THE STOCHASTIC THEORY OF CELL PROLIFERATION

BURT V. BRONK, G. J. DIENES, *and* ARTHUR PASKIN

From Queens College of City University of New York, New York 11367, and the Brookhaven National Laboratory, Upton, Long Island, New York 11973

ABSTRACT A stochastic theory of cell kinetics has been developed based on a realistic model of cell proliferation. A characteristic transit time, $\bar{t}_i$, has been assigned to each of the four states (G_1, S, G_2, M) of the cell cycle. The actual transit time, t_i, for any cell is represented by a distribution around $\bar{t}_i$ with a variance σ_i^2. Analytic and computer formulations have been used to describe the time development of such characteristics as age distribution, labeling experiments, and response to perturbations of the system by, for example, irradiation and temperature. The decay of synchrony is analyzed in detail and is shown to proceed as a damped wave. From the first few peaks of the synchrony decay one can obtain the distribution function for the cell cycle time. The later peaks decay exponentially with a characteristic decay constant, λ, which depends only on the average cell-cycle time, $\bar{T}$, and the associated variance. It is shown that the system, upon any sudden disturbance, approaches new "equilibrium" proliferation characteristics via damped periodic transients, the damping being characterized by λ. Thus, the response time of the system, $\bar{T}/\lambda$, is as basic a parameter of the system as the cell-cycle time.

I. INTRODUCTION

The experiments in which Howard and Pelc (1) studied the schedule of DNA synthesis in plant cells initiated a series of investigations into the manner in which the cell cycle is divided into functional periods. The development of labeling methods (2, 3) stimulated the use of mathematical models. Quastler in particular emphasized the importance of models (4, 5) in understanding the cell kinetics of renewal systems.

Since changes in parameters characterizing the various parts of the cycle produce complex effects in experimental observations, it is not surprising that the use of mathematical models has become necessary for the interpretation of results. What at first is surprising, in view of the extensive literature in the area, is the fact that many of the more recent investigators have not made use of earlier theoretical work. This work pertains to populations, with application both in genetics and actuarial problems, and investigations of problems in industrial replacement. The exponential age distribution for example, which has been used to obtain the per-

centages of a cell population in a given state (6), dates back at least to the first decade of this century (7–10). The decaying wave phenomena observed in various experiments involving synchronized populations of cells was anticipated by the work of Lotka on the time development of birthrate of a human population starting with ancestors born on the same date. (See reference 11 for a good review summarizing many early results.)

A large part of the development of the mathematical theory of populations seems to have been due to Lotka. The central tool used in these investigations is the renewal equation. This is an integral equation giving the average number of individuals, or alternatively their birthrate at a given time, in terms of a nonlinear integral equation involving a convolution of the same quantity at earlier times with the probability density of generation times. Lotka (12) credits Herbelot (13) with the first use of this equation within the context of the subject of "selfrenewing aggregates." The earliest reference Lotka gives to the integral equation is reference 14. Feller's (15) clear and rigorous mathematical paper, in which he investigates conditions of validity of the important methods and statements about solutions of the renewal equation, is very valuable as a guide to the literature written before 1941.

More recently, some underlying probabilistic models which lead to the renewal equation have been investigated (16–18). From the stochastic point of view one might for example attempt to calculate the *probability* of a given cell having n descendants at time t after it is born. This approach contains more information than just the *average* at time t, and may be of biological importance when a small number of cells is considered.

A concept important in the understanding of the cell cycle is the idea that the physiological age of a given cell is *not* equivalent to the amount of time since the cell was born. This idea has been incorporated into a probabilistic theory by Kendall (19), who generalized the birth-and-death process introduced by Feller (20). Kendall's approach is to break the life cycle into k independent sequential states with the probability of a cell which is in a given state at time t entering the next state by time $t + dt$ given by $e^{-\lambda t}\, dt$, where λ is a constant. The physiological age of the cell corresponds to the particular one of the k states which it occupies. When the cell leaves the kth state, it divides into two cells in the first state.

We incorporate the idea of physiological age in our model (defined in section II) by the use of four states, each with a different and independent probability function which determines the lifetime in that state. This model is similar to the one used by Williams (21) to discuss the age distributions of insects. Our ith state of cell growth corresponds in Williams' model to the ith instar or immature stage of an insects life. The age within a state is identified in the present model with the time a cell has spent in a state.

Another approach taking into consideration the difference between the physiological and chronological age of cells was put forward by Von Foerster (24). In this method a differential equation is written which expresses the assumption that

the density of cells of a given physiological age is a function of that age and the time with respect to some known initial conditions. If a small change in physiological age is equal or proportional to a small change in time, the differential equation used simply sets the total derivative of the number of cells with respect to time equal to a biologically determined loss function times the density at the same age and time (25). The stochastic nature of cell proliferation does not appear as compatible with this approach as with others, but many of the asymptotic results are similar.

In addition to the analytical approaches discussed above, modern computers allow the numerical investigation of theoretical models of cell kinetics. The strength of the numerical approach is its capability in handling complicated interactions without the use of drastic simplifying assumptions. The earliest application (known to us) of these methods of cell proliferation is a Monte Carlo study by Hoffman, Metropolis, and Gardiner (26). Others have made use of some of the analytical models discussed above combined with numerical methods to produce solutions in cases of experimental interest (28–32).

In the present paper we use both analytical and numerical methods to extend the theory in a manner suitable for comparison with experiments. For our numerical work we have developed computer programs utilizing both the Monte Carlo method and an averaging technique which is more convenient for the study of large populations. Most of the analytical results are derived using renewal theory. We have considered in detail useful quantitative formulations for effects such as decay of synchrony, kinetic behavior of labeled cells, time development of age-in-state and of waiting time-in-state distributions, asymptotic time dependent characteristics, and disturbances of the proliferating system.

The computer methods are used to test the domain of validity of various approximations, and to obtain graphs which display the characteristics of interest for particular values of the parameters. These methods should be particularly useful for comparison with experiments where the complexity of the experimental situation renders an analytic description too cumbersome.

The computer techniques have been particularly useful in enabling us to examine the response of the system to disturbances. An example of this is the description of the response of cell proliferation to the sudden change in the characteristic lifetime of one of the phases. Another example is the behavior of a system during and after exposure to radiation. We have already used the Monte Carlo technique for a quantitative description of cell response to radiation and subsequent recovery. These results, discussed elsewhere (35), are in agreement with two-dose radiation recovery experiments.

II. MODEL OF THE PROLIFERATION CYCLE

Before describing the details and the results of our calculations we shall discuss the general features of the model for the proliferation cycle. This model is designed to

provide for the fact that the cell cycle can be subdivided into a number, N, of experimentally distinguishable states. At the present time, the states of interest are determined by the condition of the cell with respect to DNA synthesis, and mitosis. In this case, N is four, and the states are G_1, S, G_2, and M.

Actually the time spent in consecutive states may or may not be independent. That is, a short-lived parent cell may tend to have long-lived daughters. Indications of such behavior have been observed (36). In the present investigation, all such correlations are neglected. The transit time of a cell through a given state is considered to be independent of its history before entering that state.

We are mainly concerned with freely growing cell cultures in generally constant environment although some sudden disturbances expressed by instantaneous changes in the parameters were also investigated. The influence of spatially neighboring cells[1] is neglected and cell death is not included in the treatment in this paper. Some interesting biological effects have clearly been neglected and the present investigation is to be regarded as a theoretical framework with which to sort out the various complicating effects to be investigated later. The algorithms used for numerical computation will be easy to modify for the purpose of relaxing such restrictions as mother-daughter independence. Analytical investigation of such cases will be guided by numerical work. Geometrical effects will be more novel, and probably a great deal more difficult to analyze.

With the above restrictions, the mathematical specification of the model is rather simple. We assign an index i to each of the four states in chronological order. The index $i = 2$, for example, indicates the state S. We assume that a density function $f_i(t)$ can be assigned to each state such that the probability of the transit time for the ith state having a value in the interval around t is given by $f_i(t)\,dt$. The probability density function, $f(t)$, for the whole cycle is then given by the 4-fold convolution of the density functions for the individual states.

Most of the results obtained apply to any regular density function. For some examples, and particularly for a comparison of numerical and analytical results, we will often find it convenient to have in mind functions approximating Gaussians[2] for our $f_i(t)$. We assume that a transit time, $\bar{t}_i$, and a variance $(\sigma_i)^2$, which are identified with the mean and second moment about the mean respectively for $f_i(t)$, can be assigned to each state. The generation time, or time spent in traversing the complete cycle, then has a mean and variance given by

$$\bar{T} = \sum_{i=1}^{4} \bar{t}_i \qquad \sigma^2 = \sum_{i=1}^{4} \sigma_i^2 \tag{1}$$

where equation 1 is independent of the functional form of the $f_i(t)$. If $f_i(t)$ is Gaussian, then $f(t)$ is also of this form, by the convolution property of Gaussians.

[1] The existence of such effects in irradiated cells, for example, has been indicated by Froese (37).

[2] Experimental generation times are reasonably well fitted by Gaussian distributions, at least for some cell lines. See, for example, references 22 and 23.

The Gaussian approximation for $f_i(t)$ cannot be strictly true, since the time spent in a state must be greater than or equal to zero. Therefore, we require

$$f_i(t) \equiv 0 \quad t < 0 \tag{2}$$

On most occasions, we use density functions with a coefficient of variation, C. V., less than one-third, where

$$(\text{C. V.})_i \equiv \sigma_i / \bar{t}_i \,. \tag{3}$$

In this case condition equation 2 has negligible effect. If $f_i(t)$ were a Gaussian for example, the fraction of the cells not within 3σ of $\bar{t}_i$, and thus affected by equation 2 is only 0.0026 of the total number traversing the state. When occasionally we make a numerical computation involving a larger C. V., we use a tailless density, satisfying equation 2, but with the parameters σ and $\bar{T}$ actually used for computation modified according to Appendix I, to give the specified C. V.

III. RENEWAL EQUATIONS

Once $f(t)$ is specified, it has been demonstrated (reference 17) that $P(t)$, the *average* number of cells present at time t, given that exactly one cell was present (just born) at time $t = 0$, is uniquely determined by the renewal equation,

$$P(t) = 2\int_0^t P(t - t')f(t')\,dt' + [1 - F(t)], \tag{4}$$

where

$$F(t) \equiv \int_0^t f(t')\,dt'. \tag{5}$$

If P_0 cells of age zero were present at $t = 0$, then the average number present at time t is given by the product $P_0 \cdot P(t)$. Taking the derivative with respect to time of equation 4 we obtain the equation for $\dot{P}$, the derivative with respect to time of the average number of cells,

$$\dot{P}(t) = f(t) + 2\int_0^t \dot{P}(t - t')f(t')\,dt'. \tag{6}$$

If we change variables we can write equation 6 in a more convenient form

$$\dot{P}(t) = f(t) + 2\int_0^t \dot{P}(t')f(t - t')\,dt'. \tag{7}$$

Substitution of the right-hand side of equation 7 for $\dot{P}(t')$ inside the integral in 7

gives

$$\dot{P}(t) = f(t) + 2\int_0^t f(t)f(t - t')\,dt' + 2^2\int_0^t\left[\int_0^t \dot{P}(t'')f(t' - t'')\,dt''\right]f(t - t')\,dt'. \quad (8)$$

By repeated iteration of this substitution, we obtain

$$\dot{P}(t) = \sum_{n=1}^{\infty} 2^{n-1}f^{(n)}(t), \quad (9)$$

where $f^{(n)}$ is given by

$$f^{(n)}(t) \equiv \int_0^t f^{(n-1)}(t')f(t - t')\,dt', \ (n = 2, \cdots)$$
$$f^{(1)}(t) \equiv f(t). \quad (10)$$

$f^{(n)}(t)\,dt$ is the probability that the sum of the lifetimes of all the ancestors of a cell in the nth generation[3] has a value between t and $t + dt$. Hence, according to the central limit theorem of probability, $f^{(n)}(t)$ approaches a Gaussian distribution for large n, regardless of the form of $f(t)$, that is

$$f^{(n)}(t) \to g_n(t) = [2\pi n\sigma^2]^{-1/2}\exp[-(t - n\bar{T})^2/2n\sigma^2]. \quad (11)$$

Since we require that f be identically zero for negative argument, the limits of integration in equation 10 may be extended to minus and plus infinity.

We will interpret $\dot{P}(t)$ as the rate of increase at time t of the number of cells present. This interpretation is valid[4] when the number of cells present in an experiment can be represented closely by a smoothly varying function, $P(t)$. In general, the requirement for validity is that the number of cells present be large. (A brief discussion of fluctuations is given in Appendix II.) It is conceptually useful to define the birthrate, $B(t)$,

$$B(t) = 2\dot{P}(t). \quad (12)$$

Since every time two daughter cells are born exactly one cell is lost, $B(t)\,dt$ gives the number of cells in the population born during the interval of time around t, per initial cell, and equation 7 becomes

$$B(t) = 2\int_0^t [B(t') + \delta(t')]f(t - t')\,dt', \qquad t > 0$$
$$B(0) = 1. \quad (13)$$

[3] The nth generation is counted starting from the zeroth generation which consists of the cells just born at $t = 0$.

[4] The reason for qualifying our remarks is that the actual number (as opposed to the average number) of cells present changes discontinuously.

The integrand in equation 13 gives the number of divisions on the average occurring at time t, due to the cells born at an earlier time t'. The Dirac delta function in the integrand accounts for the single cell just born at the initial epoch. This formulation is particularly useful for the discussion of the age distribution.

We can obtain useful quantitative information from equation 9 if we assume, for example, our Gaussian form for $f(t)$. Then

$$f(t) \equiv g_1(t) = (2\pi\sigma^2)^{-1/2} \exp\left[-(t - \bar{T})^2/2\sigma^2\right] \quad \text{for} \quad t > 0$$
$$f(t) \equiv 0 \text{ for } t \leqq 0. \tag{14}$$

In analytical approximations it is convenient to replace this truncated Gaussian by a Gaussian defined for all values of t. The error then incurred in equation 10 is given by the integral

$$\epsilon^{(n)}(t) = \int_t^\infty f^{(n-1)}(t')g_1(t - t')\, dt', \tag{15}$$

where $g_1(t)$ is defined as a Gaussian over its entire domain. This error is small provided $\sigma/\bar{T} \leqq 3$. Using the convolution property of Gaussians we then have approximately

$$\dot{P}(t) = \sum_{n=1}^{\infty} 2^{n-1} g_n(t), \tag{16}$$

where $g_n(t)$ is given by the right-hand side of equation 11 for all n. Integrating equation 16 we find for the average number of cells per unit cell initially present

$$P(t) = 1 + \sum_{n=1}^{\infty} 2^{n-1} G_n(t), \tag{17}$$

where

$$G_n(t) = \int_0^t g_n(t')\, dt'. \tag{18}$$

These are the basic results of the renewal theory which will be used and from which various approximations will be obtained for comparison with the computer calculations.

IV. NUMERICAL METHODS OF CALCULATION

We have employed two different methods for calculating the time development of various cell populations within the framework of our model. The first is a Monte Carlo method in which the life history of each individual cell and its progeny is followed. The second is a statistical method in which the numbers of cells transferred from one state to another at a given time is taken as directly proportional to

the probability of the transfer for cells of the appropriate age. According to the law of large numbers (38), the average of the fraction of the number transferred should approach the probability of transfer when the number of cells considered is large. We are, therefore, considering an average behavior by this method. These methods are very briefly described below.[5]

A. Monte Carlo Method

The Monte Carlo method is probably the most direct numerical method of simulating experiments with proliferating cells. The essential feature of the method is that we follow a probable or Monte Carlo history of each cell from birth through mitosis. The method requires a table of appropriately distributed "random" numbers and a tabulation scheme which is best described as a vertical list. Each cell is assigned a line on the list which contains two numbers. The first number gives the time (the waiting time, see section V-D) at which the cell will enter the state specified by the second number. A clock time, T_c, which is initially zero, is advanced by a time increment after each cycle of the following operation. Each line on the list is checked consecutively to determine if the time listed passes the criterion of being less than or equal to the clock time. If it does not, we proceed to the next line. If the criterion is met, then a new time is selected for the mth line according to

$$T_m' = T_m + t_{i+1} + \sigma_{i+1}z.$$

T_m is the time we find on the mth line, and z is a random number selected from a Gaussian distribution of zero mean and unit variance. If it turns out that

$$(t_{i+1} + \sigma_{i+1}z) < 0$$

then, in accordance with equation 14, that particular value for z is discarded, and another selected. If $i = 1, 2, 3$, we proceed to line $(m + 1)$. If $i = 4$, before proceeding to the next line, in addition to the previous operation, we create a new line at the bottom of the list with the same time (T_m') and state ($i = 1$) as we have just put on the mth line.

The Monte Carlo method is easy to use and simple to modify. It has a serious limitation, however, in that deviations from average behavior are large unless the number of cells considered, and hence the amount of computer memory used, is large.

B. Age Transfer Method

An iterative procedure which enables one to calculate averages directly and requires a great deal less storage than the Monte Carlo method was suggested and designed for us by our colleague, Peter Kemmey. In this method n locations are assigned to

[5] Details may be obtained from BNL Report 12383 (1968).

each state. The number assigned to the mth location of a given state represents the number of cells of age $m\Delta t$, where Δt is a unit of time short compared to the smallest $\bar{t}_i$. If we define

$$T_i(t, \tau) \equiv \frac{f_i(t + \tau)}{1 - F_i(t)}$$

as the transfer probability density, then $T_i\delta\tau$ is the probability that a cell which is of age t in state i will leave that state in about τ additional hours. During a given cycle the number for each of the n nonzero index locations for each state is multiplied by T_i evaluated at the appropriate argument for that location. The number which is in the mth age location of state i is multiplied by T_i $(m\Delta t, \delta\tau/2)\delta\tau$. During the in-

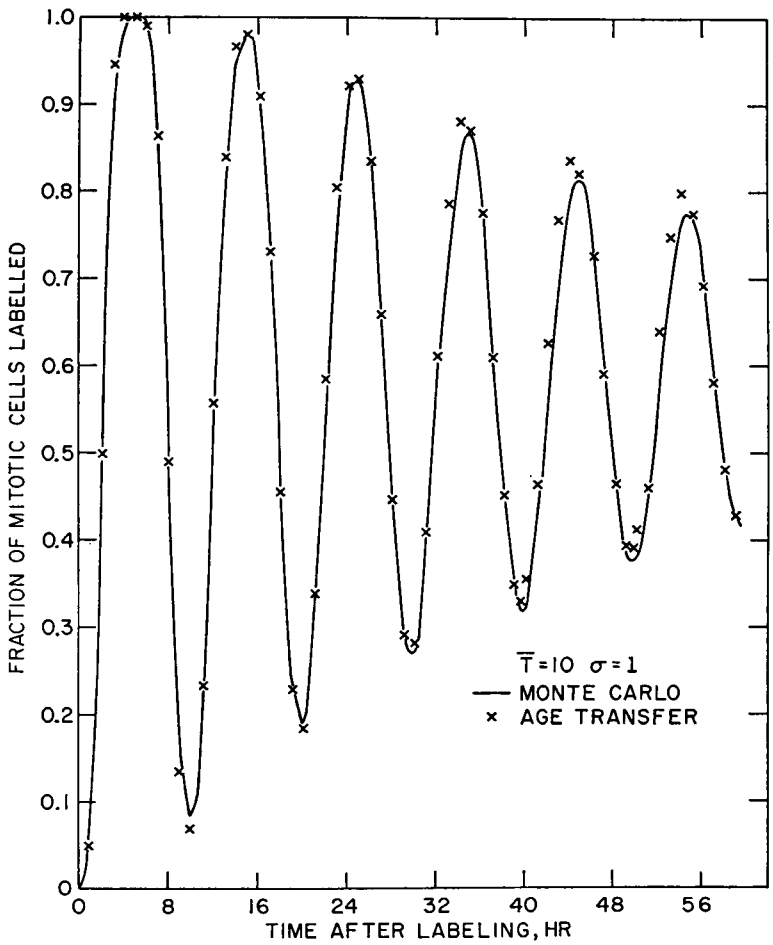

FIGURE 1 Comparison of Monte Carlo and age transfer programs in terms of labeled mitosis, $(M^*/M_0 + M^*)$, data for C.V. = 0.1. (Initial Monte Carlo population is approximately 400 cells in log-phase.)

terval of time $\delta\tau$, $P_{m,i} = (T_i \cdot \delta\tau) \cdot N_{m,i}$ cells are transferred from the mth age group of state i to the zeroth age group of state m. $N_{m,i}$ is the average number of cells of age m in state i before the transfer. Therefore, $P_{m,i}$ is added to the number which is in the zeroth location of state $(i + 1)$ and subtracted from $N_{m,i}$. The resultant of the subtraction is the new $N_{m,i}$. If i equals four, two times $P_{m,i}$ is added to the zeroth location of the first state.

Although $\delta\tau$ and Δt are conceptually different quantities it is convenient to set them equal. In our calculation we set $\delta\tau = \Delta t = 0.05$ hr. After the above operation

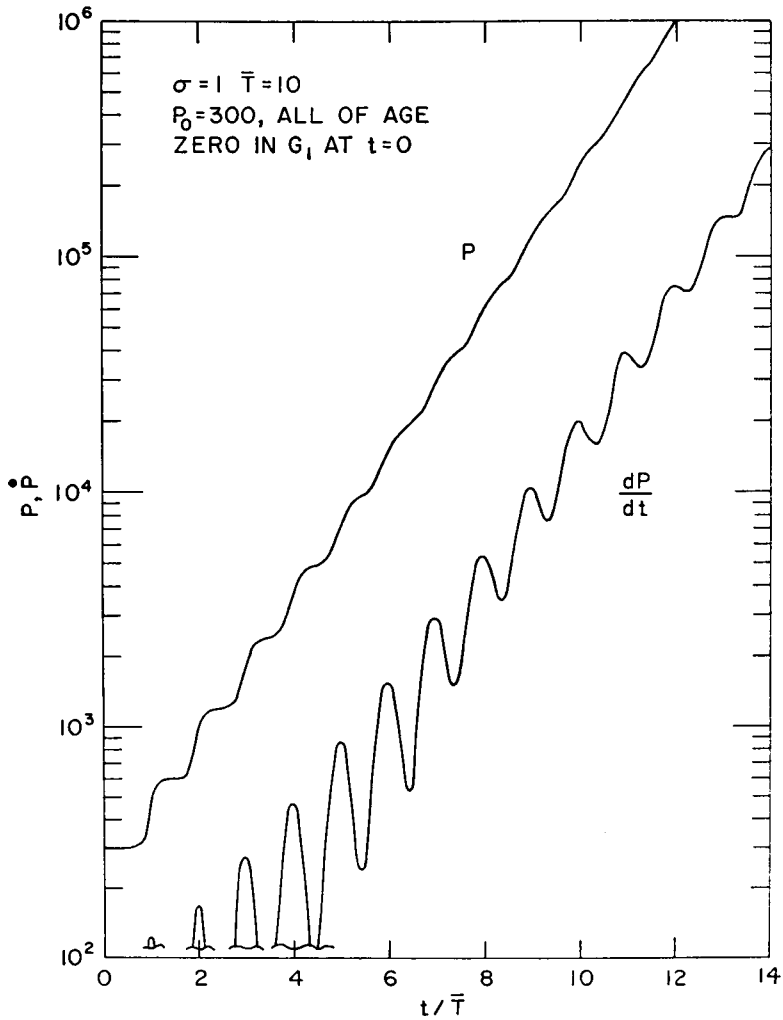

FIGURE 2 Cell population, P, and its time derivative, $\dot{P}$, vs. $t/\bar{T}$ starting with a synchronous population of age zero in G_1 at $t = 0$. $\bar{T} = 10$, $\sigma = 1$ (σ_1 : 50, σ_2 : 0.624, σ_3 : 0.50, σ_4 : 0.333).

is completed for each location, and any desired information is recorded, the number in every location is advanced into the next older location. The number of locations is chosen such that the number of cells in the nth location is negligibly small and may be discarded. The clock time is advanced by Δt after each of the above cycles.

The accuracy of the numerical methods was checked by allowing a fully synchronized population to desynchronize to log phase and computing the doubling time. The theoretical doubling times given in equation 37 were reproduced to better than three significant figures. The asymptotic percentages in the various states given in equation 61 were also checked and reproduced. A rough comparison was made of transient phenomena, as calculated by this method, against the Monte Carlo results. Satisfactory agreement between the methods was obtained as shown in Fig. 1 for per cent labeled mitosis curves.

V. APPLICATIONS OF THE MODEL

In this section a variety of applications of the model is discussed with the aid of the results of computer experiments and mathematical analysis using the renewal equations. The numerical results are checked in many cases against approximate formulae which are also developed in this section. For most of the computer experiments we have chosen parameters that simulate the behavior of a well-studied line of Chinese hamster cells (39). For this system we have taken the following transit times (in hours): l_1:1.5, l_2:6.0, l_3:1.5, l_4:1.0, $\bar{T} = 10$. The values of σ_i were varied and their values will be indicated for the various computer experiments.

A. Proliferation and Rates of Proliferation Curves

An important feature that a proliferation model must describe is the decay of synchrony for an initially synchronous population. Such results are shown in Fig. 2 (for $\sigma = 1$) starting with a population of 300 cells synchronized to be of age zero in G_1 (just born) at $t = 0$. In these graphs, P is the number of cells and $\dot{P} = dP/dt$ is its time derivative plotted against the reduced proliferation time $t/\bar{T}$. The well known "carpeted staircase" behavior is reproduced with slow approach to log-phase behavior for $\sigma = 1$ (and a much faster approach to the asymptotic behavior as σ is increased). Many peaks are seen in the derivatives for $\sigma = 1$ but only three true maxima in $\dot{P}$ exist at $\sigma = 2$. The final logarithmic slopes are given by

$$d\frac{\ln P}{dt} = \alpha = \frac{\ln 2}{\bar{T}}\left[1 + \frac{(\ln 2)\sigma^2}{2\bar{T}^2}\right]. \tag{19}$$

(This relation is derived later in section B.)

We can obtain a convenient approximation for the early maxima in $\dot{P}$ directly from the renewal equations of section III. Let us suppose that the C. V. is sufficiently small that $\epsilon^{(n)}$ given by equation 15 is negligible in comparison to the nth term of equation 15. For small n there is not much overlap between the terms of

equation 16. Hence when

$$\sqrt{n}\sigma \ll n\bar{T}$$

and t is contained in the interval

$$n\bar{T} - 2\sqrt{n}\sigma \leqq t \leqq n\bar{T} + 2\sqrt{n}\sigma,$$

which places $\dot{P}(t)$ near a maximum, then we have the approximations

$$\dot{P}(t) \approx P_0 2^{n-1} g_n(t) \tag{20}$$

and

$$P(t) \approx P_0 2^{n-1}[1 + G_n(t)] \tag{21}$$

where P_0 is the number of cells of age zero present at $t = 0$. One can also approximate P and $\dot{P}$ between the peaks (for similar conditions on σ, $\bar{T}$ and t) by placing t in the interval between $n\bar{T}$ and $(n + 1)\bar{T}$. Then

$$\dot{P}(t) \cong P_0[2^{n-1} g_n(t) + 2^n g_{n+1}(t)] \tag{22}$$

and when there is little overlap

$$P(t) \cong P_0 2^n. \tag{23}$$

For direct comparison with the maxima obtained in the computer calculations one notes that the value of $g_n(t)$ at $t = n\bar{T}$ is $1/\sqrt{2\pi n\sigma^2}$. $\dot{P}_{\max}$ is then given by

$$\dot{P}_{\max} = P_0 \frac{2^{n-1}}{\sqrt{2\pi n\sigma^2}}, \qquad n = 1, 2, \cdots \tag{24}$$

and the ratio, R, of adjacent peaks by

$$R = 2\left(\frac{n}{n+1}\right)^{1/2} \tag{25}$$

which is independent of the coefficient of variation. In this same region, P itself increases as

$$P = P_0 \tfrac{3}{2} 2^{n-1} \tag{26}$$

since at $t = n\bar{T}$ the value of $G_n(t) = \frac{1}{2}$. Thus,

$$\frac{\dot{P}_{\max}}{P} = \frac{2}{3} \frac{1}{\sqrt{2\pi n\sigma^2}}. \tag{27}$$

These relations are confirmed by the computer results for both $\sigma = 1$ (about the first 10 peaks) and $\sigma = 2$ (first three peaks) as shown in Fig. 3 where $\dot{P}_{max}/P$ is plotted against $1/\sqrt{n}$. The slopes are in excellent agreement with the theoretical slopes. For any nonzero C. V., additional terms of equation 16 (representing several generations) contribute to $\dot{P}$ after a sufficiently long time and the above approximations are then no longer appropriate.

The shapes of the curves of $\dot{P}$ vs. $t/\bar{T}$, Fig. 4, can also be interpreted by renewal

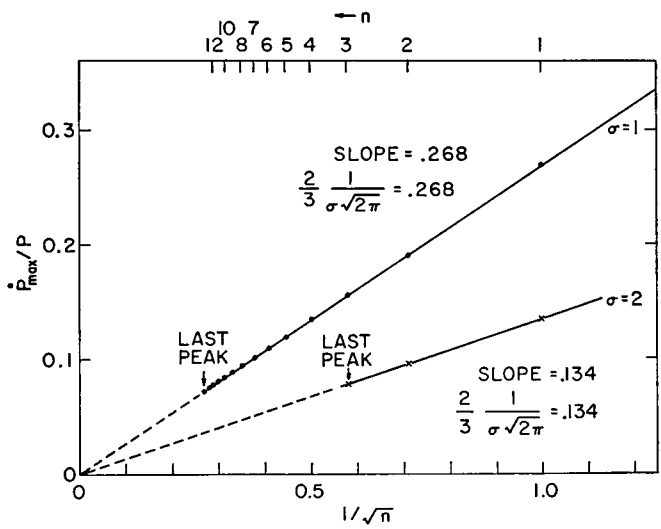

FIGURE 3 The decay of $\dot{P}_{max}/P$ vs. $1/\sqrt{n}$ for $\sigma = 1$ and $\sigma = 2$ ($\bar{T} = 10$).

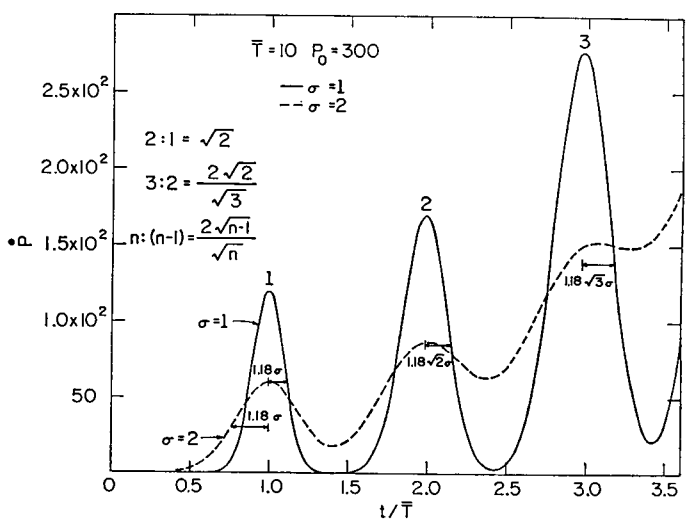

FIGURE 4 The early behavior of $\dot{P}$ for $\sigma = 1$ and $\sigma = 2$ ($\bar{T} = 10$).

theory in this early regime. According to equation 20 the curve giving rise to the first maximum is $f(t)$ itself. This is a very useful explicit result of the theory showing that the distribution function can be accurately determined from the shape of the first peak starting with a synchronous population. The higher the initial degree of synchrony the more accurately $f(t)$ will be obtained, as already pointed out by Engelberg and Hirsch (43). As illustrated in Fig. 4, the ratios of the peaks and the half width at half maximum (1.18σ) are in excellent agreement with the Gaussian used in calculation for $\sigma = 1$. There is much more overlap for $\sigma = 2$ and only the width of the first curve can be interpreted in this manner. As shown in Fig. 5, however, the overlap for $\sigma = 2$ can be taken into account as the distributions for

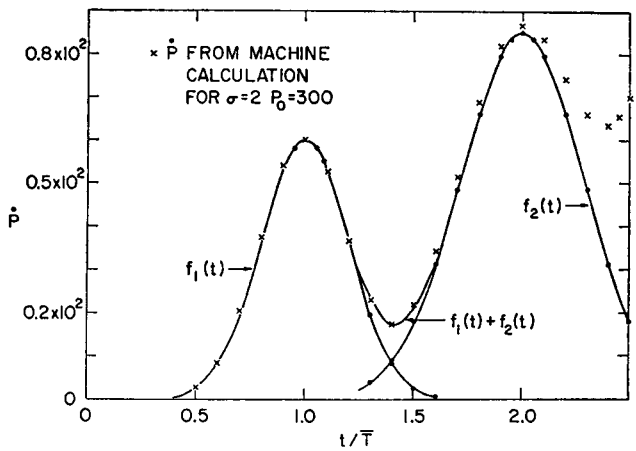

FIGURE 5 Comparison of the Gaussians used in the calculations with the shapes of the first two peaks in $\dot{P}$ ($\sigma = 2$, $\bar{T} = 10$).

the first and second peaks. The approximations used in the calculations represent the results accurately.

B. General Decay of Synchrony

In the previous section the early decay of synchrony in the proliferation rate was discussed. As overlap develops the peaks broaden and change into inflection points before they become unobservable in log-phase growth (Fig. 2). If the data are analyzed as the maxima in the quantity $(\dot{P}/P)$, these maxima persist for a larger number of cycles. The significant quantity is the maximum in $(\dot{P}/P)$ relative to the average, or longtime value, of $(\dot{P}/P)$. The data are presented as $(\dot{P}/P)_{max} - (\dot{P}/P)_{avg}$ vs. $t/\bar{T}$ in Fig. 6 for two values of σ. Past the first few peaks, where the analysis of the previous section is valid, the decay of synchrony becomes exponential (there is clearly some overlap between the two regimes) with a characteristic slope which increases rapidly with increasing σ.

This behavior was found to be generally valid for all the quantities one may use for studying the decay of synchrony. For example, the S phase of a log-phase population may be labeled with tritiated thymidine and the fraction of labeled mitotic cells scored. If M^* is the number of labeled cells in mitosis this fraction is $M^*/(M_0 + M^*)$, where M_0 is the number of unlabeled cells in mitosis. This fraction, as a function of time after labeling, comes through as a wave and the wave spreads out in subsequent cycles with $M^*/(M_0 + M^*)$ approaching a constant value in log-phase growth. An investigation of the shapes of these waves is presented later. It is the decay of maxima that is of interest here. The maxima minus the longtime average are plotted against the reduced time after labeling in Fig. 7. Exponential

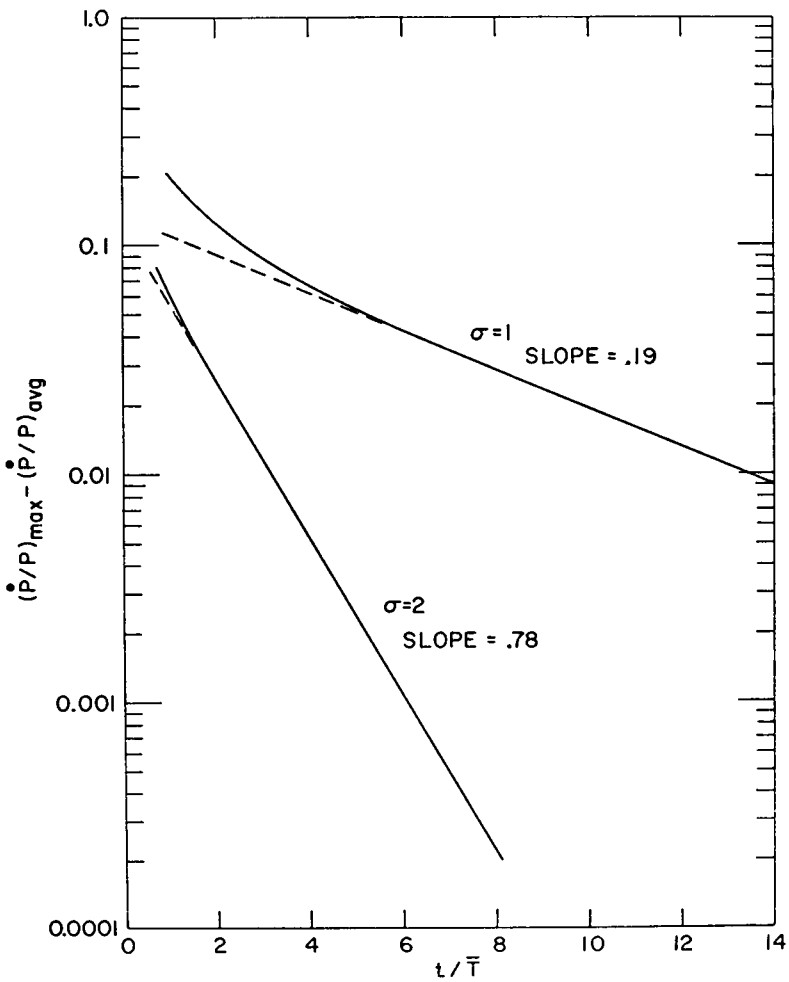

FIGURE 6 Decay of synchrony in proliferation rate, presented as $(\dot{P}/P)_{max} - (\dot{P}/P)_{avg}$ vs. the reduced proliferation time, $t/\bar{T}$ ($\bar{T} = 10$).

behavior is approached after a few cycles with the slopes increasing rapidly with increasing σ. The early curvature is in the opposite sense to that observed for $(\dot{P}/P)_{max}$ in Fig. 6. The assignment of the individual σ_i are listed in Table I.

Similar data are shown in Fig. 8 for the decay of synchrony in the fraction of S in a population which was perfectly synchronized in G_1 at $t = 0$. Detailed assignments of the parameters are listed in Table I. Note in Table I that the sense of the early curvature depends on the fractional phase composition of the cycle, as discussed in section V-D.

The most significant quantity characterizing the decay of synchrony is the exponential decay rate constant, λ. That is, all the decay processes are described by

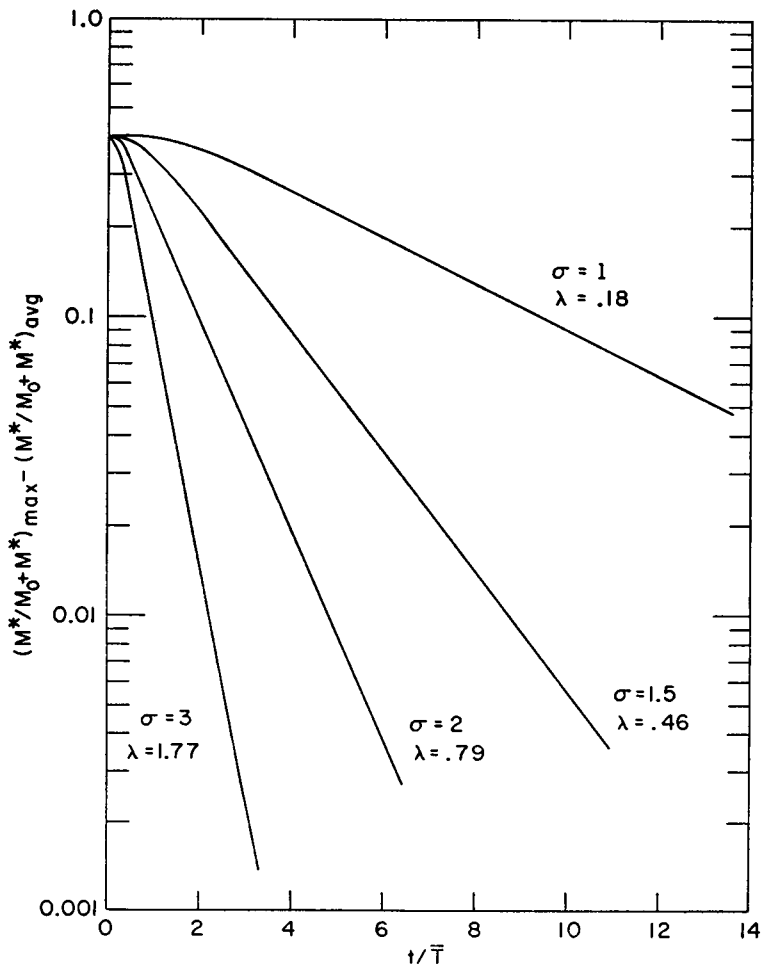

FIGURE 7 Decay of synchrony in the fraction of labeled cells in mitosis as a function of time after labeling. ($\bar{T} = 10$).

TABLE I

PARAMETERS USED IN DECAY OF SYNCHRONIZATION CALCULATIONS

λ	Obtained from*	$\bar{T}$	σ	i_1	i_2	i_3	i_4	σ_1	σ_2	σ_3	σ_4	Early curvature
0.046	S/P	10	0.5	1.5	6.0	1.5	1.0	0.25	0.25	0.25	0.25	Convex up
0.19	"	"	1.0	"	"	"	"	0.50	0.624	0.50	0.333	" "
0.28	"	"	1.2	"	"	"	"	0.50	0.910	0.50	0.333	" "
0.47	"	"	1.5	"	"	"	"	0.50	1.28	0.50	0.333	" "
0.81	"	"	2.0	"	"	"	"	0.50	1.84	0.50	0.333	" "
1.84	"	"	3.0	"	"	"	"	0.50	2.89	0.50	0.333	" "
0.80	"	"	2.0	0.5	6.0	0.5	3.0	0.25	1.84	0.25	0.70	" "
0.81	"	5	1.0	0.75	3.0	0.75	0.50	0.25	0.92	0.25	0.167	" "
0.18	$M^*/(M_0 + M^*)$	10	1.0	1.5	6.0	1.5	1.0	0.50	0.624	0.50	0.333	Convex up
0.46	"	"	1.5	"	"	"	"					" "
0.79	"	"	2.0	"	"	"	"	0.50	1.84	0.50	0.333	" "
1.77	"	"	3.0	"	"	"	"	0.50	2.89	0.50	0.333	" "
0.79	"	"	2.0	0.5	6.0	0.5	3.0	0.25	1.84	0.25	0.70	" "
0.21	M/P	10	1.0	1.5	6.0	1.5	1.0	0.50	0.624	0.50	0.333	Concave up
0.81	"	"	2.0	0.5	6.0	0.5	3.0	0.25	1.84	0.25	0.70	" "
1.77	"	"	3.0	1.5	6.0	1.5	1.0	0.50	2.89	0.50	0.333	" "
0.21	G_1/P	10	1.0	1.5	6.0	1.5	1.0	0.50	0.624	0.50	0.333	" "
0.79	S/P	10	2.0	2.5	2.5	2.5	2.5	1.0	1.0	1.0	1.0	Concave up
0.78	M/P	"	"	"	"	"	"	"	"	"	"	" "
0.79	G_1/P	"	"	"	"	"	"	"	"	"	"	" "
0.78	$M^*/(M_0 + M^*)$	"	"	"	"	"	"	"	"	"	"	" "
0.79	G_2/P	"	"	"	"	"	"	"	"	"	"	" "
0.19	$\left(\frac{dP}{dt}\right)\Big/P$	10	1.0	1.5	6.0	1.5	1.0	0.50	0.624	0.50	0.333	Concave up
0.78	"	10	2.0	"	"	"	"	0.50	1.84	0.50	0.333	" "

* This column means the maximum of the indicated quantity minus its long time average plotted as ln vs. $t/\bar{T}$.

the expression

$$X_{max} - X_{avg} = Ae^{-\lambda t/\bar{T}} \tag{28}$$

where a typical X would be the fraction in a state. Values of λ for a variety of computer experiments are collected in Table I. In these runs not only was the decay of synchrony in various phases of the cell cycle investigated but also σ_i and $\bar{t}_i$ were varied. All the data obey the relation

$$\lambda = A(\sigma/\bar{T})^2 \tag{29}$$

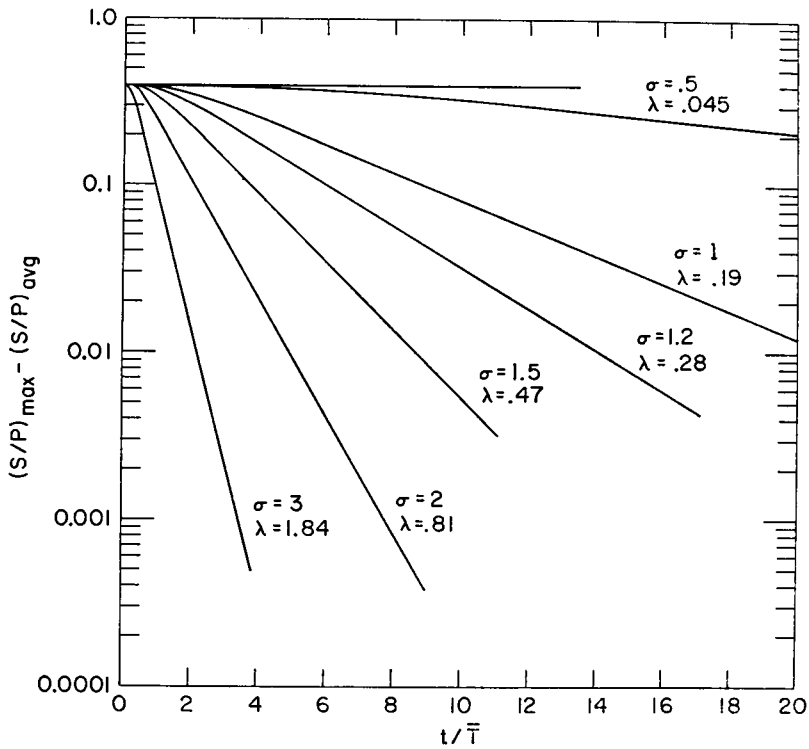

FIGURE 8 Decay of synchrony in S in a population synchronized at age zero in G_1 at $t = 0$ as a function of the reduced proliferation time. ($\bar{T} = 10$).

with $A \cong 20$. This relation is illustrated in Fig. 9, with the points taken from several experiments. All the points from Table I would fall on this line. From the computer calculations, therefore, we can come to the conclusion that the above relations are of general validity and that λ is a very important characteristic parameter of the proliferation process since it controls the rate at which disturbances (such as synchrony) are damped out. The value of λ will be obtained below from renewal theory.

In order to investigate the behavior of $\dot{P}(t)$ for large t, we can, by the argument

leading to equation 11, choose any regular function for $f(t)$ for which asymptotic solutions can be obtained. The family of functions

$$f(t) = \eta^m \frac{t^{m-1}}{(m-1)!} e^{-\eta t} \qquad t > 0$$
$$= 0 \qquad t \leqq 0 \tag{30}$$

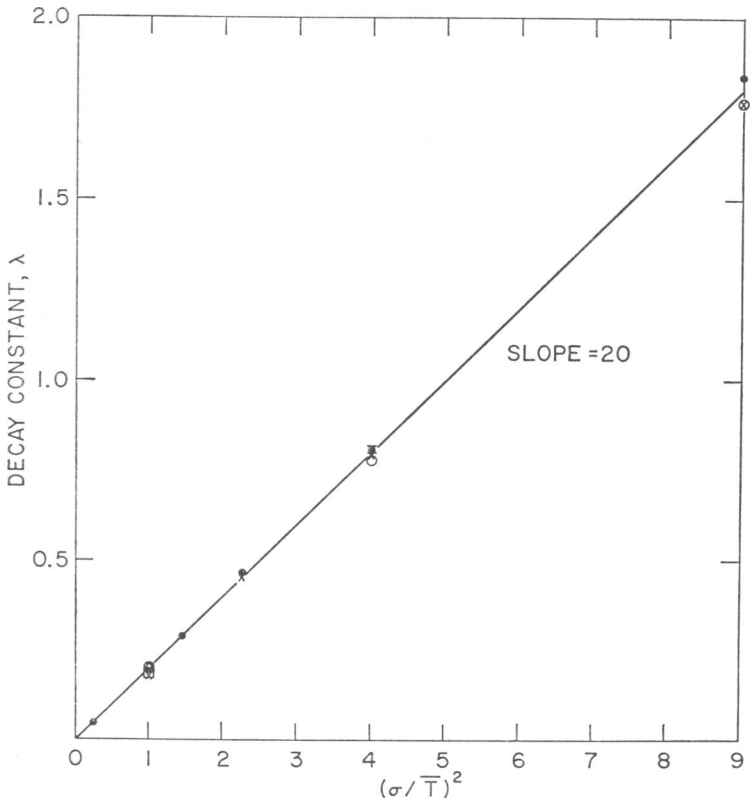

FIGURE 9 The characteristic decay rate constant, λ, as a function of $(\sigma/\bar{T})^2$ for a large variety of computer calculations. (See Table I).

leads to exact solutions for equation 7. (See, for example, reference 40, also Appendix III). This function is particularly convenient for several reasons: It gives an exact solution valid for all times and it approximates a Gaussian for large m and thus can be directly compared with the computer results which are based on a Gaussian form of $f(t)$. For the functions of equation 30 the mean and the variance are given by

$$\bar{T} = m/\eta$$
$$\sigma^2 = m/\eta^2$$
$$\text{C.V.} = \sigma/\bar{T} = (m)^{-1/2}. \tag{31}$$

Equation 30 gives the exact distribution for the sum of m random times, each of which is distributed independently, and with a probability equal to $\eta e^{-\eta t}\, dt$ of being of a magnitude in the neighborhood of t. Hence, if m is sufficiently large, we know by the central limit theorem that equation 30 will give a good approximation to a Gaussian over the contributing part of its domain. We will be interested in cases with a coefficient of variation ranging from one-twentieth to one-third. Therefore, m ranges from nine to four hundred and we are justified in comparing analytic results with numerical results obtained using a Gaussian.[6]

It is worthwhile to note that to the extent that equation 30 approximates the functions of interest, the present model gives the same probabilistic description of the total number of cells alive at a given time as does Kendall's model (19). Some later papers (27–34) take several of the m microstates to represent one of the four states we consider. For example, the last two microstates may represent mitosis. Our model gives somewhat different results for the average behavior of the fraction of cells in a given state. Our model also gives a description of the age distribution within a state which is completely different from these other models.

The exact solution of equation 7 obtained (see Appendix III) by using equation 30 is

$$\dot{P}(t) = \left[\eta m^{-1} 2^{-\frac{(m-1)}{m}}\right]\left\{\sum_{p=0}^{m-1} \exp\left[h(p, m)\eta t\right]/\exp\left[2\pi i p(m-1)/m\right]\right\} \quad (32)$$

where

$$h(p, m) \equiv [2^{1/m} \exp(2\pi i p/m) - 1].$$

For large times, those terms for which the real part of h is largest will contribute most. By making an Argand diagram (vector diagram in the complex plane) of h, it is easy to see that the real part of h decreases algebraically as the angle $(2\pi i p/m)$ moves away from zero either counterclockwise ($p = 1, 2, \cdots$) or clockwise ($p = m - 1, m - 2, \cdots$). If we write

$$\dot{P}(t) \approx [\dot{P}(t)]_1 + [\dot{P}(t)]_2 \quad (33)$$

for large t, then $[\dot{P}(t)]_1$ is the $p = 0$ term, and $[\dot{P}(t)]_2$ is the $p = 1$ plus the $p = m - 1$ term (these two terms are equal). Taking $((m-1)/m)$ as approximately one, and using equation 31 we obtain

$$[\dot{P}(t)]_1 = (1/2\bar{T}) \exp\{[2^{1/m} - 1]mt/\bar{T}\} \quad (34)$$

[6] For the same reason, a standard way of generating a table of numbers distributed according to a Gaussian, is to sum several numbers chosen from a uniform distribution. In our Monte Carlo procedure, we found that results obtained by summing ten such numbers to obtain our "Gaussian" were quite similar to those obtained from the statistical method with direct use of a Gaussian.

writing

$$[2^{1/m} - 1] = [\exp((1/m)\ln 2) - 1] \tag{35}$$

and expanding the right-hand side of equation 35 we obtain approximately

$$[\dot{P}]_1 \cong (1/2\bar{T}) \exp(\alpha t) \tag{36}$$

$$\alpha = \ln 2/T_D \; ; \; T_D = \bar{T}[1 + (\ln 2)\sigma^2/2\bar{T}^2]^{-1} \tag{37}$$

A similar expansion in the exponent of the $p = 1$ terms gives for large t,

$$[\dot{P}]_2 \cong 2[\dot{P}]_1 \exp\{(2\pi i/T_0)t\} \exp(-\lambda t/\bar{T}) \tag{38}$$

$$T_0 \equiv \bar{T}\left[1 + \frac{\sigma^2}{\bar{T}^2}(\ln 2)\right]^{-1} \tag{39}$$

$$\lambda \equiv (2\pi^2\sigma^2/\bar{T}^2). \tag{40}$$

Equation 37 gives the exponential population growth which we have already used in section V-A, in discussing the approach to asymptotic exponential behavior. Equation 38 gives an additional term which multiplies the growth term by an exponentially damped oscillating function (this general behavior was obtained by Hirsch and Engelberg (33)). The period of the oscillating function is slightly less than the generation time. One can estimate the error involved in these approximations by comparing the neglected cubic term in the expansion of the exponent relative to quadratic term. The ratio of these two terms for λ is $\ln 2\left(\frac{\sigma}{\bar{T}}\right)^2$ which is small for small C. V. (and is small for the cases treated by the computer technique).

Equation 32 is valid for all m and can be used as a guide to determine when the C. V. $= (m)^{-1/2}$ is large enough to lead to a nonoscillatory solution of equation 7. One obtains oscillations for $m = 3$, but only an exponentially growing plus an exponentially decaying term for $m = 2$. One therefore concludes that

$$\text{C. V.} \geqq (0.707) \tag{41}$$

will give nonoscillatory behavior of $\dot{P}(t)$.

Although the foregoing discussion pertains to a population initially synchronized at age zero, an arbitrary initial condition can be specified by a linear combination of such populations with different initial times. One, therefore, expects that the constants T_0 and λ will characterize many different experiments as demonstrated numerically for labeling curves and sudden disturbances (sections V-C and V-E).

C. Shapes of Labeling Curves

An example of a labeling curve was shown in Fig. 1 and the decay of the labeling curve maxima was discussed in the previous section. In this section the shapes of

labeling curves, that is $M^*/(M_0 + M^*)$ as a function of time after labeling, are discussed. As mentioned before, in a typical labeling experiment the cells in the S state of a log-phase population are labeled, in a time short compared to the generation time (flash labeling), and the fraction of labeled cells in mitosis, $M^*/(M_0 + M^*)$, is subsequently scored as a function of time after labeling. In the following, instantaneous labeling in S will be assumed. The labeled S cells are assumed to move through the proliferation cycle exactly like normal cells. That is, they move through

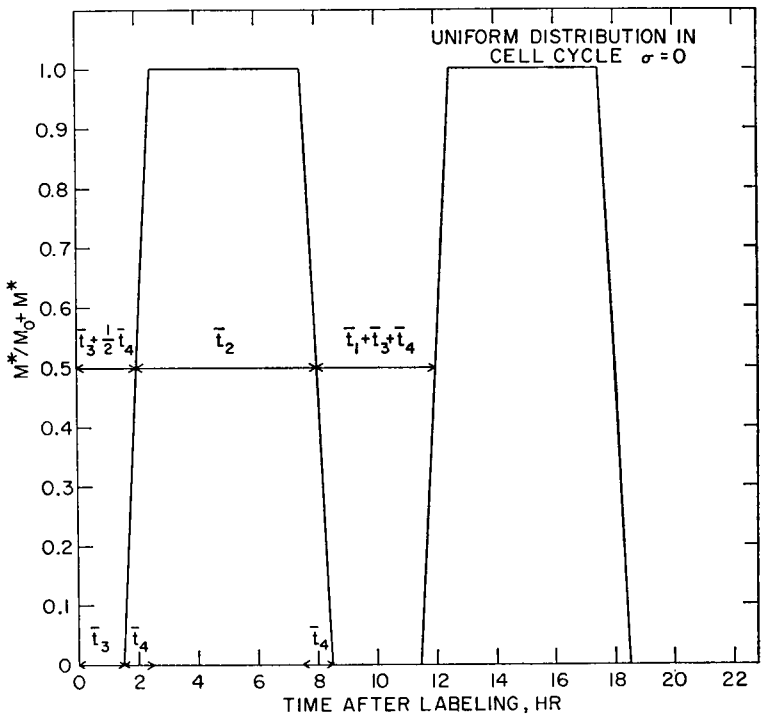

FIGURE 10 Labeling curve for a uniform distribution in S and $\sigma = 0$.

G_2 and then are observed in M as a function of time after labeling. The shape of the graph of the labeled fraction observed in M is influenced by the time distribution of cells in S and by the spread in the distribution of the residence times, and later by the spread in the total generation time. In the following we shall try to understand the influence of each of these factors.

If the cells in S are uniformly distributed with respect to time spent in S and if there is no spread in the generation time ($\sigma = 0$), then the labeling curve of Fig. 10 would result. This curve of trapezoidal shape, of course, represents a highly idealized situation which is not achievable experimentally because it is unlikely that a $\sigma = 0$ population exists experimentally. (If a $\sigma = 0$ population did exist one could arrange the initial conditions such that the population is uniformly distributed in the

cycle.) Nevertheless, this curve is useful as a starting point for discussing the distortions introduced by various factors. The transit times in the various phases of the cell cycle are indicated on the diagram of Fig. 10. This curve results from the fact that at $\sigma = 0$ the cells keep their order in "marching through" the cell cycle and,

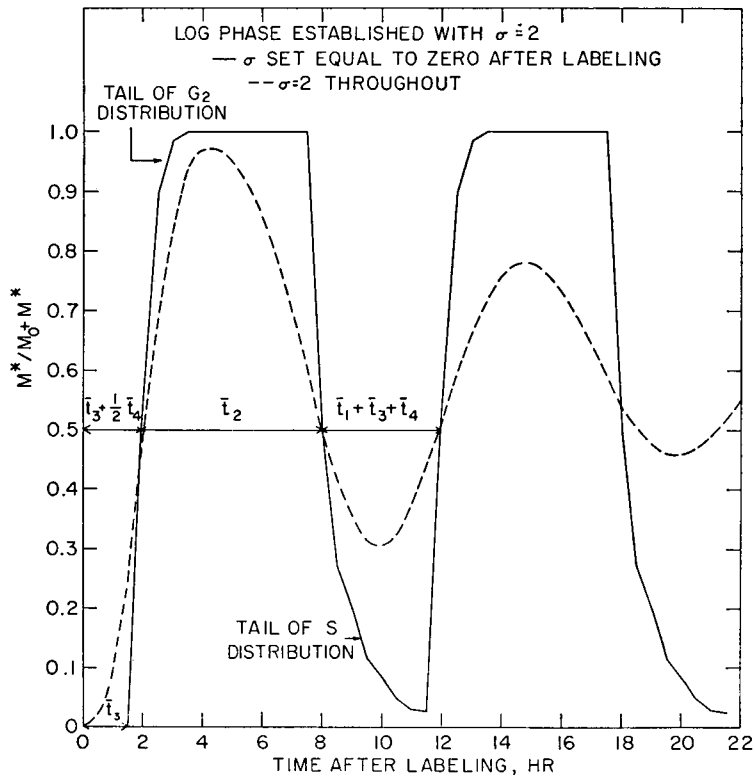

FIGURE 11 Labeling curve for a log-phase population with $\sigma = 2$, $\bar{t} = 10$ (C.V. = 0.1) Solid line: σ set equal to zero at time of labeling. Dashed line: $\sigma = 2$ throughout experiment.

since there are just as many "young" as "old" cells (uniform distribution), the rise and decay processes are completely symmetrical. The slope of the rise and decay parts of the curve depends only on the transit time in M (i.e., vertical rise if $\bar{t}_4 = 0$).

In order to illustrate the influence of the cell distribution in S, and in the other phases, the following computer experiment was carried out using the Monte Carlo method. A log-phase population was established with a given σ (spread of generation time) and generation time, $\bar{T}$. In this way, waiting-time-in-state distributions characteristic of these $\bar{T}$'s and σ's were established. The S phase was flash labeled and for the subsequent motion around the cell cycle σ was artificially set equal to zero. More precisely, each cell entered the very next state according to an asymptotic waiting time distribution. The times spent in all subsequent states are exactly

given by the t_i, however. The resultant distortion of the labeling from the trapezoidal form is, therefore, characteristic of the distribution existing at the time of labeling and is not complicated by the generation time distribution. The results of these calculations are shown by the solid line of Fig. 11 for $\sigma = 2$ (C.V. = 0.2). Two distortions of the ideal trapezoidal shape are observable. There is a rounding off of the top of the rising curve and a larger rounding off of the bottom of the descending part of the curve. Since σ is now zero, the curve will repeat identically in the second cycle. The leading edge of the waiting-time distribution in S is quite sharp (see section V-D) and, therefore, the first rounding off is caused by stragglers which were in G_2 at the time of labeling (i.e., the long time tail of the waiting-time distribution in G_2). The rounding off at the bottom is caused by the stragglers in S (i.e., by the tail of the waiting-time distribution in S).

The same computation has been carried out without changing σ at the time of labeling (i.e., $\sigma = 2$ throughout). This calculation illustrates the combined effects of the distribution discussed above and the subsequent spreading due to the σ assigned to $\bar{T}$. This is, of course, the actual experimental situation. The results are shown by the dashed line in Fig. 11. The influence of σ is clearly very strong; it produces an asymmetry in the curves and a decay in the peaks. The analytic expressions for these curves, which are developed elsewhere (42), exhibit the decaying oscillatory behavior discussed in section V-B.

D. Age-in-State and Fractions-in-State

The various age distributions are interesting both from a theoretical and experimental (43) point of view. The straightforward application of familiar probabilistic principles to our model leads to useful results concerning the time development of age-in-state or time-in-state distributions. Because of possible confusion between the general concept of age-in-state and other uses of the word age, a careful, operational definition is given of age-in-state as it will be used subsequently. We first define the *age in the cycle* as the time since the previous cell division that a given cell spends in its cell-cycle. Similarly, the *age-in-state* (e.g. age-in-mitosis, age-in-synthesis, etc.) is the time since entering the state that a given cell spends in the particular state.

A contrasting concept which is sometimes confused with the above is maturity-in-state. The *maturity-in-state* is a measure of how far the cell has progressed towards completing the requirements for leaving that state and entering the next state. For example, the amount of DNA synthesized in a given cell could be used to measure the maturity-in-S. Another illustration of our use of maturity is given in terms of our model. The index of the state of the cell (i.e. $i = 1, 2, 3, 4$ designating G_1, S, G_2, and M, respectively) is an integer-valued measure of the cells maturity in the cycle. It is emphasized that because individual cells mature in a state with randomly distributed velocities, an individual cell's age-in-state may be

different from its maturity-in-state. It should also be noted that our definition of age is quite distinct from the aging of an organism which is a measure of degenerative processes which occur as the organism gets older. The idea of age-in-state may be clarified by an operational definition. If a cell were marked at the instant of time that it entered a state, the age-in-state is the length of time between the instant of marking and the time of observation provided that the cell has not left the state.

Having defined age-in-state let us describe it in terms of our model. We define $A_i(z, t)$ as the mean number of cells in state i at time t with ages between z and $z + dz$. For i equal one, the cells with age z are those born at time $(t - z)$ which have not left the state. This may be expressed as

$$A_1(z, t) = B(t - z)[1 - F_1(z)], \tag{42}$$

where $B(t - z)$ as defined in equation 13 is the mean number of cells born into the cycle, and hence into state one, z hours before the observation time t. The second factor gives the probability of a cell spending a time longer than z hours in state one.

To obtain A_j for j greater than one, we need a general expression for the probability of a cell spending a time between t and $t + dt$ in the states preceding j. This is defined as $f(1, 2, \cdots, j - 1; t)\, dt$ and is given by the convolution of the individual densities for the transit times of the states preceding j. For example, for $j = 3$ we have

$$f(1, 2; t) = \int_0^t f_1(t')f_2(t - t')\, dt', \tag{43}$$

which for our Gaussian densities gives to good approximation

$$f(1, 2; t) = [2\pi(\sigma_1^2 + \sigma_2^2)]^{-1/2} \exp\{-[t - (t_1 + t_2)]^2/2(\sigma_1^2 + \sigma_2^2)^{1/2}\}. \tag{44}$$

Thus, we may write for $j > 1$

$$A_j(z, t) = \int_0^\infty B(t - z - t')f(1, 2, \cdots j - 1; t')\, dt'[1 - F_j(z)]. \tag{45}$$

The integral in equation 45 averages the time spent during the various sojourns in the states preceding j to give the mean number of cells entering j at z hours before the time of measurement. The factor $[1 - F_j(z)]$ is the probability of a cell remaining in state j a time longer than z.

The total number of cells in state j at time t is given by

$$P_j(t) = \int_0^\infty A_j(z, t)\, dz \tag{46}$$

for all values of j. Using this expression we arrive at the fraction of cells with age z

in the jth state as

$$a_j(z, t) \equiv A_j(z, t)/P_j(t). \tag{47}$$

A distribution related to the above which is of some interest may be obtained by dividing an n state system into two parts relative to a state j, where j is less than or equal to n. The two parts are the states preceding j, and the states with index greater than or equal to j. Then the mean number of cells which entered state j, between z and $z + dz$ hours before t, and are in state j *or* one of the states following j, is defined to be $A_{\geqq j}(z, t)$, and is given by

$$A_{\geqq j}(z, t) = \int_0^\infty B(t - z - t')f(1, 2, \cdots, j - 1; t')\, dt'[1 - F(j, \cdots, n; z)] \tag{48}$$

$$F(j, \cdots, n; z) = \int_0^z f(j, \cdots, n; t)\, dt. \tag{49}$$

The second factor in equation 48 gives the probability of a cell spending a time longer than z in state j and the states following j. To obtain A_j from equation 48, we would multiply $A_{\geqq j}$ by the probability that a cell is still in state j, after entering that state z hours ago, given that it still has not completed the cycle.

This probability is

$$[1 - F_j(z)]/[1 - F(j, \cdots, n; z)]. \tag{50}$$

Therefore we see that the age distributions for an n state system are related in a simple way to those for a two state system.

We will next derive some explicit expressions for $A_j(z, t)$ for the situation where one starts with a population synchronized at the beginning of G_1, at time $t = 0$. During the early part of the first cycle, very few new cells are born. We may therefore approximately represent the function $B(t)$ by a delta function (a spike) centered at $t = 0$. Using this in equations 42 and 45, we obtain

$$A_1(z, t) \approx [1 - F_1(z)]\delta(z - t)$$

and

$$A_j(z, t) \approx f(1, 2, \cdots, j - 1; t - z)[1 - F_j(z)]\, j > 1 \quad \text{for} \quad t \lesssim (\quad - \sigma). \tag{51}$$

For a time t large enough so that several cycles have passed since initial synchronization, we can calculate the approach to the asymptotic age distribution. We will use the following equations: 12, 33, 37, 38, 42, 45, and 47. The following approximation will also be used. Consider the integral

$$I \equiv \int_0^\infty e^{-bt} f(t)\, dt, \tag{52}$$

where $f(t)$ is a Gaussian with mean $\bar{T}$ and variance σ^2. Then when the mean $\bar{T}$ of the Gaussian is large compared with the magnitude[7] of $(2\sigma + \sigma^2 b)$, one can extend the lower limit to minus infinity and evaluate the integral in the usual way by completing the square in the exponent, giving

$$I \approx e^{-\bar{T}b+\sigma^2 b^2/2} \tag{53}$$

for $$\bar{T} \gtrsim Re[\sigma^2 b + 2\sigma].$$

Using this approximation together with the definitions 42 and 45 and the asymptotic birth rate, we obtain

$$A_j(z, t) = [e^{\alpha(t-z)}/\bar{T}][1 - F_j(z)\{b_{1j} + 2Re(b_{2j} \exp[(i\omega_0 - \lambda/\bar{T})(t - z)]\} \tag{54}$$

where $\omega_0 = [(2\pi)/T_0]$ (see equations 39 and 40).

$$b_{1j} = \exp[-\dot{l}_{bj}\alpha + \sigma_{bj}^2\alpha^2/2]$$

$$b_{2j} = \exp[-\dot{l}_{bj}\underline{\alpha} + \sigma_{bj}^2\underline{\alpha}^2/2]$$

$$\underline{\alpha} = [\alpha + i\omega_0 - (\lambda/\bar{T})] \tag{55}$$

$$\dot{l}_{bj} \equiv \sum_1^{j-1} \dot{l}_K \qquad \sigma_{bj}^2 \equiv \sum_1^{j-1} \sigma_K^2 \qquad j > 1$$

$$\equiv 0 \qquad\qquad \equiv 0 \qquad\qquad j = 1 \tag{56}$$

If we integrate equation 54 over z, we obtain

$$P_j(t) \approx [e^{\alpha t}/\alpha\bar{T}]\{(1 - e^{-\dot{l}_j\alpha+\sigma_j^2\alpha^2/2})b_{1j} + 2Re(b_{2j}(1 - e^{-\dot{l}_j\underline{\alpha}+\sigma_j^2\underline{\alpha}^2/2}) \exp[(i\omega_0 - \lambda/\bar{T})t]\}. \tag{57}$$

The average fraction of the population in the jth state at time t is given by $[P_j(t)/P(t)]$, where

$$P(t) \equiv \sum_{j=1}^{4} P_j(t)$$

$$= [e^{\alpha t}/\alpha\bar{T}]C_1[1 + Re\{C_2 \exp(i\omega_0 - (\lambda/\bar{T}))t\}] \tag{58}$$

$$C_1 \equiv \sum_{j=1}^{4} C_{1j} \qquad C_{1j} \equiv b_{1j}(1 - e^{-\dot{l}_j\alpha + \sigma_j^2\alpha^2/2})$$

$$C_2 \equiv \sum_{j=1}^{4} C_{2j}/C_1 \qquad C_{2j} \equiv 2b_{2j}(1 - e^{-\dot{l}_j\underline{\alpha} + \sigma_j^2\underline{\alpha}^2/2}). \tag{59}$$

[7] When b is imaginary one observes that the integral is approximately equal to an integral over a nearby contour on which the integrand is a real Gaussian. In this case the additional error is less in magnitude than $|e^{-(\bar{T}-b\sigma^2)^2} Im(b\sigma^2)|$.

Then for large t we have

$$[P_j(t)/P(t)] \approx [1/C_1]\{C_{1j} + Re\{[C_{2j} - C_{1j}C_2] \exp(i\omega_0 - (\lambda/\bar{T}))t\}. \quad (60)$$

If we use the fact that α^2 is small we can neglect the various terms of the form $\sigma^2\alpha^2$ in the exponent. (This is not justified for $\sigma^2\underline{\alpha}^2$ because of the nonnegligible imaginary part of $\underline{\alpha}$.) Then taking $\alpha \approx \ln 2/\bar{T}$ we obtain for t sufficiently large that the second term in equation 60 can be neglected.

$$P_j/P\,|_\infty \approx 2e^{-i_{bj}\alpha}[1 - e^{-i_j\alpha}]. \quad (61)$$

This is exactly the expression obtained (6) in the less detailed model where one simply assumes that cells of age z have proliferated exponentially a time $(t - z)$, giving a density at that age proportional to $e^{-\alpha z}$. The assumption is then made (incorrectly unless all $\sigma_j = 0$) that cells of age in the cycle between $z = \bar{t}_{bj}$, and $z = \bar{t}_{bj} + \bar{t}_j$ are in state j. Then the density is integrated between those two ages giving the above expression, the value of which is within a couple percent of the correct expression for realistic σ's.

The second term in equation 60 gives the decaying wave form of the fraction in state. If one makes a plot of the log of the maxima of $D_j = [P_j(t)/P(t) - P_j/P\,|_\infty]$ vs. time, the graph approaches a straight line after a number of cycles depending on λ which is a function of the C.V. The graph will be convex up or concave up depending on whether the intercept of the tangent to the linear portion at $t = 0$ is greater than or less than $\log[1 - P_j/P\,|_\infty]$ which gives the first maximum of D_j. The maxima always equal $[1 - P_j/P\,|_\infty]$ when $\sigma = 0$. In terms of expression 59 we have

$$R_e[(C_{2j} - C_{1j}C_2)/C_1] \gtrless [1 - P_j/P\,|_\infty] \begin{matrix} \Rightarrow \text{convex up} \\ \Rightarrow \text{concave up.} \end{matrix} \quad (62)$$

Because of the complexity of equation 62, it is easier to demonstrate this property numerically as shown in Fig. 8 for states for which the maxima decay in a convex manner (see also Table I).

Now returning to the age distribution, if we take the ratio of equations 54 and 57, we get for the asymptotic age density

$$a_j(z, t) = a_j(z)[1 + \epsilon_j(z, t) - \delta_j(t)]. \quad (63)$$

The first term of equation 63 is the time-independent asymptotic expression given by

$$a_j(z) = c_j e^{-\alpha z}[1 - F_j(z)] \quad (64)$$

where we have suppressed the time dependence of the normalization, c_j, and

$$c_j \approx \left\{ \int_0^\infty e^{-\alpha z}[1 - F_j(z)]\, dz \right\}^{-1}$$

$$= \alpha e^{\gamma_j} / (e^{\gamma_j} - 1) \tag{65}$$

where $\gamma_j = [\bar{t}_j\alpha - \sigma_j^2\alpha^2/2]$.

The third term of the second factor in equation 63 is given by the oscillation times the decaying exponential, and merely means that the exact normalization has a slight oscillation perturbing it. The second term is more interesting. It is the product of a constant times an expression given with approximate phase by

$$\epsilon_j(z, t) = \cos(\omega_0(\bar{t} - \bar{t}_{bj} - z))e^{-(\lambda/\bar{t})(t-z)}. \tag{66}$$

Now the peak of the wave is given by the point of constant phase,

$$\omega_0(t - \bar{t}_{bj} - z) = 2\pi n \tag{67}$$

or

$$t - \bar{t}_{bj} - z = nT_0.$$

This implies that the exponential factor at the peak is constant during a given cycle, so that on a given cycle we would see a wave moving across our age-in-state density toward increasing values of z.

The *waiting-time-in-state* is defined as the additional time a cell will spend in the state in which it resides at the time of measurement, before transfer to the next state. The waiting time could be measured if a cell could be marked at some time and the later time when it left the state, in which it was marked, noted. The numerical distribution of waiting times is directly available from our Monte Carlo calculation.

If the density of waiting times at time t is given by $W_j(\tau, t)$ we have

$$W_j(\tau, t) = \int_0^\infty a_j(z, t) \left\{ \frac{f_j(z + \tau)}{1 - F_j(z)} \right\} dz. \tag{68}$$

The first factor in the integrand is the fractional age density derived above.

The second factor is the probability that a cell will transfer to state $(j + 1)$ after a sojourn in state j of length about $(z + \tau)$, given that it stays in state j a time at least equal to z. The product gives the mean fraction of cells of age z which enter the $(j + 1)$st state in τ additional hours. For t large, we obtain the asymptotic solution

$$W_j(\tau) = c_j \int_0^\infty e^{-\alpha z} f_j(z + \tau)\, dz \tag{69}$$

where c_j is the normalization constant given in equation 65. We can write an explicit expression for $W_j(\tau)$ for f_j, a Gaussian.

$$W_j(\tau) = c_j e^{+\alpha\tau-\gamma_j}(1/2)[1 + \text{erf}\,[(-\tau + \gamma'_j/\alpha)/\sigma_j\sqrt{2}]$$

$$\gamma'_j = [t_j\alpha - \sigma_j^2\alpha^2], \tag{70}$$

where $$\text{erf}\,(x) \equiv \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\,dt \qquad x > 0$$

and $$\text{erf}\,(-\,|\,x\,|) \equiv -\text{erf}\,(|\,x\,|).$$

The criterion that a population must satisfy in this model to be in log phase is that it have the correct asymptotic percentages in the various states as given by equation 61 and that its waiting times-in-state be given by equation 69. Since the expression 70 is very smooth as a function of τ, we observe that the waiting time distribution is not too far off from that given by a uniform distribution, (although an exponential distribution is a better approximation). We found that if one starts the Monte Carlo program with 100 cells distributed with the correct asymptotic percentages for a C.V. of 0.2, and distributes them uniformly in each state, that within one cycle run at the same C.V. the waiting time distributions resemble closely the asymptotic distribution, and in three cycles the fractions in phase settle down to their asymptotic values.

We now consider briefly the corresponding numerical results. The age distribution in any given state is directly obtainable from the age transfer program at any state of proliferation history. Thus, one can follow directly the change in age distribution as the system changes, for example, from complete synchrony in G_1 at $t = 0$ to log-phase growth at long times. An example of such a development is shown in Fig. 12 for the age distribution in S. The fraction of cells per unit age interval is plotted as a function of age (in hours). At clock time $t = 0$ the age distribution was a delta function in G_1 centered around age zero. The various curves in Fig. 12 were taken at different clock times i.e., as a function of proliferation time (for a system with $\bar{T} = 10$). At 2 hr the distribution is very sharp. There is a significant build-up of young cells at 8 hr changing over to a double peak distribution at 10 hr. As the old cells move out (double) the age distribution at 14 hr results. As the proliferation progresses the curves become gradually smoothed out approaching a steady-state distribution (60 and 100 hr) characteristic of log-phase growth and independent of proliferation time.

The analytic approximation of equation 51 shows that the age distribution in $S(j = 2)$ during most of the first cycle is a wave described by (from equations 46, 47, and 51)

$$f_1(t - z)[1 - F_2(z)]/\int_0^t f_1(t - z)[1 - F_2(z)]\,dz,$$

where the shape is determined by the numerator, i.e., by f_1 for the first few peaks. The peaks of Fig. 12 are easily calculated from this formula by evaluating the maxima. The integral in the denominator is trivial as long as $1 - F_2(z) \cong 1$, i.e.,

for the first two to three peaks. A numerical check for the first three peaks showed that the computer results and the analytical expression are in agreement both as to location and magnitude of the peaks.

The corresponding waiting-time distributions in S are shown in Fig. 13. The

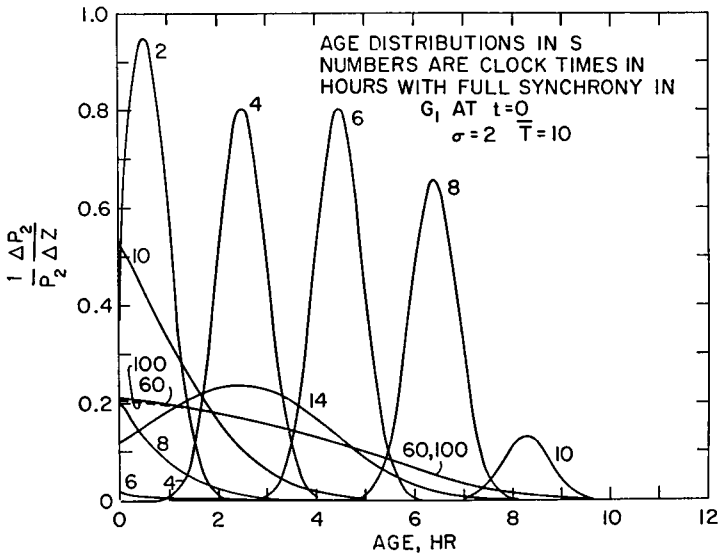

FIGURE 12 The change in the age distribution in S starting with complete synchrony of age zero in G_1 at $t = 0$. The proliferation times (clock times) are shown for each curve; 100 hr corresponds to age distribution in log-phase growth. $\bar{T} = 10$, $\sigma = 2$. (See Table I for i_i and σ_i.)

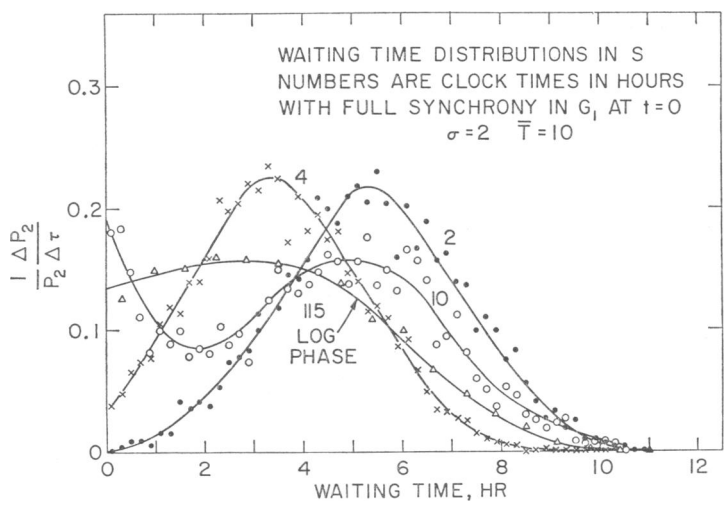

FIGURE 13 The change in the waiting-time distribution in S starting with complete synchrony in G_1 at $t = 0$. The proliferation times (clock times) are shown for each curve with the distribution corresponding to log-phase indicated. $\bar{T} = 10$, $\sigma = 2$. (See Table I for i_i and σ_i.)

waiting time is defined as the time a cell will spend in a state before leaving it. The waiting-time distribution is complementary to, but different from, the age distribution and is presented analogously, i.e., as the fraction of cells in the waiting-time interval Δt, per Δt, plotted as a function of waiting time (in hours). The waiting-time distribution data are directly obtained from the Monte Carlo program. Because of the limitation on the number of cells in this program there is considerable statistical scatter in the data so obtained. The general features are, however, clear.

The progression of the distribution in S and its approach to log phase is shown in Fig. 13. The earliest distributions are not too far from Gaussians corresponding to the sharp age distributions. A double peak structure develops at 10 hr corresponding to a similar structure in the age distribution.

Other phases behave very similar with some special features noted for G_1. The age distribution in G_1 starts out as a delta function since perfect synchrony at $t = 0$ was assumed. The corresponding waiting-time distribution is a Gaussian centered around $\bar{t}_1$. These distributions then progress similarly to the ones in S.

E. Sudden Disturbances of the Proliferation Cycle

We will now consider those changes in the conditions of proliferation which can be simulated during a computer experiment by an instantaneous alteration of parameters. Such changes of conditions may occur for a variety of experimental reasons such as sudden temperature change, change in chemical environment, selective removal, or damage of cells (35). We can easily handle with our present program sudden disturbances of the following three types: sudden changes in state of cells already in the cycle or the removal (addition) of cells from (to) states of the cycle; changes in the parameters $\bar{t}_i$ and σ_i; alteration of the age distribution existing at a given time.

A change of the first type is the easiest to consider. Several items already discussed are of this type. One example is the synchronization of a population at the beginning of G_1. Another example would be the death of all but one state of a log-phase population, for example the S state (Chinese hamster cells respond essentially this way to radiation since the S state is highly radiation resistant (35)). In these cases, the full or partial synchronization produced leads to damped oscillations in the birth rate and fraction in state. Similarly, one expects that changes of the latter two types will produce partial synchronization since cells are now in a non-asynchronous distribution with respect to the new conditions. Hence, one again expects damped periodic transients.

We consider several sudden changes in connection with two typical biological situations: (*a*) The transit time in mitosis, $\bar{t}_4$, is suddenly lengthened. (*b*) The transit time in G_1, $\bar{t}_1$, is suddenly shortened.

In situation (*a*) we simulate a condition in which the time for each cell to complete mitosis is greatly lengthened, but the cells are not blocked in the cycle. Such a situation might be achieved experimentally by a change of the temperature which seems

to affect division to a greater extent than synthesis.[8] We show the results of a sudden lengthening of $\bar{t}_4$ in Figs. 14 and 15. In these computer experiments we used the age-transfer method. A log-phase population with $\bar{T} = 10$ and $\sigma = 2$ was established prior to the sudden change which takes place at time $t = 0$. At this time, $\bar{t}_4$ is increased from 1 to 6 hr, and σ_4 is also increased. The subsequent development

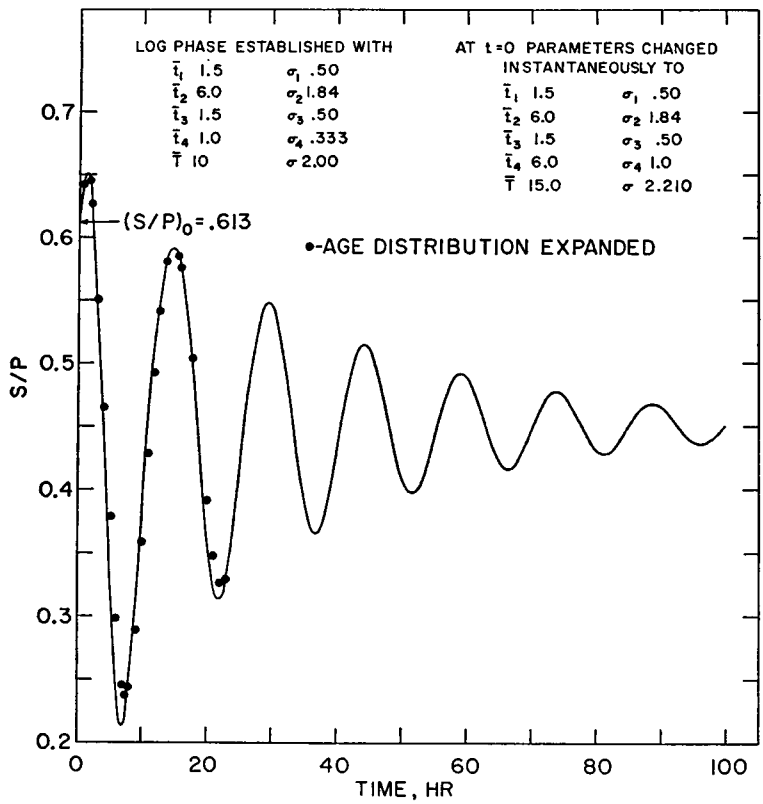

FIGURE 14 The approach of S/P to the new equilibrium fraction-in-state upon sudden lengthening (at $t = 0$ in the figure) of mitotic transit time, $\bar{t}_4$.

of S/P and M/P toward a new equilibrium (asynchronous) distribution is shown. The approach to equilibrium is via a series of damped waves as expected from our experience with decay of synchrony. During this period the new logarithmic growth rate is established via the usual carpeted staircase. The peaks of Figs. 14 and 15 decay with the rate constant, λ, determined by the σ and $\bar{T}$ after the disturbance, just as in the synchrony experiments.

The question arises as to whether the age-in-state distribution should also be

[8] The experiments of Rao and Engelberg on HeLa cells (44) indicate that the mitotic transit time is preferentially increased by a change in temperature, i.e., apparently all the parameters change but the increase in $\bar{t}_4$ is larger than that in the other transit times.

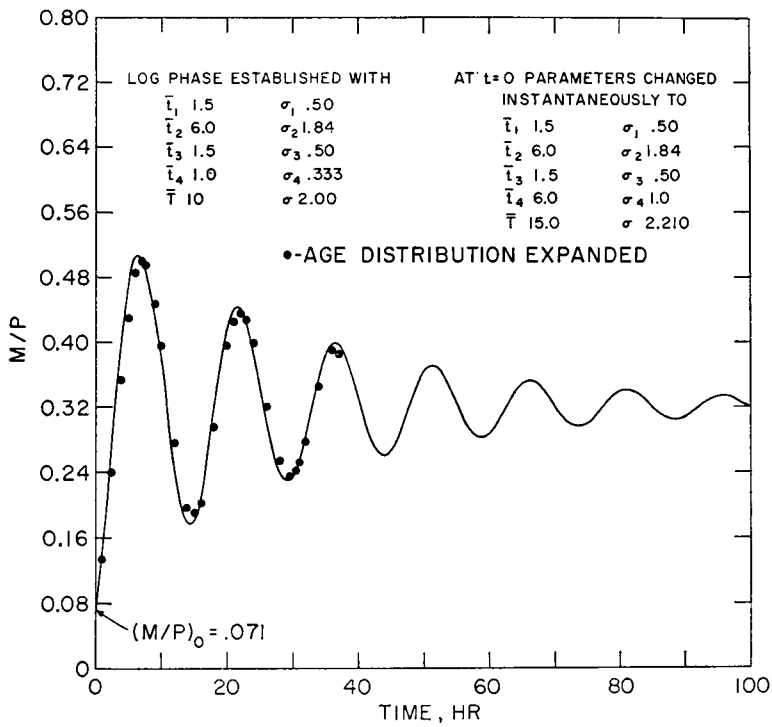

FIGURE 15 The approach of M/P to the new equilibrium fraction-in-state upon sudden lengthening (at $t = 0$ in the figure) of mitotic transit time, $\bar{t}_4$.

instantaneously altered for the state which has its parameters changed. When the sudden change amounts to an increase by a large factor of one of the transit times, the fraction of cells in the altered state immediately following the sudden change is very much smaller than the new log-phase value for that fraction. The essential feature of the experiment therefore is a gap or shortage of cells in the altered state. Hence, one expects that the age-in-state distribution of those few cells remaining in the state will not appreciably affect the experiment.

This idea was tested in the following way. A trial was run in which the age distribution was unaltered. Thus, the cells initially in M (about 7% of the total) are bunched up in the beginning of the new, 6 hr M state, with the oldest cells having the same probability of dividing that they would have after about an hour in the new M state. The results are shown by the solid lines in Figs. 14 and 15. In a second run, the time scale of the existing age distribution in M was expanded by the same factor as was the transit time. In this case the cells are now spread throughout the new M state according to the asymptotic-age-in-state distribution. The cells which were y hours old before the disturbance are $6y$ hours old after the disturbance. The results for the fraction-in-state are shown by the solid circles in Figs. 14 and 15. As expected, the two results are very similar. The first and subsequent maxima for the

spread-out run come about 0.6 hr later, and they are slightly less in magnitude than in the bunched up case.

In situation *b* we simulate what is currently thought by some to be a mechanism for triggering conditional renewal systems.[9] The idea is that a certain increase in cell volume is the signal for the cell to enter DNA synthesis which in turn leads inexorably to division. The G_1 period is ordinarily extremely long but a demand situation signals cell growth which in turn sends the cell into S and a shortened G_1 on the next cycle. In this case the number of cells in G_1 following the disturbance is in excess of the final asymptotic fraction, and these excess cells constitute the perturbation.

Let us first consider the case in which the ages of the cells scale in proportion to the change in the transit time. Then the excess in G_1 simply amounts to a partially synchronized portion superposed onto an asynchronous population. Consider the case when the transit time for G_1 is a large fraction of $\bar{T}$ prior to the disturbance and a small fraction of the new $\bar{T}$ after the disturbance. The development of the population after the disturbance should approximate that of the limiting case where $\bar{t}_1$ (before) $\rightarrow \bar{T}$ (before), and $\bar{t}_1$ (after) $\rightarrow 0$, leading to a population synchronized at the beginning of the cycle.

We tested this idea by starting a run with all cells in G_1, taken out of a log-phase population and distributed according to the asymptotic age-in-state distribution for $\sigma_1/\bar{t}_1 = 0.33$ (equation 64). Let us call this our pseudo-synchronous population. We then simulate the proliferation cycle with our Chinese hamster cell parameters ($\bar{t}_i$ = 1.5; 6.; 1.5; 1.; σ_i = 0.5; 1.84; 0.5; 0.33.). The resulting graphs were then compared with a desynchronization of our synchronized population (all cells initially less than 0.05 hr old in G_1) and found to be almost identical. For the pseudo-synchronous population we had a first maximum of 98.6% for S/P at 2.5 hr and 14.1% for M/P at 8.2 hr. A slight decrease in the time prior to the maxima occurs because in the pseudo-synchronous case many cells leave G_1 immediately rather than wait the average 1.5 hr in G_1. A slight decrease in the maxima is due to the initial spread of the pseudo-synchronous population through G_1.

The effect of altering the age distribution was explored next. The values of $\bar{t}_1$ and σ_1 prior to the disturbance were varied, but all the other parameters were kept identical prior to and after the disturbance. The age distribution for G_1 is scaled so that a cell which was y hours old in G_1 before the disturbance has an age $[\bar{t}_1(\text{after})/\bar{t}_1(\text{before})]\cdot y$ in G_1 after the disturbance. The ratio of σ_1 to $\bar{t}_1$ is kept about the same after the disturbance as before. The results are summarized in Table II. It is seen that the first maxima for S/P and M/P are gradually reduced as the initial $\bar{t}_1$ is decreased but the location of these maxima remains essentially

[9] This idea was brought to our attention by H. Johnson of the Brookhaven Medical Department. A possible example is the kidney in which cell division in the adult animal is ordinarily rather rare. However, following unilateral nephrectomy, a period of cell growth is followed by a burst of mitotic activity.

TABLE II

CHANGE IN THE FIRST MAXIMA OF S/P AND M/P AS A FUNCTION OF $\bar{t}_1$ (BEFORE)/$\bar{t}_1$ (AFTER)

$\bar{t}_1$ (before)/$\bar{t}_1$ (after)	First maxima in S/P	Location of maxima	First maxima in M/P	Location of maxima
	%	*hr*	%	*hr*
Pseudosynchronous	98.6	2.5	14.1	8.2
18/1.5	82.8	2.5	11.3	8.5
9/1.5	77.2	2.5	10.1	8.5

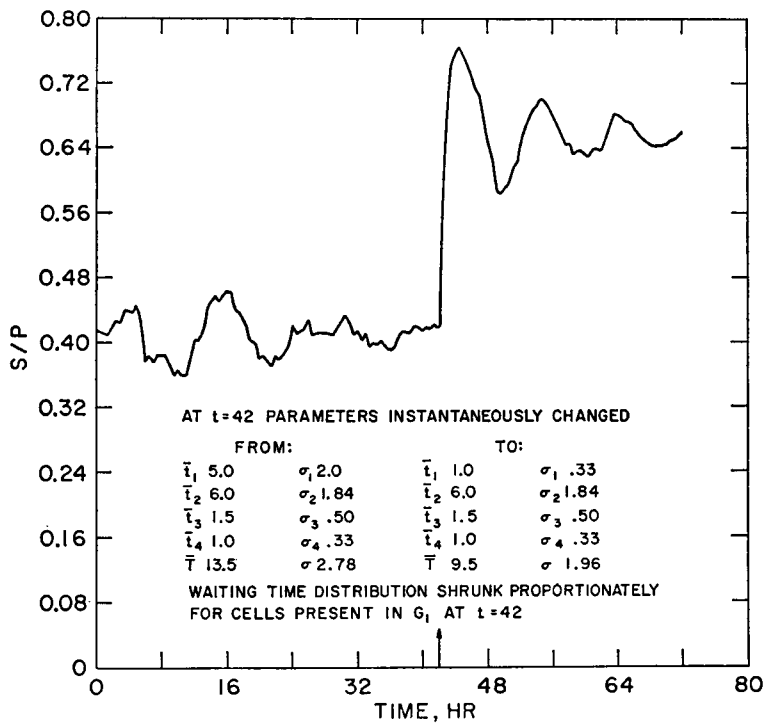

FIGURE 16 The approach of S/P to the new equilibrium fraction-in-state upon sudden shortening (at $t = 42$ in the figure) of the transit time in G_1. Age distribution of cells present in G_1 at the time of the change compressed as discussed in the text.

unchanged. The graph of S/P for a similar run is shown in Fig. 16 for the initial $\bar{t}_1 = 5.0$ hr.

In the event that the age distribution is not scaled down as described above some problems arise. First we have a biological choice. Do the cells which are overage according to the new parameters enter the next state instantaneously, or is there an initial delay until the cells "realize" that they are behind schedule?

A computational question arises as to how to handle these overage cells in the

age-transfer method, since additional provision would have to be made somehow for the first cycle following the change. Also, since the age distribution is discrete (a histogram) in the age-transfer computation, the new average transfer time must be an integral divisor (convenient multiple) of the old for a shrinkage (expansion). The computational difficulties are easily solved by switching to the Monte Carlo method. Here, any scaling in the existing age distribution is easily handled by multiplying the stored waiting time for each cell by the scale factor.

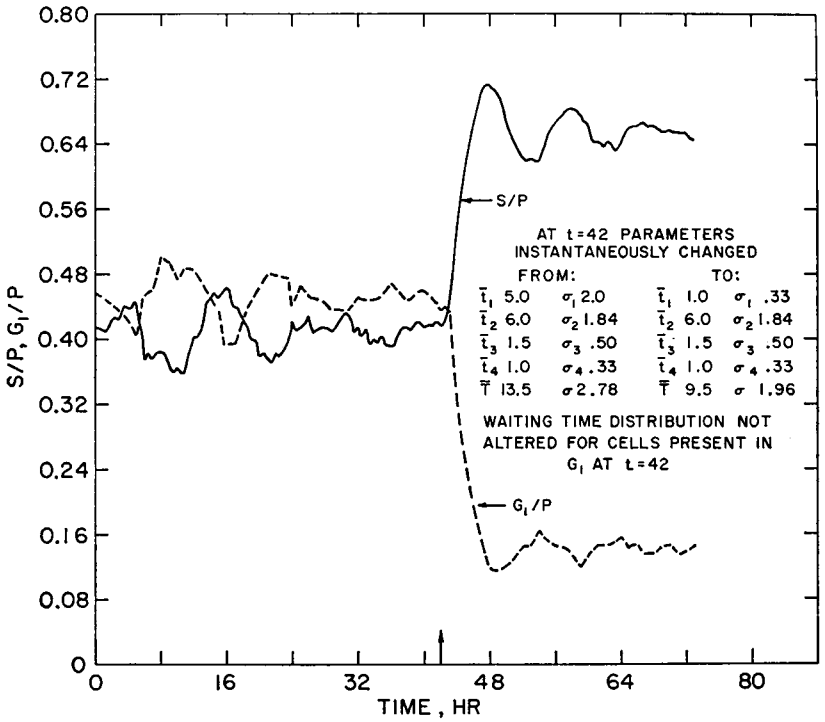

FIGURE 17 The approach of S/P and G_1/P to the new equilibrium fraction-in-state upon sudden shortening (at $t = 42$ in the figure) of the transit time in G_1. Age distribution of cells present in G_1 at the time of the change unaltered.

Consider, for example, a case where following the disturbance each cell waits, before moving to the next state, a time interval appropriate to the waiting-time distribution before the disturbance. For each subsequent move, its transit time is chosen by the new parameters (i.e., the scale factor is 1.0). In Fig. 17, the results for S/P and G_1/P are shown for this case. The change in parameters is exactly the same as that for Fig. 16, but the waiting time for the *first* postdisturbance transfer from G_1 was chosen according to the predisturbance distribution. We see that the first maximum is reduced from 76.4 to 71.2% and is delayed 3.4 hr with respect to the case when the age and waiting times are scaled at the time of the disturbance. The shape of the first peak is also affected.

VI. DISCUSSION

A stochastic theory of cell kinetics has been developed in the previous sections based on a model in which the over-all cycle is composed of several phases. The characteristic parameters of the model are the average transit times, $\bar{t}_i$, through each phase and the associated variance σ_i^2. Analytic and computer techniques have been used to investigate the properties of this model. When appropriate analytic approximations could be carried out, the results were in excellent agreement with those derived from computer experiments. The experimentally well-known cyclic behavior of cell proliferation is reproduced by the theory. A detailed analysis of synchrony decay and of labeling experiments has been carried out. The time development of age and waiting-time distributions were obtained directly from the computer experiments. The analysis of synchrony decay showed that the exponential decay is controlled by a characteristic constant, λ, which depends only on the cell cycle time, $\bar{T}$, and the over-all variance, σ^2, in the form $\lambda = 2\pi^2(\sigma/\bar{T})^2$.

Once a quantitative description of cell kinetics is at hand, one can inquire how the system responds to outside disturbances, such as changes in temperature, response to radiation, etc. The "equilibrium" state of a proliferating cell system may be defined as that corresponding to asynchronous log-phase growth. A corollary of this state is the constancy of the fraction of cells in any given phase of the cell cycle. If any of the cell parameters is altered the system will adjust to the new "equilibrium" state. Our computer studies showed that, analogous to the decay of synchrony, the system responds to a sudden disturbance by approaching the new "equilibrium" state via damped periodic transients. The rate of damping (exponential decay) is controlled by λ. Thus, the response time of the system, $\bar{T}/\lambda$, is as basic a parameter of the system as the cell cycle time since it determines the response to perturbations.

In illustrating typical cell kinetic behavior, parameters appropriate to a line of Chinese hamster cells were used. While it is apparent that the calculations reproduce the general features of the relevant experiments, no attempt was made to use cell parameters adjusted to any particular experiment. It would have been desirable to demonstrate the agreement between the model calculations and experiment, but this would have made the paper impractically long. Rather, the spirit of these calculations is to present most of the information needed for analyzing cell kinetics, with the hope that this will encourage experimenters to apply the results to their particular cell systems and to suggest new experiments to explore cell responses to disturbances.

We are indebted to Miss E. Wolfson for invaluable help with our programming problems. We are particularly grateful to Dr. P. Kemmey for suggesting and developing the age transfer method. We are indebted to many in the Medical and Biology Departments at Brookhaven for encouragement, support, and stimulating discussions, in particular to Drs. D. G. Baker, V. P. Bond, H. A. Johnson, and J. Van't Hof.

This work was supported under the auspices of the U. S. Atomic Energy Commission.

Received for publication 27 April 1968 and in revised form 25 June 1968.

APPENDIX I

Correction for Skewness

We have described the density function for the generation time as though it could be described by a Gaussian. That is, we have used

$$f(t) = \frac{1}{2\pi} \exp\,[(t - \bar{T})^2/2\sigma^2]. \tag{I.1}$$

Since the generation time can never be less than zero, we actually use for our density

$$\begin{aligned} \tilde{f}(t) &\equiv f(t) \qquad t > 0 \\ &\equiv 0 \qquad t \leqq 0. \end{aligned} \tag{I.2}$$

Let us define the average generation time and variance for $\tilde{f}$ by $\tilde{t}$ and $\tilde{\sigma}^2$. We also define

$$\Delta\bar{T} \equiv (\tilde{t} - \bar{T}) \qquad \Delta\sigma^2 \equiv (\tilde{\sigma}^2 - \sigma^2). \tag{I.3}$$

These are the quantities we must add to the nominal generation time and variance, σ and $\bar{T}$, to obtain the true generation time and variance for $\tilde{f}$. By inspection, one can see that cutting off the negative tails of the density will move the average to the right, and decrease the variance, since that quantity corresponds to a moment of inertia. Explicitly

$$\Delta\bar{T} = \int_{-\infty}^{\infty} \tilde{t}f(t)\,dt - \int_{-\infty}^{\infty} tf(t)\,dt,$$

which gives after standard changes of variables and integration,

$$\Delta\bar{T} = (\sigma/\sqrt{2\pi}) \exp\,(-\bar{T}^2/2\sigma^2) - \bar{T}/2[1 - \operatorname{erf}\,(\bar{T}/\sqrt{2}\sigma)] \tag{I.4}$$

where $\operatorname{erf}\,(z) \equiv \frac{2}{\sqrt{\pi}} \int_0^z \exp\,(-t^2)\,dt.$

Similarly,

$$\begin{aligned} \Delta\sigma^2 = &\int_0^{\infty} [t - (\bar{T} + \Delta\bar{T})]^2 f(t)\,dt \\ &- \int_{-\infty}^{\infty} [t - \bar{T}]^2 f(t)\,dt. \end{aligned} \tag{I.5}$$

Neglecting terms involving $\Delta\bar{T}$, since they are second order corrections, we obtain

$$\Delta\sigma^2 = \frac{-\sigma^2}{\sqrt{2\pi}} \left[\frac{\bar{T}}{2} \exp\,(-\bar{T}^2/2\sigma^2) + \frac{1}{2}\,(1 - \operatorname{erf}\,(\bar{T}/\sqrt{2}\sigma))\right]. \tag{I.6}$$

The proper values of $\bar{T}$ and σ to obtain a given $\tilde{t}$ and σ can be determined numerically using equations 15 and 16. For a coefficient of variation of one-third or less, we see by Table A-1 that the corrections may be neglected.

TABLE A-I

C. V.	1/3	1/2
$\Delta\sigma^2$	$-0.014\sigma^2$	$-0.131\sigma^2$
$\Delta\bar{T}$	$0.002\bar{T}$	$0.004\bar{T}$
$\bar{\sigma}$	0.993σ	0.932σ
$\bar{\imath}$	$1.002\bar{T}$	$1.004\bar{T}$

APPENDIX II

Fluctuations

Graphs of the time development of various quantities during a given experiment with a finite number of cells naturally will show fluctuations about the graph for the time development for the expected values of such quantities. We will discuss this phenomenon for the fractions-in-state in order to give an idea of how these fluctuations are exhibited within the framework of our model. To illustrate the problem, graphs of the time development of the fraction in synthesis for a Monte Carlo numerical experiment are displayed in Figs. 18 and 19 and compared with the time development of the average of the same quantity computed by the statistical method. In Fig. 18 the initial experimental population is taken to be six cells at the beginning of G_1. The C.V. of this population is 0.2. In Fig. 19 the initial experimental population is again six cells, but with a C.V. of 0.1.

Such fluctuations will be of biological interest when attempts are made to apply statistical kinematics to the development of a system such as an embryo, or the kinetics of a system such as an intestinal crypt which maintains a small number of cells. Consider, for example, the number of cells in a normal crypt. This is several hundred on the average. If a large percentage of these are damaged, as in radiation experiments, then fluctuations downward may cause whole crypts to disappear, whereas a calculation of the average number of cells surviving in a crypt indicates that the average crypt survives.

We will define two functions which will be helpful in considering fluctuations and will estimate their properties in a very rough way for the fraction in state. If the quantity of interest is $q(t)$ and its average over an ensemble of systems with identical initial conditions is $\bar{q}(t)$, then we define the time dependent variance to be

$$S^2([a]; t) \equiv [q(t) - \bar{q}(t)]^2. \tag{II.1}$$

Here $[a]$ indicates the various parameters of the ensemble and the state under consideration such as C.V., and initial population and its distribution. The second function is the fluctuation, $F\ell(t)$, which is the integral of the variance over time, and is comparable to the variation in the analysis of nonrandom functions. We define

$$[F\ell([a]; t)]^2 \equiv \int_0^t S^2([a]; t)\, dt \tag{II.2}$$

and $\langle F\ell^2([a]; t)\rangle$ is the average of $F\ell^2(t)$ over an ensemble of an infinite number of identical experiments.

Suppose the quantity $q(t)$ is taken for the present discussion to indicate the fraction of the population in state i. Let us illustrate the meaning of these functions by indicating for this case how one would obtain an estimate of $\langle F\ell^2\rangle$ from a numerical experiment. Starting with

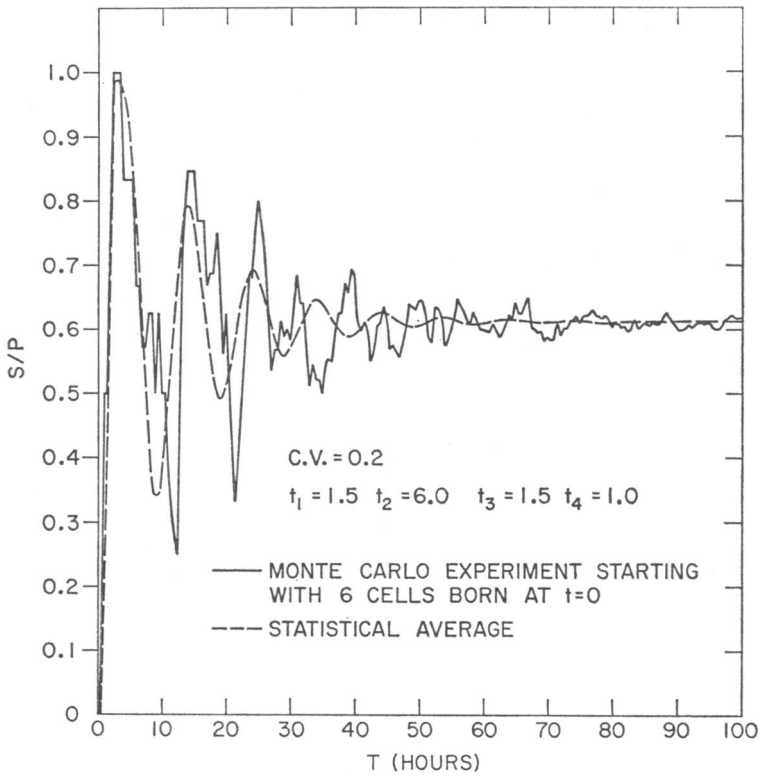

FIGURE 18 Fluctuations from the average in the time development of the synthesis fraction-in-state (S/P) for C.V. = 0.2.

an initial population numbering N_0, we run the Monte Carlo simulation m times, using different "random numbers" to generate transit times each of the m times. For each run, if we subtract the graph provided by the statistical method for $\overline{P_i(t)/P(t)}$ from the graph of that run, we can evaluate equation II.1, and do the integral in equation II.2 numerically. Taking the arithmetic average of the m integrals gives us our estimate for $\langle Ft^2 \rangle$. We can find also the variance for our estimate, as in any other experimental estimate of an average.

According to the central limit theorem of probability theory,[10] if S_n is the sum of n mutually independent random variables with a common distribution having expectation μ and variance σ^2, then the probability that the average S_n/n, deviates from the mean μ by a number less than $(B)/\sqrt{n}$ where B is a constant, equals $\Phi(B)$, where

$$\Phi(B) = 1/\sqrt{2\pi} \int_{-\infty}^{B} e^{-x^2/2}\, dx. \tag{II.3}$$

We will use this decrease in the deviation by $1/\sqrt{n}$ to estimate orders of magnitude. (For B equal 3, $\Phi(B)$ gives a 99.9% probability.) Consider for an initial population N_0 the quantity

$$S([P_i, N_0, \cdots]; t) = [P_i(t)/P(t) - [\overline{P_i(t)/P(t)}]]. \tag{II.4}$$

[10] See reference 38, p. 229.

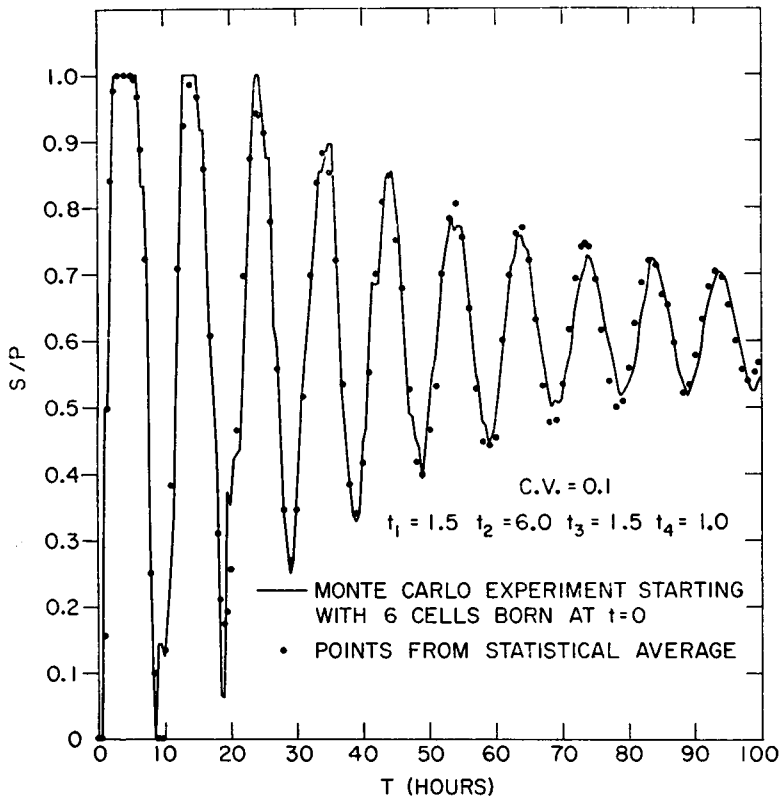

FIGURE 19 As Fig. 18 for C.V. = 0.1.

Now consider an initial population identical to the above except that it is m times as large, that is each one of the N_0 cells are replaced by m cells, all of which start with initial conditions identical to the cell they replaced. Then both the numerator and denominator of the first term in (II.4) are sums of m random variables and so, by the central limit theorem, should be of order of magnitude m times their average for N_0 cells, plus a term of order $\sqrt{m}$ times their average. The difference between the quotient of the averages and $[\overline{P_i/P}]$ should be of the same order of magnitude so that

$$S([P_i, N_0, \cdots]; t) = 0[S([P_i, N_0, \cdots]; t)/\sqrt{m}]. \tag{II.5}$$

By a similar argument,

$$Fl[mN_0, \cdots; t] \sim Fl[N_0, \cdots; t]/\sqrt{m}. \tag{II.6}$$

Let us next look at the behavior of $S^2(t)$ for large t. First we assert that the variance of $P_i(t)$ and $P(t)$ are of the same order of magnitude. If we assume this, we can find the order of magnitude of S by taking $\overline{P}_i(t)/\overline{P}(t) \cong [\overline{P_i/P(t)}]$ and looking at the similar quantity

$$\underline{S}^2(t) = [(P_i(t) - \overline{P}_i(t))/\overline{P}(t)]^2. \tag{II.7}$$

When the population of cells is large, we can take $\overline{P_i}(t)/(\overline{P}t)$ as equal to the probability that a randomly selected cell is in state P_i at time t. Then $[1 - \overline{P_i}(t)/\overline{P}(t)]$ is the probability of the cell being in some other state. If we neglect correlations between cells implied by our model then we can use the binomial distribution for the N cells present.

$$\text{Then if} \quad p(t) \equiv \overline{P_i}(t)/\overline{P}(t) \text{ and } q(t) \equiv [1 - p(t)] \tag{II.8}$$

we have

$$[P_i(t) - \overline{P_i(t)}]^2 \sim Npq$$

and

$$\underline{S}^2 \sim 0(1/N). \tag{II.9}$$

Since the number, N, of cells present increases exponentially with time, this implies that $F\ell(\infty)$ is finite.

Finally we note that when the coefficient of variation of the cycle is zero, $S(t)$ and $F\ell(t)$ are both identically zero. Qualitatively the graphs for $P_i/P(t)$ are considerably smoother for populations having C.V. = 0.1, than are the graphs for populations having C.V. = 0.2. It may be of interest to see if the fluctuation approaches zero monotonically with decreasing C.V. This is not at all obvious since the fluctuations for a smaller C.V. may last longer than for a larger C.V.

In conclusion, we have outlined an approach to understanding the fluctuations in biological quantities arising from the randomness of the transit times. These fluctuations may be quantitatively described in terms of a time-dependent variance, $S^2(t)$, and a fluctuation function, $F\ell(t)$. When we applied these quantities to the fraction-in-state, we found that S^2 varies inversely with the initial population and $F\ell(\infty)$ is finite. The graphs in Figs. 18 and 19 illustrate experimental effects of the fluctuations. These graphs and the suggested procedure following equation II.2 show that our numerical methods are adequate to estimate the fluctuations if they should be mathematically intractable in particular cases.

APPENDIX III

As we have noted in the text, the method of solution of the renewal equation,

$$\dot{P}(t) = f(t) + 2\int_0^t \dot{P}(t - t')f(t')\,dt' \tag{III.1}$$

with

$$\begin{aligned} f(t) &\equiv \eta^m \frac{t^{m-1}}{(m-1)!} - e^{\eta t} \qquad t > 0 \\ &\equiv 0 \qquad\qquad\qquad\qquad t \leqq 0 \end{aligned} \tag{III.2}$$

is well known (40, 41). Since our complete derivation is quite brief, we feel that a summary will be of value for reference. If we denote the Laplace transform of a function $f(t)$ by $L(f)$, then we use the following theorem (45). Consider two functions $f(t)$ and $g(t)$, both continuous

for $t > 0$ and identically zero for $t \leqq 0$, and both bounded above by an exponential of the form e^{bt} where b is a constant.
If the convolution, $h(t)$, is defined by

$$h(t) \equiv \int_0^t g(t')f(t - t')\,dt' \tag{III.3}$$

then

$$L[h] \equiv L[f]L[g]. \tag{III.4}$$

The functions $f^{(n)}(t)$ in the series solution, equations 9 and 10, are of this form, therefore applying equation III.4 successively to each term of equation 9 gives

$$\begin{aligned} L[\dot{P}] &= \sum_{n=1}^{\infty} 2^{n-1}\{L[f]\}^n \\ &= L[f]/[1 - 2L[f]]. \end{aligned}$$

Now using the explicit form equation III.2, we have (45)

$$L[f](z) = \eta^m/(z + \eta)^m.$$

Hence

$$L[\dot{P}](z) = 1/[(z + \eta)^m/\eta^m - 2] \tag{III.5}$$

or factoring the mth order polynomial in the denominator

$$L[\dot{P}](z) = \eta^m/\prod_{j=1}^{m}(z - r_j) \tag{III.6}$$

where r_j is the jth root of the equation

$$\begin{aligned} (z + \eta)^m - 2\eta^m &= 0 \\ r_j &= 2^{1/m}\eta \exp(2\pi i(j)/m) - \eta \\ j &= 1, 2, \cdots m. \end{aligned} \tag{III.7}$$

Expanding equation III.6 as a sum of a partial fraction we obtain

$$L[\dot{P}] = \eta^m \sum_{j=1}^{m} A_j/(z - r_j) \tag{III.8}$$

where A_j is simply found by multiplying equation III.6 by $(z - r_j)$ and allowing z to approach r_j, giving

$$\begin{aligned} A_j &= \left\{\prod_{k\neq j}^{m}(r_j - r_k)\right\}^{-1} \\ &= \left\{\eta^{m-1}2^{m-1/m} \exp(2\pi i(j)(m - 1)/m) \prod_{\substack{k\neq j \\ k=1}}^{m}[1 - \exp(2\pi i(k - j)/m)]\right\}^{-1}. \end{aligned} \tag{III.9}$$

We can simplify the form of A_j by noticing that

$$\prod_{\substack{k=1 \\ k \neq j}}^{m} [1 - \exp(2\pi i(k-j)/m)] = \lim_{z \to 1} [(z^m - 1)/(z - 1)] = m. \quad \text{(III.10)}$$

Therefore

$$A_j = \eta^{m-1} 2^{m-1/m} m \exp(2\pi i(j)(m-1)/m)$$

$$j = 1, 2, \cdots, m$$

or

$$j = 0, 1, \cdots, m-1. \quad \text{(III.11)}$$

Using the fact that the inverse Laplace transform of $1/(z - r_j)$ is $e^{r_j t}$ we obtain finally from equations III.8 and III.11,

$$\dot{P}(t) = [\eta/2^{m-1/m} m] \sum_{j=0}^{m-1} \frac{\exp [2^{1/m} e^{(2\pi i)j/m} - 1]\eta \cdot t}{\exp [(2\pi i)j(m - 1/m)]} \quad \text{(III.12)}$$

which is the derivative of the result obtained by Kendall (19) using a set of coupled differential equations for his model. Notice that $\dot{P}$ is real in form, since for every term with $j \neq 0$ or $m/2$ in the sum we also have the complex conjugate term.

REFERENCES

1. Howard, A., and S. R. Pelc. 1953. *Heredity*. **6** (Suppl.):261.
2. Hughes, W. C., V. P. Bond, G. Brecher, E. P. Cronkite, E. P. Painter, H. Quastler, and F. G. Sherman. 1958. *Proc. Natl. Acad. Sci. U.S.* **44**:476.
3. Quastler, H., and F. G. Sherman. 1959. *Exptl. Cell Res.* **17**:420.
4. Quastler, H. 1960. *Ann. N.Y. Acad. Sci.* **90**:580.
5. Quastler, H. 1963. *In* Cell Proliferation. L. F. Lamerton and R. V. M. Fry, editors. A Guiness Symposium, Blackwell Scientific Pubs. Oxford, England. 18.
6. Puck, T. T., and J. Steffen. 1963. *Biophys. J.* **3**:379.
7. Lotka, A. J. 1907. *Science*. **26**:21.
8. Lotka, A. J. 1913. *J. Wash. Acad. Sci.* **3**:241, 289.
9. Lotka, A. J., and F. R. Sharpe. 1911. *Phil. Mag.* **21**:435.
10. Lotka, A. J. 1922. *Proc. Natl. Acad. Sci. U.S.* **8**:339.
11. Lotka, A. J. 1945. *In* Essays on Growth and Form Presented to D'Arcy Wentworth Thompson. W. E. L. Clark and P. P. Madawar, editors. Oxford Univ. Press, Oxford, England. 355.
12. Lotka, A. J. 1939. *Ann. Math. Statist.* **10**:1.
13. Herbelot, L. 1909. *Bull. Trimestrivel Inst. Actuaries Francais.* **19**:293.
14. Hertz, P. 1908. *Math. Annalen.* **65**:89.
15. Feller, W. 1941. *Ann. Math. Statist.* **12**:243.
16. Bellman, R., and T. E. Harris. 1948. *Proc. Natl. Acad. Sci. U.S.* **34**:601; 1952. *Ann. Math.* **55**:280.
17. Harris, T. E. 1959. *In* The Kinetics of Cellular Proliferation. F. Stohlman, Jr., editor. Grune and Stratton, Inc., New York. 368.
18. Bartlett, M. S. 1949. *J. Statist. Soc.* (*B*). **11**:211; *also* 1946. Stochastic Processes. Notes of a course given at Univ. of North Carolina.
19. Kendall, D. G. 1948. *Biometrika*. **35**:316.
20. Feller, W. 1939. *Acta Biotheoretica*. **5**:11.
21. Williams, E. J. 1961. *Biometrics*. **17**:349.

22. PRESCOTT, D. M. 1959. *Exptl. Cell Res.* **16**:279.
23. DAWSON, K. B., H. MADOC-JONES, and E. O. FIELD. 1965. *Exptl. Cell Res.* **38**:75.
24. VON FOERSTER, H. 1959. *In* The Kinetics of Cellular Proliferation. F. Stohlman, Jr., editor. Grune and Stratton, Inc., New York. 382.
25. TRUCCO, E. 1965. *Bull. Math. Biophys.* **27**:285; 449.
26. HOFFMAN, J. G., N. METROPOLIS, and V. GARDINER. 1956. *J. Natl. Canc. Inst.* **17**:175.
27. ENGELBERG, J. 1964. *Exptl. Cell Res.* **36**:647.
28. MERKLE, T. C., R. N. STUART, and J. W. GOFFMAN. 1965. The Calculation of Treatment Schedules for Cancer Chemotherapy. Univ. of California, Liverware, Calif. Lawrence Radiation Lab. Report U.C.R.L. 14505; U.C.R.L. 14505 Part II.
29. STUART, R. N. 1966. Kinetics of Cell Growth. Univ. of California, Lawrence Radiation Lab. Report U.C.R.L. 70039.
30. HIRSCH, H. R., and J. J. ENGELBERG. 1965. *J. Theoret. Biol.* **9**:303.
31. BARRETT, J. C. 1966. *J. Natl. Cancer Inst.* **37**:443.
32. HAHN, G. M. 1966. *Biophys. J.* **6**:275.
33. HIRSCH, H. R., and J. J. ENGELBERG. 1965. *J. Theoret. Biol.* **9**:297.
34. TAKAHASHI, M. J. 1966. *J. Theoret. Biol.* **13**:202.
35. PASKIN, A., B. BRONK, and G. J. DIENES. 1968. *In* Recovery and Repair Mechanisms in Radiobiology. Brookhaven Symposia in Biology, No. 20, 169.
36. KUBITSCHEK, H. 1962. *Exptl. Cell Res.* **26**:439.
37. FROESE, G. 1965. *Intern. J. Radiation Biol.* **10**:353.
38. FELLER, W. 1950. An Introduction to Probability Theory and its Applications. John Wiley and Sons, New York. 2nd edition. **I**:141.
39. SINCLAIR, W. K. 1964. *Nature.* **203**:247.
40. FELLER, W. 1966. An Introduction to Probability Theory and its Applications. John Wiley and Sons, New York. Vol. II, Chap. XIV (particularly problem 1).
41. HIRSCH, H. R., and J. ENGELBERG. 1966. *Bull. Math. Biophys.* **28**:391.
42. BRONK, B. 1968. *J. Theoret. Biol.* In press.
43. ENGELBERG, J., and H. R. HIRSCH. 1966. *In* Cell Synchrony. I. L. Cameron and G. M. Padilla, editors. Academic Press, Inc., New York. 14.
44 RAO, P. N., and J. ENGELBERG. 1966. *In* Cell Synchrony. I. L. Cameron and G. M. Padilla, editors. Academic Press, Inc., New York. 332.
45. DETTMAN, J. W. 1965. Applied Complex Variables. Macmillan, New York. 406, 415; theorem 9.2.7.