83. On a Theorem of F. L. Spitzer and C. J. Stone

By Takesi WATANABE

Department of Applied Science, Kyushu University, Fukuoka, Japan

(Comm. by Z. SUETUNA, M.J.A., July 12, 1961)

1. **Introduction.** In their recent work [6], Spitzer and Stone have proved the following interesting theorem which was the basis of their discussion. Consider a sequence $\{c_k; k=0, \pm 1, \pm 2, \cdots\}$ satisfying the conditions:

(c.1) $$c_k \geqq 0, \quad \sum_{k=-\infty}^{\infty} c_k = 1,$$

(c.2) $$0 < \sum_{k=-\infty}^{\infty} k^2 c_k = v < +\infty,$$

(c.3) $$c_k = c_{-k},$$

(c.4) $$\text{g.c.d. } \{k; k>0, c_k>0\}=1.$$

Putting $\varphi(\theta)=\sum_{k=-\infty}^{\infty} c_k e^{ik\theta}$ and noting $2 \geqq 1-\varphi(\theta) \geqq 0$, it follows that there exists a unique sequence of polynomials $\{p_n(z)=\sum_{k=0}^{n} p_{nk} z^k; p_{nn}>0, n=0, 1, 2, \cdots\}$ satisfying

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} p_n(e^{i\theta})\overline{p_m(e^{i\theta})}[1-\varphi(\theta)]d\theta = \delta_{nm}$$

for $n, m=0, 1, 2, \cdots$.

THEOREM (Spitzer and Stone). *The relation*

$$p_{nk}-(2/v)^{\frac{1}{2}}(k/n) \to 0 \quad (n-k \to \infty)$$

holds uniformly in k and n.

In this note we shall derive a more probabilistic version of the above theorem under a weaker condition (c.3)$'$ $\sum_{k=-\infty}^{\infty} kc_k=0$ instead of (c.3). The main feature of our discussion is in full use of the general theory of Markov chains. By doing so we can prove Theorem 2.1 in [6] under (c.3)$'$ and substitute some simple probabilistic arguments for the rather complicated calculations in [6] (e.g. Lemmas 5–11).

2. **Markov chains.** We now summarize some fundamental facts on Markov chains (with discrete parameter). As to the details, we refer the reader to Chap. I of [7].

Let S be a finite or denumerable space and $T=(T(x,y); x, y \in S)$, a *substochastic matrix*[1] on S. Adding a new point e (called 'extra' point) to S, we extend T to $\tilde{S}=S \cup \{e\}$ as follows: $T(x,e)=1-\sum_{y\in S} T(x,y)$, $T(e,e)=1$ and $T(e,y)=0$ for $y \in S$. For any x in $\tilde{S}$, the new transition

1) $\sum_{y\in S} T(x,y) \leqq 1$ for every $x \in S$.

matrix $T=(T(x,y);x,y\in\tilde{S})$ determines the Markov chain $(x_t^{(x)}(w), t\in D=\{0,1,2,\cdots,+\infty\})$ whose initial distribution is the unit distribution at x, while $x_{+\infty}^{(x)}(w)=e$ with probability 1. With no loss of generality we can take the basic probability field $(W,\mathcal{B},P_x)$ in the following way. W is the set of all paths ($\tilde{S}$-valued function of t) satisfying the conditions that $w_{+\infty}=e$ and that if $w_t=e$, then $w_s=e$ for every $s\geqq t$, where w_t means the value at t of the path w. $\mathcal{B}$ is the ordinary Borel field generated by all cylinder sets in W. $P_x(\cdot)$ coincides with the joint distribution of $x_t^{(x)}(w)$. The system $(W,\mathcal{B},P_x,x\in\tilde{S})$ with the above choice for all x is called *the Markov chain over S associated with T* and is denoted simply by x_t. For any fixed $w\in W$ and $s\in D$, the stopped path w_s^- and shifted one w_s^+ are defined by $[w_s^-]_t=w_{\min(t,s)}$ $(t\neq+\infty)$, $=e(t=+\infty)$ and $[w_s^+]_t=w_{s+t}$, respectively. We define several quantities and properties concerning the Markov chain. Let A or E denote a subset of S. *The hitting time to A*, $\sigma_A(w)=\min\{t;w_t\in A\}$;[2] *the hitting probability from x to E*, $p(x,E)=P_x(\sigma_E<+\infty)$; *the Green measure* $G(x,E)=\sum_{t=0}^{\infty}P_x(w_t\in E)$ and *the harmonic measure to A*, $H_A(x,E)=P_x(w_{\sigma_A}\in E)$. The point x in S is called *recurrent*[3] if $P_x(\sigma_x(w_1^+)<+\infty)=1$ and *transient* if it is not recurrent. Since the notions of Markov times and the strong Markov property are well known, we omit their precise description.

Let A be any subset of S. Then *the restriction x_t^A of x_t to A* is defined as the Markov chain over A, $(W^A,\mathcal{B}^A,P_x^A,x\in A\cup\{e\})$, associated with $T(A)=(T(x,y);x,y\in A)$. The new measure $P_x^A(\cdot)$ corresponds to the original one $P_x(\cdot)$ in the following simple manner. Considering the transformation $x^A(w)$ from W to W^A defined by $x_t^A(w)=w_t\,(t<\sigma_{A^c})$ and $=e\,(t\geqq\sigma_{A^c})$, we have $P_x^A(\Lambda)=P_x(w;x^A(w)\in\Lambda)$ for every $\Lambda\in\mathcal{B}^A$. The hitting probability and Green measure of x_t^A are denoted by $p^A(x,E)$ and $G^A(x,E)$ respectively.

The following results to be used later are well known (see [7]) except the last two assertions.

1° *If x is recurrent and $p(x,y)>0$, y is also recurrent and $p(x,y)=p(y,x)=1$.*

2° *If y is transient, $G(x,y)=p(x,y)G(y,y)<+\infty$ for any x.*

3° *If $A\supset B$, $H_B(x,E)=\sum_{y\in A}H_A(x,y)H_B(y,E)$ for any x and E.*

4° *If $A\cap B=\phi$ and $B\supset E$,*

$$H_{A\cup B}(x,E)=H_B(x,E)-\sum_{y\in A}H_{A\cup B}(x,y)H_B(y,E)\quad \textit{for all } x. \tag{2.1}$$

Noting that $B\supset E$ and using the strong Markov property to

2) If { } is void, $\sigma_A(w)=+\infty$ conventionally.

3) In appearance this definition of recurrence is a little different to that of [7]. But in our discrete parameter case, both definitions are equivalent to each other.

σ_{A+B}, the proof is straightforward.

5° *For any* $A \subset S$, $x \in A$ *and any function* u *over* $A^c = S - A$,

(2.2) $$\sum_{y \in A^c} u(y) H_{A^c}(x, y) = \sum_{y \in A^c} \sum_{z \in A} G^A(x, z) T(z, y) u(y),$$

which must be interpreted in the sense that the existence of the one side of (2.2) *implies that of the other and the equality holds. Especially if* u *is a function over* S *and* $v(z) = \sum_{y \in S} T(z, y) |u(y)|$ *is integrable with respect to the measure* $G^A(x, \cdot)$, *the right side of* (2.2) *can be rewritten in the form*

(2.3) $$\sum_{z \in A} G^A(x, z) [\sum_{y \in S} T(z, y) u(y) - u(z)] + u(x).$$

PROOF. Putting $\sigma(w) = \sigma_{A^c}(w)$ and extending u to $A + \{e\}$ by $u(e) = 0$, we have

$$\text{the left side of (2.2)} = E_x(u(w_\sigma)) = \sum_{t=0}^{\infty} E_x(u(w_t); \sigma = t),^{4)}$$

$$E_x(u(w_0); \sigma = 0) = 0 \quad (\text{from } x \in A),$$

$$E_x(u(w_t); \sigma = t) = E_x(u((w_{t-1}^+)_1); \sigma(w_{t-1}^+) = 1, \sigma(w) > t-1)$$
$$= E_x(E_{w_{t-1}}(u(w_1); \sigma(w) = 1); \sigma(w) > t-1)$$

and therefore

$$E_x(u(w_\sigma)) = E_x\left(\sum_{t=0}^{\infty} E_{w_t}(u(w_1); \sigma = 1); \sigma > t\right),$$

which verifies (2.2). The latter half is a direct consequence of the formula $\sum_{z \in A} G^A(x, z) T(z, y) = G^A(x, y) - \delta(x, y)$[5] for every $y \in A$.

3. **Main results.** Let S be the set of all integers and $\{c_k\}$, the sequence over S satisfying the conditions (c.1), (c.2), (c.3)′ $\sum_{k=-\infty}^{\infty} k c_k = 0$ and (c.4)′ g.c.d. $\{|k|; c_k > 0\} = 1$. It is evident that (c.3)′ is much weaker than (c.3) and that (c.4)′ coincides with (c.4) if (c.3) is satisfied. We consider the Markov chain x_t corresponding to $T(k, j) = c_{j-k}$, $k, j \in S$. For such Markov chain, it is well known that $w_{t+1} - w_t$, $t = 0, 1, 2, \cdots$ are independent random variables having the same distribution $\{c_k\}$ relative to $P_k(\cdot)$ for any k and that, defining the shift transformation θ_j on W by $(\theta_j w)_t = w_t + j$, we have $P_k(\Lambda) = P_{k+j}(\theta_j \Lambda)$ for any Λ of $\mathcal{B}$. In this connection our chain x_t may be called *an additive Markov chain*. The open interval (k, j) of S means the set $\{l; l \in S, k < l < j\}$. The closed (or half open) interval of S should be understood in the same manner. Adopting this convention and the notations introduced §2, our theorem is stated as follows:

THEOREM. *Let* x_t *be the additive Markov chain defined just above. Then* $\mu = \sum_{j \geq 1} j H_{[1,\infty)}(0, j)$ *converges and the relation*

(3.1) $$p^{[0,n]}(k, n) - \mu^{-1}(k/n) \to 0 \quad (n - k \to \infty)$$

4) In general, $E_x(f(w); \Lambda)$ means the integral of $f(w)$ over the set $\Lambda \in \mathcal{B}$ relative to the measure $P_x(\cdot)$. If $P_x(\Lambda) = 1$, we shall omit Λ in the expectation.

5) $\delta(x, y) = 1$ $(x = y)$, $= 0$ $(x \neq y)$.

holds uniformly in k and n.

Before proving we prepare two lemmas, in which we have no need of the aperiodicity condition (c.4)′.

LEMMA 1. $G^{[0,\infty)}(j,j)=O(j)$.

PROOF. Consider the sequence of Markov times: $\tau_0(w)=0$, $\tau_1(w)=j^2+1+\sigma_j(w^+_{j^2+1}),\cdots,\ \tau_n(w)=\tau_{n-1}(w)+\tau_1(w^+_{\tau_{n-1}}),\cdots$. Putting $\sigma(w)=\sigma_{(-\infty,0)}(w)$, we have

$$G^{[0,\infty)}(j,j)=E_j\Big(\sum_{t\geqq 0}\chi_j(w_t);\,t<\sigma\Big)^{6)}=\sum_{n\geqq 0}E_j\Big(\sum_{t=\tau_n}^{\tau_{n+1}-1}\chi_j(w_t);\,t<\sigma\Big)$$
$$\leqq\sum_{n\geqq 0}E_j\Big(\sum_{t=\tau_n}^{\tau_{n+1}-1}\chi_j(w_t);\,\tau_n<\sigma\Big).$$

Applying the strong Markov property to τ_n and noting that $w_{\tau_n}=j$, it follows that

$$E_j\Big(\sum_{t=\tau_n}^{\tau_{n+1}-1}\chi_j(w_t);\,\tau_n<\sigma\Big)=E_j\Big[E_{w_{\tau_n}}\Big(\sum_{t=0}^{\tau_1-1}\chi_j(w_t)\Big);\ \tau_n<\sigma\Big]$$
$$=E_j\Big(\sum_{t=0}^{j^2}\chi_j(w_t)\Big)P_j(\tau_n<\sigma),$$
$$P_j(\tau_n<\sigma)=E_j[P_{w_{\tau_{n-1}}}(\tau_1<\sigma);\,\tau_{n-1}<\sigma]$$
$$=P_j(\tau_1<\sigma)P_j(\tau_{n-1}<\sigma)=[P_j(\tau_1<\sigma)]^n.$$

But from the central limit theorem we get

$$P_j(\tau_1<\sigma)\leqq P_j(j^2<\sigma)\leqq P_j(w_{j^2}\geqq 0)=P_0(w_{j^2}\geqq -j)$$
$$=P_0\Big(\frac{w_{j^2}}{j\sqrt{v}}\geqq-\frac{1}{\sqrt{v}}\Big)=\frac{1}{\sqrt{2\pi}}\int_{-\frac{1}{\sqrt{v}}}^{\infty}e^{-\frac{x^2}{2}}dx+o(1)<\alpha<1,$$

where α is a constant independent of j. Moreover the local limit theorem ([4], p. 233) shows that we can choose some constant β such that

$$P_j(w_t=j)=P_0(w_t=0)\leqq\beta(t+1)^{-\frac{1}{2}}\quad\text{for every } t\in D,$$

whence $$E_j\Big(\sum_{t=0}^{j^2}\chi_j(w_t)\Big)=\sum_{t=0}^{j^2}P_j(w_t=j)\leqq\beta\sum_{t=0}^{j^2}(t+1)^{-\frac{1}{2}}\leqq\beta(j+1).$$

Therefore $$G^{[0,\infty)}(j,j)\leqq\beta(j+1)\sum_{n\geqq 0}\alpha^n=(\beta/1-\alpha)(j+1),$$

which is what we wanted to show.

LEMMA 2. *It holds uniformly in k and n that*

$$H_{(-\infty,0)\cup(n,\infty)}(k,(n,\infty))-k/n\to 0\quad(n\to\infty).$$

PROOF. We shall give only a simple sketch of the proof because it runs along the same lines as the arguments of Lemmas 1–4 in [6], noting that the above-mentioned lemma acts as a substitute for Lemma 3 in [6]. Putting $A=[0,n]$ in 5° of §2 and using (2.3) and (c.3)′, it is shown that $\sum_{j\in[0,n]^c}jH_{[0,n]^c}(k,j)=k$, from which we get

$$H_{[0,n]^c}(k,(n,\infty))=k/n-(1/n)\sum_{j\leqq -1}jH_{[0,n]^c}(k,j)-(1/n)\sum_{j\geqq n+1}(j-n)H_{[0,n]^c}(k,j)$$
$$=k/n+I_1-I_2.$$

But by (2.2) and Lemma 1

6) χ_j is the indicator function of the one point set $\{j\}$.

$$I_1=-(1/n)\sum_{l=0}^{n}G^{[0,n]}(k,l)\sum_{j\leq -1}jc_{j-l}$$
$$\leq(\beta/1-\alpha)(1/n)\sum_{l=0}^{n}(l+1)\sum_{j\geq 1}jc_{-j-l}\to 0\quad(n\to\infty),$$

since (c.2) guarantees the convergence of $\sum_{l\geq 0}\sum_{j\geq 1}jc_{-j-l}=\sum_{j\geq 1}j(\sum_{l\geq j}c_{-l})$.[7] In the same way $I_2\to 0$ $(n\to\infty)$, using $G^{(-\infty,0]}(j,j)=O\,|j|$ which is a counterpart of Lemma 1.

PROOF OF THEOREM. It is convenient to divide our proof into several steps.

(i) $\mu<\infty$. This is a special case of Theorem 3.4 of Spitzer [5]. In fact we have

$$\mu=(v/2)^{\frac{1}{2}}\exp\left\{\sum_{t=1}^{\infty}\frac{1}{t}\left[\frac{1}{2}-P_0(w_t\geq 1)\right]\right\}<\infty.$$

(ii) $H_{[1,\infty)}(0,1)>0$.[8] We define the *right step hitting probability* $p^+(k,j)=P_k\{\sigma_j<+\infty,\ w_t<w_{t+1}$ for every $t<\sigma_j\}$ and put $S^+=\{j;\ p^+(0,j)>0\}$. It is clear that $j+j'\in S^+$ if both $j\in S^+$ and $j'\in S^+$. Therefore S^+ contains all sufficiently large multiples of $d^+=$g.c.d. of $S^+=$g.c.d. $\{j;\ j>0,\ c_j>0\}$ ([2], p. 176). In the same manner we consider the *left step hitting probability* $p^-(k,j)$, the set $S^-=\{j;\ p^-(0,j)>0\}$ and $d^-=$g.c.d.$\{-j;\ j\in S^-\}=$g.c.d.$\{-j;\ j<0,\ c_j>0\}$. For all sufficiently large n (>0), $-nd^+$ is in S^-. Since d^+ and d^- are relatively prime by (c.4)$'$, there exist some $j^+\in S^+$ and $j^-\in S^-$ such that $j^++j^-=1$. Consequently $H_{[1,\infty)}(0,1)\geq p^-(0,j^-)p^+(j^-,1)=p^-(0,j^-)p^+(0,1-j^-)=p^-(0,j^-)p^+(0,j^+)>0$. By the way we note that both S^+ and S^- are not void according to (c.2) and (c.3)$'$.

(iii) $\sum_{j\geq 1}H_{[1,\infty)}(0,j)=p(0,[1,\infty))=1$. The condition (c.3)$'$ implies that the point 0 (and therefore any point in S) is recurrent ([1], p. 2). But since $p(0,1)\geq H_{[1,\infty)}(0,1)>0$, it follows from 1° of §2 that $1=p(0,1)\leq p(0,[1,\infty))$.

(iv) $H_{[n,\infty)}(0,n)\to\mu^{-1}$ $(n\to\infty)$. Putting $A=[1,\infty)$, $B=[n,\infty)$ and $E=n$ in 3° of §2, we get

$$H_{[n,\infty)}(0,n)=\sum_{j\geq 1}H_{[1,\infty)}(0,j)H_{[n,\infty)}(j,n)=\sum_{j=1}^{n}H_{[1,\infty)}(0,j)H_{[n-j,\infty)}(0,n-j),$$

which is the well-known renewal equation. Since $H_{[1,\infty)}(0,j)$ satisfies (i)–(iii), the Feller's renewal theorem ([3], p. 286) is applicable and our assertion is verified.

(v) Noting that $p^{[0,n)}(k,n)=H_{(-\infty,0)\cup[n,\infty)}(k,n)$ and using (2.1), Lemma 2 and the above (iv), we have

$$p^{[0,n]}(k,n)=H_{[n,\infty)}(k,n)-\sum_{j\leq -1}H_{(-\infty,0)\cup[n,\infty)}(k,j)H_{[n,\infty)}(j,n)$$
$$=H_{[n-k,\infty)}(0,n-k)-\sum_{j\leq -1}H_{(-\infty,0)\cup[n,\infty)}(k,j)H_{[n-j,\infty)}(0,n-j)$$

7) In fact, $\sum_{j\geq 1}(2j+1)(\sum_{l\geq j}c_{-l})=\sum_{j\geq 1}(j^2+1)c_{-j}<+\infty$.

8) Our argument implies that x_t is irreducible, i.e. $p(k,j)>0$ for all $k,\ j\ni S$.

$$=\mu^{-1}(1-\sum_{j\leq -1} H_{(-\infty,0)\cup[n,\infty)}(k,j))+o(1)$$
$$=\mu^{-1}H_{(-\infty,0)\cup[n,\infty)}(k,[n,\infty))+o(1)=\mu^{-1}(k/n)+o(1),$$

where $o(1)$ tends to zero uniformly in k and n if $n-k\to\infty$. Thus our theorem has been proved completely.

REMARK. It is easily seen that $p^{[0,n]}(k,n)\to H_{[j,\infty)}(0,j)$ if $n-k\to j$.

4. The symmetric case. We shall show that *the theorem of* §3 *is a generalization of the Spitzer-Stone theorem stated in* §1. To see this, assuming the condition (c.3) instead of (c.3)′, we use the following facts which were established in Section 1 of [6]. (a) $G^{[0,n]}(k,j) = \sum_{r=\max(k,j)}^{n} p_{rk}p_{rj}$, where p_{rk}'s are those defined in §1. (b) There exists $u_0=\lim_{n\to\infty} p_{nn}$ and there holds $\mu=(v/2)^{\frac{1}{2}}u_0$. Then it results from (a) and 2° of §2 applied to $x_t^{[0,n]}$ that $p_{nk}p_{nn}=G^{[0,n]}(k,n)=p^{[0,n]}(k,n)G^{[0,n]}(n,n) =p^{[0,n]}(k,n)p_{nn}^2$. Therefore $p^{[0,n]}(k,n)=p_{nk}/p_{nn}$. Combining (b) and our theorem, the Spitzer-Stone theorem is immediate.

REMARK. From (a) and (c.3), it is clear that $p_{nn}=[G^{[0,n]}(n,n)]^{\frac{1}{2}} =[G^{[0,n]}(0,0)]^{\frac{1}{2}}\to[G^{[0,\infty)}(0,0)]^{\frac{1}{2}}$, so that (b) is reduced to show a probabilistic relation $E_0(w_{\sigma[1,\infty]})=2^{-\frac{1}{2}}[E_0(w_1^2)G^{[0,\infty)}(0,0)]^{\frac{1}{2}}$. But we have failed to give a simple probabilistic proof of this formula.

References

[1] K. L. Chung and W. H. J. Fuchs: On the distribution of values of sums of random variables, Mem. Amer. Math. Soc., **6**, 1–12 (1951).

[2] J. L. Doob: Stochastic Processes, New York (1953).

[3] W. Feller: An Introduction to Probability Theory and its Applications, 1, 2nd ed., New York (1957).

[4] B. V. Gnedenko and A. N. Kolmogorov: Limit Distributions for Sums of Independent Random Variables (English translation from the Russian by K.L. Chung), Cambridge, Mass. (1954).

[5] F. L. Spitzer: A Tauberian theorem and its probability interpretation, Trans. Amer. Math. Soc., **94**, 150–169 (1960).

[6] F. L. Spitzer and C. J. Stone: A class of Toeplitz forms and their application to probability theory, Ill. J. Math., **4**, 253–277 (1960).

[7] T. Watanabe: On the theory of Martin boundaries induced by countable Markov processes, Mem. Coll. Science, Univ. Kyoto, ser. A, **33**, 39–108 (1960).