ON CERTAIN FUNCTIONS USEFUL IN POPULATION EVOLUTION STUDIES

Bernard I. Spinrad

January 1979
WP-79-10

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

PREFACE

The problem of the size of a cell population, descendant from a single cell or a small group of cells, each of which is subject to the alternative fates of death or division into two new cells, is a prototype problem for a number of different types of population models. I first came across it while studying models for the penetration of mutations into a population, for example. At that time, about 20 years ago I could not find an exact solution in the literature available to me, although asymptotic properties were well known.

I recently reviewed the problem, primarily as a mathematical recreation. In so doing, I blundered into the solution. I do not know whether it has been published, either before 1958 or by 1978. This memorandum may be useful to researchers dealing with the problem or with its analysis. I write it partly for its potential value as a collection, in one place, of properties of the solutions of the stated problem. However, I must also admit that, to a certain extent, it is written to illustrate that scientific research is much more often dominated by heuristics and serendipity than by pure logical inference.

PROBLEM STATEMENT

Let $P_n(t) \equiv$ Probability that at any given time the population consists of "n" units.

$f \equiv$ Probability that a single unit will reproduce by fission into two units.

$\therefore 1-f =$ Probability that the unit will die before fissioning

$\rho \equiv 2f-1 =$ Average relative increase in population per generation ("reactivity")

$\tau \equiv$ Average time between formation of a unit and its termination by either death or fission.

With the further assumption (which is itself an approximation valid only in the mean for most populations) that the probability of death or division is constant in time for any existing unit--i.e. that the population is a renewable one--the problem may be written mathematically as

$$\tau \dot{P}_n = \frac{1+\rho}{2}(n-1)P_{n-1} - nP_n + \frac{1-\rho}{2}(n+1)P_{n+1} \qquad (1)$$

for $n \geq 1$.

Equation (1) is in principal solvable from initial conditions, for example that $P_1(0) = 1$, $P_n(0) = 0$ for $n > 1$, the problem of populations descendant from a single ancestor.

HEURISTICS AND SERENDIPITY

I actually know of no finite, general way of solving a set of equations such as (1). However, two points are notable:

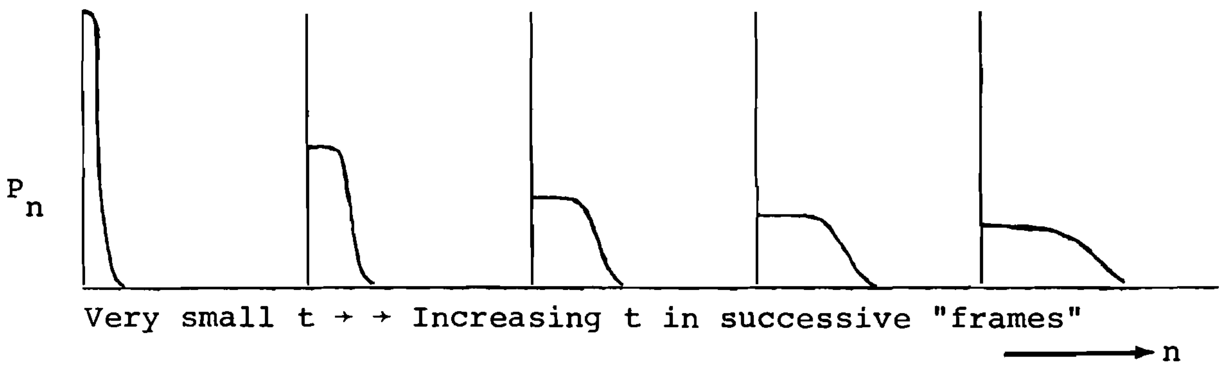
(a) Terms in "n" disappear if P_n is a constant, and the right hand side goes to zero if P_n is a constant divided by n.

(b) Both these "possible solution sets" diverge in the sense that $\sum_{n=1}^{\infty} P_n$ would increase without limit.

Thinking about the system as a multiplying one, a situation with which, as a reactor physicist, I am familiar, at any given time one must expect that the probability of n being much greater than some average expected value will be very small. The average expected value will, however, grow exponentially. Thus, I expect that the solution will, in the long run, go to a function which is approximately constant or varying as $\frac{1}{n}$, up to some value of n_o, a function of time, above which it will decrease rapidly. I expect, in other words, an evolution in time of P_n according to Sketch A or Sketch B. (In both these sketches, I have plotted P_n as a function of a continuous, rather than a discrete value of n.)

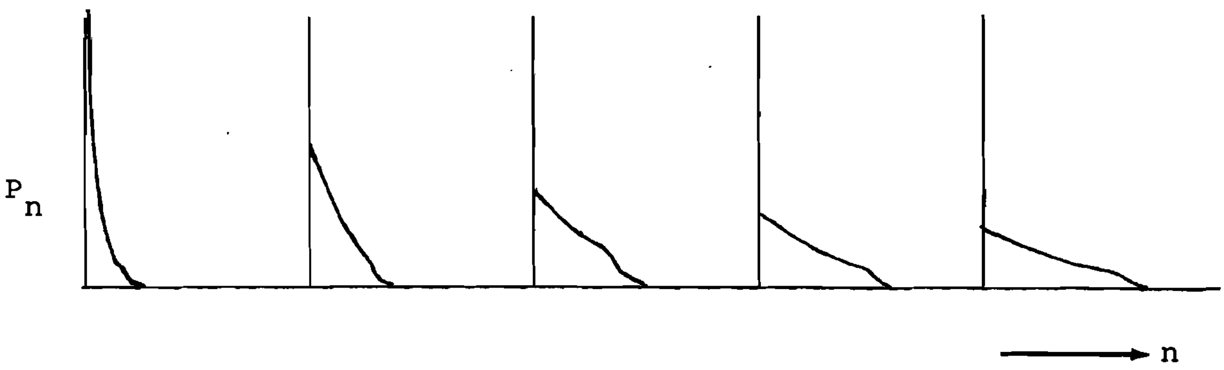
Sketch A

Time evolution (schematic) of P_n if it approaches a constant up to some n_o



Sketch B

Time evolution if P_n approaches C/n



Repeated trials of solutions of the form $f(X) + g(X)h(n)$ convinced me that this wasn't the way to get a solution; so I looked harder why this was. The answer was clear as to what was wrong: the differentiation on the left hand side of (1) was not providing any manipulation of n, and that was needed.

An easy way of manipulating n on differentiating is to write a solution as some function to the n^{th} power. Accordingly I tried solutions of the form $P_n = \frac{X(t)[Y(t)]^n}{n}$ and $P_n = u(t)[v(t)]^n$. This turned out to be pay dirt.

I first present the results of trying $P_n = XY^n/n$, which, because it is *not* the desired solution, I call $P_n^{(2)}$.

$P_n^{(2)}$, A NON-PHYSICAL SOLUTION

$$P_n^{(2)}(t) \equiv X(t)[Y(t)]^n/n \tag{2}$$

Substituting into (1)

$$\frac{\tau\dot{X}Y^n}{n} + \tau X\dot{Y}Y^{n-1} = \frac{1+\rho}{2}XY^{n-1} - XY^n + \frac{1-\rho}{2}XY^{n+1} \tag{3}$$

The only term with "n" in it is the first term on the lhs of (3). If this is to be a solution $\dot{X} = 0$. Thus, X is a constant. The rest of the equation then reduces to

$$\tau\dot{Y} = \frac{1+\rho}{2} - Y + \frac{1-\rho}{2}Y^2 \tag{4}$$

Equation (4) can be solved routinely to give

$$Y = \frac{c_1 - e^{-\rho t/\tau}}{c_1 - \frac{1-\rho}{1+\rho}e^{-\rho t/\tau}} \tag{5}$$

where "c_1" is an arbitrary constant of integration. Thus, a solution of equation (1) is the set of probabilities

$$P_n^{(2)} = c_2\left(\frac{c_1 - e^{-\rho t/\tau}}{c_1 - \frac{1-\rho}{1+\rho}\, e^{-\rho t/\tau}}\right)^n \Big/ n \qquad (6)$$

The constant c_2 is, of course, the value assigned to X.

That $P_n^{(2)}$ is non-physical can be noted by summing $P_n^{(2)}$ over all n. The result is

$$\sum_{n=1}^{\infty} P_n^{(2)}(t) = c_2 \ln \frac{1}{1 - \dfrac{c_1 - e^{-\rho t/\tau}}{c_1 - \frac{1-\rho}{1+\rho}\, e^{-\rho t/\tau}}}$$

$$= c_2\left[\frac{\rho t}{\tau} + \ln \frac{(1+\rho)c_1 - e^{-\rho t/\tau}}{2\rho}\right] \qquad (7)$$

For some t, the sum must be greater than 1, regardless of our choice of c_2. This would amount to a probability greater than unity. The only exceptions are the trivial case, $c_2 = 0$, and the non-trivial one, $c_1 = 0$. If $c_1 = 0$, we get

$$P_n^{(2)} = c_2\left(\frac{1+\rho}{1-\rho}\right)^n \Big/ n$$

whose sum also diverges for positive ρ.

$P_n^{(1)}$, A PHYSICAL SOLUTION

We now try

$$P_n^{(1)}(t) \equiv u(t)\,[v(t)]^n \qquad (8)$$

Substituting into (1)

$$\tau\dot{u}v^n + n\tau u\dot{v}v^{n-1} = \frac{1+\rho}{2}(n-1)uv^{n-1} - nuv^n + \frac{1-\rho}{2}(n+1)uv^{n+1} \qquad (9)$$

Isolating terms in "n" gives

$$\tau \dot{v} v^{n-1} = \frac{1+\rho}{2} v^{n-1} - v^{n} + \frac{1-\rho}{2} v^{n+1} \tag{10}$$

This has the same form as (4), and the same solution:

$$v = Y = \frac{c_1 - e^{-\rho t/\tau}}{c_1 - \frac{1-\rho}{1+\rho} e^{-\rho t/\tau}} \tag{11}$$

Terms remaining in (9) give rise to:

$$\tau \dot{u} v = - \frac{1+\rho}{2} u + \frac{1-\rho}{2} u v^2 \tag{12}$$

This has the solution

$$u = \frac{c_2 e^{-\rho t/\tau}}{(c_1 - e^{-\rho t/\tau})(c_1 - \frac{1-\rho}{1+\rho} e^{-\rho t/\tau})} \tag{13}$$

Then

$$P_n^{(1)} = c_2 e^{-\rho t/\tau}\left(c_1 - e^{-\rho t/\tau}\right)^{n-1}\left(c_1 - \frac{1-\rho}{1+\rho} e^{-\rho t/\tau}\right)^{-n-1} \tag{14}$$

The sum of all $P_n^{(1)}$ converges:

$$\sum_{n=1}^{\infty} P_n^{(1)} = c_2\left(\frac{1+\rho}{2\rho}\right)\Big/\left(c_1 - \frac{1-\rho}{1+\rho} e^{-\rho t/\tau}\right) \tag{15}$$

If we use a standard initial condition that, at zero time, the sum of all P_n shall be unity, we solve for c_2 to get

$$c_2 = \frac{2\rho}{1+\rho}\left(c_1 - \frac{1-\rho}{1+\rho}\right) \tag{16}$$

RELATION BETWEEN $P_n^{(1)}$ AND $P_n^{(2)}$

If we differentiate $P_n^{(2)}$ with respect to time, we get $P_n^{(1)}$, except for a normalizing factor which may be ignored since the c_2's in equations (7) and (14) are both arbitrary anyway. This suggests that further differentiations could lead to other solutions. Examination of equation (1) shows us that I should have known this from the beginning! Indeed, inverse differentiation, i.e., indefinite integration could also give valid solution of the differential equations. We shall return to this point later.

THE PROTOTYPE PROBLEM

Equation (14) has a property which enables us immediately to set c_1 for an important problem. This is the problem for which $P_1(0) = 1$ and all other $P_n = 0$ at time zero, and is thus the one whose time evolution was sketched in Sketches A and B. We note that there is a factor $(c_1 - e^{-\rho t/\tau})^{n-1}$ in $P_n^{(1)}$. At zero time, this is $(c_1 - 1)^{n-1}$, and if $c_1 = 1$ all terms vanish at $t = 0$ except for P_1. Thus, we may write

$$P_n = P_n^{(1)}(c_1=0) = \left[\frac{2\rho/(1+\rho)}{(1 - \frac{1-\rho}{1+\rho} e^{-\rho t})}\right]^2 e^{-\rho t/\tau} \cdot \left[\frac{(1 - e^{\rho t/\tau})}{(1 - \frac{1-\rho}{1+\rho} e^{-\rho t/\tau})}\right]^{n-1} \qquad (17)$$

as a solution for this important problem. We shall call this solution, P_n as given by equation (17), P_{n1}.

P_{n1} AS THE CORRECT SOLUTION

P_{n1} satisfies the differential equations, (1), and the boundary conditions of a problem which describes a population consisting initially of one unit. I suspect that this solution is unique from general theorems, but the way I fell into it makes me a little uneasy. Therefore, I then went about proving that P_{n1} led to the same moment equations as arise from summing equations (1) multiplied by powers of n. After doing this, I noted that this had to be the case; moreover, the moment equations, which are simple differential equations, may be solved one by one, as will be exhibited later. But since I know that, given sufficient moments, I can construct appropriate solutions which are as precise as I please (the physicists' faith in the power of variational methods!), for me, this proves uniqueness.

P_{n2} AND HIGHER SOLUTIONS

What about problems for which, for example, $P_1 = 0$ and $P_2 = 1$--or P_3, or any P_n? This is easy. I have already noted that the differentiation of $P_n^{(1)}$ leads to another set of solutions of (1). Specifically, then, if I start with P_{n1}, the set of functions $\dot{P}_{11}$, $\dot{P}_{21}$, $\dot{P}_{31}$... is also a solution. But $\dot{P}_{11}(0) = -\frac{1}{\tau}$ and $\dot{P}_{21}(0) = \frac{1+\rho}{2\rho}$. The equations are linear and superposable, so that the solution of the system when $P_1 = 0$, $P_2 = 1$ can be immediately written, under the defining symbolism P_{n2}, as

$$P_{n2} = \frac{2}{1+\rho}(P_{n1} + \tau\dot{P}_{n1}) \tag{18}$$

Explicitly, this results in

$$P_{n2} = \left[\frac{2\rho/(1+\rho)}{1 - \frac{1-\rho}{1+\rho}\, e^{-\rho t/\tau}}\right]^2 e^{-\rho t/\tau}\left\{\frac{2(-n+1)}{(1+\rho)}\left[\frac{1 - e^{-\rho t/\tau}}{1 - \frac{1-\rho}{1+\rho}\, e^{-\rho t/\tau}}\right]^{n-1}\right.$$
$$+ \frac{(n-1)}{(1+\rho)}\left[\frac{1 - e^{-\rho t/\tau}}{1 - \frac{1-\rho}{1+\rho}\, e^{-\rho t/\tau}}\right]^{n-2} \tag{19}$$
$$\left. + \frac{(1-\rho)(n+1)}{(1+\rho)}\left[\frac{1 - e^{-\rho t/\tau}}{1 - \frac{1-\rho}{1+\rho}\, e^{-\rho t/\tau}}\right]^{n}\right\}$$

It next becomes clear from inspection that solutions which, at time zero, have non-zero probabilities for n_o and zero probabilities for all $n > n_o$ can be written as linear combinations of P_{n1} and its first $(n_o - 1)$ derivatives.

At this point, I lost interest in solutions obtained by integrating P_{n1}; I assume they all diverge and are uninteresting.

SUMS AND SURVIVAL PROBABILITIES

At this point, we return to equation (15), with c_2 evaluated by (16) and $c_1 = 1$. In other words,

$$\sum_{n=1}^{\infty} P_{n1} = \frac{2\rho/(1+\rho)}{1 - \frac{1-\rho}{1+\rho} e^{-\rho t/\tau}} \qquad (20)$$

The sum of probabilities of state occupation for any state, initially unity, asymptotically approaches $2\rho/(1+\rho)$. This is a well known result. One point of interest is that (20) tells us how the asymptote is approached. The important point is that for all $\rho \leq 1$, $\sum_{n=1}^{\infty} P_{n1}$ asymptotically approaches a number less than one: even when ρ is close to unity, there is a finite probability that the line will become extinct, owing to the stochastic nature of the process of multiplication.

MOMENTS

Oddly enough, the population moments form a set of equations that can be solved deterministically. We define

$$M_k \equiv \sum_{1}^{\infty} P_n n^k \qquad (21)$$

as the k^{th} population moment. Thus, M^1 is the mean (expected) value of the total population, M^2 the mean-square, and so on. We may note that, according to this definition,

$$\sum_{1}^{\infty} P_n = M_o \qquad (22)$$

Returning to differential equations (1), we get, for M_o,

$$\tau \dot{M}_o = -(1-\rho)P_1 \qquad (23)$$

This is not a very useful equation, of course, unless we know $P_1(t)$. For the prototype problem, solved by P_{n1}, we do of course have $P_1(t)$, and the solution of (23) is (20).

For higher moments, we get a solvable recursive set: for M_1, the equation is

$$\tau\dot{M}_1 = \rho M_1 \tag{24a}$$

which for the standard problem is solved by

$$M_1 = e^{\rho t/\tau} \tag{24b}$$

The average population increases exponentially for positive ρ; the "chain reaction" is divergent. This is of interest because the number of "chains" keeps decreasing. In simple language, the mean population is the result of a relatively small fraction of the ancestral units multiplying strongly into the population. As $\rho \to 0$, this leads to a type of "gamble against the bank" situation: almost all ancestral units ultimately are devoid of progeny; but, because the odds are even, there will be one unit which has a great many progeny (if one has a large enough supply of chips!).

Successive moments can be found from:

$$\begin{aligned} \tau\dot{M}_2 &= 2\rho M_2 + M \\ \tau\dot{M}_3 &= 3\rho M_3 + 3M_2 + \rho M_1 \\ &\vdots \\ \tau\dot{M}_k &= \sum_{j=1}^{k} \binom{k}{j} M_{k+1-j} \left[\frac{1 + \rho + (1-\rho)(-)^2}{2}\right] \end{aligned} \tag{25}$$

Since M_1 can be solved, M_2 can be solved knowing M_1, and so forth. Specific solutions are exercises in elementary calculus.

ASYMPTOTIC PROPERTIES

As already noted, the asymptotic properties of the equation are well known, and are listed here just to bring everyting together.

For large t, the ratio $\left(\dfrac{1 - e^{\rho t/\tau}}{1 - \frac{1-\rho}{1+\rho}\, e^{-\rho t/\tau}}\right)$ can be very well represented by $\exp -(\frac{2\rho}{1+\rho}\, e^{-\rho t/\tau})$. Since for large t such terms as $(1 - e^{-\rho t/\tau})^{j}$ or $(1 - \frac{1-\rho}{1+\rho}\, e^{-\rho t/\tau})^{j}$, where j is a small integer, are essentially unity, we can reduce P_{n1} to:

$$P_{n1} \approx \left(\frac{2\rho}{1+\rho}\right)^2 e^{-\rho t/\tau} \exp -\left(\frac{2\rho n}{1+\rho}\, e^{-\rho t/\tau}\right) \tag{26}$$

Equation (26) can be manipulated more easily than (17) to get, for example, asymptotic properties of P_{n2}, etc.

For large n, we may consider n as a continuous variable in certain manipulations. For example, it can be useful to determine the probability that the population is n_o or larger. Defining this number as Q_{n_o}, we have

$$Q_n \equiv \sum_{n_o}^{\infty} P_{n1} \tag{27}$$

$$\approx \sum_{n_o}^{\infty} \left(\frac{2\rho}{1+\rho}\right)^2 e^{-\rho t/\tau} \exp\left(-\ \frac{2\rho n}{1+\rho}\, e^{-\rho t/\tau}\right) \tag{28a}$$

$$\approx \int_{n_o}^{\infty} \left(\frac{2\rho}{1+\rho}\right)^2 e^{-\rho t/\tau} \exp\left(-\ \frac{2\rho n}{1+\rho}\, e^{-\rho t/\tau}\right) dn \tag{28b}$$

$$\approx \frac{2\rho}{1+\rho} \exp\left(-\ \frac{2\rho n_o}{1+\rho}\, e^{-\rho t/\tau}\right) \tag{28c}$$

By a similar integration, the total population in groups n_o or larger can be determined as:

$$R_n = \sum_{n_o}^{\infty} n \; P_{n1} \tag{29a}$$

$$\approx e^{\rho t/\tau} \exp\left(- \frac{2\rho n_o}{1+\rho} e^{-\rho t/\tau}\right) \tag{29b}$$

Manipulation of the asymptotic properties is particularly useful when the aim is to justify treatment of large populations by continuum, rather than stochastic, models.