Dual Diffusions, Killed Diffusions, and the Age Distribution Problem in Population Genetics

S. TAVARÉ

Department of Mathematics, University of Utah, Salt Lake City, Utah 84112

A method is given for justifying some recent approaches to the definition of the age of an allele in two-allele diffusion-type population genetic models. The method uses the notion of the dual of a diffusion, and introduces naturally an interesting application of killed diffusions. Many results in the literature are shown to be special cases of this approach.

INTRODUCTION

Recently, much attention has been paid to developing methods of determining reasonable definitions for the age of an allele, given a current allele frequency, in a wide variety of population genetics models. The pioneering work in this field is due to Kimura and Ohta (1973), who derived moments of what was then taken to be the age of an allele in diffusion-type models with absorbing, or exit, boundaries. In a more general context, Levikson (1977) described a method of defining the age distribution for Markov chains with absorbing barriers by introducing return processes, which jumped to the nearest point in the interior of the state space whenever an absorbing barrier was hit. The age distribution is then defined to be the (limiting) distribution of the time that has elapsed since one of the boundaries was last visited, conditional on some current position. Levikson assumed that the return processes were positive recurrent. A useful reference to the ideas involved is provided by Watterson (1977). The results for Markov chains have been extended by Pakes (1977, 1978) in cases where the return process need not be positive recurrent. Levikson postulated how this method could be applied to diffusion processes, unifying some results of Li (1975), Kimura and Ohta (1973), and Maruyama (1974). Watterson (1977) showed how to extend the diffusion results by a reversibility-type argument, and in particular showed how the (past) age of an allele and the (future) absorption time of an allele are related in a special case of the return process considered by Levikson. He also showed how reversibility in this special case could lead to a simple solution to the 'Which allele is the oldest?' problem in the context of two allele models. Tavaré (1978a, b) showed that for recurrent processes, the age distribution of a Markov chain is just the extinction time distribution

of a related dual chain (see the next section for some details about duals). The idea is to compute the dual, or time reversed, chain generated by the stationary distribution, and kill the dual whenever one of the absorbing barriers is reached. The method of exhibiting the age in terms of a well-specified Markov chain was shown to lead to simple ways of deriving a wide variety of properties of age distributions. The restarting distributions are arbitrary, save only that the return process is positive recurrent. In most cases, the dual chain is not stochastically identical to the return process; that is, the return process is not symmetrically reversible. It follows that the (past) age distribution and (future) absorption time distribution need not be identical. We will show how certain dual diffusions lead simply to a variety of results for the age distribution of a diffusion process which is killed, or stopped, at a killing time. The age distribution is identical to the killing time distribution.

We consider in this paper diffusion processes which have two absorbing boundaries. These are assumed to correspond to two-allele genetic models, in which there are no mutation pressures. The details for models with one absorbing barrier, corresponding for example to genetic models with one-way mutation, are similar, and will not be presented here. Special cases of the results here may be found in Watterson (1977), Levikson (1977), Maruyama (1974, 1977), Kimura and Ohta (1973), Maruyama and Kimura (1975), and Narain (1978). Sawyer (1977) has rather a different approach to the age problem in the case of an infinite-allele genetic model.

1. MARKOV CHAIN METHODS AND DUALITY

Let $X = \{X_n, n \geqslant 0\}$ be a Markov chain with state space $S = \{0, 1,..., M\}$, and absorbing boundaries at 0 and M. Let $B = \{1, 2,..., M - 1\}$ be the set of transient states, and denote the transition matrix by P. We can write P in the form

$$P = \begin{pmatrix} 1 & 0 & 0 \\ \mathbf{p}_0' & Q & \mathbf{p}_M' \\ 0 & 0 & 1 \end{pmatrix}. \tag{1}$$

$\mathbf{p}_0'$ and $\mathbf{p}_M'$ are $(M - 1) \times 1$ vectors which given the one-step probabilities of entering the absorbing states. Denote the n-step transition probabilities by $(p_{ij}^{(n)})$. The return process $\bar{X} = \{\bar{X}_n, n \geqslant 0\}$ is defined by specifying how the X process is restarted whenever an absorbing boundary is reached. We will only consider the case where $\bar{X}$ jumps to a if 0 is hit, or to b if M is hit, where $a \in B$ and $b \in B$. We can then write the transition matrix $\bar{P}$ of $\bar{X}$ in the form

$$\bar{P} = \begin{pmatrix} 0 & \mathbf{e}_a & 0 \\ \mathbf{p}_0' & Q & \mathbf{p}_M' \\ 0 & \mathbf{e}_b & 0 \end{pmatrix}, \tag{2}$$

where e_i is a vector with elements equal to 0, except in the ith position, which is 1. Assuming that X is aperiodic and irreducible, X has a stationary and limiting distribution with positive elements denoted by $(\alpha_0, \alpha_1, ..., \alpha_M)$. We associate with X the dual, or time reversed process $\bar{Y}$, whose transition matrix $\bar{P}^* = (\bar{p}_{ij}^*)$ is obtained from $\bar{P}$ by the formula

$$\bar{p}_{ij}^* = \frac{\bar{p}_{ji}\alpha_j}{\alpha_i}. \tag{3}$$

Associated with $\bar{Y}$ is the dual absorbing chain Y which is obtained from $\bar{Y}$ by killing $\bar{Y}$ whenever 0 or M is reached. From (3) we see that the resulting absorbing chain has transition matrix P^* given by

$$P^* = \begin{pmatrix} 1 & 0 & 0 \\ \alpha_0 D^{-1}e_a' & D^{-1}Q'D & \alpha_M D^{-1}e_b' \\ 0 & 0 & 1 \end{pmatrix}, \tag{4}$$

where D is a diagonal matrix with elements $(\alpha_1, \alpha_2, ..., \alpha_{M-1})$.

The limiting age G_j of X, given current position $j \in B$ has distribution determined by

$$P(G_j = n) = \lim_{m \to \infty} P(X_{m-n} \in \{0, M\}, X_{m-k} \in B, 0 < k < n \mid X_m = j).$$

It is shown in Tavaré (1978a) that this distribution is the same as the absorption time distribution of Y given $Y_0 = j$, and is given by

$$P(G_j = n) = \frac{\pi_0(b)\, p_{aj}^{(n-1)} + \pi_M(a)\, p_{bj}^{(n-1)}}{\pi_0(b)\, n(a, j) + \pi_M(a)\, n(b, j)}, \qquad n \geqslant 1, \tag{5}$$

where $\pi_0(i) = 1 - \pi_M(i) = P(X$ absorbed at $0 \mid X_0 = i)$, and $n(i, j) = \sum_{n=0}^{\infty} p_{ij}^{(n)}$ is the mean sojourn time at j, given $X_0 = i$. Further, it was shown that $\alpha_j \propto \pi_0(b)\, n(a, j) + \pi_M(a)\, n(b, j), j \in B$.

As the X process evolves, we will observe occasional jumps from 0 to a or from M to b, while we observe occasional jumps from a to 0 or from b to M as the $\bar{Y}$ process evolves. The Y process coincides with $\bar{Y}$ until the first visit to 0 or M, when the Y process stops. We have shown above how the age distribution can be found from properties of the Y process. We remark that if the X process is stationary, so that $P(X_0 = j) = \alpha_j$ for $0 \leqslant j \leqslant M$, then the age distribution specified by $P(X_{m-n} \in \{0, M\}, X_{m-k} \in B, 0 < k < n \mid X_m = j)$ is independent of m, and is again given by (5).

2. DIFFUSION PROCESSES

We will use the same suggestive notation as the previous section to describe the corresponding quantities for the diffusion processes. X is a diffusion with

drift coefficient $m(x)$, diffusion coefficient $v(x)$, state space $S = [0, 1]$, and absorbing (exit) boundaries at 0 and 1. Denote the transition density of X by $p(t, y \mid x)$; that is,

$$\int_z^w p(t, y \mid x)\, dy = P(z \leqslant X(t) \leqslant w \mid X(0) = x).$$

Again, define $\pi_0(x) = P(X$ absorbed at $0 \mid X(0) = x)$. Using standard formula (cf. Goel, Richter-Dyn, 1974, p. 65), we have:

$$\pi_0(x) = 1 - \pi_1(x) = \frac{\int_x^1 G(u)\, du}{\int_0^1 G(u)\, du}, \qquad G(u) = e^{-2\int^u [m(y)/v(y)]\, dy} \qquad (6)$$

and

$$n(x, y) = \int_0^\infty p(t, y \mid x)\, dt = \frac{2\pi_0(x) \int_0^y G(u)\, du}{v(y)\, G(y)}, \qquad 0 < y \leqslant x;$$

$$= \frac{2\pi_1(x) \int_y^1 G(u)\, du}{v(y)\, G(y)}, \qquad x \leqslant y < 1. \qquad (7)$$

The corresponding $\tilde{X}$ process is obtained from X by forcing instanteous return to the point a if 0 is visited, or to the point b if 1 is visited, where $0 < a < b < 1$. It can be shown from the work of Feller (1954, p. 23) that $\tilde{X}$ has a stationary distribution $\alpha(x)$, defined by

$$\alpha(x) = k(\pi_0(b)\, n(a, x) + \pi_1(a)\, n(b, x)), \qquad (8)$$

where k is a norming constant. $\alpha(x)$ is integrable over $[0, 1]$ because the mean absorption time for X from any point in S is finite. We will later only need ratios of $\alpha(\cdot)$, so we can take $k = 1$ in (8). When $k = 1$, we use (6), (7), and (8) to obtain the formula:

$$\begin{aligned} &= n(b, x), & x \leqslant a; \\ \alpha(x) = \frac{2\pi_0(b)\, \pi_1(a) \int_0^1 G(u)\, du}{v(x)\, G(x)}, & a \leqslant x \leqslant b; \qquad (9) \\ &= n(a, x), & x \geqslant b. \end{aligned}$$

If we suppose that the $\tilde{X}$ process is stationary, we can reverse time to obtain the process $\bar{Y}$ which jumps in the following way. $\bar{Y}$ is occasionally interrupted at the point a and restarted from 0, or interrupted at b and restarted from 1. The stationary age of the process X given current position y is defined by the time taken by the $\bar{Y}$ process to be interrupted, starting from $\bar{Y}(0) = y$. We therefore want to find the interruption or killing time density of the Y process, which is derived by killing $\bar{Y}$ at the first jump.

3. THE DUAL OF X

Suppose that the transition density $p(t, y \mid x)$ of X satisfies the backward and forward differential equations given by

$$\frac{\partial p(t, y \mid x)}{\partial t} = Lp \equiv \frac{v(x)}{2} \frac{\partial^2 p(t, y \mid x)}{\partial x^2} + m(x) \frac{\partial p(t, y \mid x)}{\partial x}, \tag{10}$$

$$\frac{\partial p(t, y \mid x)}{\partial t} = \hat{L}p \equiv \frac{1}{2} \frac{\partial^2}{\partial y^2} (v(y) \, p(t, y \mid x)) - \frac{\partial}{\partial y} (m(y) \, p(t, y \mid x)), \tag{11}$$

respectively, with appropriate boundary conditions. The Y process has transition density $p^*(t, x \mid y)$ defined by

$$p^*(t, x \mid y) = \frac{p(t, y \mid x) \, \alpha(x)}{\alpha(y)}. \tag{12}$$

We can show using (10) and (11) that p^* satisfies the backward equation

$$\frac{\partial p^*(t, x \mid y)}{\partial t} = L^* p^*(t, x \mid y) - k(y) \, p^*(t, x \mid y)$$

$$\equiv \frac{v(y)}{2} \frac{\partial^2 p^*(t, x \mid y)}{\partial y^2} + \left\{ \frac{(\alpha(y) \, v(y))'}{\alpha(y)} - m(y) \right\} \frac{\partial p^*(t, x \mid y)}{\partial y} \tag{13}$$

$$+ \left\{ \frac{(\alpha(y) \, v(y))''}{2\alpha(y)} - \frac{(m(y) \, \alpha(y))'}{\alpha(y)} \right\} p^*(t, x \mid y)$$

(13) looks just like the usual backward equation, with drift coefficient denoted $m^*(\cdot)$, except for the last term. Following Dynkin (1965, p. 10, §9, §10) we can interpret $k(y)$ in the following way. Given the history of Y up to time t, and given that $\bar{Y}(t) = x$, the probability that the process is interrupted in time $[t, t + h)$ is $k(x)h + o(h)$. This is just the probability that Y is killed in $[t, t + h)$, given $Y(t) = x$.

It remains to compute the drift coefficient m^* and the killing density of Y. Noticing that $k(y) = \hat{L}\alpha/\alpha(y)$, and the fact that $n(x, y)$ satisfies the differential equation $Ln = -\delta(x - y)$, we can compute the following:

$$\begin{aligned} k(y) &= \pi_0(b)/\alpha(a), & y &= a; \\ &= 0, & y &\neq a, \quad y \neq b; \\ &= \pi_1(a)/\alpha(b), & y &= b, \end{aligned} \tag{14}$$

and

$$\begin{aligned} m^*(y) &= \pi_0(b)/\alpha(y), & y &< a; \\ &= 0, & a &\leq y < b; \\ &= -\pi_1(a)/\alpha(y), & y &\geq b. \end{aligned} \tag{15}$$

As expected, Y can only be killed from a or b, corresponding to the fact that $\overline{Y}$ can only be interrupted from a or b. This is clearly the analogue of the discrete-time behaviour discussed in section 1.

4. AGE DISTRIBUTIONS

4.1. *The Age Density*

We first establish a formula for the density of the age distribution, given current position y. This is the density of the killing time of the Y process, given $Y(0) = y$. For convenience, we will denote this random variable by Γ. From the argument after (13), we derive the density $\gamma_y(\cdot)$ of Γ as follows.

$$P(\Gamma \in [t, t + h) \mid Y(0) = y)$$
$$= p^*(t, a \mid y)\, k(a)h + p^*(t, b \mid y)\, k(b)h + o(h).$$

Dividing by h, and letting $h \to 0$ shows that

$$\gamma_y(t) = p^*(t, a \mid y)\, k(a) + p^*(t, b \mid y)\, k(b), \qquad t \geqslant 0.$$

We can simplify this using (12) and (14) to get

$$\gamma_y(t) = \frac{\pi_0(b)\, p(t, y \mid a) + \pi_1(a)\, p(t, y \mid b)}{\alpha(y)}, \qquad t \geqslant 0. \tag{16}$$

This is the form for the age density given by Levikson (1977, section 2). See also Watterson (1977, p. 186).

4.2. *Moments of Functions of the Age*

It is useful to determine the mean time spent by Y at the frequency z before the killing time; this will lead to formulae for the mean age, for example. Define $n^*(y, z)$ to be the sojourn time density of the Y process up to the killing time; thus $n^*(y, z)\, dz$ is approximately the mean time spent in the interval $[z, z + dz)$ before the killing time. We can write

$$n^*(y, z) = E\left(\int_0^{\Gamma} \delta(Y(u) - z)\, du \mid Y(0) = y\right).$$

where $\delta(x - y) = 1$ if $x = y$, and $= 0$ if $x \neq y$.

We again argue via conditional densities to see that

$$
\begin{aligned}
n^*(y, z) &= \int_0^\infty \int_0^t p^*(u, z \mid y)\{p^*(t - u, a \mid z)\, k(a) \\
&\quad + p^*(t - u, b \mid z)\, k(b)\}\, du\, dt \\
&= \int_0^\infty p^*(u, z \mid y)\, du \int_0^\infty \gamma_y(t)\, dt \\
&= \frac{\alpha(z)\, n(z, y)}{\alpha(y)},
\end{aligned}
\tag{17}
$$

where the last equality follows from (12) and (16).

In a similar way, we could derive the useful formula

$$
E\left(\int_0^\Gamma f(Y(u))\, du \mid Y(0) = y\right) = \int_0^1 \frac{f(z)\, \alpha(z)\, n(z, y)}{\alpha(y)}\, dz,
\tag{18}
$$

where $f(\cdot)$ is a bounded function on S. (cf. Maruyama, 1977, (4.72).) Some special cases of $f(\cdot)$ are $f(x) = \delta(x - z)$, which leads back to (17), and $f(x) \equiv 1$, which leads to

$$
E(\Gamma \mid Y(0) = y) = \int_0^1 \frac{\alpha(z)\, n(z, y)}{\alpha(y)}\, dz.
\tag{19}
$$

Some applications of (18) appear in section 5.

4.3. Conditioned Killing Time

Let Γ_b be the time until the process is killed at b. Note that Γ_b need not be a proper random variable. Using a method similar to that which lead to (16), we can see that Γ_b has a density $\gamma_y{}^b(\cdot)$, such that $P(\Gamma_b \in [t, t + h) \mid Y(0) = y) = \gamma_y{}^b(t)h + o(h)$, and that

$$
\begin{aligned}
\gamma_y{}^b(t) &= k(b)\, p^*(t, b \mid y) \\
&= \frac{\pi_1(a)\, p(t, y \mid b)}{\alpha(y)}, \qquad t \geqslant 0.
\end{aligned}
\tag{20}
$$

Hence we have the result

$$
\pi_1^*(y) = P(\text{killed at } b, \text{ not } a \mid Y(0) = y) = \frac{\pi_1(a)\, n(b, y)}{\alpha(y)}.
\tag{21}
$$

Using (6), (7), and (9), we can see that

$$
\begin{aligned}
&= \pi_1(a), && y \leqslant a; \\
\pi_1^*(y) &= \pi_1(y), && a \leqslant y \leqslant b; \\
&= \pi_1(b), && y \geqslant b.
\end{aligned}
\tag{22}
$$

5. APPLICATIONS AND EXAMPLES

We have derived some formulae related to the age of a diffusion process with absorbing boundaries. In this section we will consider some of the results which have applications to genetic models. Consider a single locus in a population of individuals. There are two possible alleles at this locus, denoted by A and B respectively. The Markov chains $\{X_n, n \geqslant 0\}$ given in section 1 describe the evolution of the number of A-alleles in the population, which has fixed size M. The diffusion processes are the 'approximations' to these chains, $X(t)$ now being the proportion of A alleles at time t in a population of large size M.

The usual specification of $m(x)$ and $v(x)$ is

$$m(x) = x(1 - x)\,\phi(x), \qquad v(x) = x(1 - x), \qquad (23)$$

where $\phi(x)$ is an arbitrary polynomial (cf. Watterson, 1977, 4.3). The process $\{X(t), t \geqslant 0\}$ then has absorbing boundaries at 0 and 1. We will discuss the results in (16), (17), (18), (19), and (20) in some special cases. Recall that the Y process jumps from b to 1 or from a to 0. If the jump is to 1, then the A allele is the oldest, in the sense that it has persisted in the population longer than the B allele. Thus we may interpret the probability of a jump from b to 1 (before a jump from a to 0) as the probability that the A allele is the oldest, and this is given by the function $\pi_1^*(y)$ in (20), (21) (cf. Levikson (1977), Watterson (1977)).

5.1. *Watterson-Type Models*

One type of return boundary of particular interest to geneticists is the case where the return process moves arbitrarily close to the absorbing boundary. This corresponds to an allele being reintroduced at frequency $1/M$, where M is the (large) population size. Here we are interested in the limiting case as $a \to 0$, $b \to 1$. The important part is the limiting value of $n^*(y, z)$ in (17). From (7) and (9) it is easy to see that

$$\lim_{\substack{a \to 0 \\ b \to 1}} n^*(y, z) = n(y, z). \qquad (24)$$

Hence

$$\lim_{\substack{a \to 0 \\ b \to 1}} E(\Gamma \mid Y(0) = y) = \int_0^1 n(y, z)\, dz, \qquad (25)$$

which, as noted by Watterson, is just the mean extinction time of X from $X(0) = y$. Similarly, it follows from (22) that the probability that the A allele is the oldest is given by

$$\lim_{\substack{a \to 0 \\ b \to 1}} \pi_1^*(y) = \pi_1(y).$$

5.2. *Kimura and Ohta Type Models*

Kimura and Ohta (1973), and Maruyama (1974) were more interested in the age of an allele which has current frequency y, and started from $X(0) = p$. We can describe this situation by using the special case $a = b = p$, so that the return process X jumps to p whichever of 0 and 1 is hit. From (9) and (16) we arrive at the age density

$$\gamma_y(t) = \frac{p(t, y \mid p)}{n(p, y)}, \qquad t \geqslant 0. \tag{26}$$

This is the formula which they used to compute moments of what they took to be the age distribution. We remark that this is formally equivalent to the age of a process with only one absorbing boundary as originally suggested by Levikson; cf. Watterson, (3). Here of course the interpretation is rather different. We arrive from (17) at the result

$$\lim_{a, b \to p} n^*(y, z) = \frac{n(p, z)\, n(z, y)}{n(p, y)}, \tag{27}$$

and hence

$$\lim_{a, b \to p} E(\Gamma \mid Y(0) = y) = \int_0^1 \frac{n(p, z)\, n(z, y)}{n(p, y)}\, dz. \tag{28}$$

We also have the formula for age distributions

$$\lim_{a, b \to p} E\left(\int_0^\Gamma f(Y(u))\, du \mid Y(0) = y\right) = \int_0^1 \frac{n(p, z)\, f(z)\, n(z, y)\, dz}{n(p, y)}. \tag{29}$$

The right-hand side of (29) is formally the right-hand side of Maruyama and Kimura (1975, (4)). The left-hand side in their case (and our notation) is

$$E\left(\int_0^t f(X(u))\, du \mid X(t) = y, X(0) = p\right), \tag{30}$$

which clearly depends on t. (29) shows how to interpret the formula on the right side of (29) in a time-independent way. Sawyer (1977, Sect. 4) has another way of deriving their formula. (27) has appeared implicitly in a different context in Maruyama (1977, p. 88, (5.25), (5.26)).

5.3 *One Barrier Models*

It is possible to use the method presented here to derive age distributions for X processes which have only one absorbing boundary. The crucial step is that the absorption time from a point $X(0) = x$ should have finite mean. This is almost always the case in studying population genetic models. Such models may arise from, for instance, processes with one-way mutation pressure, or

the conditioned diffusions of Ewens (1973). It is clear how these results will encompass many already in the literature, but since the details are similar we will not describe them here. For an example of how age distributions apply to X-diffusions which are already recurrent, see for example Tavaré (1978b).

5.4. *An Example of the Method*

For concreteness, we derive a variety of results for the classical genetic diffusion (random drift) models where X has drift coefficient $m(x) \equiv 0$, and diffusion coefficient $v(x) = x(1 - x)$. It is well known that in this case, $n(\cdot, \cdot)$ is given by

$$n(x, y) = \frac{2x}{y}, \quad x \leqslant y; \qquad n(x, y) = \frac{2(1 - x)}{1 - y}, \quad x \geqslant y.$$

In table 1, we give the corresponding $n^*(y, z)$ functions, in table 2 the mean allele age for given frequency y, and in table 3 some limiting cases. In table 4, we give an example of (19), in which $f(x) = 2x(1 - x)$. Table 4 thus gives results on the mean total heterozygosity of the age process, given current frequency y. Table 5 contains some limiting values, and table 6 gives the form of the probability of the A-allele being the oldest, given current frequency y.

5.5. *Some Comments*

Levikson showed that for discrete time Markov chains it was possible to express the age distribution in terms of a conditioned last visit distribution of the X process. This was extended by Tavaré (1978b) to continuous time Markov chains with arbitrary restarting distributions. It is argued in that paper that this

TABLE 1

Values of $n^*(y, z)$, $0 < z < 1$ (17)

$y < a$		$a < y < b$		$b < y$	
$\dfrac{2z(1 - y)}{y(1 - z)},$	$z < y$	$\dfrac{2z(1 - y)}{a(1 - z)},$	$z < a$	$\dfrac{2z(1 - b)}{a(1 - z)},$	$z < a$
$2,$	$y < z < a$	$\dfrac{2(1 - y)}{1 - z},$ $a < z < y$		$\dfrac{2(1 - b)}{z},$ $a < z < b$	
$\dfrac{2a}{z},$	$a < z < b$	$\dfrac{2y}{z},$	$y < z < b$	$2,$	$b < z < y$
$\dfrac{2a(1 - z)}{z(1 - b)},$	$b < z$	$\dfrac{2y(1 - z)}{z(1 - b)},$	$z < b$	$\dfrac{2y(1 - z)}{z(1 - y)},$	$z > y$

TABLE 2

Mean Age of Allele, $E(\Gamma \mid Y(0) = y)$ (19)

$$y < a: \quad -2\left[\frac{(1-y)\ln(1-y)}{y} + \frac{ab \ln b}{1-b} + a \ln a + 1\right]$$

$$a < y < b: \quad -2\left[1 + y \ln y + (1-y)\ln(1-y) + \frac{yb \ln b}{1-b} + \frac{(1-y)(1-a)\ln(1-a)}{a}\right]$$

$$y > b: \quad -2\left[\frac{y \ln y}{1-y} + (1-b)\ln(1-b) + \frac{(1-a)(1-b)\ln(1-a)}{a} + 1\right]$$

TABLE 3

Limiting Cases of $E(\Gamma \mid Y(0) = y)$ (25), (28)

$$a \to 0, b \to 1: \quad -2(y \ln y + (1-y)\ln(1-y)), \qquad 0 < y < 1^a$$

$$a, b \to p: \quad -2\left[\frac{(1-y)\ln(1-y)}{y} + \frac{p \ln p}{1-p} + 1\right], \qquad y < p$$

$$-2\left[\frac{y \ln y}{1-y} + \frac{(1-p)\ln(1-p)}{p} + 1\right], \qquad y > p^b$$

[a] Watterson, Sect. 4.3.
[b] Kimura, Ohta (1973, 11).

TABLE 4

Expected Total Heterozygosity with Dual Process with $Y(0) = y$
((18), with $f(z) = 2z(1 - z)$)

$$y < a: \quad -2a^2 - \frac{2}{3}y^2 + \frac{2}{3}a(a^2 + 3 - (b-1)^2)$$

$$a < y < b: \quad -2y^2 - \frac{2a^2}{3} + \frac{2}{3}y(a^2 + 3 - (b-1)^2)$$

$$y > b: \quad \frac{2}{3}y(2-y) - \frac{2}{3}a^2 + \frac{2}{3}b(a^2 - b^2)$$

TABLE 5

Limit Cases of Mean Total Heterozygosity

$a \to 0, b \to 1$:	$2y(1 - y)$,	$0 < y < 1$[a]
$a, b \to p$:	$2/3(p(2 - p) - y^2)$,	$y < p$
	$2/3(y(2 - y) - p^2)$,	$y > p$[b]

[a] Maruyama (1977, (4.29)).
[b] Maruyama and Kimura (1975, (21)).

TABLE 6

$$\pi_1^*(y) = P(A \text{ Allele Oldest} \mid \text{Now at } y) \quad (22)$$

	a,	$y < a$
$\pi_1^*(y) = $	y,	$a < y < b$
	b,	$y > b$

TABLE 7

Limiting Cases of $\pi_1^*(y)$

$a \to 0, b \to 1$:	$\pi_1^*(y) = y$,	$0 < y < 1$[a]
$a \to p, b \to p$:	$\pi_1^*(y) = p$,	$0 < y < 1$

[a] Watterson (1977, (16))

analogy is most easily reconciled with age distributions in the case of X processes which have only one absorbing boundary. In the case of diffusion processes, let L_y be the time of the last visit of X to the point y before absorption. Then, conditional on $L_y < \infty$, L_y has density $l_y(\cdot)$ given by

$$l_y(t) = \frac{p(t, y \mid p)}{n(p, y)}, \qquad t \geq 0.$$

This is just the same as (26). The explanation of the age for the process with $a \neq b$ as a (conditional) last visit distribution is, however, less intuitive.

6. Conclusions

In this paper, we have studied one possible way of determing the age distribution of a diffusion process with absorbing boundaries. The results presented here are an extension of those postulated by Levikson. The age in this context is defined to be the time that has elapsed since a boundary was last visited, given some 'current' position. To avoid problems involving imprecise knowledge of what 'current' means, we have assumed that the return process that generates the age distribution was itself stationary. We then have a stationary age, which also applies if the return process has been running for a long time (cf. Watterson). The time-reversal, or dual, of this return process makes occasional jumps from inside the state space to its boundaries. The age distribution is shown to be identical to the time of the first such jump. Some special cases of the results have been analysed in detail, and some previous results unified by this approach.

References

Dynkin, E. B. 1965. "Markov Processes I," Springer–Verlag, Berlin.

Dynkin, E. B. 1965. "Markov Processes II," Springer–Verlag, Berlin.

Ewens, W. J. 1969. "Population Genetics," Methuen, London.

Ewens, W. J. 1973. Conditional diffusion processes in population genetics, *Theor. Pop. Biol.* 4, 21–30.

Feller, W. 1954. Diffusion processes in one dimension, *Trans. Amer. Math. Soc.* 77, 1–31.

Goel, N. S., and Richter-Dyn, N. 1974. "Stochastic Models in Biology," Academic Press, New York.

Kimura, M., and Ohta, T. 1973. The age of a neutral mutant persisting in a finite population, *Genetics* 75, 199–212.

Levikson, B. 1977. The age distribution of Markov processes, *J. Appl. Prob.* 14, 492–506.

Li, W.-H. 1975. The first arrival time, and mean age of a deleterious mutant gene in a finite population, *Amer. J. Hum. Genet.* 27, 274–286.

Maruyama, T., and Kimura, M. 1975. Moments for sum of an arbitrary function of gene frequency along a stochastic path of gene frequency charge, *Proc. Nat. Acad. Sci.* 72, 1602–1604.

Maruyama, T. 1974. The age of an allele in a finite population, *Genet. Res.* 23, 137–143.

Maruyama, T. 1977. "Stochastic Problems in Population Genetics," Lecture Notes in Biomathematics No. 17, Springer–Verlag, Berlin.

Narain, P. 1978. On the average of a mutant in finite populations, *in* "Proceedings, I.S.I. Conference, Delhi."

Pakes, A. G. 1977. "The Age of a Markov Process," Technical Report No. 127, Dept. of Statistics, Princeton University.

PAKES, A. G. 1978. On the age distribution of a Markov chain, *J. Appl. Prob.* 15, 65–77.

SAWYER, S. 1977. On the past history of an allele now known to have frequency *p*, *J. Appl. Prob.* 14, 439–450.

TAVARÉ, S. 1978a. Time reversion and age distributions I: Discrete time Markov chains, *J. Appl. Prob.*, in press.

TAVARÉ, S. 1978b. Time reversion and age distributions II: Continuous time Markov chains, submitted.

TAVARÉ, S. 1979. Time reversion and age distributions, *J. Appl. Prob.* 11, 10–12 (abstract).

WATTERSON, G. A. 1977. Reversibility and the age of an allele, II: Two allele models with selection and mutation, *Theor. Pop. Biol.* 12, 179–196.