)}\,.$$

It is easily seen, that the function taking the value

$$\frac{h(x_0, x_1) \ldots h(x_k, x)}{h(x, x_k) \ldots h(x_1, x_0)}$$

at the point $(x_1, \ldots, x_k, x)$ is measurable on $Z_k \times X$ (as soon as it is defined). Evidently this gives measurability of $g(.)$ on G_k^X.

By the assumptions for each $x \in X$ there is a natural k and a point $z_k' = (x_1', \dots, x_k')$ such that $h_k(z_k', x) > 0$, that is $(z_k', x) \in G_k$, $x \in G_k^X$. Thus we have shown $X = \bigcup_{k=1}^{\infty} G_k^X$. But X can be written in the form of a disjoint union $X = \bigcup_{k=1}^{\infty} E_k$, where $E_k = G_k^X - \bigcup_{j=1}^{k-1} G_j^X$. Here G_k^X are open as was observed in the proof of lemma 3.1 which implies $E_k \in \Sigma$ for each k.

If we define the function $g(.)$ by (13) on each E_k, then $g(.)$ is defined everywhere on X and clearly it is measurable. Now the proof may be completed as in theorem 3.3. We define a measure μ on Σ_X by the formula $\mu(E) = \int_E g(x)\, \nu(\mathrm{d}x)$, $E \in \Sigma_X$. Then μ is σ-additive and σ-finite. If $A_k \subset E_k$, $B_l \subset E_l$ are two sets from Σ_X, by condition (K_2) we obtain

$$\int_{B_l} p(y, A_k)\, \mu(\mathrm{d}y) = \int_{B_l} \left[\int_{A_k} h(y, x)\, \nu(\mathrm{d}x)\right] \frac{h(x_0, x_1(y)) \dots h(x_l(y), y)}{h(y, x_l(y)) \dots h(x_1(y), x_0)}\, \nu(\mathrm{d}y) =$$

$$= \int_{A_k} \left[\int_{B_l} \frac{h(x_0, x_1(x)) \dots h(x_k(x), x)}{h(x, x_k(x)) \dots h(x_1(x), x_0)}\, h(x, y)\, \nu(\mathrm{d}y)\right] \nu(\mathrm{d}x) =$$

$$= \int_{A_k} g(x)\, p(x, B_l)\, \nu(\mathrm{d}x) = \int_{A_k} p(x, B_l)\, \mu(\mathrm{d}x)\,.$$

If $A, B \in \Sigma_X$ are arbitrary, then we set $A_k = A \cap E_k$, $B_l = B \cap E_l$, and summing the equations obtained for A_k, B_l we have (1); therefore the system $p^{(n)}$ is μ-self-adjoint.

The converse assertion of theorem 3.4 is obtained in a manner quite similar to the proof given in theorem 3.3.

Remark. Let us observe that condition (11) from an intuitive point of view means a certain kind of intercommunication between x_0 and an arbitrary state x. Our conditions (K_1) and (K_2), particularly the second of them, appear to be an analogy of Kolmogorov's condition in [10], or of condition (K) introduced by Kendall [7], [8].

Theorem 3.4 does not seem to be too satisfactory, because it has rather strong assumptions: separability of X and continuity of h. These assumptions are needed for the proof of lemma 3.1 which gives the key assertion. Though it seems that there are some possibilities of generalizing lemma 3.1, we have not yet succeed in doing so.

Nevertheless, for many practical cases theorem 3.4 may be quite sufficient; for example, in Sarymsakov's book [15] it is always required that the state space X is a compact set on the real line and is therefore separable, and that $h(.,.)$ is continuous. It is also seen, however, that theorem 3.4 in its present form generalizes a result by D. G. Kendall [8] for $X = \{1, 2, 3, \dots\}$. Here X is denumerable itself, the measure ν is defined $\nu\{k\} = 1$ for $k = 1, 2, 3, \dots$, and then $h(x, y) = p_{xy}$ is continuous and clearly our characterization contains that given by Kendall.

Let us still pay some attention to condition (K_2), understanding that our remark is valid also for Kendall's case of a denumerable state space. First of all it is sufficient to require (K_2) only for a single fixed point x_0 as a starting point (in contradistinction to Kendall's condition). Further from the proof the following fact is evident: denoting for each $x \in X$ by $k(x)$, $l(x)$ the least lengths of sequences $x'_1, \ldots, x'_k, x''_1, \ldots, x''_l$ satisfying (11) it is sufficient to verify condition (K_2) only for products

$$h(x_0, y'_1)\, h(y'_1, y'_2) \ldots h(y'_{k(x)}, x)\, h(x, y''_{l(x)}) \ldots h(y''_2, y''_1)\, h(y''_1, x_0)\,,$$
$$h(x_0, y'_1)\, h(y'_1, y'_2) \ldots h(y'_{k(x)}, x)\, h(x, y)\, h(y, y'''_{l(y)}) \ldots h(y'''_2, y'''_1)\, h(y'''_1, x_0)$$

for all $y'_1, \ldots, y'_{k(x)}$; $y''_1, \ldots, y''_{l(x)}$; $y'''_1, \ldots, y'''_{l(y)}$ (i.e. for products of respective least lengths), for which these products are positive. Practically e.g. if $k(x)$, $l(x)$ have an upper bound N (in the extreme case may be even $N = 0$), then it suffices to verify (K_2) for products of lengths $1, 2, \ldots, 2N + 3$.

Concluding the section let us turn for a while to the case of a finitely additive invariant measure $\lambda \in \boldsymbol{ba}(X, \Sigma)$.

An initial idea for a general Markov system $p^{(n)}$ was that the mapping T generated by it is an operator in $\boldsymbol{B}_p(\lambda)$ (see theorem 2.8); quite a similar idea can be used also in the present case of self-adjoint systems. Since $\boldsymbol{B}_2(\lambda)$ is not a Hilbert space but only a dense subset of the Hilbert space $\boldsymbol{V}^2$ (which is seen from Leader's results [12]), we shall speak of a symmetric operator instead of a self-adjoint one. By a symmetric operator T we understand an operator whose domain $\boldsymbol{D}$ is dense in a Hilbert space and for which $(Tf, g) = (f, Tg)$ for all $f, g \in \boldsymbol{D}$.

Theorem 3.5. *Let $p^{(n)}$ be a Markov system which has an invariant measure $\lambda \in \boldsymbol{ba}(X, \Sigma)$. The mapping T generated by the system $p^{(n)}$ is a symmetric operator in $\boldsymbol{B}_2(\lambda)$ if and only if*

$$\int_B p(x, A)\, \lambda(\mathrm{d}x) = \int_A p(x, B)\, \lambda(\mathrm{d}x)$$

holds for arbitrary $A, B \in \Sigma$.

Proof is similar to that given for theorem 3.1, using the fact that each function from $\boldsymbol{B}_2(\lambda)$ can be uniformly approximated by simple functions from $\boldsymbol{B}_2(\lambda)$.

3.2. Representations for transition probabilities. In a quite similar way as in section 2.2 the representations for self-adjoint systems can now be obtained.

Theorem 3.6. *Let $p^{(n)}$ be a μ-self-adjoint Markov system. Then for each set $A \in \Sigma$ for which $\mu(A) < \infty$ there is a mapping $\boldsymbol{v}(A; .)$ of the interval $[-1, 1]$ into real $\boldsymbol{L}_2(\mu)$ such that*

$$p^{(n)}(., A) = \int_{-1}^{1} t^n \boldsymbol{v}(A; \mathrm{d}t) \quad \textit{for} \quad n = 0, 1, 2, \ldots\,, \tag{14}$$

where equalities of functions in $\boldsymbol{L}_2(\mu)$ are meant, i.e. μ-almost everywhere.

Proof. It is essentially similar to the proof of theorem 2.6, but somewhat simpler, because the spectral representation of the self-adjoint operator T in the Hilbert space $\mathbf{L}_2(\mu)$ is used directly. The limits of the integral are obtained from $\|T\| \leqq 1$.

Theorem 3.7. *Let $p^{(n)}$ be a μ-self-adjoint Markov system. Then for each two sets $A, B \in \Sigma$, for which $\mu(A) < \infty$, $\mu(B) < \infty$, there is a real function $v(A, B; .)$ of bounded variation in $[-1, 1]$ such that*

$$(15) \quad \int_B p^{(n)}(x, A)\, \mu(\mathrm{d}x) = \int_A p^{(n)}(x, B)\, \mu(\mathrm{d}x) = \int_{-1}^{1} t^n v(A, B; \mathrm{d}t) \quad \textit{for} \quad n = 0, 1, 2, \ldots .$$

The proof is quite similar to that of theorem 3.6 or 2.7.

Theorem 3.8. *Let $p^{(n)}$ be a Markov system which has an invariant measure $\lambda \in \mathbf{ba}(X, \Sigma)$ and let the mapping T generated by $p^{(n)}$ be a symmetric operator in $\mathbf{B}_2(\lambda)$. Then for each two sets $A, B \in \Sigma$ there is a real function $v(A, B; .)$ of bounded variation in $[-1, 1]$ such that*

$$(16) \quad \int_B p^{(n)}(x, A)\, \lambda(\mathrm{d}x) = \int_A p^{(n)}(x, B)\, \lambda(\mathrm{d}x) = \int_{-1}^{1} t^n\, v(A, B; \mathrm{d}t) \quad \textit{for} \quad n = 0, 1, 2, \ldots .$$

Proof. A bounded symmetric operator T in $\mathbf{B}_2(\lambda)$ can be continuously extended to a self-adjoint operator in the Hilbert space $\mathbf{V}^2$. Similarly, as in the text after the theorem 2.9, we define $v^{(n)}(A; B) = \int_B p^{(n)}(x, A)\, \lambda(\mathrm{d}x)$ for $B \in \Sigma$ and we get a mapping $\mathbf{v}(A; .)$ of the interval $[-1, 1]$ into $\mathbf{V}^2$ such that

$$v^{(n)}(A; .) = \int_{-1}^{1} t^n \mathbf{v}(A; \mathrm{d}t) \quad \text{for} \quad n = 0, 1, 2, \ldots ,$$

where the equalities are understood as equalities of set functions in $\mathbf{V}^2$. Putting set B here and using for the integral on the right side its definition and Leader's theorem 17 in [12], we have the equality of both extreme terms in (16). The first equality in (16) is evident by $(T^n \chi_A, \chi_B) = (\chi_A, T^n \chi_B)$.

4. CONCLUSION

Concluding the paper let us point out that all integral representations (or in different terminology, moment representations) for the transition probabilities discussed in this paper can be immediately generalized in the following way:

Let f be a measurable function on the space X. Let us denote by $\mathbf{E}_x^{(n)}\{f\}$ the expectation of f with respect to the distribution on X after the n-th step, provided the chain started at the point x, that is

$$\mathbf{E}_x^{(n)}\{f\} = \int_X f(x_1)\, p^{(n)}(x, \mathrm{d}x_1)\,, \quad n = 0, 1, 2, \ldots .$$

For these $\mathbf{E}_x^{(n)}\{f\}$ representations similar to those for $p^{(n)}(x, E)$ can be immediately obtained by quite similar proofs. For example the generalization of theorems 2.6, 2.7 is that for $f, g \in \mathbf{L}_2(\mu)$ the representations

$$\mathbf{E}_x^{(n)}\{f\} = \int_0^{2\pi} e^{int}\mathfrak{v}(f; \mathrm{d}t) \quad \text{for } \mu\text{-almost all } x\,,$$

$$\int_X \mathbf{E}_x^{(n)}\{f\} \cdot g(x)\, \mu(\mathrm{d}x) = \int_0^{2\pi} e^{int}v(f, g; \mathrm{d}t)$$

are obtained, and similarly for $\lambda \in \boldsymbol{ba}(X, \Sigma)$. For self-adjoint systems we have the representations

$$\mathbf{E}_x^{(n)}\{f\} = \int_{-1}^{1} t^n\mathfrak{v}(f; \mathrm{d}t) \quad \text{for } \mu\text{-almost all } x\,,$$

$$\int_X \mathbf{E}_x^{(n)}\{f\} \cdot g(x)\, \mu(\mathrm{d}x) = \int_{-1}^{1} t^n v(f, g; \mathrm{d}t)$$

etc.

However, since these generalizations are quite obvious and, moreover, seem to be uninteresting, we did not study them in the paper but contented ourselves with this brief concluding remark.

In the paper some unsolved problems were touched. At the end let us recapitulate the most important of them.

1. Has every Markov system a sub-invariant σ-additive σ-finite measure? Which conditions are sufficient for the existence of such a measure? (See section 2.1, theorem 2.1, 2.2 and the text before it, theorem 2.3.)

2. If $p^{(n)}$ is a Markov system having a sub-invariant $\lambda \in \boldsymbol{ba}(X, \Sigma)$, and if T is the mapping generated by $p^{(n)}$, is then T an operator in the whole space $\mathbf{L}_2(\lambda)$? (See theorem 2.8.)

3. Under which conditions would it be possible to find representations for $p^{(n)}(x, E)$ for all x, and not only for almost all x (as was done in theorem 2.10 or in the case of a denumerable state space)?

4. In the representation formulae (4) of theorem 2.7 is the continuous part of $v(A, B; .)$ always continuous, as it is in the case of a denumerable state space? (See section 2.3.)

5. Would it be possible to generalize lemma 3.1? That is, if we have a Cartesian product $(Z \times X, \Sigma_Z \times \Sigma_X)$ and a set $G \in \Sigma_Z \times \Sigma_X$, under which conditions is it possible to define a measurable mapping $z(.)$ of the projection G^X into Z such that $(z(x), x) \in G$ for all $x \in G^X$? From this a generalization of theorem 3.4 might follow.

References

[1] *J. L. Doob:* One-parameter families of transformations. Duke Math. J. *4* (1938), 752—774.
[2] *J. L. Doob:* Stochastic processes. New York 1953.

[3] *N. Dunford, J. T. Schwartz:* Linear operators I. Interscience Publishers, New York 1958.
[4] *W. Feller:* Non-Markovian processes with the semigroup property. Annals of Math. Stat. *30* (1959), 1252—1253.
[5] *P. Halmos:* Measure theory. New York 1950.
[6] *T. E. Harris:* The existence of stationary measures for certain Markov processes. Proc. Third Berkeley Symposium, 1956, vol. II, 113—124.
[7] *D. G. Kendall:* Integral representations for Markov transition probabilities. Bull. Amer. Math. Soc. *64* (1958), 358—362.
[8] *D. G. Kendall:* Unitary dilations of Markov transition operators, and the corresponding integral representations for transition-probability matrices. Probability & Statistics, The Harald Cramér Volume, New York 1959, 139—161.
[9] *D. G. Kendall:* Unitary dilations of one-parameter semigroups of Markov transition operators, and the corresponding integral representations for Markov processes with a countable infinity of states. Proc. London Math. Soc. (*3*) *9* (1959), 417—431.
[10] *A. N. Kolmogorov:* Zur Theorie der Markoffschen Ketten. Math. Ann. *112* (1936), 155—160.
[11] *H. P. Kramer:* Symmetrizable Markov matrices. Annals of Math. Stat. *30* (1959), 149—153.
[12] *S. Leader:* The theory of L^p-spaces for finitely additive set functions. Ann. of Math. *58* (1953), 528—543.
[13] *E. Nelson:* The adjoint Markoff process. Duke Math. J. *25* (1958), 671—690.
[14] *S. Orey:* Recurrent Markov chains. Pacific J. Math. *9* (1959), 805—827.
[15] *T. A. Сарымсаков:* Основы теории процессов Маркова, Москва 1954.
[16] *B. Sz.-Nagy:* Sur les contractions de l'espace de Hilbert. Acta Sci. Math. Szeged *15* (1953—54), 87—92.
[17] *K. Yosida, S. Kakutani:* Operator theoretical treatment of Markoff's process and mean ergodic theorem. Ann. of Math. *42* (1941), 188—228.
[18] *A. Zygmund:* Trigonometric series. Cambridge 1959.

Резюме

ИНТЕГРАЛЬНЫЕ ПРЕДСТАВЛЕНИЯ ДЛЯ ВЕРОЯТНОСТЕЙ ПЕРЕХОДА МАРКОВСКИХ ЦЕПЕЙ С ОБЩЕЙ СИСТЕМОЙ СОСТОЯНИЙ

ЗБЫНЕК ШИДАК (Zbyněk Šidák), Прага

В работе исследуются представления типа

$$(1)\qquad p^{(n)}(.,A) = \int_0^{2\pi} e^{int} v(\mathrm{d}t)\,, \quad \int_B p^{(n)}(x,A)\,\mu(\mathrm{d}x) = \int_0^{2\pi} e^{int} v(\mathrm{d}t)\,,$$

$$(2)\qquad p^{(n)}(.,A) = \int_{-1}^{1} t^n v(\mathrm{d}t)\,, \quad \int_B p^{(n)}(x,A)\,\mu(\mathrm{d}x) = \int_{-1}^{1} t^n v(\mathrm{d}t)$$

вероятностей перехода $p^{(n)}$ однородных марковских цепей с дискретным временем и с общей системой состояний. Результаты и методы возникли по существу обобщением работы Д. Дж. Кендалла [8]. Много используются методы функционального анализа. Главная идея состоит в том, что известный мар-

ковский оператор T, порождённый вероятностями перехода по формуле $Tf = \int_X f(x)\, p(.,\mathrm{d}x)$ (где f некоторая функция), определяется в пространстве $\boldsymbol{L}_2$ и потом используются спектральные представления унитарного или самосопряжённого операторов в гильбертовом пространстве. Для того, чтобы T был оператором в $\boldsymbol{L}_2$, надо иметь для марковской цепи так называемую субинвариантную меру μ (что значит, для которой $\int_X p(x, E)\, \mu(\mathrm{d}x) \leqq \mu(E)$ для каждого множества E из основной σ-алгебры Σ), по которой интегрируется. Во всей работе уделяется внимание конечно-аддитивным субинвариантным мерам.

Результаты этой работы касаются не только марковских цепей, а более общих так называемых марковских систем, что есть система функций $p^{(n)}(x, E)$, где $p^{(n)}(x, .)$ вероятность, $p^{(n)}(., E)$ измеримая функция, и $p^{(n)}$ удовлетворяют уравнениям Чэпмена-Колмогорова.

В части 1 вводятся обозначения и две известные основные теоремы.

Часть 2 занимается представлениями общих марковских систем. Прежде всего доказаны некоторые теоремы о субинвариантных и инвариантных мерах, например: при довольно общих предположениях каждая неприводимая марковская система имеет нетривиальную субинвариантную σ-аддитивную меру, которая однако не должна быть σ-конечной; каждая марковская система имеет инвариантную конечно-аддитивную конечную меру. Далее выводятся теоремы о представлениях типа (1): если μ субинвариантная σ-аддитивная мера, $\mu(A) < \infty$, то первое равенство в (1) верно для $n = 0, 1, 2, \ldots$ μ-почти всюду, где $\boldsymbol{v}$ некоторое отображение $[0, 2\pi]$ в комплексный $\boldsymbol{L}_2(\mu)$; второе равенство в (1) аналогично, но здесь v численная комплекснозначная функция с ограниченным изменением; это равенство доказывается тоже для μ конечно-аддитивной. Наконец исследуются свойства непрерывности функции v из (1): при некоторых условиях v может иметь скачки только в точках $2\pi j/d$, в остальных точках она непрерывна; эта теорема применяется к периодическим цепям; дается достаточное условие, чтобы непрерывная часть v была абсолютно непрерывной.

Часть 3 занимается представлениями самосопряжённых марковских систем, то есть для которых выполняется $\int_B p(x, A)\, \mu(\mathrm{d}x) = \int_A p(x, B)\, \mu(\mathrm{d}x)$ для каждой пары множеств A, B из σ-алгебры Σ. Эквивалентное условие есть, чтобы T был самосопряжённым оператором в $\boldsymbol{L}_2(\mu)$. Если существуют плотности h вероятностей перехода, то $h(x, y) = h(y, x)$ почти всюду. Следующие две теоремы указывают два класса достаточных условий для самосопряжённости, в которых основными являются требования (K_1) или (K_2), возникшие обобщением условия Колмогорова [10], [7], [8] для счётной системы состояний. Для самосопряжённых систем верны представления типа (2), где значение символов аналогично как в (1); опять второе равенство в (2) доказано тоже для μ конечно-аддитивной.

Заключительная часть 4 отмечает возможность дальнейшего обобщения и приводит перечень наиболее интересных нерешенных проблем.