# RECURRENCE AND FIRST PASSAGE TIMES

By M. S. BARTLETT

*Received* 23 September 1952

1. *Introduction.* The theoretical problem of the probable time of recurrence of a given state of a physical system which changes according to a specified random or stochastic process is *in general* not a very tractable one, but many situations are covered by the more particular methods and formulae given in this paper. Several of these are believed new*, but others are also included for reference. We shall usually limit our attention to homogeneous processes, i.e. the transition probabilities do not depend explicitly on the time. Even so, a generally useful solution of the complete distributional problem of recurrence times would appear rather unlikely; for example, it would have to include as a special case the recurrence theory for trigonometrical or power series, problems of some complexity which have separately engaged the attention of mathematicians (see, for example, Kac (7)). Thus a manageable simplification of the complete distributional problem emerges only in particular but important classes of physical problems classifiable theoretically as Markoff processes or, somewhat more generally, as 'renewal' processes (cf. Feller (4)). In other cases, if we are content with mean values (which do not always exist), and are dealing with stationary processes with complete ergodic properties†, such as a statistical mechanical system in equilibrium, the simple but powerful arguments originally used by Smoluchowski (11) are available. We first formulate the probability problem in symbolic terms, being more concerned in this paper with general structural relations than with special points of rigour, or specific applications‡.

2. *Symbolic formulae.* Suppose we wish to study the recurrence at the discrete time $n$ (i.e. the system can be, or is, only observed at discrete intervals) of a specified state $S$, given that $S$ occurred at time zero. We shall assume at present that the conditional probability $P\{S_r \mid S_0\}$ that the system will again be in the state $S$ at any time $r$, regardless of the states occurring between the times 0 and $r$, is not always zero, and it will be convenient to denote it by $\{e_r\}$. Similarly, the further notation will be used:

$$P\{\text{not } S_r \mid S_0\} \equiv \{\bar{e}_r\} = 1 - \{e_r\},$$

$$P\{S_r, S_s \mid S_0\} \equiv \{e_r e_s\},$$

$$P\{\text{not } S_r, S_s \mid S_0\} \equiv \{\bar{e}_r e_s\} = \{(1 - e_r) e_s\} = \{e_s\} - \{e_r e_s\},$$

* For example, in the classification given in the appendix, the distributions for case B, the *conditional* recurrence and first passage distributions in all cases, and the alternative formulae for finite Markoff chains.

† Mathematically, this means (if the possible states are classified by a single quantitative variable $x$) that the characteristic functional $\mathscr{E}\left\{\exp i \int_0^\tau X(u)\, d\Theta(u)\right\}$ of the stochastic process $X(u)$ for an arbitrary function $\Theta(u)$ is determined by the time-average

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T dt \left\{\exp i \int_0^\tau x(u+t)\, d\Theta(u)\right\}$$

obtained from a single realization $x(u)$ for any $\tau$ (and $\Theta(u)$).

‡ For a discussion more particularly referring to 'random walk' problems, see Harris (6).

https://doi.org/10.1017/S0305004100028358 Published online by Cambridge University Press

etc. Then for the probability of a first recurrence at time $n$ (here we do not exclude continuation in the state from time 0 to time 1, this counting as recurrence at time 1), we have

$$\{\bar{e}_1\bar{e}_2\dots\bar{e}_{n-1}e_n\} = \{(1-e_1)(1-e_2)\dots(1-e_{n-1})e_n\}$$
$$= \{e_n\} - \sum_{r=1}^{n-1}\{e_re_n\} + \sum_{r=1}^{n-1}\sum_{s=r+1}^{n-1}\{e_re_se_n\} - \dots. \quad (1)$$

This formula in theory determines the recurrence distribution, for the probabilities on the right are theoretically available for a validly specified stochastic process. For stationary processes they will of course be independent of the choice of the time origin 0.

If the discrete time $r$ is replaced by a continuous time $t$, and we assume that the probability of $S$ occurring in the infinitesimal interval $t, t+dt$ is $[e_t]\,dt$, the formula (1) becomes replaced by the (formal) relation for the recurrence probability *density* at time $t$,

$$[e_t] - \int_0^t [e_ue_t]\,du + \int_0^t\int_u^t [e_ue_re_t]\,du\,dv - \dots = [e_t] - \int_0^t [e_ue_t]\,du + \frac{1}{2!}\int_0^t\int_0^t [e_ue_ve_t]\,du\,dv - \dots \quad (2)$$

(cf. Rice (10), § 3·4). The formal solution (2) may be useful when the number of states is non-enumerable, for example, if $S$ refers to the value of a continuous spatial co-ordinate; it is of no interest without modification if the state $S$ is one of a finite or enumerable number of states, for $\{e_t\}$ would usually not then be of the form $[e_t]\,dt$, and recurrence on the above definition would almost certainly be instantaneous. In such cases we exclude instantaneous recurrence or continuation in state $S$, as we shall see in the next section.

3. *Renewal processes.* There is a great simplification in formulae (1) and (2) if the process is of the renewal or regenerative type (this includes Markoff processes, for which *every* state has the regenerative property), where the occurrence of the specified state $S$ is sufficient to regenerate the process, probabilities being no longer dependent on past history. For then, if the process is also homogeneous, we may write

$$\{e_re_n\} = \{e_r\}\{e_{n-r}\}, \quad \{e_re_se_n\} = \{e_r\}\{e_{s-r}\}\{e_{n-s}\},$$

etc., and taking generating functions on both sides, i.e. writing

$$\Pi_R(z) = \sum_{n=1}^{\infty} z^n\{\bar{e}_1\bar{e}_2\dots\bar{e}_{n-1}e_n\}, \quad \Pi(z) \equiv \sum_{n=1}^{\infty} z^n\{e_n\},$$

we obtain (for $0 \leqslant z < 1$),

$$\Pi_R = \Pi - \Pi^2 + \Pi^3 \dots = \Pi/(1+\Pi). \quad (3)$$

This formula, which is equivalent to the relation derived by Feller in his important paper on recurrence times for renewal processes ( (4); see also his book (5), ch. 12*), can be obtained more simply, but it is often convenient to make use of the above procedure, which will be used again in the case of first passage times. It might be emphasized that formula (3) determines either $\Pi_R$ in terms of $\Pi$ if $\Pi$ is given, or $\Pi$ in terms of $\Pi_R$ if $\Pi_R$ is given; this last situation arises when $S$ refers, say, to the renewal or replacement of an article with a 'life-time' distribution $\Pi_R$.

* For greater consistency with later formulae, $\Pi(z)$ in (3) differs from the generating function $U(z)$ defined by Feller in omitting his $u_0 \equiv 1$, so that $\Pi \equiv U - 1$.

https://doi.org/10.1017/S0305004100028358 Published online by Cambridge University Press

For continuous time, in problems for which a recurrence density in the sense of equation (2) exists, we replace generating functions by Laplace transforms in $t$, so that $z^n$ is replaced by $e^{-\phi t}$. Thus for

$$M \equiv \int_0^\infty e^{-\phi u}[e_u]\,du \tag{4}$$

and similarly for $M_R$, (3) is replaced by

$$M_R = M/(1+M), \tag{5}$$

with the reciprocal relation familiar in renewal theory

$$M = M_R/(1-M_R). \tag{6}$$

As a trivial example of (5), consider the purely random occurrence of an event or state $S$ at rate $\lambda$, so that

$$M = \int_0^\infty e^{-\phi u}\lambda\,du = \lambda/\phi. \tag{7}$$

Then $M_R = 1/[1+\phi/\lambda]$, which is well known as the transform of the exponential recurrence distribution for such random events.

In the case of continuous time, but a finite or enumerable number of states, we could proceed to modify our recurrence time in (3) by removing the term $\{e_1\} = p$, say, representing continuation in the state $S$, and measure the recurrence time from $t = 1$, for all processes for which we know the state at $t = 1$ is *not* $S$. The modified generating function is

$$\Pi'_R = \left\{\frac{\Pi}{1+\Pi} - zp\right\} \Big/ \{z(1-p)\}. \tag{8}$$

For continuous time, this will now usually give a definite limiting function if we write $z = e^{-\phi\tau}$ and find the limit $M'_R(\phi)$ of $\Pi'_R$ as $\tau \to 0$. However, it is convenient to use more direct arguments. The probability $P\{S_t\} \equiv \{e_t\} \equiv p_t$ is assumed non-zero, and we extend the renewal or regenerative argument to the sequence of events: $S$ from $t = 0$ to $u$, then departure from $S$ and first return at $v$, then occurrence of $S$ at $t$ (without conditions on the state during the interval $v, t$). The chance of leaving $S$ in time $dt$ is assumed to be $\lambda\,dt$, and the recurrence distribution for the time from departure to return is given by the *density* function $f(t)$. Then we have

$$p_t = \int_{v=0}^{t}\int_{u=0}^{v} \lambda e^{-\lambda u} f(v-u)\,p_{t-v}\,du\,dv + e^{-\lambda t},$$

whence, taking Laplace transforms, we obtain

$$G(\phi) = \frac{1}{(1+\phi/\lambda)}\,M'_R G + \frac{1}{\lambda}\frac{1}{(1+\phi/\lambda)},$$

whence finally

$$M'_R = 1 + \phi/\lambda - 1/(\lambda G). \tag{9}$$

4. *Passage times.* These formulae are readily extended to passage times from one state $S$ to a *different* state $T$. Let

$$P\{S_r \mid S_0\} \equiv \{e_r\}, \quad P\{T_r \mid T_0\} \equiv \{f_r\},$$
$$P\{S_r \mid T_0\} \equiv \{g_r\}, \quad P\{T_r \mid S_0\} \equiv \{h_r\},$$

where it is now important to specify explicitly whether $S$ or $T$ is given as the initial state. Consider

($a$) the first passage from $S$ to $T$, *irrespective of the recurrence of* $S$.

https://doi.org/10.1017/S0305004100028358 Published online by Cambridge University Press

We have

$$\{(1-h_1)(1-h_2)\dots(1-h_{n-1})h_n\} = \{h_n\} - \sum_{r=1}^{n-1}\{h_r h_n\} + \sum_{r=1}^{n-1}\sum_{s=r+1}^{n-1}\{h_r h_s h_n\} - \dots. \quad (10)$$

For homogeneous regenerative processes*, we may write (10) further as

$$\{h_n\} - \sum_{r=1}^{n-1}\{h_r\}\{f_{n-r}\} + \sum_{r=1}^{n-1}\sum_{s=r+1}^{n-1}\{h_r\}\{f_{s-r}\}\{f_{n-s}\} - \dots$$

or

$$\Pi_P^{ST} = \Pi^{ST}/(1+\Pi^{TT}), \quad (11)$$

where $\Pi_P^{ST}$ is the 'first passage' generating function for passage from $S$ to $T$, and

$$\Pi^{ST} \equiv \sum_{r=1}^{\infty}\{h_r\}z^r,$$

etc. If we alternatively require

(*b*) first passage to $T$, *without recurrence of* $S$, i.e. the passage time to $T$ measured from the *last* occurrence of $S$, then the formulae are somewhat more complicated. We must now consider

$$\{(1-h_1-e_1)(1-h_2-e_2)\dots(1-h_{n-1}-e_{n-1})h_n\}$$
$$= \{h_n\} - \sum_{r=1}^{n-1}\{(h_r+e_r)h_n\} + \sum_{r=1}^{n-1}\sum_{s=r+1}^{n-1}\{(h_r+e_r)(h_s+e_s)h_n\} - \dots. \quad (12)$$

Using such relations for homogeneous regenerative processes as

$$\{h_r e_n\} = \{h_r\}\{g_{n-r}\}, \quad \{h_r e_s e_n\} = \{h_r\}\{g_{s-r}\}\{e_{n-s}\},$$

we find that

$$\Pi_Q^{ST} = \frac{\Pi^{ST}}{(1+\Pi^{TT})(1+\Pi^{SS}) - \Pi^{ST}\Pi^{TS}}, \quad (13)$$

where $\Pi_Q^{ST}$ is the 'first passage' generating function *without intermediate recurrence of* $S$.

These formulae also extend immediately to the case of continuous time with densities $[e_t]$. Thus the formula analogous to (11) is

$$M_P^{ST}(\phi) = M^{ST}/(1+M^{TT}). \quad (14)$$

It is perhaps useful in the case of (13) to check the completely analogous formula by deriving it by an alternative argument. The recurrence of $S$ from 0 to $t$ may be split up into two contingencies: (i) recurrence of $S$ *conditional on no passage through* $T$, and (ii) first passage to $T$ *conditional on no recurrence of* $S$ followed by first *unconditional* passage from $T$ to $S$. This readily gives, on taking transforms,

$$M_R^S = C_R^S + M_Q^{ST}M_P^{TS}, \quad (15)$$

where $C_R^S$ is the conditional recurrence function for $S$ as above defined.

Similarly, the passage from $S$ to $T$ may consist either of passage conditional on no recurrence of $S$, or recurrence of $S$ conditional on no passage through $T$, followed by unconditional passage. Hence also

$$M_P^{ST} = C_R^S M_P^{ST} + M_Q^{ST}. \quad (16)$$

From these two equations we find the analogue of (13),

$$M_Q^{ST} = \frac{M^{ST}}{(1+M^{TT})(1+M^{SS}) - M^{ST}M^{TS}}, \quad (17)$$

* The regenerative property is of course now assumed for both $S$ and $T$.

https://doi.org/10.1017/S0305004100028358 Published online by Cambridge University Press

and also the further formula

$$C_R^S = \frac{M^{SS}(1+M^{TT}) - M^{ST}M^{TS}}{(1+M^{TT})(1+M^{SS}) - M^{ST}M^{TS}}. \tag{18}$$

Formula (18) also of course has its analogue for discrete time,

$$\frac{\Pi^{SS}(1+\Pi^{TT}) - \Pi^{ST}\Pi^{TS}}{(1+\Pi^{TT})(1+\Pi^{SS}) - \Pi^{ST}\Pi^{TS}}, \tag{19}$$

this being alternatively derivable by the method used for (13). It should be noticed that even if the process is ergodic and $M_R^S(0)$, $M_P^{ST}(0)$ and $M_P^{TS}(0)$ are all unity (cf. § 5), the equations (15) and (16) imply that $C_R^S(0)$ and $M_Q^{ST}(0)$ are less than unity. This merely means that *any* recurrence or passage may not satisfy the conditional requirement. To convert formulae like (13) or (18) into valid conditional distributions, we may always divide by the total probability sum $\Pi_Q^{ST}(1)$ or $C_R^S(0)$ respectively.

Corresponding formulae may also be obtained if the time is continuous but the number of states enumerable. For example, for first passage regardless of intermediate recurrence we obtain the relation for $p_{ST}(t) \equiv \{h_l\}$,

$$p_{ST}(t) = \int_{v=0}^{t} \int_{u=0}^{v} \lambda_S e^{-\lambda_S u} f_{ST}(v-u)\, p_{TT}(t-v)\, du\, dv,$$

from the sequence of events: $S$ from 0 to $u$, then (unconditional) first passage to $T$ at $v$, and also occurrence of $T$ at $t$ (without conditions on the state during the interval $v$, $t$). In this equation $p_{TT}(t) \equiv \{f_l\}$, $f_{ST}(t)$ is the 'first passage' density* from $S$ to $T$, and we now write for clarity $\lambda_S \equiv \lambda$. For the transforms we obtain

$$G^{ST}(\phi) = \frac{1}{(1+\phi/\lambda_S)}\, M_P^{ST}(\phi)\, G^{TT}(\phi), \tag{20}$$

or

$$M_P^{ST}(\phi) = (1+\phi/\lambda_S)\, G^{ST}(\phi)/G^{TT}(\phi). \tag{21}$$

For conditional recurrence and passage times we may employ a similar argument to that used for the relations (15) and (16). The only change is the insertion of a lifetime in $S$ after the conditional recurrence of $S$ and a lifetime in $T$ after the conditional passage to $T$. Hence (15) and (16) are replaced by the relations

$$M'_R = C'_R + M_Q^{ST}[1+\phi/\lambda_T]^{-1} M_P^{TS}, \tag{22}$$

$$M_P^{ST} = C'_R[1+\phi/\lambda_S]^{-1} M_P^{ST} + M_Q^{ST}, \tag{23}$$

where $C'_R$ denotes the transform of the conditional recurrence distribution. Solving these equations, we find

$$M_Q^{ST} = \frac{G^{ST}}{\lambda_S[G^{SS}G^{TT} - G^{ST}G^{TS}]}, \tag{24}$$

$$C'_R = 1 + \frac{\phi}{\lambda_S} - \frac{G^{TT}}{\lambda_S[G^{SS}G^{TT} - G^{ST}G^{TS}]}. \tag{25}$$

The consistency of formulae (9), (21), (24) and (25) was checked on the following elementary example of a Markoff process with the transition rates between its only two states $S$ and $T$ given by the matrix

$$\begin{pmatrix} -1 & 2 \\ 1 & -2 \end{pmatrix}.$$

* For formal simplicity it is written as a density, though in general the distribution will have a component at zero, corresponding to a direct transition to $T$ after leaving $S$.

https://doi.org/10.1017/S0305004100028358 Published online by Cambridge University Press

Here $\{e_t\} = \frac{2}{3} + \frac{1}{3}e^{-3t}$, $\{g_t\} = 2\{h_t\} = \frac{2}{3}(1-e^{-3t})$, $\{f_t\} = \frac{1}{3} + \frac{2}{3}e^{-3t}$,

$$G^{SS} = \frac{2+\phi}{\phi(3+\phi)}, \quad G^{TS} = 2G^{ST} = \frac{2}{\phi(3+\phi)}, \quad G^{TT} = \frac{1}{\phi(3+\phi)}.$$

We find on applying the above formulae

$$M'_R = (1+\tfrac{1}{2}\phi)^{-1}, \quad C'_R = 0, \quad M_P^{ST} = M_Q^{ST} = 1,$$

which are obviously correct, as the recurrence time for $S$ is identical in this example with lifetime in $T$, no recurrence is possible without traversing $T$, and conditional and unconditional passages from $S$ to $T$ are identical and take zero time as no other state is possible.

5. *Ergodic properties of renewal processes.* As noted by Feller*, $\Pi_R$ and $\Pi$ in (3) not only exist for $0 \leqslant z < 1$, but the former also must exist for $z = 1$, though it is not necessarily then unity; in fact, the event $S$ is only *certain* to recur if $\Pi_R(1) = 1$. From (3), $\Pi_R(1) < 1$ implies that (and is implied by) $\Pi(1) < \infty$, which is thus an equivalent condition for $S$ being uncertain (sometimes then called a transient state). When $S$ is certain and not periodic then a theorem due to Erdös, Feller and Pollard (see, for example, (5), p. 244) establishes the ergodic formula

$$\lim_{r\to\infty} \{e_r\} = 1/\mu, \tag{26}$$

where $\mu$ is the mean $[\partial\Pi_R/\partial z]_{z=1}$. This formula remains true if $\mu$ is infinite, when $\{e_r\} \to 0$ as $r \to \infty$.

It is worth remarking that Feller's arguments may be extended to the other cases discussed above. In the case of continuous time with a probability density $[e_t]$, we deduce from (5) that

$$\lim_{\phi\to 0} M_R(\phi) = 1 \tag{27}$$

if $M(\phi) \to \infty$ as $\phi \to 0$, and $S$ is then certain to recur. The relevant theorem for such conditions (which automatically exclude the periodic case) is now one given by Blackwell (2), namely, that if the expected number of recurrences or renewals in the interval $(0, t)$ is $U(t)$, and $\mu$ is the mean recurrence or renewal interval, then

$$\lim_{t\to\infty} [U(t+h) - U(t)] = h/\mu \tag{28}$$

for every $h > 0$. This formula also remains true if $\mu$ is infinite. In our present case, $[e_t] \equiv dU(t)/dt$.

In the case of continuous time but an enumerable number of states, so that formula (9) holds, we note that $G(0) = \infty$ still implies $M_R(0) = 1$, and, alternatively, $G(0) = \infty$ implies $M_R(0) < 1$ (and conversely). To obtain the formula corresponding to (16) or (18), apply Blackwell's theorem to a 'renewal' defined for the immediate purpose as the complete lifetime plus recurrence of $S$. As contributions to $\{e_t\}$ can arise from renewals of $S$ in this sense occurring at a previous time $t-u$ and surviving in state $S$ to time $t$, we have the relation

$$\{e_t\} = \int_0^t e^{-\lambda_S(t-v)}\, dU(u). \tag{29}$$

* It is considered useful to summarize Feller's arguments here, so that their extension to other important cases can be the more readily indicated.

https://doi.org/10.1017/S0305004100028358 Published online by Cambridge University Press

$U(u)$ is now the 'convolution' of the lifetime and recurrence distributions, each with densities, so that $U(t)$ is still differentiable. It follows that

$$\lim_{t\to\infty}\{e_t\} = \frac{1}{\lambda_S \mu},$$

where $\mu$ is the sum of the lifetime mean $1/\lambda_S$ and the true recurrence mean $\mu_R$, i.e. $\mu = \mu_R + 1/\lambda_S$ and

$$\lim_{t\to\infty}\{e_t\} = \frac{1}{1+\lambda_S \mu_R}. \tag{30}$$

6. *Ergodic properties of Markoff processes**. With Markoff processes *every* state has the regenerative or renewal property. We consider only 'indecomposable' Markoff processes, for which every state can be reached with finite probability from any other state after some interval; processes with 'absorbing' states are thus excluded. Then if, following Feller, we consider Markoff chains with discrete time and an enumerable number of states, and assume that the state $S_i$ is certainly recurrent, we may choose intervals of minimum sufficient length to allow passage from state $S_i$ to another state $S_j$ and back to occur. Denote† $P\{S_j \text{ at } t \mid S_i \text{ at } 0\}$ by $p_{ij}(t)$, etc. Then

$$\left.\begin{aligned} p_{ii}(t+\tau_1+\tau_2) &\geqslant p_{ij}(\tau_1)\, p_{jj}(t)\, p_{ji}(\tau_2) = \alpha\beta p_{jj}(t), \\ p_{jj}(t+\tau_1+\tau_2) &\geqslant p_{ji}(\tau_2)\, p_{ii}(t)\, p_{ij}(\tau_1) = \alpha\beta p_{ii}(t), \end{aligned}\right\} \tag{31}$$

where $\alpha, \beta > 0$. The relations (32) imply that $p_{ii}(t)$ and $p_{ij}(t)$ have the same type of limiting behaviour, corresponding to the three alternatives: $S_j$ also certainly recurrent, with (i) $p_{ii}(t)$ and $p_{jj}(t)$ both tending to zero, (ii) $p_{ii}(t)$ and $p_{jj}(t)$ both tending to finite limits, (iii) $S_i$ and $S_j$ both periodic, with the same period.

Moreover, from (11), with $S \equiv S_i$, $T \equiv S_j$, since $\Pi_P^{ST}(1) \leqslant 1$, we must have $\Pi^{ST}(1) < \infty$ if $\Pi^{TT}(1) < \infty$; this is the case of uncertain recurrence. Alternatively, for certain recurrence we must have $\Pi_P^{ST}(1) = 1$, and hence $\Pi^{ST}(1)/\Pi^{TT}(1) = 1$. This implies, in the aperiodic case when $p_{jj}(t) \to 1/\mu_j$, that $p_{ij}(t)$ also tends to $1/\mu_j$. This is still true if $\mu_j = \infty$.

For a finite or enumerable number of states but continuous time, a similar argument makes use of (9), (21) and (30). The periodic case does not arise; and the intervals $\tau_1$ and $\tau_2$ in (31) may now be any non-zero intervals.

For continuous time and a continuous infinity of states where we assume densities $f_{ii}(t) \equiv [e_t]$, $f_{ij}(t) \equiv [h_t]$, etc. (excluding periodic processes), we may for certainly recurrent $S_i$ replace (31) by the relations

$$\left.\begin{aligned} \int_t^{t+h} f_{ii}(\tau+\tau_1+\tau_2+h)\, d\tau &\geqslant \iiint f_{ij}(\tau_1+u) f_{jj}(\tau+h-u-v) f_{ji}(\tau_2+v)\, du\, dv\, d\tau, \\ \int_t^{t+h} f_{jj}(\tau+\tau_1+\tau_2+h)\, d\tau &\geqslant \iiint f_{ji}(\tau_2+v) f_{ii}(\tau+h-u-v) f_{ij}(\tau_1+u)\, du\, dv\, d\tau, \end{aligned}\right\} \tag{32}$$

where the integration is over the interval $(0, h)$ for $u$ and $v$. These relations ensure the same type of limiting behaviour for $\int_t^{t+h} f_{ii}(v)\, du$ and $\int_t^{t+h} f_{jj}(v)\, du$. The formula

* For more general theoretical classification and discussion of Markoff processes, see Lévy (9), and bibliography at the end of his paper.

† The notation in this section is no longer identical with that in § 4.

https://doi.org/10.1017/S0305004100028358 Published online by Cambridge University Press

(14) ensures that, since $M_P^{ST}(0) \leqslant 1$, $M^{TT}(0) < \infty$ implies $M^{ST}(0) < \infty$. Alternatively, if $M^{TT}(0) = \infty$, so that recurrence is certain, we must have $M_P^{ST}(0) = 1$, and

$$\int_t^{t+h} f_{ij}(u)\,du \quad \text{and} \quad \int_t^{t+h} f_{ii}(u)\,du$$

have the same limit $h/\mu_j$ (which is zero if $\mu_j$ is infinite).

7. *Smoluchowski's formulae for mean recurrence times.* We shall now relinquish the assumption of 'regeneration', but assume the conditions of stationarity and ergodicity referred to in the introductory remarks (we are now no longer interested in the possibility of transient phenomena, and assume that the state $S$ is certain to recur, and after a duration of finite mean length); we may thus go back to Smoluchowski's original arguments. It will be convenient to refer to the number of separate times the state is occupied in a 'long' interval of time $T$, on the understanding that all the formulae derived will involve only ratios of such numbers, and hold in the limit as $T \to \infty$. In the case of an enumerable number of states and discrete time (or intermittent observation) denote the frequency of the state $S$ being occupied continuously for an interval length $k$ by $N_k$. Similarly, we denote the frequency for the *composite* state 'not $S$' by $M_k$.

Then we must have

$$P(S) \equiv \{e_\infty\} \equiv P = \frac{N_1 + 2N_2 + 3N_3 + \ldots}{N_1 + 2N_2 + 3N_3 + \ldots + M_1 + 2M_2 + 3M_3 + \ldots}, \tag{33}$$

$$P(S \mid S) \equiv \{e_1\} \equiv p = \frac{N_2 + 2N_3 + 3N_4 + \ldots}{N_1 + 2N_2 + 3N_3 + \ldots}, \tag{34}$$

$$1 = \frac{N_1 + N_2 + N_3 + \ldots}{M_1 + M_2 + M_3 + \ldots}, \tag{35}$$

where (35) follows from the equality in frequency of excursions from and to state $S$. We define the *lifetime* $t_1$ in state $S$, consistently with our previous usage in § 5, as the total continuous interval in that state. The mean lifetime

$$T_1 \equiv \mathscr{E}\{t_1\} = \frac{\tau(N_1 + 2N_2 + \ldots)}{N_1 + N_2 + \ldots} = \frac{\tau}{1-p}, \tag{36}$$

where $\tau$ is the time-interval between consecutive time-instants. Similarly, if the *recurrence time* $\theta_1$ is the total continuous interval out of state $S$, then*

$$\Theta_1 \equiv \mathscr{E}\{\theta_1\} = \frac{\tau(M_1 + 2M_2 + \ldots)}{M_1 + M_2 + \ldots} = T_1 \frac{1-P}{P}. \tag{37}$$

To obtain the equivalent formulae as the time becomes strictly continuous (i.e. $\tau \to 0$), we assume that $1-p$ is of the form $\lambda\tau + O(\tau)$, so that (36) becomes in the limit

$$T_1 = 1/\lambda. \tag{38}$$

Formula (37) still holds; so that

$$\Theta_1 = \frac{1}{\lambda}\frac{1-P}{P}. \tag{39}$$

In the particular case when $S$ becomes a rare state, we may write $1 - P \sim 1$ in (39).

* This definition corresponds to that used in formula (8), *not* (3); cf. also (8).

https://doi.org/10.1017/S0305004100028358 Published online by Cambridge University Press

We shall at this stage introduce two further definitions due to Smoluchowski: (i) the *further* lifetime $t_2$ of state $S$, given $S$ at some arbitrary instant; (ii) the further recurrence time $\theta_2$ of state $S$, given 'not $S$' at an arbitrary instant. Then by definition, for discrete time,

$$T_2 \equiv \mathscr{E}\{t_2\} = \frac{\tau[N_1 + (1+2)\,N_2 + (1+2+3)\,N_3 + \ldots]}{N_1 + 2N_2 + 3N_3 + \ldots}, \tag{40}$$

$$\Theta_2 \equiv \mathscr{E}\{\theta_2\} = \frac{\tau[M_1 + (1+2)\,M_2 + (1+2+3)\,M_3 + \ldots]}{M_1 + 2M_2 + 3M_3 + \ldots}. \tag{41}$$

However, we may define also the complete generating functions for $t_1$ and $t_2$ (putting for convenience $\tau = 1$)

$$\Pi_L^{(1)} = \frac{zN_1 + z^2N_2 + \ldots}{N_1 + N_2 + \ldots},$$

$$\Pi_L^{(2)} = \frac{zN_1 + (z+z^2)\,N_2 + \ldots}{N_1 + 2N_2 + \ldots}$$

$$= \frac{z}{1-z}\left[\frac{N_1 + N_2 + \ldots}{N_1 + 2N_2 + \ldots}\right] - \frac{z}{1-z}\left[\frac{zN_1 + z^2N_2 + \ldots}{N_1 + 2N_2 + \ldots}\right]$$

$$= \frac{z}{1-z}\left[\frac{1}{T_1} - \frac{\Pi_L^{(1)}}{T_1}\right],$$

or

$$T_1(1-z)\,\Pi_L^{(2)} = z(1 - \Pi_L^{(1)}). \tag{42}$$

A similar relation obviously holds between $\Pi_R^{(1)}$ and $\Pi_R^{(2)}$ for the recurrence times $\theta_1$ and $\theta_2$, i.e.

$$\Theta_1(1-z)\,\Pi_R^{(2)} = z(1 - \Pi_R^{(1)}). \tag{43}$$

In the case of a regenerative process, the transitions of the process, *given $S$ at any instant*, are independent of the state at previous instants, so that the variables $t_1$ and $t_2$ become *stochastically* equivalent. This immediately yields from (42), putting $\Pi_L^{(1)} = \Pi_L^{(2)}$,

$$\Pi_L^{(1)} = \frac{z}{z + (1-z)\,T_1}, \tag{44}$$

a well-known result. It must be stressed that the same result does *not* hold for $\theta_1$, since $\theta_1$ and $\theta_2$ are not in general stochastically equivalent even for regenerative or Markoff processes. A Markoff process does not retain its special character under grouping of its states, so that 'not $S$' cannot be treated as a single state of a Markoff process like $S$, except in the obvious and trivial exception of only two states in all (cf. Bartlett (1), where a slip on this point by Chandrasekhar (3) was noted).

When the interval $\tau$ (taken for convenience as unity in (42), (43) and (44)) tends to zero, (42) and (43) become

$$T_1\phi M_L^{(2)}(\phi) = 1 - M_L^{(1)}(\phi), \tag{45}$$

$$\Theta_1\phi M_R^{(2)}(\phi) = 1 - M_R^{(1)}(\phi), \tag{46}$$

and (44) becomes the exponential distribution

$$M_L^{(1)}(\phi) = 1/(1 + \phi T_1). \tag{47}$$

It is of interest to observe that the converse to (47) is also true, and the relation (45) can lead to the exponential distribution for $M_L^{(1)}(\phi)$ *only* if $t_1$ and $t_2$ are stochastically equivalent. For let $M_L^{(1)}(\phi) = 1/(1 + \phi T_1)$; then $M_L^{(2)}(\phi)$ is also $1/(1 + \phi T_1)$.



https://doi.org/10.1017/S0305004100028358 Published online by Cambridge University Press

8. *Recurrence distributions for finite Markoff chains.* While we have seen that the recurrence time distribution for a Markoff chain may be studied by the 'renewal' arguments introduced by Feller for this purpose, an alternative method which is sometimes convenient for direct calculation with Markoff chains with a finite, or effectively finite, number of states is to modify the probability transition matrix for the Markoff chain by prohibiting all transitions *out* from the specified state $S$. This automatically freezes the state $S$ once it recurs, and the probability of any lifetime in 'not $S$' is identified with the probability of still being in 'not $S$' for the modified process. Thus for discrete time and a matrix $\mathbf{Q}$ of transition probabilities, the modified matrix is, say,

$$\mathbf{Q}_m = \left(\begin{array}{c|c} 1 & \mathbf{x}_1' \\ \hline \mathbf{0} & \mathbf{B} \end{array}\right),$$

where for convenience the top left-hand corner corresponds to the specified state $S$. The probability column vector $\mathbf{p}_0$, given 'not $S$' at time zero (and $S$ at time $-1$ if we are dealing with $\theta_1$ rather than $\theta_2$), will have the form $\mathbf{p}_0' = (0 \,|\, \mathbf{y}')$ as a row vector (dashes with matrices and vectors denote transposes). Hence

$$\mathbf{Q}_m^n \mathbf{p}_0 = \left(\begin{array}{c|c} 1 & \mathbf{x}_n' \\ \hline \mathbf{0} & \mathbf{B}^n \end{array}\right)\left(\begin{array}{c} 0 \\ \hline \mathbf{y} \end{array}\right) = \left(\begin{array}{c} \mathbf{x}_n'\mathbf{y} \\ \hline \mathbf{B}^n\mathbf{y} \end{array}\right),$$

and the cumulative distribution is $1-[\mathbf{B}^n\mathbf{y}]$, where $[\mathbf{y}]$ denotes (in the present context) the *sum* of the elements of the vector $\mathbf{y}$. The generating function is thus given by

$$\Pi_R = \sum_{n=1}^{\infty} [(\mathbf{B}^{n-1} - \mathbf{B}^n)\,\mathbf{y}]\, z^n = \left[\frac{(1-\mathbf{B})z}{1-\mathbf{B}z}\,\mathbf{y}\right], \tag{48}$$

where $\Pi_R$ is $\Pi_R^{(1)}$ corresponding to $\theta_1$, or $\Pi_R^{(2)}$ corresponding to $\theta_2$, according to the choice of $\mathbf{y}$ ($\Pi_R$ of § 3 is identical with $\Pi_R^{(1)}$). In the case of only two states, let $\mathbf{B} = 1-\lambda$, and, with $\mathbf{y} = 1$, we confirm the trivial solution

$$\Pi_R^{(1)} = \Pi_R^{(2)} = \lambda z/\{1-(1-\lambda)z\}.$$

Similarly for continuous time, let the transition matrix be of the form

$$\mathbf{Q} = \mathbf{I} + \mathbf{R}\delta t + o(\delta t),$$

where $\mathbf{I}$ is the unit matrix. Let $\mathbf{R}_m = \left(\begin{array}{c|c} 0 & \mathbf{x}' \\ \hline \mathbf{0} & \mathbf{C} \end{array}\right)$.

The cumulative recurrence distribution becomes $1-[e^{\mathbf{C}t}\mathbf{y}]$, and its generating function

$$M_R(\phi) = \int_0^{\infty} e^{-\phi u}[-\mathbf{C}\,e^{\mathbf{C}u}\mathbf{y}]\,du = \left[\frac{\mathbf{C}}{\mathbf{C}-\phi}\,\mathbf{y}\right]. \tag{49}$$

Again, for only two states, and $\mathbf{C} = -\lambda$,

$$M_R^{(1)}(\phi) = M_R^{(2)}(\phi) = 1/(1+\phi/\lambda).$$

In order that the distribution $M_R(\phi)$ reduces to the exponential distribution, we saw that $\theta_1$ and $\theta_2$ must be stochastically equivalent. Denote the original matrix $\mathbf{R}$ by

$$\mathbf{R} = \left(\begin{array}{c|c} a & \mathbf{x}' \\ \hline \mathbf{b} & \mathbf{C} \end{array}\right),$$

https://doi.org/10.1017/S0305004100028358 Published online by Cambridge University Press

where it should be noted that the vector $\mathbf{y}$ in (49) will for $\theta_1$ be proportional to $\mathbf{b}$, and for $\theta_2$ be equivalent to the asymptotic distribution of the individual 'not $S$' states, given by the column latent vector of the zero latent root of $\mathbf{R}$ (in view of our ergodic assumptions, there is only one zero root of $\mathbf{R}$). Hence the required condition is

$$\left(\begin{array}{c|c} a & \mathbf{x}' \\ \hline \mathbf{b} & \mathbf{C} \end{array}\right)\left(\begin{array}{c} \mu \\ \hline \mathbf{b} \end{array}\right) = \left(\begin{array}{c} a\mu + \mathbf{x}'\mathbf{b} \\ \hline \mu\mathbf{b} + \mathbf{C}\mathbf{b} \end{array}\right) = \left(\begin{array}{c} 0 \\ \hline \mathbf{0} \end{array}\right),$$

whence $\mathbf{b}$ must be a latent vector of $\mathbf{C}$ with corresponding latent root $-\mu = \mathbf{x}'\mathbf{b}/a$. A simple example is

$$\mathbf{R} = \begin{pmatrix} -2 & 2 & 4 \\ 1 & -4 & 1 \\ 1 & 2 & -5 \end{pmatrix},$$

for which $M_R(\phi) = 1/(1+\frac{1}{3}\phi)$. The more general type of recurrence time distribution for finite Markoff chains will obviously be a sum of exponential distributions if $\mathbf{C}$ has simple roots; if multiple roots exist, the expression $[e^{\mathbf{C}t}\mathbf{y}]$ will involve components of the form $Q(t)\,e^{-\nu t}$, where $Q(t)$ is a polynomial in $t$ (the distribution is thus expressible as a sum of $\chi^2$ distributions).

9. *Mean recurrence times in the case of a continuous infinity of states.* We saw with renewal processes that cases where the finite probability of $S$ must be replaced by a density function are best handled separately, and this applies also with Smoluchowski's methods for obtaining mean values, as the concept of lifetime in a state breaks down. In the case of a continuous variable with continuous range, it is usual to consider the rate of change $\dot{X}$ of $X$, assuming this to exist in any appropriate well-defined sense, so that the simultaneous probability distribution $f(\xi,\eta)\,d\xi\,d\eta$ of $X$ and $\dot{X}$ exists. As the process is assumed stationary, $f$ will be independent of $t$. For the density at $X = x$ *in time*, we then readily find, at least formally,

$$\int_{-\infty}^{\infty} |\,\eta\,|\, f(x,\eta)\,d\eta, \tag{50}$$

and a corresponding mean recurrence time $\Theta_1$ given by its reciprocal (see, for example, Rice **(10)**, § 3·3, where precise conditions are quoted under which (50) holds).

For example, for a normal process, as $X$ and $\dot{X}$ are necessarily uncorrelated in stationary processes, we have

$$f(\xi,\eta) = \frac{1}{2\pi\sigma_1\sigma_2}\exp\left\{-\frac{1}{2}\left(\frac{\xi^2}{\sigma_1^2}+\frac{\eta^2}{\sigma_2^2}\right)\right\},$$

where, as we may always represent a normal stationary process by

$$X(t) = \int_{-\infty}^{\infty} e^{it\omega}\,dZ(\omega),$$

where $Z(\omega)$ is an additive process,

$$\sigma_1^2 = \int_{-\infty}^{\infty} d\mathscr{E}\{Z(\omega)\,Z^*(\omega)\} = \int_{-\infty}^{\infty} dF(\omega),$$

say, and $\sigma_2^2 = \int_{-\infty}^{\infty} \omega^2 dF(\omega)$ when $\dot{X}$ exists in the mean-square sense. This readily gives

$$\Theta_1 = \frac{\pi}{\sigma_\omega}\,e^{\frac{1}{2}x^2/\sigma_1^2}, \tag{51}$$



https://doi.org/10.1017/S0305004100028358 Published online by Cambridge University Press

where $\sigma_\omega$ is the standard deviation of the spectral distribution of $X$. The above method breaks down for a normal Markoff process, for which $\dot{X}$ does not exist and $\sigma_\omega$ is infinite, but for Markoff processes we have seen that other and more complete methods are available*.

As an extension of (51), the more general form for $f(\xi, \eta)$ may be assumed:

$$\left\{\exp \sum_{r+s>2}\left(-\frac{\partial}{\partial\xi}\right)^r\left(-\frac{\partial}{\partial\eta}\right)^s \frac{\kappa_{rs}}{r!\,s!}\right\}\left\{\frac{1}{2\pi\sigma_1\sigma_2}\exp\left[-\frac{1}{2}\left(\frac{\xi^2}{\sigma_1^2}+\frac{\eta^2}{\sigma_2^2}\right)\right]\right\},$$

where for a stationary process $\kappa_{21} = \mathscr{E}\{X^2(t)\,\dot{X}(t)\} = 0$, $\kappa_{31} = \mathscr{E}\{X^3(t)\,\dot{X}(t)\} = 0$, etc. We find by straightforward integration in the case of near-normal processes, retaining terms up to the fourth order, an average density of crossing $X = x$ given by

$$\frac{\sigma_\omega}{\pi}e^{-\frac{1}{2}x^2/\sigma_1^2}\left\{1+\frac{\gamma_{30}}{6}\left(\frac{x^3}{\sigma_1^3}-3\frac{x}{\sigma_1}\right)+\frac{\gamma_{40}}{24}\left(\frac{x^4}{\sigma_1^4}-\frac{6x^2}{\sigma_1^2}+3\right)+\tfrac{1}{2}\gamma_{12}\frac{x}{\sigma_1}+\tfrac{1}{4}\gamma_{22}\left(\frac{x^2}{\sigma_1^2}-1\right)-\frac{\gamma_{04}}{24}\right\}, \quad (52)$$

where
$$\gamma_{30} = \mathscr{E}\{X^3\}/\sigma_1^3, \quad \gamma_{40} = \mathscr{E}\{X^4\}/\sigma_1^4 - 3, \quad \gamma_{12} = \mathscr{E}\{X\dot{X}^2\}/(\sigma_1\sigma_2^2),$$
$$\gamma_{22} = \mathscr{E}\{X^2\dot{X}^2\}/(\sigma_1^2\sigma_2^2) - 1, \quad \gamma_{04} = \mathscr{E}\{\dot{X}^4\}/\sigma_2^4 - 3.$$

In particular, if the marginal distribution of $X$ is normal (as appears true, for example, if $X$ denotes turbulent velocity in a fluid), $\gamma_{30} = \gamma_{40} = 0$, and (52) reduces to

$$\frac{\sigma_\omega}{\pi}e^{-\frac{1}{2}x^2/\sigma_1^2}\left\{1+\tfrac{1}{2}\gamma_{12}\frac{x}{\sigma_1}+\tfrac{1}{4}\gamma_{22}\left(\frac{x^2}{\sigma_1^2}-1\right)-\frac{\gamma_{04}}{24}\right\}. \quad (53)$$

The practical problem of determining complete recurrence distributions, as distinct from average recurrence times, still remains in the case of non-regenerative processes, the general formulae (1) and (2), while valid in principle, not being very convenient. We have also not considered, even in the case of Markoff processes, some other types that may arise in practice (we are not concerned with pathological cases). If there is a continuum of states classified by a single quantitative variable $x$, the process may contain quiescent lifetimes in a particular enumerable set of states before proceeding 'smoothly' until it reaches other such states. This is a mixture of cases B and C (see the appendix for classification), but it would still be quite feasible to separate off the finite lifetimes as in case B (in fact, if we are only interested in such 'end-states' the theory for case B still applies). Another possibility perhaps more liable to arise is the case of discontinuous transitions between *any* two of a continuum of states each of which has a finite lifetime. The notion of recurrence in the strict sense is now excluded, for not all states can be visited under such conditions; it has to be replaced by a weaker definition before any non-trivial results can be obtained.

* *Note added in proof.* The normal Markoff (including the normal additive) process is not strictly classifiable under case *C* of the Appendix, having an infinite density in time owing to what might be termed its 'jitter' properties. The earlier formulae may, however, fairly readily be adapted, e.g. (14) reduces to the ratio $M^{ST}/M^{TT}$, which is then independent of the scale conversion factor from space to time densities. Alternatively one may proceed from the diffusion equation (for which, see (3)).

https://doi.org/10.1017/S0305004100028358 Published online by Cambridge University Press

## APPENDIX

(*Index to formulae*)

| | Case A (Discrete time; non-zero probability for state) | Case B (Continuous time; non-zero probability for state) | Case C (Continuous time; finite density for state) |
|---|---|---|---|
| | Distribution transforms for regenerative processes | | |
| Lifetime | (44) | (47) | 1 |
| Recurrence | (3) (or (8)*) | (9)* | (5) |
| Conditional recurrence | (19) | (25)* | (18) |
| First passage | (11) | (21)* | (14) |
| Conditional first passage | (13) | (24)* | (17) |
| | Alternative formulae for finite Markoff chains | | |
| Recurrence | (48)* | (49)* | — |
| | Mean values for stationary ergodic processes | | |
| Lifetime | (36) | (38) | 0 |
| Recurrence | (37)* | (39)* | (50) particular case |

* Excluding continuation in same state.

## REFERENCES


(1) Bartlett, M. S. Recurrence times. *Nature, Lond.*, 165 (1950), 727.

(2) Blackwell, D. A renewal theorem. *Duke math. J.* 15 (1948), 145.

(3) Chandrasekhar, S. Stochastic problems in physics and astronomy. *Rev. mod. Phys.* 15 (1943), 1.

(4) Feller, W. Fluctuation theory of recurrent events. *Trans. Amer. math. Soc.* 67 (1949), 98.

(5) Feller, W. *An introduction to probability theory and its applications*, vol. 1 (New York, 1950).

(6) Harris, T. E. First passage and recurrence distributions. *Trans. Amer. math. Soc.* 73 (1952), 471.

(7) Kac, M. On the distribution of values of trigonometric sums with linearly independent frequencies. *Amer. J. Math.* 65 (1943), 609.

(8) Kac, M. On the notion of recurrence in discrete stochastic processes. *Bull. Amer. math. Soc.* 53 (1947), 1002.

(9) Lévy, P. Systèmes markoviens et stationnaires: cas dénombrable. *Ann. sci. Éc. norm. sup., Paris*, (3), 68 (1951), 327.

(10) Rice, S. O. Mathematical analysis of random noise. *Bell. Syst. tech. J.* 23 (1944), 282 and 24 (1945), 46.

(11) Smoluchowski, M. v. Molekulartheoretische Studien über Umkehr thermodynamisch irreversibler Vorgänge und über Wiederkehr abnormaler Zustände. *S.B. Akad. Wiss. Wien* (2*a*), 124 (1915), 339.


University of Manchester

https://doi.org/10.1017/S0305004100028358 Published online by Cambridge University Press