A NON-MARKOVIAN MODEL FOR CELL POPULATION GROWTH: TAIL BEHAVIOR AND DURATION OF THE GROWTH PROCESS

By Mathisca C. M. de Gunst[1] and Willem R. van Zwet

Free University of Amsterdam, and University of Leiden and University of North Carolina

De Gunst has formulated a stochastic model for the growth of a certain type of plant cell population that initially consists of n cells. The total cell number $N_n(t)$ as predicted by the model is a non-Markovian counting process. The relative growth of the population, $n^{-1}(N_n(t) - n)$, converges almost surely uniformly to a nonrandom function X. In the present paper we investigate the behavior of the limit process $X(t)$ as t tends to infinity and determine the order of magnitude of the duration of the process $N_n(t)$. There are two possible causes for the process N_n to stop growing, and correspondingly, the limit process $X(t)$ has a derivative $X'(t)$ that is the product of two factors, one or both of which may tend to zero as t tends to infinity. It turns out that there is a remarkable discontinuity in the tail behavior of the processes. We find that if only one factor of $X'(t)$ tends to zero, then the rate at which the limit process reaches its final limit is much faster and the order of magnitude of the duration of the process N_n is much smaller than when both occur approximately at the same time.

1. Biological background. Much of the research in plant cell biotechnology is directed at biosynthesis of secondary metabolites in plant cell cultures [Morris, Scragg, Stafford and Fowler (1986)]. Control of the productivity of these cell cultures in multiliter vessels in industry requires detailed knowledge of the kinetics of growth, division, differentiation and product formation of cells grown under different environmental conditions. However, our understanding of these kinetics is still very incomplete. In collaboration with K. R. Libbenga of the Department of Plant Molecular Biology at the University of Leiden, we have developed a mathematical model for the division, differentiation and population growth of plant cells in a liquid medium. This model is based on the presently available experimental knowledge of the behavior of individual cells, and takes into account the influence

of the depletion of two components of the medium that are indispensable for the growth and division of the cells.

Before we formulate the model in Section 2, let us first describe the biological background. We start at the level of a single plant cell, which is transferred at time $t = 0$ to a fresh liquid medium containing the substances needed for the growth, division and survival of the cell. The cell will go through a sequence of events called the cell cycle, which starts at time 0 and ends with the division of the cell. The cell cycle is illustrated in the diagram of Figure 1. The cycle starts with the G_1-phase (G for gap) during which the biosynthetic activity of the cell proceeds at a high rate. The S-phase (synthesis) that follows starts when DNA synthesis begins and ends when the DNA content of the cell nucleus has doubled and the chromosomes have replicated. The cell then enters the G_2-phase (another gap), which continues until the final M-phase (mitosis), which is the brief period of actual cell division. During the M-phase, the biosynthetic activity of the cell proceeds very slowly and increases again rapidly after division as the two new cells enter the G_1-phase of their cell cycles. Together, the G_1-, S- and G_2-phases are also called the interphase.

It is a well verified fact that the duration of the cell cycle varies considerably, even among cells of the same type under the same external conditions. Most of the variability is observed in the length of the G_1-phase; the remainder of the cycle time shows far less variation. Moreover, it is known that the G_1-phase tends to last longer if the supply of certain nutrients is reduced; the
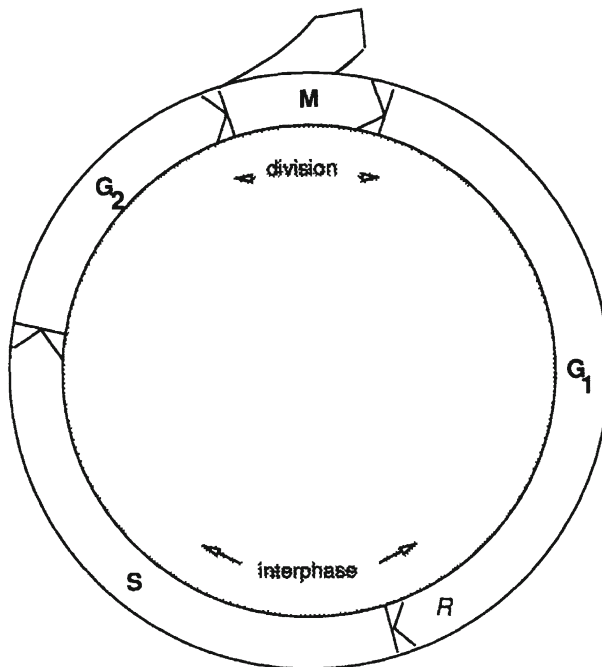

FIG. 1.   The four successive phases of a typical cell cycle.

duration of the other parts of the cycle is scarcely affected [cf. Alberts, Bray, Lewis, Raff, Roberts and Watson (1989), page 745].

Another well established fact is that once the cell has left the G_1-phase, it is committed to complete the cell cycle regardless of environmental conditions [Alberts, Bray, Lewis, Raff, Roberts and Watson (1989), page 745]. This indicates the existence of a point of no return—often called the restriction point $\mathscr{R}$—somewhere late in the G_1-phase. It is thought that, as the cell approaches the point $\mathscr{R}$, it must "wait" for some endogenous trigger or stimulus that moves it past $\mathscr{R}$, and that under fixed external conditions the probability per unit time of passing $\mathscr{R}$ is roughly constant. In a mathematical model this would correspond to an exponentially distributed waiting time for a stimulus to occur, and if the remainder of the length of the G_1-phase—like that of the other phases—is almost constant, then the total duration of the cell cycle would be the sum of a constant time and an exponentially distributed one. This hypothesis was advanced in a seminal paper by Smith and Martin (1973) and verified on the basis of experimental data [see also Shields (1977)]. Later authors have criticized this so-called transition probability model [cf. Nelson and Green (1981)], but at present the existence of the restriction point seems to be firmly established. Others [Brooks, Bennett and Smith (1980), Castor (1980) and Cooper (1982)] have proposed alternative probability distributions for the duration of the cell cycle, incorporating more than one waiting time, for instance, but it seems difficult to distinguish between these models on the basis of the existing experimental data. At this time the transition probability model appears to be the accepted theory in the biological literature [Alberts, Bray, Lewis, Raff, Roberts and Watson (1989), pages 733 and 746].

We have already noted that the average duration of the G_1-phase increases as the supply of nutrients is reduced, and when no nutrients are present, cells cannot pass the restriction point at all. Hence we shall assume in our model that the parameter of the exponential waiting time for a stimulus is an increasing function of the concentration of nutrients, which tends to zero as the concentration does. The remaining part of the cycle is not affected by the concentration of nutrients. Of course, a cell also consumes nutrients, especially during the first part of the G_1-phase leading up to the restriction point $\mathscr{R}$. In our model we shall telescope this process and assume for simplicity that a cell only consumes a fixed amount of nutrient at the time it receives the stimulus to pass $\mathscr{R}$.

A plant cell that takes part in the cycling process is usually small and spherical, with its nucleus positioned at the center. However, if one watches a population of plant cells grow by cell division, one also notices after some time the presence of larger, more stretched out cells, with nuclei close to the cell wall. These cells are in an early stage of differentiation and do not divide. Such a differentiating cell most probably resides in the G_1-phase, before the restriction point $\mathscr{R}$, in a so-called quiescent, or G_0-state [Alberts, Bray, Lewis, Raff, Roberts and Watson (1989), page 750]. It is possible for such a cell to restart its cycle, but it needs a much more powerful trigger to do so

than the one needed for passing $\mathscr{R}$: it has to dedifferentiate first before it resumes its cell cycle. Transfer of a cell to a fresh liquid medium is apparently such a trigger and will make the cell start its cycle almost at once.

Although the underlying biochemical mechanisms are far from known, there are strong indications [Bayliss (1985); Trewavas (1985)] that plant hormones play a crucial role in determining differentiation: the higher their concentration in the medium, the larger the proportion of cells that are actively cycling, and the smaller the proportion of cells that will never divide but will differentiate instead. In the absence of hormones, there will be no cycling cells. We shall model this phenomenon by assuming that at the time of cell division, the two new cells independently become cycling cells (type A cells) with probability P or differentiating cells (type B cells) with probability $(1 - P)$. Here P is assumed to be an increasing function of the hormone concentration in the medium at the time of division, which vanishes as the hormone concentration does. A cell also takes up hormones, and we shall assume that a fixed amount of hormone is used up by each cell at the time of its division.

Having described the behavior of a single cell, we now turn to the behavior of a population of plant cells in a liquid medium. Such populations can occur either as batch cultures or as continuous cultures [Street (1973)]. In either case, the culture consists of isolated cells—or very small cell aggregates—that remain dispersed as they grow in the liquid medium. This is achieved by continuous stirring of the fermentor in which the cells grow. A culture is started by the transfer of a certain number of cells to a fresh medium containing known quantities of nutrients and hormones. In contrast to a continuous culture, a batch culture does not have any inflow of fresh medium or outflow of culture. As such, the batch culture is the appropriate system to study the growth of the number of cells of a population and to investigate the influence of the different components of the medium on the population growth. In what follows we shall, therefore, restrict our attention to plant cells in batch culture.

The transfer of the cells to a fresh medium at time 0 triggers all cells to start their cycles almost at once, and we shall, therefore, assume that at time 0 all cells are of type A and that their cell cycles have been synchronized. If the amount of nutrient were kept constant or varied over time in a nonrandom fashion, it would be reasonable to assume that the duration of the cycles of different cells would be independent. However, in batch culture, the concentration of nutrient decreases at the random times when stimuli arrive, and cells compete for the available nutrient. This creates a complicated type of dependence between the division times for different cells.

Similarly, the hormone concentration decreases at the times of cell division. It follows that the cell population will ultimately stop growing, either because the nutrient is exhausted and no more stimuli can occur, or because the hormone concentration has become so low that the population of cycling cells can no longer be sustained. In the model this occurs when the probability P of becoming an A cell has fallen below 0.5.

In batch culture, cell death is observed only at the beginning when cells are transferred to a fresh medium, and at a time when the cell density has been very high for a considerable period. The effect of the former is removed by simply never counting these dead cells, whereas the latter occurs in practice only after the growth of the population has stopped. We shall, therefore, assume that cell death does not occur.

On the basis of this biological description we shall build a mathematical model for the growth of a plant cell population in batch culture in Section 2. In this model we shall study the duration of the growth process (i.e., the time until the population stops growing) when the initial population size n is large. We shall find that the duration is usually proportional to $\log n$. However, if the nutrient is exhausted at approximately the same time as the hormone concentration becomes too low to sustain the process, the duration is proportional to $n^{1/2} \log n$. Similar results are proved for an appropriately defined limit process.

2. The mathematical model. Let us turn this biological description into a mathematical model for the growth of a population of plant cells in batch culture. We start at time $t = 0$ with n cells. Because we intend to consider the growth of this population as n tends to infinity, we use n as an index throughout. The Smith–Martin model tells us that the duration of a cell cycle is of the form $(W + c)$, where W is a random waiting time for a stimulus to arrive and $c > 0$ is a constant. In a constant environment W has an exponential distribution with parameter λ (i.e., with expectation $1/\lambda$) and stimuli arrive independently for different cells. To fix thoughts, we assume that the cell cycle starts with the exponential waiting time for the stimulus and that the cell divides a constant time c after receiving the stimulus.

At time t, there will be $N_n(t)$ cells, of which $N_{An}(t)$ are A cells (i.e., cycling cells). Of these $N_{An}(t)$ A cells, $N_{An}^0(t)$ cells are at time t waiting for a stimulus to arrive, whereas the remaining $N_{An}(t) - N_{An}^0(t)$ A cells are somewhere in the time span of length c between arrival of a stimulus and division, and will, therefore, divide before or at time $(t + c)$. Thus

$$(2.1) \quad \begin{aligned} N_{An}(t) - N_{An}^0(t) &= N_n(t + c) - N_n(t) \quad \text{or} \\ N_{An}^0(t) &= N_{An}(t) - \left(N_n(t + c) - N_n(t)\right). \end{aligned}$$

At time $t = 0$, all cells are of type A and at the beginning of their cycle, so

$$(2.2) \quad N_n(0) = N_{An}(0) = N_{An}^0(0) = n.$$

We shall also need normalized versions of the three processes defined so far, and we write

$$(2.3) \quad \begin{aligned} X_n(t) &= n^{-1}(N_n(t) - n), \\ X_{An}(t) &= n^{-1}N_{An}(t), \\ X_{An}^0(t) &= n^{-1}N_{An}^0(t). \end{aligned}$$

At the time of a cell division, the two new cells independently become A cells with probability P and B cells with probability $(1 - P)$. This probability P is an increasing function of the hormone concentration immediately before the division, and in batch culture this concentration decreases as time goes on. Suppose that the amount of hormone at time $t = 0$ equals $[nb_h]$, with b_h a positive constant, and that an amount 1 is used up at each division. Here $[x]$ denotes the largest integer less than or equal to x. Immediately before the ith division, the amount of hormone is $([nb_h] - (i - 1))$, and hence the probability of a cell becoming an A cell at the ith division equals

$$(2.4) \qquad P_{in} = P\left(\frac{[nb_h] - i + 1}{n}\right), \qquad i = 1, 2, \ldots,$$

where P is increasing on $[0, \infty)$ and $P(u) = 0$ for $u \le 0$. According to Monod kinetics, which is the standard model for these biochemical processes [see, for instance, Roels (1983)], P is given by

$$(2.5) \qquad P(u) = \begin{cases} 0, & u \le 0, \\ \dfrac{u}{a_h + u} = 1 - \dfrac{a_h}{a_h + u}, & u > 0, \end{cases}$$

where a_h denotes a positive constant. Note that $P_{in} = 0$ for $i \ge [nb_h] + 1$. Because P is nonnegative, nondecreasing and concave, one easily verifies that for $m = 1, 2, \ldots, [nb_h]$,

$$(2.6) \qquad \left|\sum_{i=1}^{m} P_{in} - n\int_0^{m/n} P(b_h - u)\, du\right| \le P(b_h) = \frac{b_h}{a_h + b_h} \le 1.$$

Let $Z_n = (Z_{1n}, Z_{2n}, \ldots)$ denote a random sequence, where $Z_{1n}, Z_{2n}, \ldots$ are independent and Z_{in} has a binomial distribution with parameters 2 and P_{in}. Here Z_{in} models the number of A cells created at the ith division, and hence

$$(2.7) \qquad \begin{aligned} N_{An}(t) &= 2n - N_n(t) + \sum_{i=1}^{N_n(t)-n} Z_{in} \\ &= n + \sum_{i=1}^{N_n(t)-n} (Z_{in} - 1). \end{aligned}$$

In view of (2.1), for $t > c$,

$$(2.8) \qquad N_{An}^0(t - c) = 2n - N_n(t) + \sum_{i=1}^{N_n(t-c)-n} Z_{in}.$$

Note that, conditional on Z_n, $\{N_{An}(s): s \le t\}$ depends only on $\{N_n(s): s \le t\}$, but $\{N_{An}^0(s): s \le t\}$ depends on $\{N_n(s): s \le (t + c)\}$.

The parameter λ of the exponential waiting time for a stimulus is an increasing function of the amount of substrate (or nutrients) present, and in batch culture this concentration decreases over time. Suppose that the amount of substrate at time $t = 0$ equals $[nb_s]$, for a positive constant b_s, and that an

amount 1 is used up for each stimulus. At time t, $(N_n(t) - n)$ divisions have taken place, and this is also the number of stimuli that have arrived before or at time $(t - c)$. Hence, the amount of substrate at time $(t - c)$ equals

$$[nb_s] - (N_n(t) - n) = n\left(\frac{[nb_s]}{n} - X_n(t)\right).$$

Thus, the time dependent rate at which a stimulus arrives at time $(t - c)$ is an increasing function of $\{[nb_s]/n\} - X_n(t)$, say $Q(\{[nb_s]/n\} - X_n(t))$. According to Monod kinetics once more, Q is defined by

(2.9) $$Q(u) = \begin{cases} 0, & u \leq 0, \\ \dfrac{u}{d(a_s + u)} = \dfrac{1}{d} - \dfrac{a_s}{d(a_s + u)}, & u > 0, \end{cases}$$

where d and a_s are positive constants. Note that the rate becomes zero after $[nb_s]$ stimuli have arrived.

The amount of substrate $n(\{[nb_s]/n\} - X_n(t))$ that is present at time $(t - c)$ will remain unchanged until the random time $(\tau - c)$ when the first stimulus after $(t - c)$ arrives. Thus, $(\tau - c)$ is distributed as the minimum of the times of the first event in independent Poisson processes with intensity $Q(\{[nb_s]/n\} - X_n(t))$. For the $N_{An}^0(t - c)$ waiting A cells that are already present at time $(t - c)$, the corresponding Poisson processes start at time $(t - c)$, and for A cells created after time $(t - c)$ the processes start at the time of their creation. Thus, given Z_n, the conditional intensity of the stimulus process at time $(t - c)$ equals the left-continuous version of $N_{An}^0(t - c)Q(\{[nb_s]/n\} - X_n(t))$. Because a stimulus at time $(t - c)$ corresponds to a cell division at time t, it follows that, *conditional on* Z_n, the process $\{N_n(t) - n: t \geq 0\}$ is a counting process with the left-continuous version of

(2.10) $$\Lambda_{Zn}(t) = \begin{cases} 0, & 0 \leq t < c, \\ N_{An}^0(t - c)Q\left(\dfrac{[nb_s]}{n} - X_n(t)\right), & t \geq c, \end{cases}$$

as its *conditional intensity*. Together with the distribution of Z_n given before, this determines our mathematical model for the growth of a plant cell population in batch culture.

The process N_n stops growing for one of two entirely different reasons: either A cells become extinct or the rate at which the stimuli arrive becomes zero. Thus the process $N_n(t)$ reaches its final value at the first time t when either $N_{An}(t) = 0$ or $N_n(t) = n + [nb_s]$. Note that $N_{An}^0(t - c) = 0$ is not sufficient for N_n to stop growing at time t, because new A cells may be born between time $(t - c)$ and t. If T_n denotes the random time of the final cell division, then by (2.7),

$$T_n = \inf\left\{t: \sum_{i=1}^{N_n(t)-n} (Z_{in} - 1) = -n\right\} \wedge \inf\{t: N_n(t) = n + [nb_s]\},$$

where $(a \wedge b)$ means the smaller of a and b [similarly, the larger of a and b will be denoted by $(a \vee b)$]. The random level that N_n has reached by then equals

$$(2.11) \qquad N_n(T_n) = n + M_n \wedge [nb_s],$$

where

$$(2.12) \qquad M_n = \inf\left\{m: \sum_{i=1}^{m} (Z_{in} - 1) = -n\right\} \geq n.$$

Clearly this final level $N_n(T_n)$ depends only on Z_n and b_s. The process stops because $N(T_n) = n + [nb_s]$ or $N_{An}(T_n) = 0$, or both, depending on whether $M_n > [nb_s]$ or $M_n < [nb_s]$ or $M_n = [nb_s]$, respectively. An alternative expression for T_n is

$$(2.13) \qquad T_n = \inf\{t: N_n(t) = n + M_n \wedge [nb_s]\},$$

which shows that given Z_n, T_n is the time at which the counting process N_n reaches a fixed level.

Suppose that, for large n, the process $X_n(t) = (N_n(t) - n)/n$ is close to a deterministic function $X(t)$ in $D[0, \infty)$, the space of right-continuous, $\mathbb{R}$-valued functions on $[0, \infty)$ with left-hand limits everywhere. Then by (2.8), (2.10) and the fact that

$$n^{-1} \sum_{1}^{m} Z_{in} \sim n^{-1} \sum_{1}^{m} 2P_{in} \sim 2\int_{0}^{m/n} P(b_h - u)\, du$$

by (2.6), we find that $\Lambda_{Zn}(t)/n$ will be close to $F(t, X)$, where $F: [0, \infty) \times D[0, \infty) \to \mathbb{R}$ is defined by

$$
\begin{aligned}
(2.14) \qquad F(t, x) = &\left\{1 - x(t) + 2\int_{0}^{x(t-c)} P(b_h - u)\, du\right\} \\
&\times Q(b_s - x(t))1_{[c, \infty)}(t).
\end{aligned}
$$

Here $a_n \sim b_n$ means that the quotient of a_n and b_n tends to 1 (in probability) as n tends to infinity. Thus it seems plausible that, if a deterministic limit X of the processes X_n exists, it should satisfy the equation

$$(2.15) \qquad x(t) = \begin{cases} 0, & 0 \leq t < c, \\ \int_{c}^{t} F(s, x)\, ds, & t \geq c. \end{cases}$$

It is shown in De Gunst (1989) that (2.15) has a unique solution X in $D[0, \infty)$. This function X is continuous, nonnegative, nondecreasing and bounded on $[0, \infty)$, and differentiable on (c, ∞) with a continuous, positive and bounded derivative. Hence $X(t)$ tends to a finite limit $X(\infty)$ as t tends to infinity, and in view of (2.14) and (2.15), it follows that $X'(t)$ also tends to a limit, which must necessarily be zero:

$$
\begin{aligned}
(2.16) \qquad &X(\infty) = \lim_{t \to \infty} X(t) < \infty, \\
&\lim_{t \to \infty} X'(t) = 0.
\end{aligned}
$$

For biologically plausible values of the parameters, the graph of X exhibits alternating intervals of slow and rapid increase, which level off as time

progresses. This reflects the synchronism of cells at time $t = 0$, which is gradually destroyed by the variability of the cycle times. Figure 2 shows a graph of X that was obtained by fitting a numerical solution of (2.15) to the only experimental data that are available so far, and that are also shown in this figure. Figure 2, as well as a statistical analysis of the experimental data, shows that for appropriate parameter values, the function X describes actual batch culture growth quite well. A detailed description of the experimental procedures, the statistical analysis of the data, and a further discussion of the relevance of the results can be found in De Gunst, Harkes, Val, Van Zwet and Libbenga (1990).

Having defined X as the nonrandom counterpart of X_n, we proceed to define the counterparts of the other processes in (2.3) by

(2.17)
$$X_A(t) = 1 - X(t) + 2\int_0^{X(t)} P(b_h - u)\, du,$$

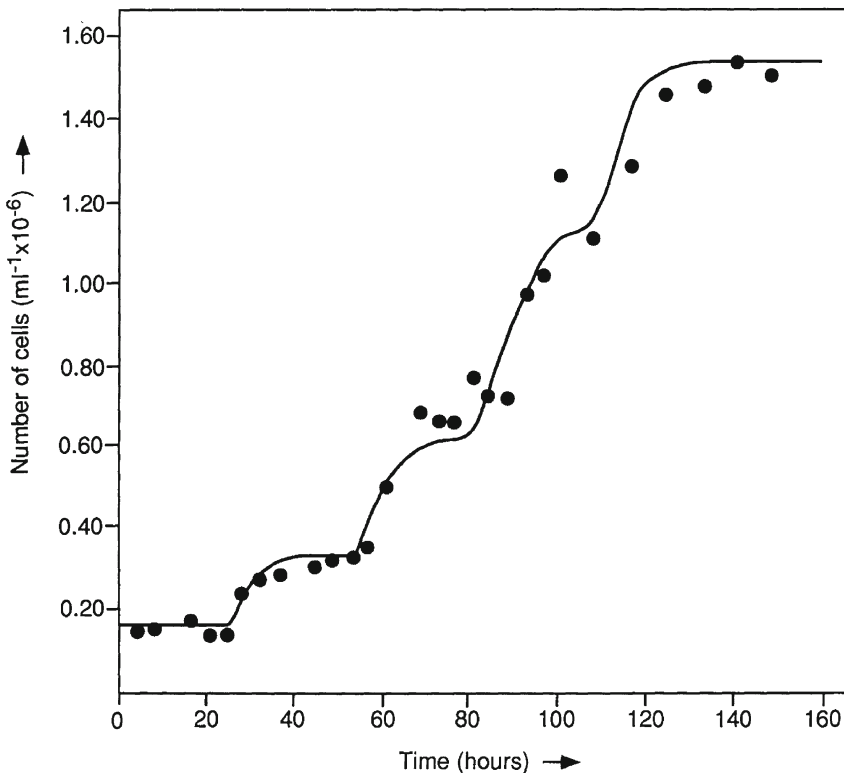$$X_A^0(t - c) = 1 - X(t) + 2\int_0^{X(t-c)} P(b_h - u)\, du,$$

FIG. 2.   *Experimental data of the growth of a batch culture of tobacco cells (dots). The curve through the data was fitted using numerical solutions of (2.10). The parameter values are* $n = 1.625 \times 10^8 l^{-1}$, $c = 26h$, $d = 4$ h, $b_s = 29.9$, $k_s = 2.2 \times 10^{-11}$ mol, $y_s = 5.4 \times 10^{10}$ mol^{-1}, $b_h = 4.7$, $k_h = 9.7 \times 10^{-17}$ mol, $y_h = 3.4 \times 10^{15}$ mol^{-1}.

in analogy with (2.7) and (2.8). Thus, by (2.10), (2.14) and (2.15) we have the two corresponding expressions

$$(2.18) \qquad \Lambda_{Zn}(t) = nX^0_{An}(t-c)Q(\{[nb_s]/n\} - X_n(t)) \quad \text{for } t \geq c,$$

$$(2.19) \qquad X'(t) = X^0_A(t-c)Q(b_s - X(t)) \qquad\qquad \text{for } t \geq c,$$

for the conditional intensity of $(N_n - n)$ and the derivative of X.

In De Gunst and Van Zwet (1992) it is shown that indeed X_n converges in probability to X at a rate of $n^{-1/2}$ uniformly on $[0, \infty)$. Moreover, there is an exponential bound for the tail probability. The same holds for X_{An} and X_A, as well as for X^0_{An} and X^0_A. Theorem 2.1 summarizes these results.

THEOREM 2.1. *Let X be the solution of (2.15). Then there exist positive numbers A and α, such that for $n = 1, 2, \ldots$ and $x \geq 0$,*

$$(2.20) \qquad \mathscr{P}\left(\sup_{t \geq 0} |X_n(t) - X(t)| \geq x \right) \leq A \exp\{ -\alpha x^2 n \},$$

and hence

$$(2.21) \qquad \sup_{t \geq 0} |X_n(t) - X(t)| = \mathscr{O}_{\mathscr{P}}(n^{-1/2}).$$

The same conclusions hold if X_n and X are replaced by X_{An} and X_A or X^0_{An} and X^0_A, respectively.

To simplify our notation in what follows, we introduce a function ψ defined on $[0, \infty)$ by

$$(2.22) \qquad \psi(v) = 1 - v + 2\int_0^v P(b_h - u)\, du.$$

Note that

$$(2.23) \qquad X_A(t) = \psi(X(t)) \quad \text{for all } t.$$

By (2.5), P is nondecreasing and vanishes for negative values of its argument, and hence ψ is concave on $[0, \infty)$ with $\psi(0) = 1$ and $\lim_{v \to \infty} \psi(v) = -\infty$. Define a new parameter γ as the solution of

$$(2.24) \qquad \psi(\gamma) = 0.$$

Obviously, γ is uniquely determined and ψ is positive (negative) to the left (right) of γ. Apart from our choice (2.5) for the function P, γ depends only on b_h. By (2.22) and (2.24), $\gamma \geq 1$.

To complete our notation we define

$$(2.25) \qquad X_A(\infty) = \lim_{t \to \infty} X_A(t) = \lim_{t \to \infty} X^0_A(t) = \psi(X(\infty)),$$

$$p = P(b_h - X(\infty)),$$

$$(2.26) \qquad q = Q(b_s - X(\infty)),$$

$$q' = Q'(b_s - X(\infty)),$$

where Q' denotes the derivative of Q.

We have already noted that the process N_n will stop because one or both of the factors in (2.18) vanish. Similarly, for the deterministic limit function X, (2.16), (2.19) and (2.25) imply that

$$(2.27) \qquad X_A(\infty)Q(b_s - X(\infty)) = \psi(X(\infty))Q(b_s - X(\infty)) = 0,$$

so that we have three different cases: either $X(\infty) = b_s$ and hence $q = Q(b_s - X(\infty)) = 0$ or $X_A(\infty) = \psi(X(\infty)) = 0$ or both. An alternative way to express this is to write

$$(2.28) \qquad X(\infty) = \inf\{v: \psi(v) \le 0\} \wedge b_s = \gamma \wedge b_s,$$

which is the analogue of (2.11) and (2.12).

In the present paper, we shall investigate two closely related issues: the behavior of $X(t)$ for large values of t and the duration of the process N_n, that is, the time T_n to the final cell division. These issues concern the tail behavior of the processes X and X_n, respectively, the former being an (easier) deterministic version of the latter. We shall show that, depending on the values of the biological parameters, there is a remarkable discontinuity in this tail behavior of X and X_n.

Let us consider more closely the three cases that (2.27) allows and classify their occurrence in terms of the parameters γ and b_s.

$\gamma > b_s$:

(2.29)(i) By (2.28), $X(\infty) = b_s$ and $q = Q(b_s - X(\infty)) = 0$.

On the other hand, $X_A(\infty) = \psi(X(\infty)) = \psi(b_s) > 0$;

$\gamma < b_s$:

By (2.28), $X(\infty) = \gamma$ and $q = Q(b_s - X(\infty)) = Q(b_s - \gamma) > 0$.

(2.29)(ii) On the other hand, $X_A(\infty) = \psi(X(\infty)) = \psi(\gamma) = 0$.

Note that in this case $p = P(b_h - X(\infty)) = P(b_h - \gamma) < \frac{1}{2}$,

because (2.22) implies $0 = \psi(\gamma) \ge 1 + \gamma\{2P(b_h - \gamma) - 1\}$;

$\gamma = b_s$:

By (2.28), $X(\infty) = b_s = \gamma$ and $q = Q(b_s - X(\infty)) = 0$.

(2.29)(iii) However, in this case, $X_A(\infty) = \psi(X(\infty)) = \psi(\gamma) = 0$.

As in the previous case, $p = P(b_h - b_s) = P(b_h - \gamma) < \frac{1}{2}$.

We shall show that if $\gamma \neq b_s$,

$$X(\infty) - X(t) \sim Ae^{-at} \quad \text{as } t \to \infty,$$
$$\frac{T_n}{\log n} \to \frac{1}{a} \quad \text{in probability as } n \to \infty$$

and give expressions for a in case $(\gamma - b_s)$ is positive or negative. If $\gamma = b_s$,

however, we find a very different behavior:

$$X(\infty) - X(t) \sim \frac{1}{at} \quad \text{as } t \to \infty,$$

$$\frac{T_n}{n^{1/2} \log n} \quad \text{is of exact order 1 in probability as } n \to \infty$$

and, again, we give an expression for a in this case.

The biological interest of these results lies in the fact that for a certain balance ($\gamma = b_s$) between the initial amounts of substrate [nb_s] and hormone [nb_h], the model predicts a much longer duration of the growth process than for other parameter values. This phenomenon, which is indeed a very essential aspect of the model, lends itself to experimental verification. Because the experiments considered here are very expensive and time consuming, this verification has not yet been carried out. Another aspect of these results that is of some practical importance is that some of the secondary metabolites are known to be synthesized only at the end of the cell culture's growth process, and one would, therefore, like to avoid values of γ close to b_s.

From a mathematical point of view, the results of this paper are rather more delicate than those of Theorem 2.1. Though the main feature of Theorem 2.1 is the uniformity in t, the theorem still provides very little information about the behavior of $X_n(t)$ and $X(t)$ for very large t, which is needed here. There seem to be few results on the duration of processes similar to the one we study. Kurtz (1982) discusses a case where the limit process X reaches its ultimate value in finite time; Barbour (1975) and Nagaev and Mukhomor (1975) study the duration of an epidemic. The problems that these authors face are very different from ours.

In Section 3, we prove the results on the behavior of $X(\infty) - X(t)$ for large t. A result on a class of differential equations that plays a key role in this analysis is given in Appendix A. In Section 4 we tackle the estimation of T_n. A maximal inequality and a fluctuation inequality that are needed in Section 4 are given in Appendix B.

3. Tail behavior of X. In this section we investigate the behavior of $X(t)$ for large values of t. Our starting point will be expression (2.19): For $t \geq c$,

$$X'(t) = X_A^0(t-c)Q(b_s - X(t)) = \left(X_A(\infty) + \left[X_A^0(t-c) - X_A(\infty) \right] \right)$$
$$\times \left(q + \left[Q(b_s - X(t)) - q \right] \right),$$

and Taylor expansion of the terms in square brackets. By (2.17) and because both P and Q have bounded derivatives of every order, we find

$$X_A^0(t-c) - X_A(\infty) = X(\infty) - X(t) - 2\int_{X(t-c)}^{X(\infty)} P(b_h - u)\, du$$

$$(3.1) \qquad\qquad = (X(\infty) - X(t)) - 2p(X(\infty) - X(t-c))$$

$$+ \mathscr{O}\left((X(\infty) - X(t-c))^2\right),$$

$$(3.2) \quad Q(b_s - X(t)) - q = q'(X(\infty) - X(t)) + \mathscr{O}\big((X(\infty) - X(t))^2\big).$$

Here p, q and q' are defined in (2.26). By (2.27),

$$(3.3) \qquad\qquad\qquad\qquad X_A(\infty)q = 0.$$

It follows that

$$
\begin{aligned}
(3.4) \quad X'(t) =\ & (q + q'X_A(\infty))(X(\infty) - X(t)) \\
& - 2pq(X(\infty) - X(t - c)) \\
& + q'(X(\infty) - X(t))^2 \\
& - 2pq'(X(\infty) - X(t))(X(\infty) - X(t - c)) \\
& + \mathscr{O}\big(q(X(\infty) - X(t - c))^2 + X_A(\infty)(X(\infty) - X(t))^2 \\
& \qquad + (X(\infty) - X(t))(X(\infty) - X(t - c))^2\big).
\end{aligned}
$$

LEMMA 3.1.

$$X(\infty) - X(t - c) = \mathscr{O}(X(\infty) - X(t)) \quad as\ t \to \infty.$$

If $\gamma = b_s$, then

$$X(t) - X(t - c) = \mathscr{O}\big((X(\infty) - X(t))^2\big) \quad as\ t \to \infty.$$

PROOF. By (3.3), $X_A(\infty)q = 0$. If $q = 0$, then $X(\infty) = b_s$. By (3.1), $X_A^0(t - c) \le X(\infty) + X_A(\infty) \le 2b_s$, and as Q' decreases, $Q(b_s - X(t)) \le Q'(0)(X(\infty) - X(t))$. Hence $X'(t) \le 2b_sQ'(0)(X(\infty) - X(t))$ for $t > c$ and

$$\int_{t-c}^{t} \frac{X'(s)}{X(\infty) - X(s)}\, ds = \log\left(\frac{X(\infty) - X(t - c)}{X(\infty) - X(t)}\right) \le 2cb_sQ'(0),$$

so that $X(\infty) - X(t - c) \le e^{2cb_sQ'(0)}(X(\infty) - X(t))$ for $t \ge 2c$.

If $X_A(\infty) = 0$, then (3.1) implies that $0 \le X_A^0(t - c) \le X(\infty) - X(t)$, whereas $Q(b_s - X(t)) \le Q(b_s)$ as Q is increasing. Repeating the foregoing argument, we see that $X(\infty) - X(t - c) \le e^{cQ(b_s)}(X(\infty) - X(t))$ for $t \ge 2c$.

Finally, if $\gamma = b_s$, then we are in case (2.29)(iii) and $X_A(\infty) = q = 0$. Now (3.1) and (3.2) imply that for $t \ge c$,

$$X'(t) \le q'(X(\infty) - X(t))^2 + \mathscr{O}\big((X(\infty) - X(t))^3\big).$$

Integrating over $(t - c, t)$, we find

$$X(t) - X(t - c) = \mathscr{O}\big((X(\infty) - X(t - c))^2\big) = \mathscr{O}\big((X(\infty) - X(t))^2\big). \quad \square$$

Theorems 3.1, 3.2 and 3.3 deal with the three essentially different cases (2.29)(i)–(iii) that we discussed in Section 2. We note that ψ and γ are defined in (2.22) and (2.24).

THEOREM 3.1. *Let $\gamma > b_s$ and define $a = \psi(b_s)/(da_s) > 0$. Then $X(\infty) = b_s$ and there exists a positive number A, such that*

$$(3.5) \qquad\qquad \lim_{t \to \infty} e^{at}(X(\infty) - X(t)) = A.$$

PROOF. This is case (2.29)(i), so that $X(\infty) = b_s$, $X_A(\infty) = \psi(b_s) > 0$, $q = 0$ and $q' = Q'(0) = 1/(da_s)$. Hence (3.4) and Lemma 3.1 imply that

$$(3.6) \qquad X'(t) = q' X_A(\infty)(X(\infty) - X(t)) + \mathscr{O}\big((X(\infty) - X(t))^2\big).$$

The theorem follows from Lemma A.1 in Appendix A with $v = 0$, $w = a = q' X_A(\infty) = \psi(b_s)/(da_s)$ and $f(t) = X(\infty) - X(t)$. $\square$

THEOREM 3.2. Let $\gamma < b_s$ and define $a \in (0, Q(b_s - \gamma)]$ as the unique solution of $Q(b_s - \gamma) - a = 2P(b_h - \gamma)Q(b_s - \gamma)e^{ac}$. Then $X(\infty) = \gamma$ and there exists a positive number A such that

$$(3.7) \qquad \lim_{t \to \infty} e^{at}(X(\infty) - X(t)) = A.$$

PROOF. This is case (2.29)(ii), so that $X(\infty) = \gamma$, $X_A(\infty) = 0$, $q = Q(b_s - \gamma) > 0$ and $p = P(b_h - \gamma) < \frac{1}{2}$. Together with Lemma 3.1, this implies that (3.4) reduces to

$$(3.8) \qquad \begin{aligned} X'(t) &= q(X(\infty) - X(t)) - 2pq(X(\infty) - X(t - c)) \\ &\quad + \mathscr{O}\big((X(\infty) - X(t))^2\big). \end{aligned}$$

The theorem follows from Lemma A.1 in Appendix A with $v = 2pq = 2P(b_h - \gamma)Q(b_s - \gamma)$, $w = q = Q(b_s - \gamma)$ and $f(t) = X(\infty) - X(t)$. Note that, because $p = P(b_h - \gamma) < \frac{1}{2}$, we have indeed $v < w$. $\square$

THEOREM 3.3. Let $\gamma = b_s$ and define $a = (1 - 2P(b_h - b_s))/(da_s) > 0$. Then $X(\infty) = b_s$ and

$$(3.9) \qquad \lim_{t \to \infty} at(X(\infty) - X(t)) = 1.$$

PROOF. We are now in case (2.29)(iii), so that $X(\infty) = b_s = \gamma$, $X_A(\infty) = \psi(\gamma) = 0$, $q = Q(0) = 0$, $q' = Q'(0) = 1/(da_s) > 0$ and $p = P(b_h - b_s) < \frac{1}{2}$. Hence $a = q'(1 - 2p)$ is positive. Together with Lemma 3.1, this implies that (3.4) reduces to

$$(3.10) \qquad X'(t) = a(X(\infty) - X(t))^2 + \mathscr{O}\big((X(\infty) - X(t))^3\big).$$

Dividing by $(X(\infty) - X(t))^2$ and integrating, we find for $c < s < t$,

$$\frac{1}{X(\infty) - X(t)} - \frac{1}{X(\infty) - X(s)} = a(t - s) + \mathscr{O}((t - s)(X(\infty) - X(s))).$$

Dividing by at and then letting first t and then s tend to infinity, we find that $(at(X(\infty) - X(t)))^{-1}$ tends to 1. $\square$

4. Duration of the growth process. We now turn to the duration T_n of the growth process N_n. To simplify our notation we shall write

$$(4.1) \quad b_{sn} = \frac{[nb_s]}{n} \in \left(b_s - \frac{1}{n}, b_s\right], \qquad b_{hn} = \frac{[nb_h]}{n} \in \left(b_h - \frac{1}{n}, b_h\right]$$

throughout this section. In view of (2.28) and the fact that $\gamma \geq 1$ on the one hand and (2.11) and (2.12) on the other, we know that $X(\infty) \geq 1 \wedge b_s$ and $X_n(T_n) \geq 1 \wedge b_{sn}$. For $\varepsilon \in (0, 1 \wedge b_s)$ we may, therefore, define $\tau(\varepsilon)$ and $\tau_n(\varepsilon)$ by

(4.2) $$X(\infty) - X(\tau(\varepsilon)) = \varepsilon,$$

(4.3) $$\tau_n(\varepsilon) = \inf\{t \geq 0 : X_n(T_n) - X_n(t) \leq \varepsilon\},$$

with the convention that $\tau_n(\varepsilon) = 0$ if $X_n(T_n) < \varepsilon$.

LEMMA 4.1. *There exist positive numbers A and α such that for $0 < \varepsilon < 1 \wedge b_s$ and $n = 1, 2, \ldots$,*

(4.4) $$\mathscr{P}\left(\tau_n(\varepsilon) > \tau\left(\frac{\varepsilon}{2}\right)\right) \leq A \exp\{-\alpha\varepsilon^2 n\}.$$

PROOF.

$$\mathscr{P}\left(\tau_n(\varepsilon) > \tau\left(\frac{\varepsilon}{2}\right)\right) \leq \mathscr{P}\left(X_n(T_n) - X_n\left(\tau\left(\frac{\varepsilon}{2}\right)\right) > \varepsilon\right)$$

$$\leq \mathscr{P}\left(\sup_{t \geq 0}|X_n(t) - X(t)| > \frac{\varepsilon}{4}\right)$$

because $X(\infty) = X(\tau(\varepsilon/2)) + \varepsilon/2$ and $X_n(T_n) = X_n(\infty)$. Hence (4.4) follows from Theorem 2.1. □

Lemma 4.1 ensures that $\tau_n(\varepsilon)$ is bounded except on a set of exponentially small probability for every fixed $\varepsilon > 0$. Of course, T_n tends to infinity in probability as $n \to \infty$ in view of Theorem 2.1 and the results of Section 3. For our study of the first order asymptotic behavior of T_n, any bounded contribution to T_n will be irrelevant, and we may, therefore, study $T_n - \tau_n(\varepsilon)$ instead. This implies that we need only take the times needed for the final $[\varepsilon n]$ cell divisions into account, for arbitrarily small positive ε.

As in Section 3, the cases $\gamma > b_s$, $\gamma < b_s$ and $\gamma = b_s$ are essentially different and we shall discuss these cases in three separate subsections.

4.1. *The case $\gamma > b_s$.* If $\gamma > b_s$ we are in case (2.29)(i), so $X(\infty) = b_s$, $X_A(\infty) = \psi(b_s) > 0$, $q = Q(0) = 0$ and $q' = Q'(0) = 1/(da_s)$. Theorem 2.1 implies that, except on a set Ω_n^c of negligible probability for large n, $X_{An}^0(T_n) = X_{An}(T_n) = X_{An}(\infty)$ will be close to $X_A^0(\infty) = X_A(\infty) = \psi(b_s) > 0$, and the same is true for $X_{An}^0(t - c)$ for sufficiently large t. In particular, $X_{An}(T_n) > 0$ on Ω_n. Because the process X_n stops when either $X_{An} = 0$ or $X_n = b_{sn}$, we must have $X_n(T_n) = b_{sn}$ and hence $Q(b_{sn} - X_n(t))$ is approximately equal to $q'(X_n(T_n) - X_n(t))$ for large t. It follows from (2.10) that on Ω_n, $\Lambda_{Zn}(t)$ is close to $na(X_n(T_n) - X_n(t))$ for large t, where $a = \psi(b_s)/(da_s) > 0$. But this means that, going back in time from T_n, the times between the last $[\varepsilon n]$ consecutive cell divisions are approximately independent and exponentially distributed random variables with means $1/a, 1/(2a), 1/(3a), \ldots, 1/([\varepsilon n]a)$,

329

provided that $\varepsilon > 0$ is sufficiently small. This implies that $(T_n - \tau_n(\varepsilon))$, which is the sum of these variables, is asymptotic to $(\log n)/a$, and by Lemma 4.1 this remains true for T_n.

In the remainder of this subsection we make this argument rigorous. In Lemma 4.2 we study the appropriate sum of independent exponential random variables. Theorem 4.1 provides a precise statement of the result and in its proof we fill in the gaps in the heuristic argument given before.

LEMMA 4.2. *Let $V_1, V_2, \ldots$ be i.i.d. random variables that are exponentially distributed with mean 1. For $\varepsilon \in (0, 1]$ and $m \geq 1/\varepsilon$, define*

$$(4.5) \qquad S_m(\varepsilon) = \sum_{j=1}^{[\varepsilon m]} \frac{V_j}{j}.$$

Then for $\varepsilon \in (0, 1]$, $m \geq 1/\varepsilon$ and $x \geq 0$,

$$\mathscr{P}\big(|S_m(\varepsilon) - \log m| \geq x\big) \leq \frac{3}{\varepsilon} \exp\Big\{-\frac{x}{2}\Big\}.$$

PROOF. Writing $S_m(\varepsilon) = S$ we have

$$Ee^{S/2} = \prod_{j=1}^{[\varepsilon m]} \frac{j}{j - \frac{1}{2}} \leq 2 \prod_{j=2}^{[\varepsilon m]} \Big(\frac{j}{j-1}\Big)^{1/2} \leq 2(\varepsilon m)^{1/2},$$

$$\mathscr{P}(S \geq \log m + x) \leq \frac{Ee^{S/2}}{\exp\{\frac{1}{2}\log m + \frac{1}{2}x\}} \leq 2\varepsilon^{1/2}e^{-x/2},$$

$$Ee^{-S} = \prod_{j=1}^{[\varepsilon m]} \frac{j}{j+1} \leq \frac{1}{\varepsilon m},$$

$$\mathscr{P}(S \leq \log m - x) \leq \frac{Ee^{-S}}{\exp\{-\log m + x\}} \leq \frac{1}{\varepsilon}e^{-x}.$$

As $\varepsilon^{1/2} \leq \varepsilon^{-1}$, the lemma is proved. □

THEOREM 4.1. *Let $\gamma > b_s$ and define $a = \psi(b_s)/(da_s) > 0$ as in Theorem 3.1. Then, for every $\delta > 0$ there exist positive numbers A and α such that for $n = 1, 2, \ldots$ and $0 \leq x \leq n$,*

$$(4.6) \qquad \mathscr{P}\Big(\Big|T_n - \frac{\log n}{a}\Big| \geq \delta \log n + x\Big) \leq A \exp\{-\alpha x\},$$

and hence

$$(4.7) \qquad \frac{T_n}{\log n} \to \frac{1}{a} \quad \text{in probability.}$$

PROOF. Choose $\varepsilon \in (0, 1 \wedge b_s \wedge X_A(\infty))$ and consider the event

$$\Omega_n = \big\{|X_{An}(T_n) - X_A(\infty)| \leq \varepsilon\big\}.$$

In view of (2.3), (2.7), (2.11) and (2.12) this event is measurable with respect to the σ-algebra $\mathscr{F}_{Z_n}$ generated by $Z_n = (Z_{1n}, Z_{2n}, \ldots)$. Because $X_{An}(T_n) = X_{An}(\infty)$, Theorem 2.1 ensures that

$$(4.8) \qquad\qquad \mathscr{P}(\Omega_n^c) \leq A_1 \exp\{-\alpha_1 \varepsilon^2 n\}$$

for positive A_1 and α_1.

As $\varepsilon < X_A(\infty)$, we have $X_{An}(T_n) > 0$ and hence $X_n(T_n) = b_{sn}$ on Ω_n. Also, by (2.7), (2.8) and (4.3),

$$\sup_{t \geq \tau_n(\varepsilon)} X_{An}^0(t - c) \leq \sup_{t \geq \tau_n(\varepsilon)} X_{An}(t)$$

$$= X_{An}(T_n) + \sup_{t \geq \tau_n(\varepsilon)} (X_{An}(t) - X_{An}(T_n))$$

$$\leq X_{An}(T_n) + X_n(T_n) - X_n(\tau_n(\varepsilon)) \leq X_{An}(T_n) + \varepsilon$$

and

$$\inf_{t \geq \tau_n(\varepsilon) + c} X_{An}^0(t - c) = X_{An}(T_n) + \inf_{t \geq \tau_n(\varepsilon)} (X_{An}^0(t) - X_{An}^0(T_n))$$

$$\geq X_{An}(T_n) - 2(X_n(T_n) - X_n(\tau_n(\varepsilon))) \geq X_{An}(T_n) - 2\varepsilon.$$

It follows that on Ω_n,

$$\sup_{t \geq \tau_n(\varepsilon)} X_{An}^0(t - c) \leq X_A(\infty) + 2\varepsilon,$$

$$\inf_{t \geq \tau_n(\varepsilon) + c} X_{An}^0(t - c) \geq X_A(\infty) - 3\varepsilon.$$

On Ω_n, $X_n(T_n) = b_{sn}$, and hence (2.9) and (4.3) imply that for $t \geq \tau_n(\varepsilon)$,

$$\frac{1}{d(a_s + \varepsilon)}(b_{sn} - X_n(t)) \leq Q(b_{sn} - X_n(t)) \leq \frac{1}{da_s}(b_{sn} - X_n(t)).$$

Define $a = X_A(\infty)/(da_s) = \psi(b_s)/(da_s)$ as in the statement of the theorem and choose $\delta \in (0, a^{-1})$. Combining (2.10) and the inequalities derived so far, we find that by taking $\varepsilon > 0$ sufficiently small, we can make sure that on Ω_n,

$$(4.9) \qquad \Lambda_{Zn}(t) \leq \frac{a}{1 - \delta a} n(b_{sn} - X_n(t)) \quad \text{for } t \geq \tau_n(\varepsilon),$$

$$(4.10) \qquad \Lambda_{Zn}(t) \geq \frac{a}{1 + \delta a} n(b_{sn} - X_n(t)) \quad \text{for } t \geq \tau_n(\varepsilon) + c.$$

Conditionally on $Z_n = z$ with $\{Z_n = z\} \subset \Omega_n$, the process $(N_n(t) - n)$ is a counting process with intensity $\Lambda_{Zn}(t)$, which is bounded above and below by (4.9) and (4.10). In view of the argument in the first paragraph of this subsection, this implies that conditionally on $Z_n = z$ with $\{Z_n = z\} \subset \Omega_n$,

$(T_n - \tau_n(\varepsilon))$ is stochastically larger than a sum of $[\varepsilon n]$ independent exponential random variables with means $(1 - \delta a)/(aj)$, $j = 1, 2, \ldots, [\varepsilon n]$, or equivalently, than $((1 - \delta a)/a)S_n(\varepsilon)$ with $S_n(\varepsilon)$ as in Lemma 4.2 with $m = n$. Hence by (4.8),

$$\mathscr{P}\left(T_n - \frac{\log n}{a} \le -\delta \log n - x\right)$$

$$\le \mathscr{P}\left(T_n - \tau_n(\varepsilon) \le \frac{1 - \delta a}{a} \log n - x\right)$$

$$\le \mathscr{P}\left(S_n(\varepsilon) \le \log n - \frac{ax}{1 - \delta a}\right) + A_1 \exp\{-\alpha_1 \varepsilon^2 n\}.$$

Similarly, (4.10) implies that conditionally on $Z_n = z$ with $\{Z_n = z\} \subset \Omega_n$, $(T_n - \tau_n(\varepsilon))$ is stochastically smaller than $((1 + \delta a)/a)S_n(\varepsilon) + c$, and hence by Lemma 4.1 and (4.8),

$$\mathscr{P}\left(T_n - \frac{\log n}{a} \ge \delta \log n + x\right)$$

$$\le \mathscr{P}\left(T_n - \tau_n(\varepsilon) \ge \frac{1 + \delta a}{a} \log n + x - \tau\left(\frac{\varepsilon}{2}\right)\right)$$

$$\quad + A_2 \exp\{-\alpha_2 \varepsilon^2 n\}$$

$$\le \mathscr{P}\left(S_n(\varepsilon) \ge \log n + \frac{a}{1 + \delta a}\left(x - \tau\left(\frac{\varepsilon}{2}\right) - c\right)\right)$$

$$\quad + A_3 \exp\{-\alpha_3 \varepsilon^2 n\}$$

for appropriate positive A_2, A_3, α_2 and α_3. Combining these results with Lemma 4.2, we find that for $n = 1, 2, \ldots$ and $\tau(\varepsilon/2) + c \le x \le n$,

$$\mathscr{P}\left(\left|T_n - \frac{\log n}{a}\right| \ge \delta \log n + x\right)$$

$$\le \mathscr{P}\left(|S_n(\varepsilon) - \log n| \ge \frac{a}{1 + \delta a}\left(x - \tau\left(\frac{\varepsilon}{2}\right) - c\right)\right) + \frac{A}{2} \exp\{-\alpha n\}$$

$$\le \frac{3}{\varepsilon} \exp\left\{\frac{a}{1 + \delta a}\left(\tau\left(\frac{\varepsilon}{2}\right) + c - x\right)\right\} + \frac{A}{2} \exp\{-\alpha n\}$$

$$\le A \exp\{-\alpha x\}$$

for appropriately chosen positive A and α. If $A \ge \exp\{\alpha(\tau(\varepsilon/2) + c)\}$, then this bound remains valid for $0 \le x < \tau(\varepsilon/2) + c$ and the proof of (4.6) is complete. For $x = \delta \log n$, (4.6) yields (4.7). $\square$

4.2. *The case* $\gamma < b_s$. If $\gamma < b_s$ we are in case (2.29)(ii), so $X(\infty) = \gamma < b_s$, $X_A(\infty) = 0$, $q = Q(b_s - \gamma) > 0$ and $p = P(b_h - \gamma) < \frac{1}{2}$. Theorem 2.1 implies

that, except on a set $\tilde{\Omega}_n^c$ of negligible probability for large n, $X_n(T_n) = X_n(\infty)$ will be close to $X(\infty) = \gamma$, and the same is true for $X_n(t)$ for sufficiently large t. In particular, $X_n(T_n) < b_{sn}$ on $\tilde{\Omega}_n$, and hence the process X_n must stop because $X_{An}(T_n) = 0$. Moreover, on $\tilde{\Omega}_n$, $Q(b_{sn} - X_n(t))$ is close to q for large t and the probabilities P_{in} of A cells are approximately equal to p for the final $[\varepsilon n]$ cell divisions if $\varepsilon > 0$ is small [cf. (2.4)]. It follows that on $\tilde{\Omega}_n$ and for small $\varepsilon > 0$, $(T_n - \tau_n(\varepsilon))$ is approximately equal to the duration of a process that starts at time $\tau_n(\varepsilon)$ with $N_{An}(\tau_n(\varepsilon))$ A cells—of which $N_{An}^0(\tau_n(\varepsilon))$ are waiting for a stimulus—and has a fixed stimulus rate $q > 0$ and a fixed probability of A cells $p < \frac{1}{2}$. Both $N_{An}(\tau_n(\varepsilon))$ and $N_{An}^0(\tau_n(\varepsilon))$ are of exact order n with probability close to 1, because $X_{An}(\tau_n(\varepsilon))$ and $X_{An}^0(\tau_n(\varepsilon))$ are close to $X_A(\tau(\varepsilon))$ and $X_A^0(\tau(\varepsilon))$, which are positive.

Let $S(p, q, c)$ be the duration of a process that starts with a single waiting A cell and has a fixed stimulus rate $q > 0$ and a fixed probability of A cells $p < \frac{1}{2}$. In Lemma 4.3 we show that the right tail of the distribution of $S(p, q, c)$ behaves like that of an exponential distribution with mean a^{-1}, where a is the solution of $q - a = 2pqe^{ac}$. Hence $(T_n - \tau_n(\varepsilon))$ is approximately distributed as the maximum of a (random) number M of independent exponentially distributed random variables with mean a^{-1}, and this number M is of exact order n. But this means that $(T_n - \tau_n(\varepsilon)) \sim (\log n)/a$, and by Lemma 4.1 we also have $T_n \sim (\log n)/a$.

In the remainder of this subsection we first prove Lemma 4.3 concerning the distribution of $S(p, q, c)$. Theorem 4.2 provides a precise formulation of the result for T_n and some additional details will be found in the proof of this theorem.

Thus in Lemma 4.3 we consider the following situation. At time $t = 0$ there is a single A cell waiting for a stimulus. A cells independently receive a stimulus after an exponential waiting time with mean $(\tilde{q})^{-1}$ and divide a constant time c later. With each division the new cells independently become A cells with probability $\tilde{p}$ and B cells with probability $(1 - \tilde{p})$. B cells do not divide. Let $S(\tilde{p}, \tilde{q}, c)$ denote the time until the final division.

LEMMA 4.3. *Suppose that $\tilde{q} > 0$ and $0 \leq \tilde{p} < \frac{1}{2}$ and define $\tilde{a} \in (0, \tilde{q}]$ as the unique solution of $\tilde{q} - \tilde{a} = 2\tilde{p}\tilde{q}e^{\tilde{a}c}$. Then there exists a positive number $\tilde{A}$ such that*

(4.11)
$$\lim_{t \to \infty} e^{\tilde{a}t} \mathscr{P}\big(S(\tilde{p}, \tilde{q}, c) > t\big) = \tilde{A}.$$

PROOF. Writing $f(t) = \mathscr{P}(S(\tilde{p}, \tilde{q}, c) > t)$ we find for $t > c$,

$$f(t) = e^{-\tilde{q}(t-c)} + 2\tilde{p}(1 - \tilde{p})\tilde{q}\int_0^{t-c} e^{-\tilde{q}s} f(t - c - s)\, ds$$

$$+ \tilde{p}^2 \tilde{q}\int_0^{t-c} e^{-\tilde{q}s}\Big[2f(t - c - s) - (f(t - c - s))^2\Big]\, ds$$

$$= e^{-\tilde{q}(t-c)} + 2\tilde{p}\tilde{q}\int_0^{t-c} e^{-\tilde{q}s} f(t-c-s)\,ds$$

$$- \tilde{p}^2\tilde{q}\int_0^{t-c} e^{-\tilde{q}s}\big(f(t-c-s)\big)^2\,ds$$

$$= e^{-\tilde{q}(t-c)}\left[1 + 2\tilde{p}\tilde{q}\int_0^{t-c} e^{\tilde{q}s} f(s)\,ds - \tilde{p}^2\tilde{q}\int_0^{t-c} e^{\tilde{q}s}\big(f(s)\big)^2\,ds\right].$$

Multiplying by $e^{\tilde{q}(t-c)}$, differentiating and dividing again by $e^{\tilde{q}(t-c)}$, we obtain for $t > c$,

$$f'(t) + \tilde{q}f(t) = 2\tilde{p}\tilde{q}f(t-c) - \tilde{p}^2\tilde{q}\big(f(t-c)\big)^2$$

or

$$f'(t) = 2\tilde{p}\tilde{q}f(t-c)\left(1 - \frac{\tilde{p}}{2}f(t-c)\right) - \tilde{q}f(t).$$

Obviously, f is strictly decreasing and $f' < 0$ on $[c,\infty)$. Because $0 \le p < \frac{1}{2}$, we have $\lim_{t\to\infty} f(t) = 0$ and we may apply Lemma A.1 in Appendix A with $v = 2\tilde{p}\tilde{q}$ and $w = \tilde{q}$ to complete the proof. $\square$

THEOREM 4.2. *Let* $\gamma < b_s$ *and define* $a \in (0, Q(b_s - \gamma)]$ *as the unique solution of* $Q(b_s - \gamma) - a = 2P(b_h - \gamma)Q(b_s - \gamma)e^{ac}$ *as in Theorem 3.2. Then, for every* $\delta > 0$ *there exist positive numbers* A *and* α *such that for* $n = 1, 2, \ldots$ *and* $0 \le x \le n$,

$$(4.12) \qquad \mathscr{P}\left(\left|T_n - \frac{\log n}{a}\right| \ge \delta \log n + x\right) \le A \exp\{-\alpha x\},$$

and hence

$$(4.13) \qquad \frac{T_n}{\log n} \to \frac{1}{a} \quad \text{in probability.}$$

PROOF. Choose $\varepsilon \in (0, 1 \wedge (b_s - \gamma))$ and define the event

$$\tilde{\Omega}_n = \{|X_n(T_n) - X(\infty)| \le \varepsilon\}.$$

By (2.3), (2.11) and (2.12, $\tilde{\Omega}_n$ is measurable with respect to the σ-algebra $\mathscr{F}_{Zn}$ generated by $Z_n = (Z_{1n}, Z_{2n}, \ldots)$, and by Theorem 2.1,

$$(4.14) \qquad \mathscr{P}\big(\tilde{\Omega}_n^c\big) \le A_1 \exp\{-\alpha_1 \varepsilon^2 n\}$$

for positive A_1 and α_1. On $\tilde{\Omega}_n$, $X_n(T_n) \le X(\infty) + \varepsilon = \gamma + \varepsilon < b_s$ and hence $X_n(T_n) < b_{sn}$ for $n \ge n_0$. It follows that for $n \ge n_0$, $X_{An}(T_n) = 0$ on $\tilde{\Omega}_n$.

As $q = Q(b_s - \gamma) > 0$ and $p = P(b_h - \gamma) < \frac{1}{2}$, we can choose $\varepsilon' \in (0, q \wedge (\frac{1}{2} - p))$ and define

$$(4.15) \qquad \begin{aligned} &p_1 = (p - \varepsilon') \vee 0 \in \left[0, \tfrac{1}{2}\right), &&p_2 = p + \varepsilon' \in \left(0, \tfrac{1}{2}\right), \\ &q_1 = q - \varepsilon' > 0, &&q_2 = q + \varepsilon' > 0. \end{aligned}$$

By taking $\varepsilon > 0$ sufficiently small we can obviously ensure that on $\tilde{\Omega}_n$ and for $n \geq n_1$ and $t \geq \tau_n(\varepsilon)$,

$$
(4.16) \qquad p_1 \leq P\left(b_{hn} - X_n(t) + \frac{1}{n}\right) \leq p_2,
$$

$$
q_1 \leq Q(b_{sn} - X_n(t)) \leq q_2.
$$

Consider a process $X_n^+(t)$ obtained by modifying $X_n(t)$ as follows: For $t > \tau_n(\varepsilon)$ the rate $Q(b_{sn} - X_n(t))$ and the probability $P(b_{hn} - X_n(t) + 1/n)$ are replaced by q_1 and p_2, respectively, and all A cells present at time $\tau_n(\varepsilon)$ are replaced by A cells waiting for a stimulus. On $\tilde{\Omega}_n$ and for $n \geq n_1$, (4.16) implies that X_n^+ is obtained from X_n by adding a random number of A cells and increasing the length of the cell cycle of a number of cells by a random amount, and as a result the duration T_n^+ of X_n^+ is stochastically larger than T_n on $\tilde{\Omega}_n$ for $n \geq n_1$. Moreover, $(T_n^+ - \tau_n(\varepsilon))$ is distributed as the maximum of $N_{An}(\tau_n(\varepsilon))$ independent and identically distributed random variables, each distributed as $S(p_2, q_1, c)$ discussed in Lemma 4.3. As $N_{An}(\tau_n(\varepsilon)) \leq n(1 + b_s)$, (4.14) and Lemma 4.1 yield for $n \geq n_1$,

$$
\mathscr{P}(T_n \geq t) \leq \mathscr{P}(T_n^+ \geq t) + \mathscr{P}(\tilde{\Omega}_n^c)
$$

$$
\leq \mathscr{P}(T_n^+ - \tau_n(\varepsilon) \geq t - \tau(\varepsilon/2)) + A_2 \exp\{-\alpha_2 \varepsilon^2 n\}
$$

$$
\leq n(1 + b_s)\mathscr{P}(S(p_2, q_1, c) \geq t - \tau(\varepsilon/2)) + A_2 \exp\{-\alpha_2 \varepsilon^2 n\}
$$

for positive A_2 and α_2, and all t.

Define a as in the statement of the theorem as the solution of $q - a = 2pqe^{ac}$. Similarly, suppose that $\bar{a}$ satisfies $q_1 - \bar{a} = 2p_2 q_1 e^{\bar{a}c}$. As $p_2 > p$ and $q_1 < q$, we have $\bar{a} < a$ and $\bar{a} \uparrow a$ as ε' in (4.15) tends to zero. Hence, for every $\delta > 0$, we can choose $\varepsilon' > 0$ sufficiently small to ensure that $a^{-1} \leq \bar{a}^{-1} \leq a^{-1} + \delta$. By Lemma 4.3 we find that for $n \geq n_1$ and $2\tau(\varepsilon/2) \leq x \leq n$,

$$
\mathscr{P}\left(T_n - \frac{\log n}{a} \geq \delta \log n + x\right)
$$

$$
(4.17) \qquad \leq n(1 + b_s)\mathscr{P}\left(S(p_2, q_1, c) \geq \left(\frac{1}{a} + \delta\right)\log n + x - \tau(\varepsilon/2)\right)
$$

$$
+ A_2 \exp\{-\alpha_2 \varepsilon^2 n\} \leq n(1 + b_s)\overline{A} \exp\left\{-\log n - \frac{\bar{a}x}{2}\right\}
$$

$$
+ A_2 \exp\{-\alpha_2 \varepsilon^2 n\} \leq A_3 \exp\{-\alpha_3 x\}
$$

for positive $\overline{A}$, A_3 and α_3. Obviously an appropriate choice of A_3 will guarantee the validity of this bound for all n and $0 \leq x \leq n$.

We may also modify the process $X_n(t)$ by replacing $Q(b_{sn} - X_n(t))$ and $P(b_{hn} - X_n(t) + 1/n)$ by q_2 and p_1 for $t > \tau_n(\varepsilon)$, and simply removing all A cells that have received a stimulus before or at time $\tau_n(\varepsilon)$. On $\tilde{\Omega}_n$ and for $n \geq n_1$, (4.16) obviously implies that the duration T_n^- of this new process $X_n^-(t)$ is stochastically smaller than T_n. Moreover, $(T_n^- - \tau_n(\varepsilon))$ is distributed

as the maximum of $N_{An}^0(\tau_n(\varepsilon))$ independent copies of $S(p_1, q_2, c)$. As $X'(t) > 0$ for all $t > c$, (2.14), (2.15) and (2.17) imply that for every $\varepsilon > 0$ there exists $\eta > 0$ such that $X_A^0(t) \geq 2\eta$ for $0 \leq t \leq \tau(\varepsilon/2)$. It follows from Theorem 2.1 and Lemma 4.1 that

$$\mathscr{P}\big(N_{An}^0(\tau_n(\varepsilon))\big) \leq \eta n\big) \leq A_4 \exp\{-\alpha_4 n\}$$

for positive A_4 and α_4. Hence, for $n \geq n_1$, (4.14) yields

$$\mathscr{P}(T_n \leq t) \leq \mathscr{P}(T_n^- \leq t) + \mathscr{P}\big(\tilde{\Omega}_n^c\big)$$
$$\leq \big[\mathscr{P}(S(p_1, q_2, c) \leq t)\big]^{\eta n} + A_5 \exp\{-\alpha_5 n\}$$

for positive A_5 and α_5.

Define $\tilde{a}$ as the solution of $q_2 - \tilde{a} = 2p_1 q_2 e^{\tilde{a}c}$. As $p_1 \leq p$ and $q_2 > q$, we have $\tilde{a} > 0$ and $\tilde{a} \downarrow a$ as ε' in (4.15) tends to zero. Hence, for every $\delta > 0$ we can choose $\varepsilon' > 0$ sufficiently small to ensure that $a^{-1} - \delta \leq \tilde{a}^{-1} \leq a^{-1}$. By Lemma 4.3 we find that for $n \geq n_1$ and $0 \leq x \leq n$,

$$\mathscr{P}\left(T_n - \frac{\log n}{a} \leq -\delta \log n - x\right)$$

$$(4.18) \qquad \leq \left[\mathscr{P}\left(S(p_1, q_2, c) \leq \left(\frac{1}{a} - \delta\right)\log n - x\right)\right]^{\eta n} + A_5 \exp\{-\alpha_5 n\}$$

$$\leq \left[1 - \tilde{A}\exp\{-\log n + \tilde{a}x\}\right]^{\eta n} + A_5 \exp\{-\alpha_5 \eta\}$$

$$\leq \exp\{-\tilde{A}\eta e^{\tilde{a}x}\} + A_5 \exp\{-\alpha_5 n\} \leq A_6 \exp\{-\alpha_6 x\}$$

for positive $\tilde{A}$, A_6, and α_6. Obviously the bound will hold for all n and $0 \leq x \leq n$ for an appropriate choice of A_6. Together, (4.17) and (4.18) prove (4.12). Taking $x = \delta \log n$ in (4.12) we complete the proof of the theorem. $\square$

4.3. *The case* $\gamma = b_s$. If $\gamma = b_s$ we are in case (2.29)(iii), so $X(\infty) = b_s = \gamma$, $X_A(\infty) = 0$, $q = Q(0) = 0$ and $p = P(b_h - b_s) < \frac{1}{2}$. The process X_n may stop because $X_n(T_n) = b_{sn}$ or $X_{An}(T_n) = 0$, and in contrast to the two previous cases, neither of these possibilities can be ruled out with large probability. We shall, therefore, have to deal with both possibilities and our approach will combine the main elements of the proofs in the two previous subsections.

In the cases $\gamma > b_s$ and $\gamma < b_s$, either $(b_{sn} - X_n(T_n))$ or $X_{An}(T_n)$ equals zero and the other one of these two quantities is of exact order 1 with high probability. Because now both $b_s - X(\infty) = 0$ and $X_A(\infty) = 0$, the latter part of this statement is no longer true and we shall have to assess the exact order of magnitude of the nonzero quantity among $(b_{sn} - X_n(T_n))$ and $X_{An}(T_n)$. In view of the complicated dependence of these two random variables, some care is needed here. We shall proceed by bounding the nonzero variable in terms of $Z_n = (Z_{1n}, Z_{2n}, \ldots)$ in Lemma 4.4, and then showing that this implies that it is of exact order $n^{-1/2}$ in probability in Corollary 4.1.

LEMMA 4.4. *Let* $\gamma = b_s$. *If* $X_n(T_n) = b_{sn}$, *then*

(4.19)
$$\frac{1}{n}\left|\sum_{i=1}^{[nb_s]}(Z_{in} - 2P_{in})\right| - \frac{3}{n} \le X_{An}(T_n)$$

$$\le \frac{1}{n}\left|\sum_{i=1}^{[nb_s]}(Z_{in} - 2P_{in})\right| + \frac{3}{n}.$$

If $X_{An}(T_n) = 0$, *then*

(4.20)
$$\frac{1}{n}\left|\sum_{i=1}^{[nb_s]}(Z_{in} - 2P_{in})\right| - \frac{3}{n}$$

$$\le b_{sn} - X_n(T_n)$$

$$\le b_s\left\{\frac{1}{n}\max_{1 \le m \le [nb_s]}\left|\sum_{i=1}^{m}(Z_{in} - 2P_{in})\right| + \frac{2}{n}\right\}.$$

PROOF. Because $\gamma = b_s$, we have by (2.22) and (2.24),

(4.21)
$$2\int_0^{b_s}P(b_h - u)\,du = b_s - 1.$$

Because P is nondecreasing, this implies for $0 \le v \le b_s$,

(4.22)
$$2\int_{b_s-v}^{b_s}P(b_h - u)\,du \le \frac{(b_s - 1)v}{b_s}.$$

If $X_n(T_n) = b_{sn}$ we use (2.7), (4.21) and straightforward algebra to obtain

$$X_{An}(T_n) = \frac{1}{n}\sum_{i=1}^{[nb_s]}(Z_{in} - 2P_{in}) + 2\left(\frac{1}{n}\sum_{i=1}^{nb_{sn}}P_{in} - \int_0^{b_{sn}}P(b_h - u)\,du\right)$$

(4.23)
$$+ \int_{b_{sn}}^{b_s}(1 - 2P(b_h - u))\,du$$

$$= \frac{1}{n}\sum_{i=1}^{[nb_s]}(Z_{in} - 2P_{in}) + \mathscr{R},$$

where $|\mathscr{R}| \le 3/n$ by (2.6) and because $0 \le P(u) \le 1$ for all u and $0 \le b_s - b_{sn} < 1/n$. Because $X_{An}(T_n)$ is nonnegative, it also equals the absolute value of the expression on the right in (4.23), and (4.19) follows.

If $X_{An}(T_n) = 0$, we again use (2.7) and (4.21) to obtain

(4.24)
$$b_{sn} - X_n(T_n) = 2\int_0^{b_s}P(b_h - u)\,du$$

$$- \frac{1}{n}\sum_{i=1}^{nX_n(T_n)}Z_{in} - (b_s - b_{sn}).$$

Hence by (4.22),

$$b_{sn} - X_n(T_n) \le \frac{b_s - 1}{b_s}(b_s - X_n(T_n)) + 2\int_0^{X_n(T_n)} P(b_h - u)\, du$$

$$-\frac{1}{n}\sum_{i=1}^{nX_n(T_n)} Z_{in} - (b_s - b_{sn})$$

$$\le \frac{b_s - 1}{b_s}(b_{sn} - X_n(T_n)) - \frac{1}{n}\sum_{i=1}^{nX_n(T_n)}(Z_{in} - 2P_{in})$$

$$- 2\left(\frac{1}{n}\sum_{i=1}^{nX_n(T_n)} P_{in} - \int_0^{X_n(T_n)} P(b_h - u)\, du\right)$$

or

$$b_{sn} - X_n(T_n) \le b_s\left\{-\frac{1}{n}\sum_{i=1}^{nX_n(T_n)}(Z_{in} - 2P_{in})\right.$$

$$\left. - 2\left(\frac{1}{n}\sum_{i=1}^{nX_n(T_n)} P_{in} - \int_0^{X_n(T_n)} P(b_h - u)\, du\right)\right\},$$

and the inequality on the right in (4.20) follows from (2.6).

To prove the lower bound in (4.20) we start once more with (4.24) and write

$$b_{sn} - X_n(T_n) - \frac{1}{n}\sum_{i=nX_n(T_n)+1}^{[nb_s]} Z_{in}$$

$$= 2\int_0^{b_s} P(b_h - u)\, du - \frac{1}{n}\sum_{i=1}^{[nb_s]} Z_{in} - (b_s - b_{sn})$$

$$= -\frac{1}{n}\sum_{i=1}^{[nb_s]}(Z_{in} - 2P_{in}) - \mathscr{R}$$

with $\mathscr{R}$ as in (4.23) so that $|\mathscr{R}| \le 3/n$. Because $0 \le Z_{in} \le 2$ for all i and n, this yields

$$\frac{1}{n}\left|\sum_{i=1}^{[nb_s]}(Z_{in} - 2P_{in})\right| - \frac{3}{n} \le \left|b_{sn} - X_n(T_n) - \frac{1}{n}\sum_{i=nX_n(T_n)+1}^{[nb_s]} Z_{in}\right|$$

$$\le b_{sn} - X_n(T_n),$$

which completes the proof of the lemma. $\square$

As before, let $\mathscr{F}_{Zn}$ denote the σ-algebra generated by $Z_n = (Z_{1n}, Z_{2n}, \ldots)$.

COROLLARY 4.1. *Let $\gamma = b_s$. For $0 < b < B$, define the events*

(4.25) $$\Omega_{1n} = \{b \le n^{1/2}X_{An}(T_n) \le B\},$$

(4.26) $$\Omega_{2n} = \{b \le n^{1/2}(b_{sn} - X_n(T_n)) \le B\}.$$

These events are $\mathscr{F}_{Zn}$-measurable and for every $\varepsilon > 0$ there exist $0 < b < B$ such that for sufficiently large n,

$$(4.27) \qquad \mathscr{P}(\Omega_{1n} \cup \Omega_{2n}) \geq 1 - \varepsilon.$$

PROOF. $\mathscr{F}_{Zn}$-measurability follows from (2.11), (2.12) and (2.7). Obviously, $b_{sn} - X_n(T_n) = 0$ on Ω_{1n} and $X_{An}(T_n) = 0$ on Ω_{2n}. Hence (4.27) follows from Lemma 4.4, the central limit theorem and Lemma B.1 in Appendix B. $\square$

So we have shown that both $(b_{sn} - X_n(T_n))$ and $X_{An}(T_n) = X_{An}^0(T_n)$ are either equal to zero or of exact order $n^{-1/2}$ in probability. However, to analyze the process $N_n(t)$ for large t we shall have to determine the exact order of magnitude of both factors $Q(b_{sn} - X_n(t))$ and $N_{An}^0(t - c) = nX_{An}^0(t - c)$ of the conditional intensity $\Lambda_{Zn}(t)$ of the process [cf. (2.10)]. The factor $Q(b_{sn} - X_n(t))$ is monotone in t and our knowledge concerning $(b_{sn} - X_n(t))$ will suffice. Determining the exact order of $X_{An}^0(t - c)$ for large t is a more delicate matter. In Lemma 4.5 we establish an asymptotic expression for $X_{An}^0(t - c)$ for large n and t in terms of $X_{An}(T_n)$ and $(X_n(T_n) - X_n(t))$. A key step in obtaining this expression is to show that $(X_n(t) - X_n(t - c))$—and hence the difference between $X_{An}^0(t - c)$ and $X_{An}(t)$—is negligible for our purposes.

LEMMA 4.5.   *Let* $\gamma = b_s$, *so* $p = P(b_h - b_s) < \frac{1}{2}$. *Then for every* $D > 0$,

$$(4.28) \qquad \sup_{t \geq \tau_n(Dn^{-1/2})+c} (X_n(t) - X_n(t - c)) = \mathscr{O}_{\mathscr{P}}\left(\frac{\log n}{n}\right),$$

$$(4.29) \qquad \sup_{t \geq \tau_n(Dn^{-1/2})+c} \Big| X_{An}^0(t - c) - \{X_{An}(T_n) \\ + (1 - 2p)(X_n(T_n) - X_n(t))\}\Big| = \mathscr{O}_{\mathscr{P}}(n^{-3/4}).$$

PROOF.   Take $\varepsilon > 0$. In (4.25) and (4.26) we choose $0 < b < B$ so that (4.27) holds for sufficiently large n. On $\Omega_n^* = \Omega_{1n} \cup \Omega_{2n}$ we have for $t \geq \tau_n(Dn^{-1/2})$,

$$X_{An}^0(t - c) = X_{An}(T_n) + (X_n(T_n) - X_n(t)) - \sum_{i = nX_n(t-c)+1}^{nX_n(T_n)} Z_{in}$$

$$\leq (B + D)n^{-1/2},$$

$$Q(b_{sn} - X_n(t)) \leq Q'(0)(b_{sn} - X_n(t)) \leq \frac{(B + D)}{da_s}n^{-1/2},$$

because $Q'(0) = 1/(da_s)$. By (2.10) this implies that on Ω_n^*,

$$\sup_{t \geq \tau_n(Dn^{-1/2})} \Lambda_{Zn}(t) \leq \lambda = \frac{(B + D)^2}{da_s}.$$

Conditionally on $Z_n = z$ with $\{Z_n = z\} \subset \Omega_n^*$, N_n is a counting process with intensity $\Lambda_{Zn}(t) \leq \lambda$. Hence, if $\Pi(t)$ denotes a unit Poisson process and k a positive integer,

$$
\mathscr{P}\left(\sup_{t \geq \tau_n(Dn^{-1/2})+c} \left(X_n(t) - X_n(t - c) \right) \geq \frac{2k}{n} \right)
$$

$$
\leq \mathscr{P}\left(\sup_{\substack{t \geq c \\ \Pi(\lambda t) \leq Dn^{1/2}}} \left(\Pi(\lambda t) - \Pi(\lambda(t - c)) \right) \geq 2k \right) + \varepsilon
$$

$$
\leq \mathscr{P}\left(\sup_{\substack{t \geq \lambda c \\ \Pi(t) \leq Dn^{1/2}}} \left(\Pi(t) - \Pi(t - \lambda c) \right) \geq 2k \right) + \varepsilon
$$

$$
\leq \frac{Dn^{1/2}}{k} \exp\{2\lambda c - k\} + \varepsilon,
$$

by Lemma B.2 in Appendix B. Because $\lambda = \lambda(\varepsilon)$ is finite for every $\varepsilon > 0$, this proves (4.28).

By (2.7) and (2.8),

$$
(4.30) \quad
\begin{aligned}
X_{An}^0(t - c) &- \{X_{An}(T_n) + (1 - 2p)(X_n(T_n) - X_n(t))\} \\
&= -\frac{1}{n} \sum_{i = nX_n(t-c)+1}^{nX_n(T_n)} Z_{in} - \frac{1}{n} \sum_{i = nX_n(t)+1}^{nX_n(T_n)} (Z_{in} - 2P_{in}) \\
&\quad - 2\left(\frac{1}{n} \sum_{i = nX_n(t)+1}^{nX_n(T_n)} (P_{in} - p) \right).
\end{aligned}
$$

Because $0 \leq Z_{in} \leq 2$, the first term on the right in (4.30) is bounded in absolute value by $2(X_n(t) - X_n(t - c)) = \mathscr{O}_{\mathscr{P}}((\log n)/n)$ uniformly for $t \geq \tau_n(Dn^{-1/2}) + c$ by (4.28). To deal with the next term we note that $nX_n(T_n) \leq [nb_s]$ and that for $t \geq \tau_n(Dn^{-1/2}) + c$, $nX_n(t) \geq [nb_s] - ([nb_s] - nX_n(T_n)) - [Dn^{1/2}] = [nb_s] - \mathscr{O}_{\mathscr{P}}(n^{1/2})$ by Corollary 4.1. Application of Lemma B.1 in Appendix B for $M = \mathscr{O}(n^{1/2})$ yields that the second term on the right in (4.30) is $\mathscr{O}_{\mathscr{P}}(n^{-3/4})$ uniformly for $t \geq \tau_n(Dn^{-1/2}) + c$. Finally, (2.4), (2.5) and Corollary 4.1 imply that uniformly for $t \geq \tau_n(Dn^{-1/2}) + c$,

$$
\begin{aligned}
0 &\leq \frac{1}{n} \sum_{i = nX_n(t)+1}^{nX_n(T_n)} (P_{in} - p) \\
&\leq (X_n(T_n) - X_n(t))(P(b_{hn} - X_n(t)) - P(b_h - b_s)) \\
&\leq Dn^{-1/2}P'(b_h - p_s)\left(\frac{1}{n} + b_{sn} - X_n(T_n) + Dn^{-1/2} \right) = \mathscr{O}_{\mathscr{P}}(n^{-1}).
\end{aligned}
$$

Together with (4.30), these estimates establish (4.29) and the lemma. $\square$

We are now in a position to determine the exact order of magnitude of T_n. Roughly speaking, we shall argue that if $X_n(T_n) = b_{sn}$, the situation is similar to the one in Theorem 4.1, the main difference being that in the present case $X_{An}^0(t - c)$, and hence $\Lambda_{Zn}(t)$, is smaller by a factor $n^{1/2}$ for large t. This implies that T_n is of order $n^{1/2} \log n$ rather than $\log n$ as is the case in Theorem 4.1. Similarly, if $X_{An}(T_n) = 0$, the situation is comparable to that

in Theorem 4.2, but now $Q(b_{sn} - X_n(t))$ is smaller by a factor $n^{1/2}$ for large t. Again the conclusion is that T_n is of order $n^{1/2} \log n$ instead of $\log n$ as in Theorem 4.2. The basic reason underlying all of this is that in the case $\gamma = b_s$ we have $X_A(\infty) = b_s - X(\infty) = 0$, and hence the nonzero quantity among $X_{An}(T_n)$ and $(b_{sn} - X_n(T_n))$ is of order $n^{-1/2}$ by Corollary 4.1. As we already noted, this is essentially different from what happens if $\gamma > b_s$ or $\gamma < b_s$, when either $X_A(\infty)$ or $(b_s - X(\infty))$ is positive and hence either $X_{An}(T_n)$ or $(b_{sn} - X_n(T_n))$ is of exact order 1.

THEOREM 4.3.   *Let $\gamma = b_s$. Then for every $\varepsilon > 0$ there exist positive numbers a and A such that for $n = 2, 3, \dots,$*

$$(4.31) \qquad \mathscr{P}\big(an^{1/2} \log n \le T_n \le An^{1/2} \log n\big) \ge 1 - \varepsilon,$$

and hence

$$(4.32) \qquad \frac{T_n}{n^{1/2} \log n} \text{ is of exact order 1 in probability.}$$

PROOF.   Take $\varepsilon > 0$ and define Ω_{1n} and Ω_{2n} as in (4.25) and (4.26) with $0 < b < B$ such that (4.27) holds for sufficiently large n. By Lemma 4.1, Theorem 3.3 and Lemma 4.5, we can also choose positive numbers C and D, and an event $\overline{\Omega}_n$ with $\mathscr{P}(\overline{\Omega}_n) \ge 1 - \varepsilon$ and such that on $\overline{\Omega}_n$,

$$(4.33) \qquad \tau_n\big(Dn^{-1/2}\big) \le \tau\left(\frac{D}{2}n^{-1/2}\right) \le Cn^{1/2},$$

$$(4.34) \qquad \begin{aligned} \sup_{t \ge \tau_n(Dn^{-1/2})+c} & \big|X_{An}^0(t - c) \\ & - \{X_{An}(T_n) + (1 - 2p)(X_n(T_n) - X_n(t))\}\big| \\ & \le Cn^{-3/4}. \end{aligned}$$

Note that on $\Omega_{1n} \cup \Omega_{2n}$ we also have the trivial inequality

$$(4.35) \qquad \begin{aligned} X_{An}^0(t - c) &\le X_{An}(t) \le X_{An}(T_n) + (X(T_n) - X_n(t)) \\ & \le (B + D)n^{-1/2} \end{aligned}$$

for $t \ge \tau_n(Dn^{-1/2})$.

Given $Z_n = z$ with $\{Z_n = z\} \subset \Omega_{1n}$, $(N_n(t) - n)$ is a counting process with conditional intensity $\Lambda_{Zn}(t)$. On $\Omega_{1n} \cap \overline{\Omega}_n$, (4.25), (4.34) and (4.35) imply that for sufficiently large n,

$$\Lambda_{Zn}(t) \geq (bn^{1/2} - Cn^{1/4})Q'(b_s)(b_{sn} - X_n(t))$$

$$\text{for } t \geq \tau_n(Dn^{-1/2}) + c,$$

(4.36)

$$\Lambda_{Zn}(t) \leq (B + D)n^{1/2}Q'(0)(b_{sn} - X_n(t))$$

$$\text{for } t \geq \tau_n(Dn^{-1/2}).$$

By redefining Λ_{Zn} on the subset of Ω_{1n}, where these inequalities do not hold, we can ensure that (4.36) is satisfied on Ω_{1n} while changing the process only on a subset of $\Omega_{1n} \cap \overline{\Omega}_n^c$. In the proof of Theorem 4.1 we now replace $\tau_n(\varepsilon)$ by $\tau_n(Dn^{-1/2})$ and (4.9) and (4.10) by (4.36), and repeat the argument following (4.10) to conclude that $(T_n - \tau_n(Dn^{-1/2}))$ is stochastically bounded above and below by two constant multiples of $n^{1/2}S_m(\varepsilon)$, where $S_m(\varepsilon)$ is as in Lemma 4.2 with $\varepsilon m = Dn^{1/2}$. Application of Lemma 4.2 yields the existence of positive numbers a_1 and A_1 such that

(4.37)
$$\mathscr{P}\big(a_1 n^{1/2} \log n \leq T_n - \tau_n(Dn^{-1/2}) \leq A_1 n^{1/2} n^{1/2} \log n | \Omega_{1n}\big)$$
$$\geq 1 - \varepsilon - \mathscr{P}\big(\overline{\Omega}_n^c | \Omega_{1n}\big)$$

for sufficiently large n.

On Ω_{2n}, (4.26) implies that for $t \geq \tau_n(Dn^{-1/2})$,

(4.38)
$$q_1 n^{-1/2} \leq Q(b_{sn} - X_n(t)) \leq q_2 n^{-1/2}$$

for positive $q_1 = Q'(b_s)b$ and $q_2 = Q'(0)(B + D)$. Arguing as in the proof of Theorem 4.2, we find that (4.35) and (4.38) imply that on Ω_{2n}, $(T_n - \tau_n(Dn^{-1/2}))$ is stochastically smaller than the maximum of $(B + D)n^{1/2}$ independent copies of $S(p_2, q_1 n^{-1/2}, c)$, with $S(\tilde{p}, \tilde{q}, c)$ as in Lemma 4.3. Moreover, $S(p_2, q_1 n^{-1/2}, c)$ is distributed as $n^{1/2}S(p_2, q_1, cn^{-1/2})$, which is stochastically smaller than $n^{1/2}S(p_2, q_1, c)$.

On the other hand, on Ω_{2n}, $(T_n - \tau_n(Dn^{-1/2}))$ is stochastically larger than the maximum of $N_{An}^0(\tau_n(Dn^{-1/2}))$ independent copies of $S(p_1, q_2 n^{-1/2}, c)$. Also $S(p_1, q_2 n^{-1/2}, c)$ is stochastically larger than $n^{1/2}S(p_1, q_2, 0)$ and on $\Omega_{2n} \cap \overline{\Omega}_n$, $N_{An}^0(\tau_n(Dn^{-1/2})) \geq bn^{1/2} - Cn^{1/4}$ by (4.34). As in the proof of Theorem 4.2, we apply Lemma 4.3 to these upper and lower bounds and find that there exist positive numbers a_2 and A_2 such that

(4.39)
$$\mathscr{P}\big(a_2 n^{1/2} \log n \leq T_n - \tau_n(Dn^{-1/2}) \leq A_2 n^{1/2} \log n | \Omega_{2n}\big)$$
$$\geq 1 - \varepsilon - \mathscr{P}\big(\overline{\Omega}_n^c | \Omega_{2n}\big)$$

for sufficiently large n.

Because $\mathscr{P}(\overline{\Omega}_n^c) \leq \varepsilon$ and $\mathscr{P}(\Omega_{1n} \cup \Omega_{2n}) \geq 1 - \varepsilon$, (4.33), (4.37) and (4.39) ensure the validity of (4.31) for large n and, therefore, trivially for all $n \geq 2$. Because (4.32) is merely a restatement of (4.31), this completes the proof of the theorem. $\square$

APPENDIX A

LEMMA A.1. *For real numbers $c > 0$ and $0 \le v < w$, let $f: [0, \infty) \to (0, \infty)$ be continuously differentiable on (c, ∞) with derivative $f' < 0$ and $\lim_{t \to \infty} f(t) = 0$, and suppose that as $t \to \infty$,*

(A.1)
$$f'(t) = vf(t - c)(1 + \mathcal{O}(f(t - c))) \\ - wf(t)(1 + \mathcal{O}(f(t - c))).$$

Then the equation $w - a = ve^{ac}$ has a unique solution $a \in (0, w]$ and there exists a positive number A such that

(A.2)
$$\lim_{t \to \infty} e^{at} f(t) = A.$$

PROOF. Because $w - x > ve^{cx}$ for $x = 0$, ve^{cx} is nonnegative and nondecreasing in x and $(w - x)$ decreases strictly to 0 as $x \uparrow w$, the equation $w - a = ve^{ac}$ does indeed have a unique solution $a \in (0, w]$. Note that $a = w$ if $v = 0$.

Take $\varepsilon = (w - v)/4$. As $f(t) \to 0$ for $t \to \infty$, there exists $t_0 > c$ such that for $t \ge t_0$,

$$f'(t) \le (v + \varepsilon)f(t - c) - (w - \varepsilon)f(t),$$

and hence, for $t \ge t_0$,

$$f(t) = -\int_t^\infty f'(u)\, du \ge -(v + \varepsilon)\int_{t-c}^\infty f(u)\, du + (w - \varepsilon)\int_t^\infty f(u)\, du$$

$$= (w - v - 2\varepsilon)\int_t^\infty f(u)\, du - (v + \varepsilon)\int_{t-c}^t f(u)\, du$$

$$\ge \frac{w - v}{2}\int_t^\infty f(u)\, du - wcf(t - c).$$

As a result,

(A.3) $$\int_t^\infty f(u)\, du \le \frac{2}{w - v}\{f(t) + wcf(t - c)\} \to 0 \quad \text{as } t \to \infty$$

and we have shown that f is integrable.

The lemma is now trivial for $v = 0$. We have

$$\frac{f'(t)}{f(t)} = -w + \mathcal{O}(f(t - c)),$$

and for $s, t \to \infty$,

$$\log\left(\frac{e^{wt}f(t)}{e^{ws}f(s)}\right) = \int_s^t \left(\frac{f'(u)}{f(u)} + w\right) du = \mathcal{O}\left(\int_{s-c}^{t-c} f(u)\, du\right) = o(1).$$

Because $a = w$ in this case, this proves (A.2).

We may therefore assume that $0 < v < w$ and $0 < a < w$, and that

(A.4) $$f'(t) = vf(t - c) - wf(t) + \mathcal{O}((f(t - c))^2).$$

Define

(A.5) $$g(t) = e^{at}f(t), \qquad t \geq 0.$$

Rewriting (A.4) in terms of g with the aid of the equation $w - a = ve^{ac}$, we find for $t > c$,

(A.6) $$g'(t) = (w - a)(g(t - c) - g(t)) + \mathcal{O}(f(t - c)g(t - c)).$$

Hence, for an appropriate constant $C > 0$ and $t > c$,

(A.7)
$$(w - a)[(1 - \varepsilon(t))g(t - c) - g(t)]$$
$$\leq g'(t) \leq (w - a)[(1 + \varepsilon(t))g(t - c) - g(t)],$$

where $\varepsilon(t) = Cf(t - c)$.

For $k = 1, 2, \ldots$, define

$$m_k = \min_{(k-1)c \leq t \leq kc} g(t), \qquad M_k = \max_{(k-1)c \leq t \leq kc} g(t), \qquad \varepsilon_k = \varepsilon(kc).$$

Choose k_0 so that $\varepsilon_{k_0} \leq \frac{1}{2}$. For $k \geq k_0$ and $kc \leq t \leq (k + 1)c$, we have $\varepsilon_k \leq \frac{1}{2}$ and by (A.7),

(A.8)
$$- (w - a)[g(t) - (1 - \varepsilon_k)m_k]$$
$$\leq g'(t) \leq (w - a)[(1 + \varepsilon_k)M_k - g(t)].$$

For $t = kc$, both $[g(t) - (1 - \varepsilon_k)m_k]$ and $[(1 + \varepsilon_k)M_k - g(t)]$ are positive and the inequalities (A.8) for $g'(t)$ ensure that both remain so throughout the interval $kc \leq t \leq (k + 1)c$. However, this implies that for $k \geq k_0$,

(A.9) $$m_{k+1} \geq (1 - \varepsilon_k)m_k, \qquad M_{k+1} \leq (1 + \varepsilon_k)M_k.$$

Because $\varepsilon_k \leq \frac{1}{2}$ for $k \geq k_0$, we find that for every $k \geq k_0$,

(A.10)
$$\prod_{r=k}^{\infty} (1 - \varepsilon_r) \geq \exp\left\{-2 \sum_{r=k}^{\infty} \varepsilon_r\right\} \geq \exp\left\{-\frac{2C}{c} \int_{(k-2)c}^{\infty} f(u)\, du\right\},$$
$$\prod_{r=k}^{\infty} (1 + \varepsilon_r) \leq \exp\left\{\sum_{r=k}^{\infty} \varepsilon_r\right\} \leq \exp\left\{\frac{C}{c} \int_{(k-2)c}^{\infty} f(u)\, du\right\}.$$

In view of (A.3), this yields the existence of a sequence $\delta_k \downarrow 0$ and $M > m > 0$ such that for every $k \geq k_0$,

$$\inf_{r \geq k} m_r \geq m_k \prod_{r=k}^{\infty} (1 - \varepsilon_r) \geq m_k(1 - \delta_k) \geq m > 0,$$

$$\sup_{r \geq k} M_r \leq M_k \prod_{r=k}^{\infty} (1 + \varepsilon_r) \leq M_k(1 + \delta_k) \leq M < \infty.$$

It follows that for every $k \geq k_0$,

(A.11)
$$0 < m \leq m_k(1 - \delta_k) \leq \liminf_{t \to \infty} g(t) \leq \limsup_{t \to \infty} g(t)$$
$$\leq M_k(1 + \delta_k) \leq M < \infty.$$

Let us look at (A.8) somewhat more carefully. Because $g(t) - (1 - \varepsilon_k)m_k > 0$ and $(1 + \varepsilon_k)M_k - g(t) > 0$ for $t \in [kc, (k+1)c]$, we find that for $kc \le t \le (k+1)c$ and $k \ge k_0$,

$$\log\left(\frac{g(t) - (1 - \varepsilon_k)m_k}{g(kc) - (1 - \varepsilon_k)m_k}\right) = \int_{kc}^{t} \frac{g'(u)}{g(u) - (1 - \varepsilon_k)m_k}\, du \ge -(w - a)c,$$

and after replacing $g(t)$ by m_{k+1},

(A.12) $m_{k+1} \ge (1 - \varepsilon_k)m_k + e^{-(w-a)c}\left[g(kc) - (1 - \varepsilon_k)m_k\right].$

Similarly, the right-hand inequality in (A.8) ensures that for $k \ge k_0$,

(A.13) $M_{k+1} \le (1 + \varepsilon_k)M_k - e^{-(w-a)c}\left[(1 + \varepsilon_k)M_k - g(kc)\right],$

and subtracting (A.12) from (A.13) we see that for $k \ge k_0$,

$$M_{k+1} - m_{k+1} \le (1 - e^{-(w-a)c})\left[(1 + \varepsilon_k)M_k - (1 - \varepsilon_k)m_k\right]$$
$$\le \beta(M_k - m_k) + 2M\varepsilon_k,$$

where $\beta = 1 - \exp\{-(w - a)c\} \in (0, 1)$. By iterating this inequality, we find that for $r \ge 1$,

$$M_{k_0 + r} - m_{k_0 + r} \le \beta^r\left(M_{k_0} - m_{k_0}\right) + 2M \sum_{\nu=1}^{r} \beta^{r-\nu}\varepsilon_{k_0 + \nu - 1}$$

$$\le \beta^r M + M \sum_{j=[(r+1)/2]}^{\infty} \beta^j + 2M \sum_{j=k_0 + [r/2]}^{\infty} \varepsilon_j$$

$$\le \beta^r M + \frac{M}{1 - \beta}\beta^{r/2} + \frac{2MC}{c}\int_{1/2(2k_0 + r - 5)c}^{\infty} f(u)\, du.$$

As $0 < \beta < 1$, (A.3) ensures that

(A.14) $\lim_{k \to \infty} (M_k - m_k) = 0,$

and because $\delta_k \downarrow 0$, (A.11) and (A.14) show that $g(t)$ tends to a positive and finite limit as $t \to \infty$. In view of (A.5), the proof is complete. $\square$

APPENDIX B

LEMMA B.1. *If* $X_1, X_2, \ldots$ *are independent bounded random variables,* $0 \le X_j \le a$ *for* $j = 1, 2, \ldots$, *then for all* $M \in \mathbb{N}$ *and every* $x \ge 0$,

(B.1) $\mathscr{P}\left(\max_{1 \le m \le M}\left|\sum_{j=1}^{m}(X_j - EX_j)\right| \ge x\right) \le 4\exp\left\{-\frac{2x^2}{9a^2M}\right\}.$

PROOF. The lemma follows from Theorem 2 in Hoeffding (1963) combined with Lévy's inequality [Shorack and Wellner (1986), page 844]. $\square$

LEMMA B.2. *If $\Pi(t)$ is a unit Poisson process, a and b are positive numbers and $m \in \mathbb{N}$, then*

$$(B.2) \qquad \mathscr{P}\left(\sup_{\substack{t \geq b \\ \Pi(t) \leq a}} (\Pi(t) - \Pi(t - b)) \geq 2m \right) \leq \frac{a}{m} \exp\{2b - m\}.$$

PROOF. Let $0 = Y_0 < Y_1 < Y_2 < \cdots$ be the consecutive jump times of Π. Then

$$\mathscr{P}\left(\sup_{\substack{t \geq b \\ \Pi(t) \leq a}} (\Pi(t) - \Pi(t - b)) \leq 2m - 1 \right)$$

$$= \mathscr{P}\left(\min_{2m - 1 \leq k \leq a} (Y_k - Y_{k - (2m - 1)}) \geq b \right)$$

$$\geq \mathscr{P}\left(\min_{1 \leq r \leq [a/m]} (Y_{rm} - Y_{(r - 1)m}) \geq b \right)$$

$$= \{\mathscr{P}(\Pi(b) \leq m - 1)\}^{[a/m]}.$$

Because $\Pi(b)$ has a Poisson distribution with mean b, we have $\mathscr{P}(\Pi(b) \geq m)$ $\leq \exp\{2b - m\}$ and hence for $a \geq m$ and $m \geq 2b$,

$$\mathscr{P}\left(\sup_{\substack{t \geq b \\ \Pi(t) \leq a}} (\Pi(t) - \Pi(t - b)) \leq 2m - 1 \right)$$

$$\geq (1 - \exp\{2b - m\})^{a/m} \geq 1 - \frac{a}{m} \exp\{2b - m\},$$

which proves (B.2) for $a \geq m$ and $m \geq 2b$. For $a < m$ or $m < 2b$, (B.2) is trivially true. $\square$

REFERENCES

ALBERTS, B., BRAY, D., LEWIS, J., RAFF, M., ROBERTS, K. and WATSON, J. D. (1989). *Molecular Biology of the Cell*, 2nd ed. Garland, New York.

BARBOUR, A. D. (1975). The duration of a closed stochastic epidemic. *Biometrika* **62** 477–482.

BAYLISS, M. W. (1985). The regulation of the cell division cycle in cultured plant cells. In *The Cell Division Cycle in Plants* (J. A. Bryant and D. Francis, eds.). Cambridge Univ. Press.

BROOKS, R. F., BENNETT, D. C. and SMITH, J. A. (1980). Mammalian cells need two random transitions. *Cell* **19** 493–504.

CASTOR, F. A. L. (1980). A G_1 rate model accounts for cell-cycle kinetics attributed to 'transition probability.' *Nature* **287** 857–859.

COOPER, S. (1982). The continuum model: statistical implications. *J. Theoret. Biol.* **94** 783–800.

DE GUNST, M. C. M. (1989). *A Random Model for Plant Cell Population Growth. CWI Tract* **58**. Math. Centrum, Amsterdam.

DE GUNST, M. C. M., HARKES, P. A. A., VAL, J., VAN ZWET, W. R. and LIBBENGA, K. R. (1990). Modelling the growth of a batch culture of plant cells: a corpuscular approach. *Enzyme Microb. Technol.* **12** 61–71.

DE GUNST, M. C. M. and VAN ZWET, W. R. (1992). A non-Markovian model for cell population growth: speed of convergence and central limit theorem. *Stochastic Process. Appl.* **41** 297–324.

HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30.

KURTZ, T. G. (1982). Representation and approximation of counting processes. In *Advances in Filtering and Optimal Stochastic Control Lecture Notes in Control and Inform. Sci.* **42** 177–191. Springer, New York.

MORRIS, P., SCRAGG, A. H., STAFFORD, A. and FOWLER, M. W., EDS. (1986). *Secondary Metabolism in Plant Cell Cultures*. Cambridge Univ. Press.

NAGAEV, A. V. and MUKHOMOR, T. P. (1975). A limit distribution of the duration of an epidemic. *Theory Probab. Appl.* **20** 805–818.

NELSON, S. and GREEN, P. J. (1981). The random transition model of the cell cycle. A critical review. *Cancer Chemother. Pharmacol.* **6** 11–18.

ROELS, J. A. (1983). *Energetics and Kinetics in Biotechnology*. North-Holland, Amsterdam.

SHIELDS, R. (1977). Transition probability and the origin of variation in the cell cycle. *Nature* **267** 704–707.

SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.

SMITH, J. A. and MARTIN, L. (1973). Do cells cycle? *Proc. Nat. Acad. Sci. U.S.A.* **70** 1263–1267.

STREET, H. E. (1973). *Plant Tissue and Cell Culture. Botanical Monographs* **11**. Blackwell, Oxford.

TREWAVAS, A. J. (1985). Growth substances, calcium and the regulation of cell division. In *The Cell Division Cycle in Plants*. (J. A. Bryant and D. Francis, eds.). Cambridge Univ. Press.

DEPARTMENT OF MATHEMATICS
FREE UNIVERSITY
DE BOELELAAN 1081 A
1081 HV AMSTERDAM
THE NETHERLANDS

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF LEIDEN
P.O. BOX 9512
2300 RA LEIDEN
THE NETHERLANDS
AND
DEPARTMENT OF STATISTICS
UNIVERSITY OF NORTH CAROLINA
CHAPEL HILL, NORTH CAROLINA 27599-3260