On the semi-groups generated by Kolmogoroff's differential equations.

By Tosio KATO

§ 1. Introduction.

Let $P(t)=\big(p_{jk}(t)\big)$, $0 \leq t < \infty$, be the matrix of transition pro-babilities of a temporally homogeneous Markoff process with denumer-able states. $P(t)$ has the properties:

(1.1)
$$p_{jk}(t) \geq 0,$$

(1.2)
$$\sum_{k=1}^{\infty} p_{jk}(t) = 1,$$

(1.3)
$$p_{ik}(s+t) = \sum_{j=1}^{\infty} p_{ij}(s)\, p_{jk}(t).$$

It is also known[1] that, under certain general conditions, $p_{jk}(t)$ satisfy the following Kolmogoroff's differential equations:

(1.4)
$$dp_{ik}(t)/dt = \sum_{j=1}^{\infty} p_{ij}(t)\, a_{jk},$$

(1.4')
$$dp_{ik}(t)/dt = \sum_{j=1}^{\infty} a_{ij} p_{jk}(t),$$

together with the initial condition

(1.5)
$$\lim_{t \downarrow 0} p_{jk}(t) = p_{jk}(0) = \delta_{jk} = \begin{cases} 1, & j=k, \\ 0, & j \neq k. \end{cases}$$

Here a_{jk} are constants satisfying the conditions

---

1) For example Feller [6], Doob [4], Hille [7].

$$-a_{jj}=a_j\geq 0;\qquad a_{jk}\geq 0,\qquad j\neq k,$$

(1.6)

$$\sum_{k=1}^{\infty} a_{jk}=0,\qquad \text{that is,}\qquad \sum_{k\neq j} a_{jk}=a_j.$$

Conversely let a_{jk} be arbitrary constants subject only to the re-striction (1.6). The integration of these differential equations in this case was treated by Feller [5] some time ago.[2] In the present paper we shall consider the same problem by means of the semi-group theory developed by Hille [7, 8] and Yosida [9, 10]. It will be convenient, however, to state here some of our results without referring to the language of the semi-group theory.

I. There exists at least one solution $P(t)=\left(p_{jk}(t)\right)$ satisfying (1.1), (1.3), (1.4), (1.4') and (1.5). Among these solutions there is a "mini-mum" one in a sense to be specified later (Lemma 9), which is uniquely determined and satisfies

(1.2') $$\sum_{k=1}^{\infty} p_{jk}(t)\leq 1.$$

To express the condition that the equality sign holds in (1.2'), it is convenient to introduce the matrices $\left(b_{jk}^{(n)}(\lambda)\right)$ defined successively by

(1.7)

$$b_{jk}^{(1)}(\lambda)=b_{jk}(\lambda)=\begin{cases}0, & j=k,\\ a_{jk}/(\lambda+a_j), & j\neq k,\end{cases}\qquad \lambda>0,$$

$$b_{jk}^{(n+1)}(\lambda)=\sum_{i=1}^{\infty} b_{ji}^{(n)}(\lambda)\,b_{ik}(\lambda).$$

Each element $b_{jk}^{(n)}(\lambda)$ is non-negative and the series appearing in this definition are convergent. The limits

(1.8) $$\lim_{n\to\infty} \sum_{k=1}^{\infty} b_{jk}^{(n)}(\lambda)=b_j(\lambda)$$

exist, $b_j(\lambda)$ and $\lambda^{-1}[1-b_j(\lambda)]$ are non-increasing functions of λ.

II. With these notations, our minimum solution satisfies

(1.9) $$\lim_{t\to\infty} \sum_{k=1}^{\infty} p_{jk}(t)=1-\lim_{\lambda\downarrow 0} b_j(\lambda),$$

---

2) Cf. further Doob [2, 3, 4] and Feller [6].

(1.10) $$\int_0^\infty \left[\sum_{k=1}^\infty p_{jk}(t) \right] dt = \lim_{\lambda \downarrow 0} \lambda^{-1} [1 - b_j(\lambda)].$$

In the last equality the value ∞ is permitted. In particular, (1.2) holds for a particular j if and only if $b_j(\lambda) = 0$ for every $\lambda > 0$, or what is equivalent, for some $\lambda > 0$.

III. In order that the minimum solution may satisfy (1.2) for all j, it is necessary and sufficient that one of the following equivalent conditions is satisfied:

(i) For every $\lambda > 0$, $b_j(\lambda) = 0$ for all j.

(ii) For every $\lambda > 0$, there is no vector $(\zeta_1, \zeta_2, \cdots)$ with $0 < \sup |\zeta_j| < \infty$ satisfying the equations

(1.11) $$\sum_{k=1}^\infty a_{jk} \zeta_k = \lambda \zeta_j, \qquad j = 1, 2, \cdots.$$

In each of the statements (i), (ii), "every $\lambda > 0$" may be replaced by "some $\lambda > 0$". If these conditions are satisfied, the solution $(p_{jk}(t))$ is unique at least in a restricted sense[3] (Theorem 3).

In the following sections §§ 2, 3 we shall prove some theorems concerning the semi-groups generated by the differential equations (1.4), (1.4'), from which the above results are easily deduced. In the last section § 4, we shall treat some examples illustrating the use of these results.

It should be remarked that our theorems and their proofs are given in general terms of operator theory so that they are valid in any abstract (L) space,[4] provided that Lemmas 1 and 2 are taken as assumptions. The only exception is the last assertion $G \subset A$ of Theorem 1 which is proved by direct reference to concrete matrix representation. Thus our theorems may be applied to the integration of integro-differential equations connected with Markoff processes with more than denumerable states such as treated by Feller [5] by a different method.

---

3) The uniqueness of the solution is not proved in the usual sense, and it is not clear if our minimum solution is identical with that obtained by Feller [5], though this is quite plausible. Also it is not verified if the condition III is equivalent to the corresponding condition of Feller [5].

4) Birkhoff [1], p. 254.

§ 2. Main theorem.

Roughly speaking, our problem is equivalent to constructing a semi-group of operators generated by the operator with the matrix (a_{jk}). For this purpose we take $\mathfrak{X}=(l)$, the set of vectors $x=(\xi_1, \xi_2, \cdots)$ with the norm $\|x\|=\sum|\xi_k|$, as the basic Banach space[5] (cf. Hille [7], p. 426). In $\mathfrak{X}$ the matrix (a_{jk}) defines a linear operator A in a natural way. The domain $\mathfrak{D}[A]$ consists of all $x=(\xi_k)\in\mathfrak{X}$ such that

$$\eta_k=\sum_{j=1}^{\infty}\xi_j\,a_{jk}\,,\qquad k=1,2,\cdots,$$

are absolutely convergent and $\sum|\eta_k|<\infty$; Ax is then defined to be equal to $y=(\eta_k)$.

The operator A thus defined is not necessarily an infinitesimal generator of a semi-group. A complete solution to our problem from the viewpoint of semi-group theory would require the determination of all the contractions of A which generate semi-groups. However, we can give here only a restricted solution to this problem.

Let $\{x_j\}$ be the canonical base of $\mathfrak{X}$, that is, $x_j=(0,\cdots,0,\,1,\,0,\cdots)$, only the j-th component being equal to 1. Then it follows easily from (1.6) that $x_j\in\mathfrak{D}[A]$, $j=1,2,\cdots$. Let $\mathfrak{D}_0$ be the linear manifold spanned by $x_1,x_2,\cdots$, $\mathfrak{D}_0\in\mathfrak{D}[A]$. We define the operator A_0 as the contraction of A with $\mathfrak{D}[A_0]=\mathfrak{D}_0$.

We can now state our main theorem.

THEOREM 1. *There is at least one positive semi-group[6] of operators on $\mathfrak{X}$ of which the infinitesimal generator is an extension of A_0. Among these semi-groups there is a minimum one,[7] $\{P(t)\}$, $0\leq t<\infty$, which is uniquely determined. $\{P(t)\}$ is a contraction semi-group[8] with the generator G such that $A\supset G\supset A_0$.*

---

5) (l) may be either real or complex. For definiteness' sake we shall assume hereafter that it is real. It will be noted that the semi-group theory can be developed in a real Banach space as well as in a complex one, at least so far as concerns those results used in the present paper, although it appears to be customary in the literature to assume complex Banach spaces.

6) In the present paper we mean by a "semi-group" always a one-parameter semi-group of bounded linear operators $U(t)$ on $\mathfrak{X}$ which is defined and strongly continuous for $0\leq t<\infty$ with $U(0)=1$ (cf. Hille [8], Yosida [9]). It is positive if $U(t)\geq 0$, $0\leq t<\infty$.

7) In the sense of order relation. See Lemma 9.

8) $\{U(t)\}$ is a contraction semi-group if $\|U(t)\|\leq 1$, $0\leq t<\infty$.

Before giving the proof of this theorem, we shall show that it implies the result I of § 1. Let $p_{jk}(t)$ be the k-th component of the vector $P(t) x_j$. Then (1.1), (1.2'), (1.3) and (1.5) follow from the fact that $\{P(t)\}$ is a positive, contraction semi-group. (1.4) and (1.4') are direct consequences of the relations (cf. Yosida [9])

$$dP(t)x_i/dt = GP(t)x_i = P(t)Gx_i$$

where d/dt means strong derivative, for $x_i \in \mathfrak{D}_0 \subset \mathfrak{D}[G]$, $P(t)\mathfrak{D}[G] \subset \mathfrak{D}[G]$ and $A \supset G$. (Note that the matrices are to be multiplied in the reverse order to the corresponding operators!) It will be noted that the absolute convergence of the series on the right side of (1.4) is not directly clear, but it is a consequence of general theory.

PROOF OF THEOREM 1. We shall now give the proof of Theorem 1 in the form of several lemmas. First we make some preliminary comments on the order relation in $\mathfrak{X}$. In $\mathfrak{X}$ there are defined positive vectors $x \geq 0$, the totality of which will be denoted by $\mathfrak{X}_+$. Consequently there are also defined positive linear operators $T \geq 0$ which transform positive vectors of $\mathfrak{D}[T]$ into positive vectors. The sum and the product of two positive, bounded linear operators[9] are positive, and $0 \leq S \leq T$ implies $\|S\| \leq \|T\|$. Further we make frequent use of the relation

$$(2.1) \qquad \|x+y\| = \|x\| + \|y\| \qquad \text{for} \qquad x, y \in \mathfrak{X}_+ .$$

In particular it follows from (2.1) that, whenever $0 \leq y_1 \leq y_2 \leq \cdots$ with $\sup \|y_n\| < \infty$, there is a $y \in \mathfrak{X}_+$ such that $y_n \to y$ strongly. And hence, if there is a sequence S_n of bounded linear operators such that $0 \leq S_1 \leq S_2 \leq \cdots$ with $\sup \|S_n\| < \infty$, then there is a bounded linear operator $S \geq 0$ such that $S_n \to S$ strongly.

We now introduce two auxiliary operators H and K. $-H$ is defined by the diagonal part of (a_{jk}); more precisely, let $\mathfrak{D}$ be the totality of $x = (\xi_k) \in \mathfrak{X}$ such that $\sum a_k |\xi_k| < \infty$, and define $Hx = (a_k \xi_k)$ with domain $\mathfrak{D}[H] = \mathfrak{D}$. We note that $x_j \in \mathfrak{D}$ and hence $\mathfrak{D}_0 \subset \mathfrak{D}$. Thus $\mathfrak{D}$ is dense in $\mathfrak{X}$. We write $\mathfrak{D}_+ = \mathfrak{D} \cap \mathfrak{X}_+$ and $\mathfrak{D}_{0+} = \mathfrak{D}_0 \cap \mathfrak{X}_+$. The following lemma is now easily proved.

---

9) By a bounded linear operator we mean one with domain $\mathfrak{X}$ unless otherwise stated.

LEMMA 1. *a)* *For every* $\lambda > 0$, $(\lambda I + H)^{-1}$ *is a bounded,[9] linear, positive operator with* $\|(\lambda I + H)^{-1}\| \leq \lambda^{-1}$. *b)* H *is a closed, linear, positive operator with domain* $\mathfrak{D}$ *dense in* $\mathfrak{X}$. *c)* H *is the smallest closed extension of* H_0, *the contraction of* H *to domain* $\mathfrak{D}_0$. *In other words, for each* $z \in \mathfrak{D}$ *there is a sequence* $z_n \in \mathfrak{D}_0$ *such that* $z_n \to z$, $H_0 z_n = H z_n \to H z$, *where we can take* $z_n \in \mathfrak{D}_{0+}$ *if* $z \in \mathfrak{D}_+$ *(for instance set* $z_n = (\zeta_1, \zeta_2, \cdots, \zeta_n, 0, 0, \cdots)$ *for* $z = (\zeta_1, \zeta_2, \zeta_3, \cdots))$.

K is defined by the non-diagonal part of (a_{jk}). We note that by (1.6)

$$\sum_k \sum_{j \neq k} |\xi_j \, a_{jk}| \leq \sum_j |\xi_j| \sum_{k \neq j} a_{jk} = \sum_j a_j |\xi_j| = \|Hx\|$$

for every $x = (\xi_k) \in \mathfrak{D}$. We can therefore define the operator K by stipulating $\mathfrak{D}[K] = \mathfrak{D}$ and $Kx = y = (\eta_k)$ with $\eta_k = \sum_{j \neq k} \xi_j \, a_{jk}$.

LEMMA 2. K *is a positive linear operator and* $\|Kx\| \leq \|Hx\|$ *for* $x \in \mathfrak{D}$, $\|Kx\| = \|Hx\|$ *for* $x \in \mathfrak{D}_+$.

This is obvious from the above inequality. Also we have

$$(2.2) \qquad\qquad A \supset -H + K \supset A_0,$$

and

LEMMA 3. $B(\lambda) = K(\lambda I + H)^{-1}$ *is a bounded,[9] linear, positive operator with* $\|B(\lambda)\| \leq 1$ *and* $0 \leq B(\mu) \leq B(\lambda)$ *for* $0 < \lambda < \mu$.

PROOF. The first part is clear from Lemmas 1 and 2, for $\|K(\lambda I + H)^{-1} x\| \leq \|H(\lambda I + H)^{-1} x\| \leq \|x\|$ by $0 \leq H(\lambda I + H)^{-1} = I - \lambda(\lambda I + H)^{-1} \leq I$. The second part follows from $B(\lambda) - B(\mu) = (\mu - \lambda) B(\lambda)(\mu I + H)^{-1} \geq 0$.

We now construct the semi-group $\{P(t)\}$ by a limiting procedure. For each r with $0 \leq r < 1$ we define the operator $G_r = -H + rK$ with $\mathfrak{D}[G_r] = \mathfrak{D}$. Then the resolvent $R_r(\lambda) = (\lambda I - G_r)^{-1}$ exists for $\lambda > 0$ and is a bounded[9] positive operator, for we have

$$(2.3) \qquad R_r(\lambda) = (\lambda I + H - rK)^{-1} = (\lambda I + H)^{-1}[I - rB(\lambda)]^{-1}$$

$$= (\lambda I + H)^{-1} \sum_{n=0}^{\infty} r^n B(\lambda)^n,$$

the series converging absolutely and every term being positive (Lemmas 1 and 3). Moreover, we have for every $x \in \mathfrak{D}_+$

$$||(\lambda I - G_r)x|| = ||(\lambda I + H - rK)x|| \geq ||(\lambda I + H)x|| - r||Kx||$$

$$= \lambda ||x|| + ||Hx|| - r||Kx|| \geq \lambda ||x||$$

by Lemmas 1, 2 and (2.1). This implies $||R_r(\lambda)y|| \leq \lambda^{-1}||y||$ for every $y \in \mathfrak{X}_+$ and hence for every $y \in \mathfrak{X}$. Thus we have

(2.4) $R_r(\lambda) \geq 0$, $||R_r(\lambda)|| \leq \lambda^{-1}$, $\lambda > 0$.

G_r is therefore the infinitesimal generator of a positive contraction semi-group[6] $\{P_r(t)\}$, $0 \leq t < \infty$, according to Hille [8] and Yosida [9]. Furthermore, since $R_r(\lambda)$ is for fixed λ an increasing function (in the sense of order relation) of r as is seen from (2.3), it follows from the inversion formula (Hille [8]) $P_r(t) = \lim_{n\to\infty} [nt^{-1}R_r(nt^{-1})]^n$ that $P_r(t)$ is for fixed t an increasing function of r. But as $||P_r(t)|| \leq 1$ for $0 \leq r < 1$, there exists for each $t \geq 0$ a bounded linear operator $P(t)$ such that $P_r(t) \to P(t)$, $r \uparrow 1$, strongly (see the remark above). Obviously we have $P(t) \geq 0$, $||P(t)|| \leq 1$, $P_r(t) \leq P(t)$. Actually we have

LEMMA 4. $\{P(t)\}$, $0 \leq t < \infty$, *is a positive contraction semi-group.*[6]

PROOF. The relations $P(s+t) = P(s)P(t)$ and $P(0) = I$ are the limits of the corresponding relations for $P_r(t)$. Since $P_r(t)$ is strongly continuous in t, the limit $P(t)$ is at any rate strongly measurable, and hence strongly continuous for $0 < t < \infty$ (Hille [7], p. 183). It remains only to be shown that $P(t)$ is strongly continuous at $t = 0$. Let $\varepsilon > 0$ and $x \in \mathfrak{X}_+$. Since $P_0(t)$ is strongly continuous at $t = 0$, there is a $\delta > 0$ such that $||P_0(t)x - x|| < \varepsilon/2$ for $0 \leq t < \delta$. For such t we have $||P_r(t)x - P_0(t)x|| = ||P_r(t)x|| - ||P_0(t)x|| \leq ||x|| - ||P_0(t)x|| \leq ||x - P_0(t)x|| < \varepsilon/2$, where we have used (2.1) for $P_0(t)x \geq 0$ and $P_r(t)x - P_0(t)x \geq 0$. Hence $||P_r(t)x - x|| \leq ||P_r(t)x - P_0(t)x|| + ||P_0(t)x - x|| < \varepsilon/2 + \varepsilon/2 = \varepsilon$. Since this is true for every $r < 1$ as long as $t < \delta$, we can let $r \uparrow 1$ and obtain $||P(t)x - x|| < \varepsilon$ for $0 \leq t < \delta$. This shows that $P(t)x \to x$, $t \downarrow 0$, for every $x \in \mathfrak{X}_+$, and hence for every $x \in \mathfrak{X}$, that is, $P(t)$ is strongly continuous at $t = 0$.

LEMMA 5. $P_r(t) \to P(t)$, $r \uparrow 1$, *holds strongly and uniformly in each finite interval of t.*

PROOF. The proof is the same as in the classical theorem of Dini. If the convergence $P_r(t)x \to P(t)x$ were not uniform in some finite interval of t and for some $x \in \mathfrak{X}_+$, there would exist sequences r_n and t_n such

that $t_n \to t_0$, $r_n \uparrow 1$, $\| P(t_n)x - P_{r_n}(t_n)x \| \geq \varepsilon_0 > 0$. Since the vector in $\| \quad \|$ and $P_{r_n}(t_n)x$ are both positive, we have by (2.1) $\| P(t_n)x \| - \| P_{r_n}(t_n)x \|$ $= \| P(t_n)x - P_{r_n}(t_n)x \| \geq \varepsilon_0$. On the other hand we have $0 \leq P_{r_m}(t_n)x$ $\leq P_{r_n}(t_n)x$ for $m < n$, so that $\| P_{r_m}(t_n)x \| \leq \| P_{r_n}(t_n)x \| \leq \| P(t_n)x \| - \varepsilon_0$. Letting $n \to \infty$ for fixed m and making use of the strong continuity of $P_{r_m}(t)$ and $P(t)$ already proved, we get $\| P_{r_m}(t_0)x \| \leq \| P(t_0)x \| - \varepsilon_0$, which contradicts the fact that $P_{r_m}(t_0)x \to P(t_0)x$. Thus we have proved that the convergence $P_r(t)x \to P(t)x$ is uniform for each $x \in \mathfrak{X}_+$, and hence also for each $x \in \mathfrak{X}$.

LEMMA 6. *Let G be the infinitesimal generator of the semi-group $\{P(t)\}$ and let $R(\lambda) = (\lambda I - G)^{-1}$ be its resolvent. $R(\lambda)$ exists for $\lambda > 0$ and $\| R(\lambda) \| \leq \lambda^{-1}$, $0 \leq R_r(\lambda) \leq R(\lambda)$, $0 \leq r < 1$, and $R_r(\lambda) \to R(\lambda)$, $r \uparrow 1$, strongly for each $\lambda > 0$.*

PROOF. That $R(\lambda)$ exists for $\lambda > 0$ and $\| R(\lambda) \| \leq \lambda^{-1}$ is a direct consequence[10] of Lemma 4. For every $x \in \mathfrak{X}$ we have[10] $R(\lambda)x - R_r(\lambda)x$ $= \int_0^\infty \exp(-\lambda t) [P(t)x - P_r(t)x] dt$. Dividing the integral $\int_0^\infty$ into two parts $\int_0^a + \int_a^\infty$ with sufficiently large a and making use of the uniform convergence $P_r(t)x \to P(t)x$ in $0 \leq t \leq a$ (Lemma 5) and the uniform boundedness $\| P_r(t) \| \leq 1$, $\| P(t) \| \leq 1$, it follows easily that $R_r(\lambda)x \to R(\lambda)x$ strongly for $r \uparrow 1$. That $R_r(\lambda) \leq R(\lambda)$ follows at once from the above expression or, more directly, from the fact that $R_r(\lambda)$ is an increasing function of r as we have already noted.[11]

LEMMA 7. *For every $\lambda > 0$, $R^{(n)}(\lambda) \equiv (\lambda I + H)^{-1} \sum_{m=0}^{n} B(\lambda)^m \to R(\lambda)$, $n \to \infty$, strongly.*

PROOF. It follows from (2.3) and Lemma 6 that

$$(2.5) \qquad R_r^{(n)}(\lambda) \equiv (\lambda I + H)^{-1} \sum_{m=0}^{n} r^m B(\lambda)^m \leq R_r(\lambda) \leq R(\lambda) .$$

10) Hille [7], [8].

11) Obviously we can prove directly the existence of $R(\lambda) = \lim_{r \uparrow 1} R_r(\lambda)$ and then define $P(t)$ through $R(\lambda)$. However, the proof of the fact that $R(\lambda)$ thus defined is the resolvent of a closed linear operator G with dense domain is not very simple. For this reason we have chosen the indirect way presented here.

Letting $r \uparrow 1$ we get $R^{(n)}(\lambda) \leq R(\lambda)$. But as $R^{(n)}(\lambda)$ is increasing with n, strong $\lim R^{(n)}(\lambda) = R'(\lambda)$ exists by the remark above and $R'(\lambda) \leq R(\lambda)$. On the other hand, we have $R_r^{(n)}(\lambda) \leq R^{(n)}(\lambda) \leq R'(\lambda)$ and hence $R_r(\lambda) = \lim_{n \to \infty} R_r^{(n)}(\lambda) \leq R'(\lambda)$, $R(\lambda) = \lim_{r \uparrow 1} R_r \leq R'(\lambda)$ by (2.3) and Lemma 6. Thus we have $R(\lambda) = R'(\lambda)$.

LEMMA 8. *G is a closed extension of* $-H + K$ *and, a fortiori, of* A_0.

PROOF. Since $B(\lambda) = K(\lambda I + H)^{-1}$ by definition, $R^{(n)}(\lambda)$ can be written as $R^{(n)}(\lambda) = (\lambda I + H)^{-1} + (\lambda I + H)^{-1} [\sum_{m=0}^{n-1} B(\lambda)^m] K(\lambda I + H)^{-1} = (\lambda I + H)^{-1} + R^{(n-1)}(\lambda) K(\lambda I + H)^{-1}$. Hence we have, for any $x \in \mathfrak{D}$, $R^{(n)}(\lambda)(\lambda I + H)x = x + R^{(n-1)}(\lambda) Kx$. The limit $n \to \infty$ gives by Lemma 7 $R(\lambda)(\lambda I + H - K)x = x$. This shows that $\mathfrak{D}[G]$ (=the range of $R(\lambda)$) contains x and that $(\lambda I - G)x = (\lambda I + H - K)x$. In other words $G \supset -H + K$. G is closed since it is the infinitesimal generator of a semi-group.

LEMMA 9. *If there is a semi-group[6]* $\{P'(t)\}$, $0 \leq t < \infty$, *such that* $P'(t) \geq 0$ *and its infinitesimal generator* G' *is an extension of* A_0, *then* $P'(t) \geq P(t)$, $0 \leq t < \infty$.

PROOF. First we note that G' is also an extension of $-H + K$. In fact, for each $z \in \mathfrak{D}$ there is a sequence $z_n \in \mathfrak{D}_0$ such that $z_n \to z$, $Hz_n \to Hz$ (Lemma 1). It follows that $\|K(z_n - z)\| \leq \|H(z_n - z)\| \to 0$, $A_0 z_n = (-H + K)z_n \to (-H + K)z$. But as G' is a closed extension of A_0, we have $z \in \mathfrak{D}[G']$ and $G'z = (-H + K)z$, that is, $G' \supset -H + K$. Now the resolvent $R'(\lambda) = (\lambda I - G')^{-1}$ exists for sufficiently large λ and is a bounded linear operator.[10] We have $R'(\lambda) - R_r(\lambda) = R'(\lambda)(G' - G_r)R_r(\lambda)$, for $R_r(\lambda)\mathfrak{X} = \mathfrak{D} \subset \mathfrak{D}[G']$ by what is just proved. But as $G' - G_r = -H + K - (-H + rK) = (1 - r)K$ in $\mathfrak{D}$, we have $R'(\lambda) - R_r(\lambda) = (1 - r)R'(\lambda) KR_r(\lambda) \geq 0$, for $R'(\lambda)$, K, and $R_r(\lambda)$ are all positive. It follows that $R'(\lambda) \geq R_r(\lambda)$ and hence, by letting $r \uparrow 1$, $R'(\lambda) \geq R(\lambda)$. This in turn implies that $P'(t) \geq P(t)$ by the inversion formula used above.

LEMMA 10. *G is a contraction of* A.

PROOF. It follows from the definition of $R^{(n)}(\lambda)$ and $B(\lambda)$ (Lemmas 3 and 7) that

(2.6) $$(\lambda I + H)^{-1} KR^{(n)}(\lambda) = R^{(n+1)}(\lambda) - (\lambda I + H)^{-1}.$$

Letting $n \to \infty$ and noting Lemma 7, we get $(\lambda I + H)^{-1} KR^{(n)}(\lambda) \to R(\lambda)$ $-(\lambda I + H)^{-1}$ strongly. Let $y \in \mathfrak{X}_+$ and set $R(\lambda)y = x$, $R^{(n)}(\lambda)y = x^{(n)}$.

Then we have $(\lambda I+H)^{-1}Kx^{(n)}\to x-(\lambda I+H)^{-1}y$. Also we have $0\leq x^{(1)}$ $\leq x^{(2)}\leq\cdots$ and $x^{(n)}\to x$. In terms of components, these imply

$$(\lambda+a_k)^{-1}\sum_{j\neq k}\xi_j^{(n)}\,a_{jk}\to\xi_k-(\lambda+a_k)^{-1}\eta_k$$

or

$$\sum_{j\neq k}\xi_j^{(n)}\,a_{jk}\to(\lambda+a_k)\,\xi_k-\eta_k\,,\qquad\qquad k=1,2,\cdots,$$

and $0\leq\xi_j^{(1)}\leq\xi_j^{(2)}\leq\cdots\leq\xi_j^{(n)}\to\xi_j$. Since $a_{jk}\geq0$ for $j\neq k$, we obtain

$$\sum_{j\neq k}\xi_j\,a_{jk}=(\lambda+a_k)\xi_k-\eta_k\qquad\text{or}\qquad\sum_j\xi_j\,(\lambda\delta_{jk}-a_{jk})=\eta_k\,,$$

the series being absolutely convergent. This shows that $x\in\mathfrak{D}[A]$ and $(\lambda I-A)x=y$. Thus we have shown that $R(\lambda)y\in\mathfrak{D}[A]$ and $(\lambda I-A)$ $R(\lambda)y=y$ for $y\in\mathfrak{X}_+$, and hence, for all $y\in\mathfrak{X}$. Since $R(\lambda)=(\lambda I-G)^{-1}$, this proves that $A\supset G$.

Summing up the results of these lemmas, the proof of Theorem 1 is complete.

§ 3. Conditions for norm-preservation.

In this section we shall obtain some conditions that $\|P(t)x\|=\|x\|$ hold for a particular $x\in\mathfrak{X}_+$, or for all $x\in\mathfrak{X}_+$.

THEOREM 2. *For every* $x\in\mathfrak{X}_+$ *and* $\lambda>0$, *the limit* $b(\lambda;x)=\lim\limits_{n\to\infty}$ $\|B(\lambda)^n x\|$ *exists. For fixed* λ, $b(\lambda;x)$ *is a positive-linear function of* x. *For fixed* x, *it is either identically zero or a strictly positive, non-increasing function of* λ. $\|P(t)x\|$ *is also a non-increasing function of* t, *and*

$$(3.1)\qquad\qquad\lim_{t\to\infty}\|P(t)x\|=\|x\|-\lim_{\lambda\downarrow0}b(\lambda;x)\,,\qquad x\in\mathfrak{X}_+\,.$$

$\lambda^{-1}[\|x\|-b(\lambda;x)]$ *is also a strictly positive, decreasing function of* λ *for* $0\neq x\in\mathfrak{X}_+$, *and*

$$(3.2)\qquad\qquad\int_0^\infty\|P(t)x\|\,dt=\lim_{\lambda\downarrow0}\lambda^{-1}[\|x\|-b(\lambda;x)]\,,\qquad x\in\mathfrak{X}_+\,,$$

where the value of both sides may be ∞.

COROLLARY. *Let* $x\in\mathfrak{X}_+$. *In order that* $\|P(t)x\|=\|x\|$ *for* $0\leq t$ $<\infty$, *it is necessary that* $B(\lambda)^n x\to0$, $n\to\infty$, *for every* $\lambda>0$, *and it is sufficient that this is true for some* $\lambda>0$.

PROOF. We base our proof on the formulas

$$(3.3) \qquad \| R(\lambda)x \| = \int_0^\infty \exp(-\lambda t) \| P(t)x \| dt, \qquad x \in \mathfrak{X}_+,$$

which follows from the fundamental formula[10] $R(\lambda)x = \int \exp(-\lambda t)$ $P(t)x dt$ by making use of (2.1), and

$$(3.4) \qquad I + KR^{(n)}(\lambda) = (\lambda I + H)R^{(n)}(\lambda) + B(\lambda)^{n+1} \quad \Big(= \sum_{m=0}^{n+1} B(\lambda)^m \Big)$$

which is a direct consequence of the definitions of $R^{(n)}(\lambda)$ and $B(\lambda)$. Since each term of (3.4) is positive, we have for $x \in \mathfrak{X}_+$

$$\| x \| + \| KR^{(n)}(\lambda)x \| = \lambda \| R^{(n)}(\lambda)x \| + \| HR^{(n)}(\lambda)x \| + \| B(\lambda)^{n+1}x \|.$$

Since $x \in \mathfrak{X}_+$ implies $R^{(n)}(\lambda)x \in \mathfrak{D}_+$, we have $\| KR^{(n)}(\lambda)x \| = \| HR^{(n)}(\lambda)x \|$ by Lemma 2, and these terms cancel out. On letting $n \to \infty$ and noting that $R^{(n)}(\lambda)x \to R(\lambda)x$ (Lemma 7), we thus obtain

$$(3,5) \qquad \| x \| = \lambda \| R(\lambda)x \| + \lim_{n \to \infty} \| B(\lambda)^n x \| = \lambda \| R(\lambda)x \| + b(\lambda; x), \qquad x \in \mathfrak{X}_+.$$

Also it follows from (3.3) and (3.5) that

$$(3.6) \qquad \| x \| - b(\lambda; x) = \lambda \int_0^\infty \exp(-\lambda t) \| P(t)x \| dt,$$

$$x \in \mathfrak{X}_+.$$

$$(3.7) \qquad b(\lambda; x) = \lambda \int_0^\infty \exp(-\lambda t)(\| x \| - \| P(t)x \|)dt.$$

That $b(\lambda; x)$ is positive-linear in x is clear from (3.7). If $b(\lambda; x) = 0$ for some $\lambda > 0$, then (3.7) shows that $\| P(t)x \| = \| x \|$ must hold for all t, for $\| P(t)x \| \leq \| x \|$ in any case and $\| P(t)x \|$ is a continuous function of t. Then (3.7) implies in turn that $b(\lambda; x) = 0$ for all $\lambda > 0$. Thus $b(\lambda; x)$ is either identically zero or strictly positive for every $\lambda > 0$. Since $0 \leq B(\mu) \leq B(\lambda)$ for $\lambda < \mu$ by Lemma 3, it follows that $0 \leq B(\mu)^n x \leq B(\lambda)^n x$ for $x \in \mathfrak{X}_+$ and hence $\| B(\mu)^n x \| \leq \| B(\lambda)^n x \|$, $b(\mu; x) \leq b(\lambda; x)$, that is, $b(\lambda; x)$ is a non-increasing function of λ. Since $0 \leq b(\lambda; x) \leq \| x \|$, $\lim_{\lambda \downarrow 0} b(\lambda; x)$ exists and lies between 0 and $\| x \|$.

(3.1) now follows immediately from (3.6) if we note that $\| P(t)x \|$ is a non-increasing function of t by $\| P(t+s)x \| = \| P(s)P(t)x \| \leq \| P(t)x \|$.

(3.6) also shows that $\lambda^{-1}[\|x\| - b(\lambda; x)]$ is a decreasing function of λ for $x \neq 0$, and on letting $\lambda \downarrow 0$ we get (3.2). This completes the proof of Theorem 2.

We note that the assertions of II of §1 follow immediately from Theorem 2 and its Corollary; we have only to note that $b(\lambda; x_j) = b_j(\lambda)$.

THEOREM 3. *In order that $\{P(t)\}$ be a transition semi-group,[12] it is necessary and sufficient that one of the following equivalent conditions be satisfied:*

(i) *For every $\lambda > 0$, $B(\lambda)^n \to 0$, $n \to \infty$, strongly.*

(ii) *For every $\lambda > 0$, the characteristic equation $A_0^* x^* = \lambda x^*$ has no solution $x^* \neq 0$, where A_0^* is the adjoint of A_0 defined in $\mathfrak{X}^*$, the conjugate space of $\mathfrak{X}$.*

(iii) *For every $\lambda > 0$, the range of $\lambda I - A_0$ is dense in $\mathfrak{X}$.*

In each of these statements, "every $\lambda > 0$" may be replaced by "some $\lambda > 0$". When these conditions are satisfied, $\{P(t)\}$ is the only semi-group of which the infinitesimal generator is an extension of A_0.

PROOF. That (i) is a necessary and sufficient condition and that "every" may be replaced by "some" in (i) follow immediately from the Corollary to Theorem 2. Also it is clear that (ii) and (iii) are equivalent for any fixed λ, since $(\lambda I - A_0)\mathfrak{D}_0$ is dense in $\mathfrak{X}$ if and only if there is no $x^* \neq 0$ of $\mathfrak{X}^*$ which is orthogonal to $(\lambda I - A_0)\mathfrak{D}_0$.

We next show that (i) and (iii) are equivalent for any fixed λ. (3.4) shows that $B(\lambda)^n \to 0$ is equivalent to $(\lambda I + H - K)R^{(n)}(\lambda)x \to x$ for each $x \in \mathfrak{X}$. Since $R^{(n)}(\lambda)x \in \mathfrak{D}$, this implies that $(\lambda I + H - K)\mathfrak{D}$ is dense in $\mathfrak{X}$. But $(\lambda I - A_0)\mathfrak{D}_0 = (\lambda I + H - K)\mathfrak{D}_0$ is dense in $(\lambda I + H - K)\mathfrak{D}$ as is easily seen by taking the sequence $z_n \to z$ used in the proof of Lemma 9. Hence $(\lambda I - A_0)\mathfrak{D}_0$ is dense in $\mathfrak{X}$ if $B(\lambda)^n \to 0$.

Conversely let $(\lambda I - A_0)\mathfrak{D}_0$ be dense in $\mathfrak{X}$. Then $\left(I - B(\lambda)\right)\mathfrak{X}$ is dense a fortiori, for $[I - B(\lambda)]\mathfrak{X} = [I - B(\lambda)](\lambda I + H)\mathfrak{D} = (\lambda I + H - K)\mathfrak{D} \supset (\lambda I - A_0)\mathfrak{D}_0$. Set $B_n(\lambda) = (n+1)^{-1}[I + B(\lambda) + \cdots + B(\lambda)^n]$. Then we have $B_n(\lambda)[I - B(\lambda)]x = (n+1)^{-1}[I - B(\lambda)^{n+1}]x \to 0$, $n \to \infty$, for each $x \in \mathfrak{X}$ by $\|B(\lambda)\| \leq 1$. Since $[I - B(\lambda)]\mathfrak{X}$ is dense in $\mathfrak{X}$ and $B_n(\lambda)$ are uniformly bounded by $\|B_n(\lambda)\| \leq 1$, we see that $B_n(\lambda) \to 0$ strongly. But as all $B(\lambda)^m \geq 0$ and $\|B(\lambda)^n x\| = \|B(\lambda)^{n-m} B(\lambda)^m x\| \leq \|B(\lambda)^m x\|$, we have

---

12) That is, $\|P(t)x\| = \|x\|$ for all $x \in \mathfrak{X}_+$ and $0 \leq t < \infty$.

$||B_n(\lambda)x||=(n+1)^{-1}\sum\limits_{m=0}^{n}||B(\lambda)^m x||\geqq||B(\lambda)^n x||$ provided $x\in\mathfrak{X}_+$. Hence we have $B(\lambda)^n x\to 0$ for every $x\in\mathfrak{X}_+$ and this implies that $B(\lambda)^n\to 0$ strongly. Thus we have shown that (i) and (iii) are equivalent.

The last assertion of Theorem 3 follows from (iii). (iii) implies that the domain of $(\lambda I-A_0)^{-1}$ is dense in $\mathfrak{X}$. If $G>A_0$ is an infinitesimal generator of a semi-group,[6] the resolvent $(\lambda I-G)^{-1}$ must be a bounded[9] extension of $(\lambda I-A_0)^{-1}$ at least for sufficiently large λ. Hence it is uniquely determined and the same is true for G and the semigroup itself. This completes the proof of Theorem 3.

The proposition III of §1 is only a translation of Theorem 3 into matrix language. In fact, since $B(\lambda)^n$ are uniformly bounded and $\{x_j\}$ form a fundamental set of $\mathfrak{X}$, $B(\lambda)^n\to 0$ holds strongly if and only if $B(\lambda)^n x_j\to 0$, $j=1,2,\cdots$. (1.11) is the expression of $x^*[(\lambda I-A_0)x_j]=0$ if we set $x^*=(\zeta_1,\zeta_2,\cdots)$ with $||x^*||=\sup|\zeta_j|$.

§ 4. Examples.

We here discuss some simple examples to illustrate the results stated in §1. For brevity we shall speak of a *regular problem* if the conditions of III (or those of Theorem 3) are satisfied and a *singular problem* otherwise.

EXAMPLE 1. *Let* $\sup a_j=a<\infty$. Then it follows easily that $\sum b_{jk}^{(n)}(\lambda)\leqq[a/(\lambda+a)]^n$, that is, $||B(\lambda)||\leqq a/(\lambda+a)$ and $||B(\lambda)^n||\leqq[a/(\lambda+a)]^n$. Hence $b_j(\lambda)=0$, $j=1,2,\cdots$, and *the problem is regular.*

EXAMPLE 2. *Let* $a_{jk}=0$ *for* $k>j$ (triangular matrix with zeros above the main diagonal. The " death process "[13] belongs to this case). Then we have $b_{jk}(\lambda)=0$ for $k\geqq j$ and hence $b_{jk}^{(n)}(\lambda)=0$ for $k\geqq j-n+1$. For fixed j, we have therefore $b_{jk}^{(n)}(\lambda)=0$ for all k if n is sufficiently large. Thus $b_j(\lambda)=0$ and *the problem is regular.*

EXAMPLE 3. *Let* $a_{j,j+1}=a_j=-a_{jj}$, all other $a_{jk}=0$ (the " birth process ").[13] Then we have $b_{j,j+1}(\lambda)=a_j/(\lambda+a_j)$, all other $b_{jk}(\lambda)=0$. It follows that $b_{j,j+n}^{(n)}(\lambda)=\prod\limits_{m=0}^{n-1}a_{j+m}/(\lambda+a_{j+m})$, all other $b_{jk}^{(n)}(\lambda)$ being zero.

Thus $b_j(\lambda)=0$ and (1.2) holds for a particular j if and only if $\sum\limits_{m=0}^{\infty}a_{j+m}^{-1}$ $=\infty$ (where we have to set $a_m^{-1}=\infty$ whenever $a_m=0$). Thus *the pro-*

---

13) Cf. Feller [6], Chap. 17.

blem is regular if and only if either $a_k=0$ for infinitely many k or the series $\sum_{n=N}^{\infty} a_n^{-1}$ diverges for such N that $a_n>0$ for $n\geq N$.[13]

Let us consider the case $\sum_{m=0}^{\infty} a_{m+j}^{-1}<\infty$ in more detail. In this case we have $b_j(\lambda)=\lim\limits_{n\to\infty} b_{j,j+n}^{(n)}(\lambda)=[\prod\limits_{m=0}^{\infty}(1+\lambda a_{j+m}^{-1})]^{-1}=1-\lambda\sum\limits_{m=0}^{\infty} a_{j+m}^{-1}+O(\lambda^2)$ for small λ. It follows from (1.9) and (1.10) that

$$\lim_{t\to\infty}\sum_{k=1}^{\infty} p_{jk}(t)=0, \qquad \int_0^{\infty}\left[\sum_{k=1}^{\infty} p_{jk}(t)\right]dt=\sum_{m=0}^{\infty} a_{j+m}^{-1}.$$

EXAMPLE 4. *Let $a_{jk}=0$ for $k>j+1$* (this example includes Examples 2 and 3 and also the " birth and death process "[13] as special cases). The explicit calculation of $b_j(\lambda)$ seems to be difficult. But we can apply the criterion (1.11), which can be solved in this case successively as

$$\zeta_{j+1}=a_{j,j+1}^{-1}\left[-\sum_{k=1}^{j-1} a_{jk}\zeta_k+(\lambda+a_j)\zeta_j\right]$$

(for simplicity we assume that $a_{j,j+1}>0$ for all j, though the contrary case offers no difficulty). Making use of (1.6) this can be written as

(4.2) $\zeta_{j+1}-\zeta_j=\lambda a_{j,j+1}^{-1}\zeta_j+a_{j,j+1}^{-1}\sum\limits_{k=1}^{j-1} a_{jk}(\zeta_j-\zeta_k).$

Thus all ζ_j are determined up to the factor ζ_1, which may be assumed to be ≥ 0. If $\zeta_1>0$, we see from (4.2) by induction that $0<\zeta_1<\zeta_2<\cdots$. On adding (4.2) from $j=1$ to $j=n$, we get

$$\xi_{n+1}-\zeta_1=\lambda\sum_{j=1}^{n} a_{j,j+1}^{-1}\zeta_j+\sum_{k=1}^{n-1}\left[\sum_{j=k+1}^{n} a_{j,j+1}^{-1} a_{jk}(\zeta_j-\zeta_k)\right].$$

If the problem is to be singular, we must have $\sup \zeta_n<\infty$ and hence $\sum\limits_{j=1}^{\infty} a_{j,j+1}^{-1}<\infty$ and $\sum\limits_{j=k+1}^{\infty} a_{j,j+1}^{-1}a_{jk}<\infty$ for each fixed k. Again, (4.2) implies $\zeta_{j+1}-\zeta_j\geq a_{j,j+1}^{-1}\left(\sum\limits_{k=1}^{j-1} a_{jk}\right)(\zeta_j-\zeta_{j-1})$ so that it is necessary for $\sup \zeta_n<\infty$ that $a_{j,j+1}^{-1}\sum\limits_{k=1}^{j-1} a_{jk}<1$ for infinite number of j. Thus we see that *the problem is regular if at least one of the following conditions is satisfied:*

$$\sum_{j=1}^{\infty} a_{j,j+1}^{-1} = \infty\ ; \qquad \sum_{j=2}^{\infty} a_{j,j+1}^{-1} a_{j1} = \infty\ ; \qquad \sum_{j=3}^{\infty} a_{j,j+1}^{-1} a_{j2} = \infty\ ,$$

(4.3) $\cdots\cdots\cdots\cdots$;

$$a_{j,j+1}^{-1} \sum_{k=1}^{j-1} a_{jk} \geq 1 \quad \text{for all sufficiently large } j .$$

On the other hand, we have from (4.2) $\zeta_{j+1} \leq \left[1 + \lambda\, a_{j,j+1}^{-1} + a_{j,j+1}^{-1} \sum_{k=1}^{j-1} a_{jk} \right] \zeta_j$. It follows that $\sup \zeta_n < \infty$ and hence that *the problem is singular if both of the following conditions are satisfied :*

(4.4) $\displaystyle\sum_{j=1}^{\infty} a_{j,j+1}^{-1} < \infty \qquad and \qquad \sum_{j=1}^{\infty} \left[a_{j,j+1}^{-1} \sum_{k=1}^{j-1} a_{jk} \right] < \infty .$

Of course there is a gap between the conditions (4.3) and (4.4) so that we have not obtained a necessary and sufficient condition for regularity, even in the case of simple birth and death process where $a_{j,j-1}$, a_{jj} and $a_{j,j+1}$ are the only non-zero elements.

Department of Physics, University of Tokyo.

Bibliography

[1] G. Birkhoff, *Lattice Theory*. Amer. Math. Soc. Coll. Publ. Vol. 25, Rev. Ed., New York, 1948.

[2] J. L Doob, *Topics in the theory of Markoff chains*, Trans. Amer. Math. Soc. 52 (1942), 37-64.

[3] J. L. Doob, *Markoff chains—denumerable case*, Trans. Amer. Math., Soc. 58 (1945), 455-473.

[4] J. L. Doob, *Stochastic Processes*, New York, 1953.

[5] W. Feller, *On the integro-differential equations of purely discontinuous Markoff processes*, Trans. Amer. Math. Soc. 48 (1940), 488-515; Errata, ibid. 58 (1945), 474.

[6] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. I, New York, 1950.

[7] E. Hille, *Functional Analysis and Semi-groups*, Amer. Math. Soc. Coll. Publ. Vol. 31, New York, 1948.

[8] E. Hille, *On the generation of semi-groups and the theory of conjugate functions*, Kungl. Fysiogr. Sällsk. Lund Förhandl. Bd. 21, Nr. 14, 1951.

[9] ·K. Yosida, *On the differentiability and the representation of one-parameter semi-group of linear operators*, J. Math. Soc. Japan 1 (1948), 15-21.

[10] K. Yosida, *An operator-theoretical treatment of temporally homogeneous Markoff processes*, J. Math. Soc. Japan 1 (1949), 244-253.