J. Appl. Prob. **43,** 500–509 (2006)
Printed in Israel

ON THE TIME BEHAVIOUR OF OKAZAKI FRAGMENTS

KRZYSZTOF BARTOSZEK [*] AND
WOJCIECH BARTOSZEK,[*] [**] *Gdańsk University of Technology*

Abstract

We find explicit analytical formulae for the time dependence of the probability of the number of Okazaki fragments produced during the process of DNA replication. This extends a result of Cowan on the asymptotic probability distribution of these fragments.

Keywords: DNA replication; quasi-renewal equation; Okazaki fragment

2000 Mathematics Subject Classification: Primary 60K05
Secondary 92B05; 92C45

In a simplified model of DNA replication, Cowan [2] obtained an asymptotic probability distribution for the number of small fragments of DNA produced when the process attains equilibrium. Such fragments are called Okazaki fragments. The reader is referred to [2]–[5] for biological background and details. Let us denote by $N_t(\omega)$ the number of Okazaki fragments at the instant $t \geq 0$. This is not a deterministic function, but rather a stochastic process with nonnegative (integer) values. Let $g_i(t) = \mathrm{P}(N_t = i)$. Assuming that so-called 'primers' appear according to a Poisson process with intensity λ, it can be proved (see [2], [3], and [6]) that the functions g_i, $i = 0, 1, \dots$, satisfy the following system of (quasi-renewal) equations:

$$\begin{aligned} g_0(t) &= \mathrm{e}^{-\lambda t} + \int_0^{at} g_0(t-y)\lambda \mathrm{e}^{-\lambda y}\,\mathrm{d}y, \\ g_i(t) &= h_i(t) + \int_0^{at} g_i(t-y)\lambda \mathrm{e}^{-\lambda y}\,\mathrm{d}y, \qquad i = 1, 2, \dots. \end{aligned} \tag{1}$$

(Readers unfamiliar with the concept of a primer are referred to [3] or [5] for a brief introduction.) The value of the constant a, $0 < a < 1$, follows from the model and the functions h_i, $i = 0, 1, \dots$, are as follows:

$$h_i(t) = \begin{cases} \mathrm{e}^{-\lambda t}, & i = 0, \\ \displaystyle\int_{at}^{t} g_{i-1}(t-y)\lambda \mathrm{e}^{-\lambda y}\,\mathrm{d}y, & i = 1, 2, 3, \dots. \end{cases}$$

A natural question arises as to whether such a system has a (unique) solution. If it does we may try to find formulae for the g_i. In his approach in [2], Cowan used the method developed earlier by Piau [6] in his studies of quasi-renewal equations and presented recurrence relationships for $g_i = \lim_{t\to\infty} g_i(t)$, $i = 0, 1, \dots$. It appears that the g_i form a probabilistic distribution on

Received 11 November 2005; revision received 6 January 2006.
[*] Postal address: Department of Mathematics, Gdańsk University of Technology, ul. Narutowicza 11/12, 80-952 Gdańsk Wrzeszcz, Poland.
[**] Email address: bartowk@mifgate.mif.pg.gda.pl

the set of nonnegative integers. In proving that the above system has a unique solution, Cowan considered the functions $g_i(t)$ to be integrable on each compact subset of $\mathbb{R}_+$. By applying the Laplace transform to $g_i(t)$ and using the Euler identity (see [1, p. 19]), Cowan finally calculated the generating function, the first moment, and the variance of the limit distribution g_i.

In our approach the $g_i(t)$ are considered to be bounded, continuous functions on $\mathbb{R}_+$. We directly prove the existence and uniqueness of such solutions and, as a side effect, obtain explicit formulae for the $g_i(t)$. By C_{B} we denote the Banach lattice of all bounded, continuous, real-valued functions on $\mathbb{R}_+$ equipped with the supremum norm $\|f\|_{\sup} = \sup_{t\in\mathbb{R}_+} |f(t)|$. We also introduce the Banach lattices $C_{\mathrm{B},u}$ of all real-valued, continuous functions f on finite intervals $[0, u]$, $u > 0$, with the same supremum norm (restricted to $t \in [0, u]$). Given a function $f \in C_{\mathrm{B}}$, we define

$$R(f)(t) = \int_0^{at} f(t-s)\lambda \mathrm{e}^{-\lambda s}\,\mathrm{d}s.$$

Clearly R is a positive, linear operator on C_{B}. It is not hard to see that

$$R(\mathbf{1}_{[0,u]} f)(t) = \mathbf{1}_{[0,u]}(t) R(f)(t).$$

In other words, R leaves $C_{\mathrm{B},u}$ invariant and its restriction $R \restriction C_{\mathrm{B},u}$ may therefore be simply denoted R. Note that the operator norm of this restriction is $\|R \restriction C_{\mathrm{B},u}\| = 1 - \mathrm{e}^{-au}$.

Given a function $h \in C_{\mathrm{B}}$, we define an affine operator

$$T_h \colon C_{\mathrm{B}} \to C_{\mathrm{B}} \quad \text{by} \quad T_h(f)(t) = h + R(f)(t).$$

We note that T_h also acts on $C_{\mathrm{B},u}$ and that, for all $f_1, f_2 \in C_{\mathrm{B},u}$,

$$\|T_h(f_1) - T_h(f_2)\|_{\sup} \le (1 - \mathrm{e}^{-au})\|f_1 - f_2\|_{\sup}.$$

By the Banach fixed-point theorem,

$$T_h^n(f) \to f_{*,u} \quad \text{uniformly on } [0, u],$$

where $f_{*,u}$ is a unique fixed-point of $T_h \restriction C_{\mathrm{B},u}$. Clearly there exists a unique $f_* \in C_{\mathrm{B}}$ (of course the limit depends on the control function h) such that $f_{*,u} = f_* \restriction [0, u]$ and, moreover, for every $f \in C_{\mathrm{B}}$, $T^n(f) \to f_*$ uniformly on every compact subset of R_+. We easily find that

$$T_h^n(f)(t) = \sum_{k=0}^{n-1} R^k(h)(t) + R^n(f)(t).$$

Notice that if we let $h = \mathrm{e}^{-\lambda t}$ then $g_0 = f_*$.

We have just proved that the solution to the equation

$$g_0(t) = \mathrm{e}^{-\lambda t} + \int_0^{at} g_0(t-y)\lambda \mathrm{e}^{-\lambda y}\,\mathrm{d}y$$

does exist and is unique. Moreover, it may be obtained as the limit $\lim_{n\to\infty} T^n_{\mathrm{e}^{-\lambda t}}(f)$, where $f \in C_{\mathrm{B}}$ is arbitrary. Clearly, for each $f \in C_{\mathrm{B}}$, we have $\|R^n(f)\|_{\sup} \to 0$. The following is a similar result.

Lemma 1. *For each $i = 0, 1, \dots$, the only solution to (1) has the form $g_i = \sum_{k=0}^{\infty} R^k h_i$, where the series converges uniformly on every compact subset of $\mathbb{R}_+$ and is strictly increasing if we start with a positive function $f \in C_{\mathrm{B}}$.*

The next lemma is a step towards finding explicit solutions to these equations. Its proof is omitted, as it is a straightforward exercise.

Lemma 2. *For any nonnegative integer k and nonnegative real numbers α and λ,*

$$R^k(t^\alpha \mathrm{e}^{-\lambda t}) = \lambda^k \bigg(\prod_{j=1}^{k} \frac{1-b^{\alpha+j}}{\alpha+j}\bigg) t^{\alpha+k}\mathrm{e}^{-\lambda t}.$$

By substituting $\alpha = 0$ into this we obtain the following corollary.

Corollary 1. *For every $k = 0, 1, \dots,$ we have*

$$R^k(\mathrm{e}^{-\lambda t}) = \prod_{j=1}^{k}(1-b^j)\frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t}.$$

We are now in a position to provide an explicit formula for the function g_0 and its limit at infinity. It should be noted that (3) has appeared before, in [2].

Proposition 1. *We have*

$$g_0(t) = \sum_{k=0}^{\infty}\prod_{j=1}^{k}(1-b^j)\frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t} \tag{2}$$

and

$$g_0 = \lim_{t\to\infty} g_0(t) = \prod_{j=1}^{\infty}(1-b^j). \tag{3}$$

Proof. Equation (2) is a direct application of Corollary 1 and Lemma 1. The second formula follows from two observations: the product $\prod_{j=0}^{k}(1-b^j)$ decreases in k to $\prod_{j=0}^{\infty}(1-b^j)$, and as $t \to \infty$ the Poisson measure

$$p_{\lambda t} = \mathrm{e}^{-\lambda t}\sum_{k=0}^{\infty}\frac{(\lambda t)^k}{k!}\delta_k$$

tends weakly to δ_∞, where δ_k denotes the Dirac delta measure.

The following theorem provides explicit formulae for the functions g_i and h_i, for $i = 0, 1, 2, \dots$.

Theorem 1. *We have*

$$g_i(t) = \sum_{n_0=0,\dots,n_i=0}^{\infty} \frac{\prod_{j=1}^{n_0+\cdots+n_i+i}(1-b^j)\prod_{k=1}^{i} b^{n_1+\cdots+n_k+k}}{\prod_{k=1}^{i}(1-b^{n_1+\cdots+n_k+k})}\frac{(\lambda t)^{n_0+\cdots+n_i+i}}{(n_0+\cdots+n_i+i)!}\mathrm{e}^{-\lambda t}$$

and

$$h_i(t) = \sum_{n_1=0,\dots,n_i=0}^{\infty} \frac{\prod_{j=1}^{n_1+\cdots+n_i+i}(1-b^j)\prod_{k=1}^{i} b^{n_1+\cdots+n_k+k}}{\prod_{k=1}^{i}(1-b^{n_1+\cdots+n_k+k})}\frac{(\lambda t)^{n_1+\cdots+n_i+i}}{(n_1+\cdots+n_i+i)!}\mathrm{e}^{-\lambda t}.$$

Proof. We have already discussed the case $i = 0$ (see Corollary 1). Applying the induction method, let us assume that the formula holds for $i - 1$. Elementary calculus yields

$$\begin{aligned}\int_{at}^{t}(t-s)^{n_0+\cdots+n_{i-1}+i-1}\mathrm{e}^{-\lambda(t-s)}\lambda\mathrm{e}^{-\lambda s}\,\mathrm{d}s &= \lambda\int_{at}^{t}(t-s)^{n_0+\cdots+n_{i-1}+i-1}\mathrm{e}^{-\lambda t}\,\mathrm{d}s \\ &= \lambda\mathrm{e}^{-\lambda t}\left(-\frac{(t-s)^{n_0+\cdots+n_{i-1}+i}}{(n_0+\cdots+n_{i-1}+i)!}\right)\Bigg|_{at}^{t} \\ &= \lambda\mathrm{e}^{-\lambda t}\frac{(t-at)^{n_0+\cdots+n_{i-1}+i}}{(n_0+\cdots+n_{i-1}+i)!} \\ &= \lambda\mathrm{e}^{-\lambda t}\frac{b^{n_0+\cdots+n_{i-1}+i}t^{n_0+\cdots+n_{i-1}+i}}{(n_0+\cdots+n_{i-1}+i)!}.\end{aligned}$$

In order to keep our proof compact we make the following abbreviations:

$$L_i = \left(\prod_{j=1}^{n_0+\cdots+n_i+i}(1-b^j)\right)\prod_{k=1}^{i}\frac{b^{n_1+\cdots+n_k+k}}{1-b^{n_1+\cdots+n_k+k}}\frac{1}{(n_0+\cdots+n_i+i)!},$$

$$\Lambda_i(t) = \frac{(\lambda t)^{n_0+\cdots+n_{i-1}+i}}{(n_0+\cdots+n_{i-1}+i)!}\mathrm{e}^{-\lambda t}.$$

Now

$$\begin{aligned}h_i(t) &= \int_{at}^{t} g_{i-1}(t-s)\lambda\mathrm{e}^{-\lambda s}\,\mathrm{d}s \\ &= \int_{at}^{t}\sum_{n_0=0,\dots,n_{i-1}=0}^{\infty} L_{i-1}\frac{\lambda^{n_0+\cdots+n_{i-1}+i-1}(t-s)^{n_0+\cdots+n_{i-1}+i-1}}{(n_0+\cdots+n_{i-1}+i-1)!}\mathrm{e}^{-\lambda(t-s)}\lambda\mathrm{e}^{-\lambda s}\,\mathrm{d}s \\ &= \sum_{n_0=0,\dots,n_{i-1}=0}^{\infty} L_{i-1}\frac{\lambda^{n_0+\cdots+n_{i-1}+i-1}}{(n_0+\cdots+n_{i-1}+i-1)!}\int_{at}^{t}(t-s)^{n_0+\cdots+n_{i-1}+i-1}\mathrm{e}^{-\lambda(t-s)}\lambda\mathrm{e}^{-\lambda s}\,\mathrm{d}s \\ &= \sum_{n_0=0,\dots,n_{i-1}=0}^{\infty} L_{i-1}\frac{\lambda^{n_0+\cdots+n_{i-1}+i-1}}{(n_0+\cdots+n_{i-1}+i-1)!}\lambda\mathrm{e}^{-\lambda t}\frac{b^{n_0+\cdots+n_{i-1}+i}t^{n_0+\cdots+n_{i-1}+i}}{(n_0+\cdots+n_{i-1}+i)!} \\ &= \sum_{n_0=0,\dots,n_{i-1}=0}^{\infty} L_{i-1}b^{n_0+\cdots+n_{i-1}+i}\frac{(\lambda t)^{n_0+\cdots+n_{i-1}+i}}{(n_0+\cdots+n_{i-1}+i)!}\mathrm{e}^{-\lambda t} \\ &= \sum_{n_0=0,\dots,n_{i-1}=0}^{\infty} L_{i-1}(1-b^{n_0+\cdots+n_{i-1}+i})\frac{b^{n_0+\cdots+n_{i-1}+i}}{1-b^{n_0+\cdots+n_{i-1}+i}}\frac{(\lambda t)^{n_0+\cdots+n_{i-1}+i}}{(n_0+\cdots+n_{i-1}+i)!}\mathrm{e}^{-\lambda t} \\ &= \sum_{n_0=0,\dots,n_{i-1}=0}^{\infty}\prod_{j=1}^{n_0+\cdots+n_{i-1}+i}(1-b^j)\frac{\prod_{k=1}^{i-1}b^{n_1+\cdots+n_k+k}}{\prod_{k=1}^{i-1}(1-b^{n_1+\cdots+n_k+k})}\frac{b^{n_1+\cdots+n_{i-1}+n_0+i}}{1-b^{n_1+\cdots+n_{i-1}+n_0+i}}\Lambda_i(t).\end{aligned}$$

https://doi.org/10.1239/jap/1152413737 Published online by Cambridge University Press

By renaming the index n_0 as n_i in the above summation, we obtain

$$h_i(t) = \sum_{n_1=0,\dots,n_i=0}^{\infty} \prod_{j=1}^{n_1+\cdots+n_i+i} (1-b^j) \frac{\prod_{k=1}^{i} b^{n_1+\cdots+n_k+k}}{\prod_{k=1}^{i}(1-b^{n_1+\cdots+n_k+k})} \frac{(\lambda t)^{n_1+\cdots+n_i+i}}{(n_1+\cdots+n_i+i)!} \mathrm{e}^{-\lambda t}.$$

Applying Lemmas 1 and 2 yields

$$\begin{aligned}
g_i(t) &= \sum_{n_0=0}^{\infty} R^{n_0}(h_i)(t) \\
&= \sum_{n_0=0}^{\infty} \sum_{n_1=0,\dots,n_i=0}^{\infty} \frac{\prod_{j=1}^{n_1+\cdots+n_i+i}(1-b^j)\prod_{k=1}^{i} b^{n_1+\cdots+n_k+k}}{\prod_{k=1}^{i}(1-b^{n_1+\cdots+n_k+k})} \\
&\qquad \times R^{n_0}\bigg(\frac{(\lambda t)^{n_1+\cdots+n_i+i}}{(n_1+\cdots+n_i+i)!}\mathrm{e}^{-\lambda t}\bigg) \\
&= \sum_{n_0=0,\dots,n_i=0}^{\infty} \frac{\prod_{j=1}^{n_1+\cdots+n_i+i}(1-b^j)\prod_{k=1}^{i} b^{n_1+\cdots+n_k+k}}{\prod_{k=1}^{i}(1-b^{n_1+\cdots+n_k+k})} \frac{\lambda^{n_1+\cdots+n_i+i}}{(n_1+\cdots+n_i+i)!} \\
&\qquad \times \lambda^{n_0} \prod_{k=1}^{n_0} \frac{1-b^{n_1+\cdots+n_i+i+k}}{n_1+\cdots+n_i+i+k} t^{n_0+\cdots+n_i+i}\mathrm{e}^{-\lambda t} \\
&= \sum_{n_0=0,\dots,n_i=0}^{\infty} \frac{\prod_{j=1}^{n_0+\cdots+n_i+i}(1-b^j)\prod_{k=1}^{i} b^{n_1+\cdots+n_k+k}}{\prod_{k=1}^{i}(1-b^{n_1+\cdots+n_k+k})} \frac{(\lambda t)^{n_0+\cdots+n_i+i}}{(n_0+\cdots+n_i+i)!}\mathrm{e}^{-\lambda t}.
\end{aligned}$$

The results of the theorem now follow by induction.

The next theorem describes the asymptotics.

Theorem 2. *For each* $i > 0$*, we have*

$$\begin{aligned}
g_i = \lim_{t\to\infty} g_i(t) &= \prod_{j=1}^{\infty}(1-b^j) \sum_{n_1=0,\dots,n_i=0}^{\infty} \prod_{k=1}^{i} \frac{b^{n_1+\cdots+n_k+k}}{1-b^{n_1+\cdots+n_k+k}} \\
&= \prod_{j=1}^{\infty}(1-b^j) \sum_{n_1=1,\dots,n_i=1}^{\infty} \prod_{k=1}^{i} \frac{b^{n_1+\cdots+n_k}}{1-b^{n_1+\cdots+n_k}}.
\end{aligned}$$

Proof. Note that

$$g_i(t) = \sum_{n_1=0,\dots,n_i=0}^{\infty} \prod_{k=1}^{i} \frac{b^{n_1+\cdots+n_k+k}}{1-b^{n_1+\cdots+n_k}} \sum_{n_0=0}^{\infty} \prod_{j=1}^{n_0+\cdots+n_i+i} (1-b^j) \frac{(\lambda t)^{n_0+\cdots+n_i+i}}{(n_0+\cdots+n_i+i)!}\mathrm{e}^{-\lambda t}.$$

Using the same argument as in the proof of Proposition 1 for fixed values $n_1, \dots, n_i \geq 0$, we obtain

$$\begin{aligned}\lim_{t\to\infty} \sum_{n_0=0}^{\infty} \prod_{j=1}^{n_0+\cdots+n_i+i} (1-b^j) \frac{(\lambda t)^{n_0+\cdots+n_i+i}}{(n_0+\cdots+n_i+i)!} \mathrm{e}^{-\lambda t} \\ = \prod_{j=1}^{\infty}(1-b^j) \lim_{t\to\infty}\biggl(1 - \sum_{k=0}^{n_1+\cdots+n_i+i-1} \frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t}\biggr) \\ = \prod_{j=1}^{\infty}(1-b^j).\end{aligned}$$

The claim (in its first form) now follows from the Lebesgue convergence theorem and the fact that the series

$$\sum_{n_1=0,\dots,n_i=0}^{\infty} \prod_{k=1}^{i} \frac{b^{n_1+\cdots+n_k+k}}{1-b^{n_1+\cdots+n_k}}$$

converges absolutely. Changing each summation to start at 1 instead of 0 yields the claim in its second form.

An application of the geometric formula yields the following corollary.

Corollary 2. *For each $i = 1, 2, \dots,$ we have*

$$g_i = \prod_{j=1}^{\infty}(1-b^j) \sum_{n_1=1,\dots,n_i=1}^{\infty} \prod_{k=1}^{i} \sum_{l=1}^{\infty} b^{(n_1+\cdots+n_k)l}.$$

We will reduce the above multiple series to a simpler recurrence expression. Note that

$$\begin{aligned}\sum_{n_1=1,\dots,n_i=1}^{\infty} \frac{b^{n_1+\cdots+n_i}}{1-b^{n_1+\cdots+n_i}} \frac{b^{n_2+\cdots+n_i}}{1-b^{n_2+\cdots+n_i}} \cdots \frac{b^{n_i}}{1-b^{n_i}} \\ = \sum_{m_i=i}^{\infty} \frac{b^{m_i}}{1-b^{m_i}} \sum_{m_{i-1}=i-1}^{m_i-1} \frac{b^{m_{i-1}}}{1-b^{m_{i-1}}} \sum_{m_{i-2}=i-2}^{m_{i-1}-1} \frac{b^{m_{i-2}}}{1-b^{m_{i-2}}} \cdots \sum_{m_1=1}^{m_2-1} \frac{b^{m_1}}{1-b^{m_1}}.\end{aligned}$$

For given natural numbers i and r, $r \geq i$, we define

$$\Psi_{i,r}(b) = \sum_{m_{i-1}=i-1}^{r-1} \frac{b^{m_{i-1}}}{1-b^{m_{i-1}}} \sum_{m_{i-2}=i-2}^{m_{i-1}-1} \frac{b^{m_{i-2}}}{1-b^{m_{i-2}}} \cdots \sum_{m_1=1}^{m_2-1} \frac{b^{m_1}}{1-b^{m_1}}.$$

Clearly, for $s \geq i+1$, we have

$$\Psi_{i+1,s}(b) = \sum_{r=i}^{s-1} \frac{b^r}{1-b^r} \Psi_{i,r}(b),$$

where we have set $\Psi_{1,r}(b) \equiv 1$ for all $r \geq 1$.

https://doi.org/10.1239/jap/1152413737 Published online by Cambridge University Press

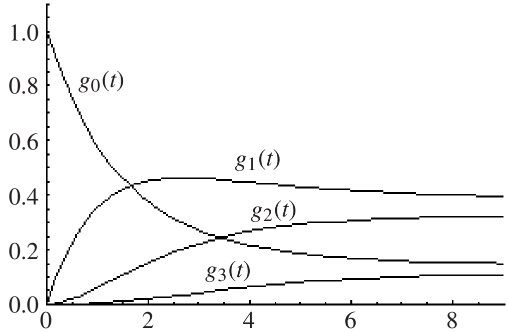

FIGURE 1.

We are now in a position to present the promised recursion formula for g_i.

Proposition 2. *For each natural number i, we have*

$$g_i = \prod_{j=1}^{\infty}(1-b^j)\sum_{m=i}^{\infty}\frac{b^m}{1-b^m}\Psi_{i,m}(b).$$

Remark 1. In [2] another representation for g_i can be found:

$$g_i = \sum_{m=i}^{\infty}(-1)^{m-i}\binom{m}{i}\prod_{k=1}^{m}\frac{b^k}{1-b^k}.$$

The above formulae were used to evaluate the values $g_0, \dots, g_{10}$ for $\lambda = 1$ and $b = 0.6$ (programming in C):

$$\begin{aligned} g_0 &= 0.143\,129\,331\,5359, & g_1 &= 0.385\,218\,306\,6464,\\ g_2 &= 0.326\,933\,548\,7938, & g_3 &= 0.120\,484\,777\,3561,\\ g_4 &= 0.022\,025\,159\,9091, & g_5 &= 0.002\,144\,129\,3616,\\ g_6 &= 0.000\,115\,947\,2975, & g_7 &= 0.000\,003\,576\,6460,\\ g_8 &= 0.000\,000\,064\,0275, & g_9 &= 0.000\,000\,000\,6727,\end{aligned}$$
$$g_{10} = 0.000\,000\,000\,004.$$

Furthermore, we include a diagram (see Figure 1), produced using MATHEMATICA®, which contains sketches of the functions $g_0(t)$, $g_1(t)$, $g_2(t)$, and $g_3(t)$. We display them to give a general idea of what these functions look like: no formal numerical analysis or error evaluation was performed. As before, $\lambda = 1$ and $b = 0.6$.

To finish the paper we will prove that $\mathfrak{g} = \{g_i\}_{i=0}^{\infty}$ defines a probability distribution on the positive integers (i.e. that $\sum_{i=0}^{\infty} g_i = 1$), and find its moments. For the convenience of the reader and completeness of the paper, we include all details (some ideas are adopted from [3] and [5]). Let us write

$$n_k(t) = \mathrm{E}(N_t^k), \qquad k = 0, 1, \dots, \ t \geq 0.$$

Since $0^0 = 1$, we have $n_0(t) \equiv 1$. Clearly, for every t and k, the moments $n_k(t)$ exist (notice that for a fixed $t \geq 0$ the process N_t is dominated by the classical Poisson process). The existence of $\lim_{t\to\infty} n_1(t) < \infty$ implies that the distribution $\mathfrak{g}$ is nondegenerate (i.e. that $\sum_{i=0}^{\infty} g_i = 1$). More generally,

$$\lim_{t\to\infty} n_{k+1}(t) < \infty$$

implies that the kth moment of $\mathfrak{g}$ is finite. We will find a formula for $n_k(t)$. Our approach is direct and requires solving linear differential equations. Let T denote the time we have to wait before the first primer appears. We begin (cf. [3]) with

$$\begin{aligned}
n_k(t) &= \int_0^{at} \mathrm{E}(N_t^k \mid T = s)\lambda \mathrm{e}^{-\lambda s}\,\mathrm{d}s + \int_{at}^{t} \mathrm{E}(N_t^k \mid T = s)\lambda \mathrm{e}^{-\lambda s}\,\mathrm{d}s \\
&= \int_0^{at} \mathrm{E}(N_{t-s}^k)\lambda \mathrm{e}^{-\lambda s}\,\mathrm{d}s + \int_{at}^{t} \mathrm{E}(N_{t-s}+1)^k \lambda \mathrm{e}^{-\lambda s}\,\mathrm{d}s \\
&= \int_0^{at} n_k(t-s)\lambda \mathrm{e}^{-\lambda s}\,\mathrm{d}s + \int_{at}^{t} \sum_{j=0}^{k} \binom{k}{j} n_j(t-s)\lambda \mathrm{e}^{-\lambda s}\,\mathrm{d}s \\
&= \int_0^{at} n_k(t-s)\lambda \mathrm{e}^{-\lambda s}\,\mathrm{d}s + \int_{at}^{t} n_k(t-s)\lambda \mathrm{e}^{-\lambda s}\,\mathrm{d}s + \sum_{j=0}^{k-1} \binom{k}{j} \int_{at}^{t} n_j(t-s)\lambda \mathrm{e}^{-\lambda s}\,\mathrm{d}s \\
&= \int_0^{t} n_k(s)\lambda \mathrm{e}^{-\lambda t+\lambda s}\,\mathrm{d}s + \sum_{j=0}^{k-1} \binom{k}{j} \int_{at}^{t} n_j(t-s)\lambda \mathrm{e}^{-\lambda s}\,\mathrm{d}s \\
&= \mathrm{e}^{-\lambda t} \int_0^{t} n_k(s)\lambda \mathrm{e}^{\lambda s}\,\mathrm{d}s + \sum_{j=0}^{k-1} \binom{k}{j} \int_{at}^{t} n_j(t-s)\lambda \mathrm{e}^{-\lambda s}\,\mathrm{d}s. \qquad (4)
\end{aligned}$$

It follows from this representation that the functions n_k (since they are measurable) belong to $C^{\infty}(\mathbb{R}_+)$. By differentiating both sides we obtain

$$\begin{aligned}
n_k'(t) &= -\lambda \mathrm{e}^{-\lambda t} \int_0^{t} n_k(s)\lambda \mathrm{e}^{\lambda s}\,\mathrm{d}s + \mathrm{e}^{-\lambda t} n_k(t)\lambda \mathrm{e}^{\lambda t} \\
&\quad + \sum_{j=0}^{k-1} \binom{k}{j} \left[n_j(0)\lambda \mathrm{e}^{-\lambda t} - a n_j(t-at)\lambda \mathrm{e}^{-\lambda at} + \int_{at}^{t} n_j'(t-s)\lambda \mathrm{e}^{-\lambda s}\,\mathrm{d}s \right].
\end{aligned}$$

Note that $n_j(0) = 1$ if $j = 0$ and $n_j(0) = 0$ for $j \geq 1$. It follows that

$$\begin{aligned}
n_k'(t) &= -\lambda \mathrm{e}^{-\lambda t} \int_0^{t} n_k(s)\lambda \mathrm{e}^{\lambda s}\,\mathrm{d}s + \lambda n_k(t) + \lambda \mathrm{e}^{-\lambda t} \\
&\quad + \sum_{j=0}^{k-1} \binom{k}{j} \left[\int_{at}^{t} n_j'(t-s)\lambda \mathrm{e}^{-\lambda s}\,\mathrm{d}s - a n_j((1-a)t)\lambda \mathrm{e}^{-\lambda at} \right]. \qquad (5)
\end{aligned}$$

Using (4), we obtain

$$\mathrm{e}^{-\lambda t} \int_0^{t} n_k(s)\lambda \mathrm{e}^{\lambda s}\,\mathrm{d}s = n_k(t) - \sum_{j=0}^{k-1} \binom{k}{j} \int_{at}^{t} n_j(t-s)\lambda \mathrm{e}^{-\lambda s}\,\mathrm{d}s.$$

https://doi.org/10.1239/jap/1152413737 Published online by Cambridge University Press

Substituting this into (5) yields

$$n_k'(t) = -\lambda\Big[n_k(t) - \sum_{j=0}^{k-1}\binom{k}{j}\int_{at}^{t} n_j(t-s)\lambda e^{-\lambda s}\,ds\Big]$$
$$+\lambda n_k(t) + \lambda e^{-\lambda t} + \sum_{j=0}^{k-1}\binom{k}{j}\Big[\int_{at}^{t} n_j'(t-s)\lambda e^{-\lambda s}\,ds - a n_j((1-a)t)\lambda e^{-\lambda at}\Big]$$
$$= \lambda^2\sum_{j=0}^{k-1}\binom{k}{j}\int_{at}^{t} n_j(t-s)e^{-\lambda s}\,ds + \lambda e^{-\lambda t}$$
$$+\lambda\sum_{j=0}^{k-1}\binom{k}{j}\Big[\int_{at}^{t} n_j'(t-s)e^{-\lambda s}\,ds - a n_j(bt)e^{-\lambda at}\Big]. \tag{6}$$

Using the above recursion scheme and the induction principle, we easily obtain the following result.

Corollary 3. *There exist constants $\beta_{k,j} > 0$, $C_k > 0$, and $\alpha_{k,j}$ and a natural number L_k such that, for each $k \geq 0$, we have*

$$n_k(t) = \sum_{j=1}^{L_k}\alpha_{k,j}e^{-\beta_{k,j}t} + C_k.$$

In particular,

$$\lim_{t\to\infty} n_k(t) = \lim_{t\to\infty}\sum_{i=0}^{\infty} i^k g_i(t) = \sum_{i=0}^{\infty} i^k g_i = C_k < \infty$$

(all moments of the asymptotic distribution $\mathfrak{g}$ are finite).

Setting $k = 1$ in (6) yields

$$n_1'(t) = \lambda^2\int_{at}^{t} e^{-\lambda s}\,ds + \lambda e^{-\lambda t} - a\lambda e^{-\lambda at} = \lambda(1-a)e^{-\lambda at}.$$

It follows that

$$n_1(t) = \int \lambda(1-a)e^{-\lambda at}\,dt = -\frac{1-a}{a}e^{-\lambda at} + C.$$

Clearly $C = (1-a)/a$, since $\lim_{t\to 0^+} n_1(t) = 0$. As a result we obtain the next corollary.

Corollary 4. *For all $t \geq 0$, we have $n_1(t) = (1-a)(1-e^{-\lambda at})/a$. It follows that*

$$\mu = \sum_{i=0}^{\infty} i g_i = \lim_{t\to\infty} n_1(t) = \frac{1-a}{a}.$$

In order to find the second moment and variance, we set $k = 2$ in (6). After several elementary calculations, we obtain

$$n_2'(t) = \lambda\frac{(1-a)(2-a)}{a}e^{-\lambda at} - 2\lambda\frac{(1-a)^2}{a}e^{-\lambda a(2-a)t}.$$

Integrating the last equation and taking into account the fact that $n_2(0) = 0$ for all $t \geq 0$ yields our final corollary.

https://doi.org/10.1239/jap/1152413737 Published online by Cambridge University Press

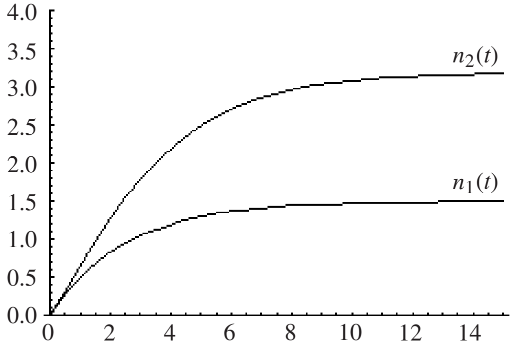

FIGURE 2.

Corollary 5. *We have*

$$n_2(t) = \frac{2(1-a)^2}{a^2(2-a)} \mathrm{e}^{-\lambda a(2-a)t} - \frac{(1-a)(2-a)}{a^2} \mathrm{e}^{-\lambda a t} + \frac{(1-a)(a^2-2a+2)}{a^2(2-a)}.$$

It follows that

$$\lim_{t\to\infty} n_2(t) = \frac{(1-a)(a^2-2a+2)}{a^2(2-a)}.$$

In particular,

$$\operatorname{var}(\mathfrak{g}) = \lim_{t\to\infty} (n_2(t) - n_1(t)^2) = \frac{1-a}{1-(1-a)^2}.$$

In Figure 2 we display the graphs of the functions $n_1(t)$ and $n_2(t)$ for $\lambda = 1$ and $a = 1-b = 0.4$.

References

[1] ANDREWS, G. (1976). *The Theory of Partitions* (Encyclopaedia Math. Appl. **2**). Addison-Wesley, Reading, MA.

[2] COWAN, R. (2001). A new discrete distribution arising in a model of DNA replication. *J. Appl. Prob.* **38,** 754–760.

[3] COWAN, R. (2003). Stochastic models for DNA replication. In *Handbook of Statistics*, Vol. 21, eds D. N. Shanbhag *et al.*, North-Holland, Amsterdam, pp. 137–166.

[4] COWAN, R. AND CHIU, S. N. (1994). A stochastic model of fragment formation when DNA replicates. *J. Appl. Prob.* **31,** 301–308.

[5] LACHAL, A. (2003). Some probability distributions in modeling DNA replication. *Ann. Appl. Prob.* **13,** 1207–1230.

[6] PIAU, D. (2000). Quasi-renewal estimates. *J. Appl. Prob.* **37,** 269–275, 1171–1172.