APPLICATION OF RECURRENT SERIES IN RENEWAL THEORY

By Alfred J. Lotka

Metropolitan Life Insurance Company

Summary. The application of integral equations to renewal theory in population analysis and problems of industrial replacement is beset with certain difficulties which have been particularly discussed by W. Feller (these *Annals* 1941 vol. 12 pp. 243–267). Some of these difficulties are avoided if the data of the problem are introduced into the analysis directly in the discontinuous form (tabulated by class intervals) in which they are usually supplied in a concrete case. A numerical example based on population statistics is presented, illustrating how, using discontinuous data, a recurrent series takes the place of the integral equation, and a finite exponential series appears in place of the Heaviside expansion of the previous solution. There is close analogy with the procedure previously presented, but with factorial moments appearing in place of ordinary moments.

The fundamental data being given for values of the replacement function at discrete intervals only, some question arises as to the applicability of the solution as an "interpolation" formula for non-integral values of the time t, and as to the effects of subdividing the class interval of the original data.

In the actual computation of the factorial moments a shift of origin by one-half class interval becomes necessary. An algorithm for effecting this shift is presented.

1. Methodology: Alternatives Available.

All application of mathematics to concrete situations involves a greater or less degree of conventionalisation, a substitution, "in place of intractable reality, of an ideal upon which it is possible to operate."[1]

This conventionalisation may be only such as to do little violence to the concrete data, as for example when, dealing with a large population, we treat the number $N(t)$ of individuals at time t as a continuous variable, knowing perfectly well that strictly speaking it varies by jumps of one unit at a time.[2]

In dealing with any particular concrete case there may be considerable choice as to the mode in which the conventionalisation or idealisation is carried out, and the particular place or step in the scheme at which it is introduced. A good illustration of this is met in the treatment of renewal theory, as applied to human populations or other biological or industrial aggregates.

The majority of authors who have dealt with the subject have set up their fundamental equations in terms of continuous variables. Many have gone fur-

[1] *Nature*, Vol. 110 (1922), p. 764.

[2] If the population is subject to extreme variation in numbers, such that $N(t)$ passes through small values, this disregard of their discontinuity may not be permissible.

ther than this in the process of conventionalisation and have assumed for the renewal function (net reproductivity) some more or less appropriate mathematical expression, such as a Charlier or a Pearson [1] frequency distribution, and have, wherever possible, carried out by standard methods the integrations involved.

Others, while retaining the formulation of the fundamental equations in continuous (infinitesimal) form, have made no specific assumptions regarding the analytical form of the renewal function, and have carried out the numerical integration by one of the established methods available for the approximate integration of arbitrary functions.

But there has also been a minority of authors who deemed it most appropriate, since the data of the problem are actually furnished in tabular (and hence discontinuous) form, to apply from the start discontinuous methods in formulating the fundamental equation for the problem. This equation then defines a recurrent series.

The most recent and also the most concise exposition of this approach to the problem is a paper by W. Dobbernack and G. Tietz presented at the Twelfth International Congress of Actuaries, 1940, *Proceedings*, vol. 4, p. 233. These authors, however, do not give any numerical application, and in consequence certain aspects of the analysis are not touched upon by them. A more detailed presentation, including numerical applications, was given by the late S. D. Wicksell[3] who, however, used only roughly approximate data (an over-all average net reproductivity for ages 20 to 44) and also introduced certain linear interpolations which would not be appropriate with more exact data, and which become unnecessary in the numerical operations if moments are introduced as indicated in what follows.

The purpose of the present paper is to exhibit this modification of the method of recurrent series, and at the same time to illustrate its relation to the method which proceeds in terms of a continuous variable, leading to an integral equation.

The $B(t - a)$ women born in the calendar year $(t - a)$, that is, between the times $(t - \frac{1}{2} - a)$ and $(t + \frac{1}{2} - a)$, will be a years old some time during the calendar year t, that is, between $t - \frac{1}{2}$ and $t + \frac{1}{2}$. If their births were evenly distributed over the year $t - a$, so will their birthdays of age a be over the year t, and their average age during that year will be a and the average number of survivors to that age during the year t will be approximately $B(t - a)p(a)$, where $p(a)$ is the probability, at birth, of surviving to age a. If the annual female reproductive rate, counting daughters only, is $m(a)$ at age a, then the $B(t - a)$-$p(a)$ survivors will, during the calendar year t, give birth to $B(t - a)p(a)m(a)$ daughters. If $B(t)$ is the total number of births of daughters in the calendar year t, then evidently, for positive values of t,

$$B(t) = \sum_{1}^{\omega} B(t - a)p(a)m(a), \tag{1}$$

[3] [2]; see also [3].

or, to simplify the notation,

$$B(t) = \sum_{1}^{\omega} c_a B(t - a). \tag{2}$$

Equation (1) or (2) defines a recurrent series of the general form

$$B(t) = c_1 B(t - 1) + c_2 B(t - 2) + \cdots + c_\omega B(t - \omega), \tag{3}$$

where some of the coefficients c may be zero and where ω denotes the upper limit of the reproductive period.

The trial substitution

$$B(t) = Qx^{-t} \tag{4}$$

in (3) gives

$$1 = c_1 x + c_2 x^2 + \cdots + c_\omega x^\omega. \tag{5}$$

The substitution (4) therefore satisfies (3) provided that x is a root of the equation (5) of degree ω for x; and the same is evidently true for the more general substitution

$$B(t) = \sum_{j=1}^{j=\omega} Q_j x_j^{-t}, \tag{6}$$

where x_j, with $j = 1, 2, \cdots \omega$, are the ω roots of (5).

Equation (5) leaves the ω coefficients Q_j indeterminate. In general they appear as arbitrary constants. In any concrete application they may be determined by "initial" conditions; that is, in order to make the problem determinate, it is necessary to be given the values of $B(t)$ for ω successive integral values of t, or some equivalent data.

While, for convenience in description, the analysis has been developed in terms of the year as time unit, the formulae are evidently independent of this choice of unit, provided that the unit employed is adequate for practical application.

Whatever the unit employed, for the direct application of (1) and (3) to a concrete case it is necessary to have the data in such form that values of $p(a)m(a)$ are known for integral values of a. The pertinent statistics do not usually come in that form, the fertility being usually known only for five year age groups, and though it may be sufficient for practical purposes to regard these quinquennial values as representing $p(a)m(a)$ for the midpoint of the group, this yields $p(a)\cdot m(a)$ for fractional values of a, as measured in five year units. We may then proceed as follows: putting

$$x = 1 + y \tag{7}$$

in (5) this becomes

$$
\begin{aligned}
1 = {} & \{c_1 + c_2 + c_3 + \cdots + c_\omega\} \\
& + \{c_1 + 2c_2 + 3c_3 + \cdots + \omega c_\omega\} y \\
& + \left\{ c_2 + 3c_3 + 6c_4 + \cdots + \frac{\omega(\omega - 1)}{2!} c_\omega \right\} y^2 \\
& + \left\{ c_3 + 4c_4 + 10c_5 + \cdots + \frac{\omega(\omega - 1)(\omega - 2)}{3!} c_\omega \right\} y^3 \qquad (8) \\
& + \cdots \\
& + \{c_\omega\} y^\omega \\
= {} & \sum_{h=0}^{h=\omega} \sum_{k=0}^{k=\omega-h} \binom{h+k}{h} c_{h+k} y^h .
\end{aligned}
$$

In application to a particular population, we shall usually have the condition

$$c_a = 0 \quad \text{for} \quad a = 1, 2, \cdots < \alpha$$

where α is the lower limit of the reproductive period.

The expressions in brackets (coefficients of successive powers of y) will be recognized as cumulations S_h of the values of the function c_a, summed backwards to the "diagonal" element c_h, where h is the exponent of y. In terms of moments m of the function c_a, equation (8) can be written

$$1 = m_0 + m_1 y + \frac{m_2 - m_1}{2!} y^2 + \frac{m_3 - 3m_2 + 2m_1}{3!} y^3 + \cdots + c_\omega y^\omega \qquad (9)$$

or, using the symbol $m_{[h]}$ to denote the hth *factorial* moment, equation (9) takes the simple form

$$1 = \sum_{h=0}^{h=\omega} \frac{m_{[h]}}{h!} y^h . \qquad (10)$$

In these expressions the moments m_h and $m_{[h]}$ are those taken about $a = 0$. Actually, the net reproduction rates are given for "semi-values" of a, that is, for values of a which are odd multiples of $\frac{1}{2}$ (using five years as the time unit). By cumulation of these given values moments m'_h and $m'_{[h]}$ about $a = -\frac{1}{2}$ are obtained.[4] From the latter the corresponding functions of the moments about $a = 0$ are obtained by the transformation formulae[5]

$$
\begin{aligned}
\frac{m_{[h]}}{h!} &= \sum_{k=0}^{k=h} \frac{(-\frac{1}{2})^{[k]}}{k!} \cdot \frac{m'_{[h-k]}}{(h-k)!}, \\
S_h &= \sum_{k=0}^{k=h} \frac{(-\frac{1}{2})^{[k]}}{k!} S'_{h-k} .
\end{aligned} \qquad (11)
$$

[4] In these cumulations zero values of c_a for $0 < a < \alpha$ must not be omitted.

[5] In accordance with a customary notation the symbol $(-\frac{1}{2})^{[k]}$ denotes the continued product $-\frac{1}{2}(-\frac{1}{2} - 1)(-\frac{1}{2} - 2) \ldots (-\frac{1}{2} - k + 1)$. In the computation of successive terms, in the sums in the right-hand member of (11), by appropriately laying out the work, advantage is taken of the fact that values of $(-\frac{1}{2})^{[k]}/k!$ for $k = 2, 3 \ldots$ are obtained each from the preceding by multiplying successively by $\frac{3}{4}, \frac{5}{6}, \frac{7}{8}$, etc., and taking care of the sign, so that fractions with complicated numerators and denominators are avoided.

It will be recalled that in the treatment of the problem of replacement by means of an integral equation,[6] a solution in the form

$$B(t) = \Sigma Q'_j x_j^{-t} = \Sigma Q'_j e^{r_j t}, \tag{12}$$

is obtained, in which the exponential coefficients r_j are the roots of the equation

$$1 = \int_\alpha^\omega e^{-ra} p(a)m(a)\, da = \int_\alpha^\omega x^a p(a)m(a)\, da, \tag{13}$$

i.e.

$$1 = m_0 - m_1 r + \frac{m_2}{2!} r^2 - \frac{m_3}{3!} r^3 + \cdots = \sum_{h=0}^{h=\infty} (-1)^h \frac{m_h}{h!} r^h, \tag{14}$$

in close analogy to equation (10) for y, with the distinction however that in (10) the *factorial* moments take the place of the ordinary moments of (14), and that the series in (10) is finite, terminating at the term in y^ω. There is also an important difference between the characteristic equation (13) and its analogue (5), namely that (5) may admit of negative roots for x, whereas (13) does not admit negative values for x.

2. The constants Q. These are determined by initial conditions, as follows. Equation (2) can be written

$$\begin{aligned} B(t) &= \sum_{a=t}^{a=\omega} c_a B(t-a) + \sum_{a=1}^{a=t-1} c_a B(t-a) \\ &= F(t) + \sum_{a=1}^{a=t-1} c_a B(t-a), \end{aligned} \tag{15}$$

with

$$\text{and} \quad \begin{aligned} F(t) &= \sum_{a=t}^{a=\omega} c_a B(t-a) \qquad & 0 < t \leq \omega \\ F(t) &= 0 & t > \omega \end{aligned} \tag{16}$$

The values of $B(t)$ being given for integral values of t, from $t = -(\omega - 1)$ to $t = 0$, it can be shown that[7]

$$Q_j = \frac{\sum_{t=1}^{t=\omega} F(t) x_j^t}{\sum_{a=1}^{a=\omega} a c_a x_j^a}. \tag{17}$$

[6] For a discussion of the limits of applicability of this method See [4].

[7] The reasoning is essentially the same as in the treatment of the problem by integral equations. See [5] and [2, p. 39 et seq.].

In the special case that we are tracing the progeny of an initial population all born at the same time, say $B(0)$ births occurring at $t = 0$, so that

$$B(-1) = B(-2) = \cdots = B(-[\omega - 1]) = 0 \tag{18}$$

the expression for Q_j, in view of (5), reduces to a particularly simple form. For if we write the summation in equation (16) in expanded form, we have

$$\begin{array}{ll} F(1) = & c_1B(0) + c_2B(-1) + c_3B(-2) + c_4B(-3) + \cdots + c_\omega B(-\overline{\omega - 1}), \\ F(2) = & \qquad c_2B(0) + c_3B(-1) + c_4B(-2) + \cdots + c_\omega B(-\overline{\omega - 2}), \\ F(3) = & \qquad\qquad c_3B(0) + c_4B(-1) + \cdots + c_\omega B(-\overline{\omega - 3}) \\ \vdots & \qquad\qquad\qquad \ddots \qquad \vdots \\ F(\omega) = & \qquad\qquad\qquad\qquad c_\omega B(0). \end{array} \tag{19}$$

If now $B(-1), \cdots, B(-\overline{\omega - 1})$, all vanish, then

$$\sum_1^\omega F(t)x^t = B(0)\{c_1x + c_2x^2 + c_3x^3 + \cdots + c_\omega x^\omega\} \tag{20}$$

$$= B(0) \text{ by (5).} \tag{21}$$

Hence,

$$Q_j = \frac{B(0)}{\sum_{a=1}^{a=\omega} ac_a x_j^a}. \tag{22}$$

In particular

$$B(0) = \sum_{j=1}^{j=\omega} Q_j = B(0) \sum_{j=1}^{j=\omega} \frac{1}{\sum ac_a x_j^a} \tag{23}$$

so that

$$\sum_{j=1}^{j=\omega} \frac{1}{\sum ac_a x_j^a} = 1. \tag{24}$$

The constant $B(0)$ here evidently functions essentially as an arbitrary unit of annual births, and may with this understanding simply be put $= 1$, thereby simplifying the notation. This has been done in what follows, where convenient, especially in the table of constants, Table 3 of the numerical illustration.

The denominator in (17) or (23) can be evaluated for any root x_j of (5) by direct summation if the coefficients c_a are given or have been computed (as indicated below) for integral values of a; or, in a manner similar to that employed in passing from equation (5) to (8), the denominator can be expressed in terms of the corresponding roots $y_j = x_j - 1$ of (8) or (10), the cumulations of c_a being replaced by cumulations of ac_a. With the denominator so expressed, the constants Q_j take the form, in obvious analogy to equation (9):

$$Q_j = \frac{\sum_{t=1}^{t=\omega} x_j^t F(t)}{m_1 + m_2 y_j + \frac{m_3 - m_2}{2!} y_j^2 + \frac{m_4 - 3m_3 + 2m_2}{3!} y_j^3 + \cdots} \tag{25}$$

The alternative procedure, to which reference was made in the preceding paragraph, is to operate upon the moments $m_{[h]}$ (taken about the origin O) by a process the inverse of cumulation—which we might term *decumulation*—and in this way to obtain from them the coefficients c_a. The polynomial $\Sigma ac_a x_j^a$ can then be evaluated directly.

The decumulation is readily carried out by an algorithm which suggests itself from the schedule of cumulation. Analytically the relation between the two processes is expressed by the reciprocal sets of transformation formulae:

Cumulation

$$\frac{m_{[h]}}{h!} = \sum_{k=0}^{k=\omega-h} \binom{h+k}{h} c_{h+k} = S_h. \tag{26}$$

Decumulation

$$c_h = \sum_{k=0}^{k=\omega-h} (-1)^k \binom{h+k}{h} \frac{m_{[h+k]}}{(h+k)!}. \tag{27}$$

3. Constants Q associated with complex roots $x = e^{-u+iv}$.

The complex roots x_j give rise to oscillatory terms which, in the special case of the progeny of a cohort of $B(0)$ births, take the form[8]

$$\frac{2B(0)e^{-ut}}{G^2 + H^2} \{G \cos vt - H \sin vt\}, \tag{28}$$

where

$$G = \sum_{a=1}^{a=\omega} ac_a e^{-ua} \cos va \tag{29}$$

and

$$H = \sum_{a=1}^{a=\omega} ac_a e^{-ua} \sin va. \tag{30}$$

These constants may be evaluated directly in this form, or, putting $y = \xi + i\eta$ in the denominator of equation (25), they can be expressed in terms of the roots y_j and the factorial moments obtained by cumulation of ac_a.[9]

---

[8] The development of these formulae is analogous to that followed in the treatment of the problem by integral equations. See [6]; for the more general case see also [7].

[9] The procedure in this case will be analogous to that followed in the development of equations (90) and (91) in [6].

(a) *Numerical Illustration.* For convenience and to furnish the opportunity for comparison, the same data (United States 1920) were here employed as in the writer's earlier publications in which the problem was treated by the application of an integral equation.

(b) *Cumulation for values of m_h.* The two operations, of (1) cumulating the values of c_a given for semi-values of a; and (2) allowing in the cumulated results

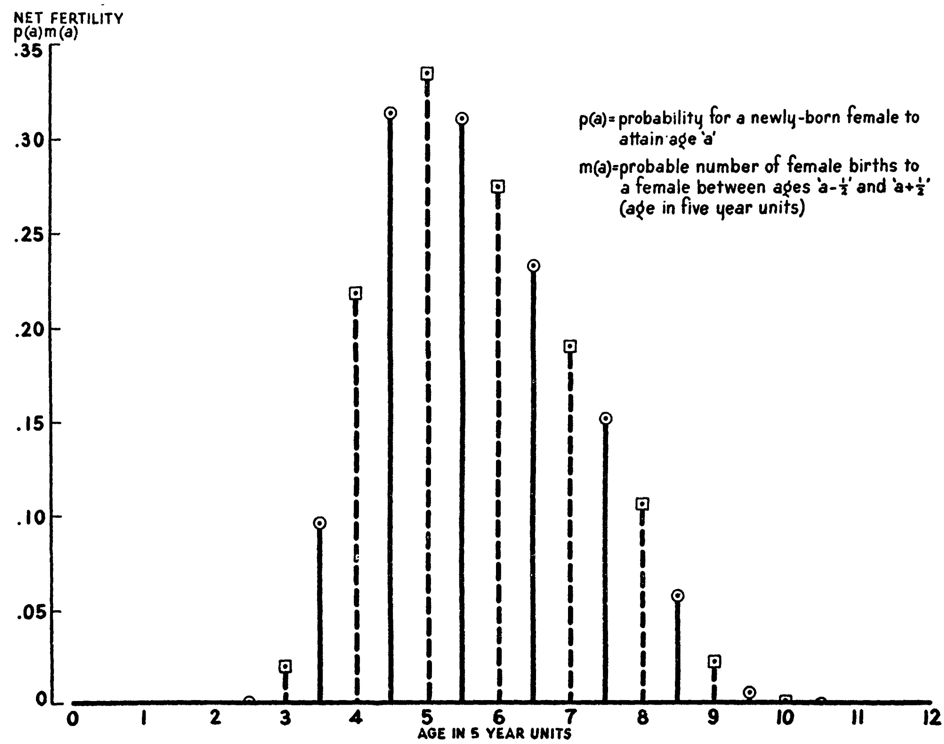

FIG. 1. Net Fertility $p(a)m(a)$ White females, United States, 1920

The verticals drawn in full and centered at mid-ages represent the original data; those drawn in dashed lines and centered at integral ages are interpolated.

for a shift of origin from $a = -\frac{1}{2}$ to $a = 0$, can be conducted in one schedule as in Table 1. Cumulation is first carried out in the usual manner from the bottom line to the diagonal, with the result appearing immediately below the diagonal. From here on the procedure is as in the following example: Starting at the lower right hand corner, we find

$$.00780 \times (-\tfrac{1}{2}) = -.00390$$
$$.12770 \times (-\tfrac{1}{2}) = -.06385 \qquad -.06385 \times (-\tfrac{3}{4}) = .04789$$
$$.97395 \times (-\tfrac{1}{2}) = -.48698 \qquad -.48698 \times (-\tfrac{3}{4}) = .36254$$
$$.36254 \times (-\tfrac{5}{6}) = -.30437.$$

TABLE 1

Computation schedule for values of $\frac{m_{[j]}}{j!} = S_j$ *of net productivity function* $p(a)m(a) = c_a$ *for integral values of age* a.*

a in 5-year units	c_a	$m_{[0]}$	$m_{[1]}$	$m_{[2]}/2!$	$m_{[3]}/3!$	$m_{[4]}/4!$	$m_{[5]}/5!$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		1.16635	6.64127	16.64550	24.34106	23.16864	15.05650
			−.58318	.43738	−.36448	.31892	−.28703
0–1	.00000	1.16635	7.22445	−3.61223	2.70917	−2.25764	1.97544
1–2	.00000	1.16635	6.05810	19.82035	−9.91018	7.43264	−6.19387
2–3	.00040	1.16635	4.89175	13.76225	31.90655	−15.95328	11.96496
3–4	.09630	1.16595	3.72540	8.87050	18.14430	33.62800	−16.81400
4–5	.31255	1.06965	2.55945	5.14510	9.27380	15.48370	24.41095
5–6	.31025	.75710	1.48980	2.58565	4.12870	6.20990	8.92725
6–7	.23170	.44685	.73270	1.09585	1.54305	2.08120	2.71735
7–8	.15090	.21515	.28585	.36315	.44720	.53815	.63615
8–9	.05795	.06425	.07070	.07730	.08405	.09095	.09800
9–10	.00615	.00630	.00645	.00660	.00675	.00690	.00705
10–11	.00015	.00015	.00015	.00015	.00015	.00015	.00015

a in 5-year units	$m_{[6]}/6!$	$m_{[7]}/7!$	$m_{[8]}/8!$	$m_{[9]}/9!$	$m_{[10]}/10!$	$m_{[11]}/11!$	Factor
(1)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	6.72500	1.99717	.36404	.03483	.00127	.00001	
	.26311	−.24432	.22905	−.21633	.20551	−.19617	−21/22
0–1	−1.77790	1.62974	−1.51333	1.41875	−1.33993	1.27293	−19/20
1–2	5.41964	−4.87768	4.47121	−4.15184	3.89235	−3.67611	−17/18
2–3	−9.97080	8.72445	−7.85201	7.19768	−6.68356	6.26584	−15/16
3–4	12.61050	−10.50875	9.19516	−8.27564	7.58600	−7.04414	−13/14
4–5	−12.20548	9.15411	−7.62843	6.67488	−6.00739	5.50677	−11/12
5–6	12.38595	−6.19298	4.64474	−3.87062	3.38679	−3.04811	−9/10
6–7	3.45870	4.31260	−2.15630	1.61723	−1.34769	1.17923	−7/8
7–8	.74135	.85390	.97395	−.48698	.36524	−.30437	−5/6
8–9	.10520	.11255	.12005	.12770	−.06385	.04789	−3/4
9–10	.00720	.00735	.00750	.00765	.00780	−.00390	−1/2
10–11	.00015	.00015	.00015	.00015	.00015	.00015	

* Figures immediately below the diagonal, obtained by cumulation from the bottom upward of the data in Column 2, are factorial moments about $a = -\frac{1}{2}$. Figures in the top line are factorial moments about $a = 0$. For use of factors in the last column see text.

The several columns are thus completed, and by addition, in each column, of the item immediately below the diagonal, and of all the items above the diagonal, the figures in the top line are obtained. These are the coefficients of equation (10) for y.

(c) *Decumulation.* While it is not necessary to carry out the decumulation, since the entire computation can, if desired, be carried out in terms of y's and m's, there is a considerable interest in noting the values c_a for integral values of a which result from the decumulations of the m's. These, together with the original values for semi-values of a, are shown in Table 2 and Fig. 1.

TABLE 2

Values of $c_a = p(a)m(a)$

(1) *for semi-values of a; original data.*

(2) *for integral values of a; computed by cumulation of original data, shift of origin, and decumulation.*

a 5-year units	c_a	a 5-year units	c_a	a 5-year units	c_a
0.0	0	4.0	.21781	8.0	.10607
0.5	0	4.5	.31255	8.5	.05795
1.0	0*	5.0	.33400	9.0	.02268
1.5	0	5.5	.31025	9.5	.00615
2.0	0*	6.0	.27427	10.0	.00116
2.5	.00040	6.5	.23170	10.5	.00015
3.0	.02073*	7.0	.18963	11.0	.00001
3.5	.09630	7.5	.15090		

* The value of c_2 came out negative, namely −.00570, and the value of c_1 came out +.00014. In the computation of $\Sigma a c_a x^t$ these two values were arbitrarily adjusted to zero, and c_3 was diminished from .02118 to .02073 to make the total $\sum_{i=1}^{11} c_a = 1.16635$, summing only for integral values of a.

4. The roots of equations (5) and (8).

From the prior study already cited, the real positive and three pairs of complex roots for r of the characteristic equation

$$\int_\alpha^\omega x^a\, p(a)m(a)\, da = \int_\alpha^\omega e^{-ra} p(a)m(a)\, da = 1 \tag{31}$$

were known. These were used to indicate the approximate location of the roots of (5) or (8), and more exact values were then obtained by Newton's method of successive approximation. Table 3 shows the values of u, v, etc., corresponding to the new roots

$$\begin{aligned} y &= x - 1 \\ &= e^{-u+iv} - 1 \end{aligned}$$

obtained through equations (8) or (10); and, for comparison the corresponding values obtained in the previous publication from equation (13).[10] The same table also exhibits the remaining roots and values of the constants Q, G, H.

TABLE 3

Constants of the series solution (6) of equation (3), corresponding to the five real and three pairs of complex roots of the characteristic equation (5)

(United States, white females, 1920)

Constants[(1)]	Five Real Roots					Three Pairs of Complex Roots		
	A. Computed on basis of recurrent series							
u	.02714*	−1.764†	−3.812†	−17.1†	−94.3†	−.19800	−.44720	−.47587
v	0.	0.	0.	0.	0.	1.06498	1.57000	2.40490
G	5.64467	7.73354	−1255.04	(2)	(2)	5.28093	10.45809	7.73103
H	0.	0.	0.	0	0	3.03239	−3.66726	2.00874
$G/(G^2+H^2)$	.17716	.12931	−.00080	(2)	(2)	.14241	.08515	.12117
$H/(G^2+H^2)$	0.	0.	0.	0	0	.08177	−.02986	.03148
	B. Computed on basis of integral equation‡							
u	.02714					−.1930	−.43655	−.4902
v	0.					1.0724	1.5771	2.44245
G	5.64514					5.15351	10.22495	7.40154
H	0.					2.98757	−3.72741	3.45312
$G/(G^2+H^2)$	.17715					.14525	.08620	.11095
$H/(G^2+H^2)$	0.					.08420	−.03135	.05175

(1) t in five year units

(2) Not computed

* $u_o = \log_e x_o = -\log_e .97322 = .02714$

† Values of x

‡ See [6, p. 899]

To determine the remaining four roots, the product of the factors $(y - y_1)(y - y_2) \cdots (y - y_7)$ was divided out of the polynomial of equation (10), rejecting the remainder and leaving a fourth degree equation

$$y^4 + 120\,y^3 + 2590\,y^2 + 14617\,y + 23118 = 0$$

In the subsequent work it turned out that the roots of this were all real, and they were computed by obvious methods. Their values are also shown in Table 3. For the two numerically largest roots great accuracy was not attempted. They introduce terms with very rapid damping and presumably very small values of Q.[11]

[10] The divergence is due in part to details of computation. In the earlier publication the curve of fertility $m(a)$ was smoothed by the method of translation, with a Gaussian distribution as basis. In the method here presented the raw data were used without smoothing, except such as is inherent in the process of the calculation described.

[11] At any rate, $Q_{10} + Q_{11}$ must be small, since $Q_1 + \ldots + Q_9 = 1.00313$, and according to (24), with the convention that $B(O) = 1$, the sum of all the Q_j must be equal to unity.

As a check, in order to be assured that no serious error was introduced in neglecting the remainder after dividing out the factors $(y - y_i)$ up to $(y - y_7)$, the product $\prod_{i=1}^{11} (y - y_i)$ was computed and, after multiplying by a factor to make the absolute terms agree (.16635), was compared with the polynomial of (10). As a further indication, the coefficients of the product Π were "decumulated" to obtain values of coefficients of the corresponding polynomial in x, to

TABLE 4

Coefficients of Powers of y in Equation (10) and in the Product $\prod_1^{11}(y - y_i)$; *Also Coefficients of Powers of x in Equation (5)*

a	Coefficients of y^a		Coefficients of x^a in Equation (5) Found by Decumulation	
	In Equation (10)	In $\prod_1^{11}(y - y_i)$	Of Column (2)	Of Column (3)
(1)	(2)	(3)	(4)	(5)
0	.16635	.16635	−1.00000	−.99915
1	6.64127	6.64072	+.00014*	+.00065
2	16.64550	16.64782	−.00057*	−.00432
3	24.34106	24.24197	.02118*	.02398
4	23.16840	23.18070	.21781	.21774
5	15.05650	15.07338	.33400	.33354
6	6.7250	6.73812	.27427	.27474
7	1.99717	2.00316	.18963	.18882
8	.36404	.36555	.10607	.10641
9	.03483	.03501	.02268	.02276
10	.00127	.00128	.00116	.00117
11	.00001	.00001	.00001	.00001

* In computing the denominator of Q according to (22) the values of the coefficients c_1 and c_2 were arbitrarily made zero and the value of c_3 (age 15) was adjusted to .02073 to retain the total $\sum_1^{11} c_j = 1.16635$.

compare with values of c_a. The results are shown in Table 4. In view of the fact that the (numerically) highest roots were determined only in first approximation, the agreement is satisfactory.

It is to be noted that instead of applying the solution (6) to compute values of $B(t)$, these latter can, of course, also be obtained directly, by carrying forward step by step the original recurrent series; or, alternatively, the births in successive generations can be computed step by step and the total births obtained by addition. The advantage of the solution (6) is that it enables one, if desired, to obtain $B(t)$ for any value of t without having to compute $B(t)$ for all intervening values of t; also, the solution in an exponential series gives a better idea of the general nature of the process, as well as a direct indication of its asymptotic course for large values of t, when the first term $Q_0 x_0^{-t}$ with the positive real root x_0 dominates all others. However this may be, it is interesting to compare

TABLE 5

Synopsis of Results of Computation of $B(t)$ as $\Sigma\ Qx^{-t}$, Column (8), and as $\Sigma B_n(t)$, Column (9), where $B_n(t)$ = Births per Unit of Time in nth Generation at Time t. (Time Unit = 5 years)

	$A = Qx^{-t} = x^{-t}\frac{1}{G}$			$A = \frac{2e^{ut}}{G^2+H^2}\left\{G \cos vt - H \sin vt\right\}$			
x* or u#	.97322*	−1.764*	−3.81208*	−.19800#	−.44720#	−.47587#	ΣA
v / t	0	0	0	1.06498	1.57000	2.40490	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0	17,716	12,931	−80	28,482	17,030	24,234	100,313
1	18,204	−7,330	21	−415	3,828	−13,781	527
2	18,704	4,156	−6	−19,498	−6,959	3,329	−274
3	19,219	−2,356	1	−15,222	−1,572	2,256	2,326
4	19,748	1,336		1,022	2,844	−3,362	21,588
5	20,291	−757		11,057	646	2,223	33,460
6	20,850	429		8,102	−1,162	−749	27,470
7	21,423	−243		−1,001	−265	−169	19,745
8	22,013	138		−6,248	475	445	16,823
9	22,619	−78		−4,294	109	−344	18,012
10	23,241	44		792	−194	145	24,028
11	23,880	−25		3,519	−45	−1	27,328
12	24,538	14		2,265	79	−55	26,841
13	25,213	−8		−568	18	51	24,706
14	25,907	5		−1,976	−32	−26	23,878
15	26,619	−3		−1,188	−8	4	25,424
16	27,352	1		385	13	6	27,757
17	28,105	−1		1,106	3	−7	29,206
18	28,878	1		620	−5	4	29,498
19	29,673			−251	−1	−1	29,420
20	30,489			−617	2	−1	29,874
21	31,328			−321	1	1	31,008
22	32,191			160	−1	−1	32,349
23	33,076			343			33,419

TABLE 5—*Continued*

t	$B_n(t)$						
	$\Sigma B_n(t)$	*Generations, n*					
		(1)	(2)	(3)	(4)	(5)	(6)
	(9)	(10)	(11)	(12)	(13)	(14)	(15)
0	100,000						
1	0						
2	0						
3	2,072	2,072					
4	21,781	21,781					
5	33,398	33,398					
6	27,472	27,429	43				
7	19,866	18,963	903				
8	16,735	10,607	6,128				
9	17,954	2,268	15,685	1			
10	24,033	116	23,889	28			
11	27,361	1	27,022	338			
12	26,878		24,905	1,973			
13	24,696		18,481	6,214	1		
14	23,851		10,980	12,858	13		
15	25,410		5,345	19,941	124		
16	27,759		2,050	25,030	679		
17	29,219		526	26,316	2,377		
18	29,506		76	23,527	5,897	6	
19	29,414		5	18,092	11,271	46	
20	29,862			12,041	17,579	242	
21	31,000			6,906	23,191	903	
22	32,348			3,381	26,442	2,523	2
23	33,423			1,397	26,426	5,583	17

the result of the computation by means of the exponential series, carried out as set forth above, with the corresponding results of the computation of births in successive generations. This comparison is exhibited in Table 5.

It will be seen that the agreement is good except for the second to fourth items, where perhaps the omission of the terms contributed by the numerically highest roots makes itself felt.

5. Discussion.

(a) *The real roots of the characteristic equation* (5). It can be shown [8] that only one of the real roots for x can be positive, and that the absolute value of any other root must be greater than the positive real root.

The negative real roots which make their appearance in the numerical example call for special comment. Practically, the "higher" negative roots are of little importance, at any rate in this example—first because the constants Q with which they are associated are relatively small; second because large absolute values of negative roots imply rapid damping, so that corresponding terms Qx^{-t} very soon become negligible as t increases. Thirdly, the determination of these roots would be subject to a wide range of uncertainty, corresponding to the large percentage fluctuations or errors of determination of the values of the functions $p(a)m(a) = c_a$ at the upper end of the reproductive period.

But in theory these negative real roots suggest some pertinent questions. One wonders what would happen to them if the data were given, say, for single years of age, instead of 5-year groups. Instead of an equation of eleventh degree we would then have one of 55th degree. Furthermore, in those cases in which it may be permissible to pass to the limit, so that an integral equation takes the place of (2), negative roots for x would seem to be excluded as they would make the integral in (13) meaningless.

A problem of perhaps little practical importance but of some theoretical interest may arise here, to which reference has also been made by P.H. Leslie in a recent article in *Biometrika*,[12] in connection with a different procedure.

(b) *Effect of finer subdivisions of histogram of* $p(a)m(a)$. The effect of this on equation (5) for x is not obvious at sight, since new coefficients would be inserted *between* previous terms. The effect is more easily understood from a consideration of equation (8) for y. Here finer subdivisions would introduce new terms only beyond the last term originally present. The original terms would not be changed at all *in form*, and those involving only lower moments would be changed but little in *numerical value*, provided that the original histogram were not so coarse as to give inappropriate values even for these lower moments.

The result, then, of finer subdivision of the histogram, would be to change the computed values of the lower roots only in minor degree. But the four negative real roots, depending in considerable measure on the higher terms of (5) or (8), would presumably be materially altered, and might perhaps give place to further complex roots. In any case they would be followed by new roots even more remote from practical significance than the original eleven.

(c) *The result as an interpolation formula*. Strictly speaking, the solution (6) of (2) is applicable only for integral values of t. In particular, terms arising out of the negative real roots of (5) for x are obviously not adapted to furnish interpolated values of $B(t)$ for fractional values of t, since fractional powers of

[12] See [9] and [10]. For a brief summary and analysis of Leslie's paper [9] see a review signed with the initials WGB in the *Jourl Inst. of Actuaries Student's Soc.*, Vol. 4 (1946), Part II. The first application of the matrix method to these problems seems to be due to H. Bernardelli, "Population Waves," *Jour. of Burma Research Soc.*, Vol. 31 (1941), Part I, p. 1.

negative quantities in general are complex. Over the range of t where the first real root together with the three parts of complex roots adequately describe the process under discussion, these terms alone are, in this sense and to this extent, suitable for interpolation, disregarding the terms corresponding to the other negative roots.

Even less suitable for interpolation purposes, it would seem, would be terms arising from further negative roots that might be introduced by a finer subdivision of the histogram of original data. If we suppose this subdivision carried to great lengths, and if negative roots still appeared under these circumstances, they would give rise to rapidly oscillating positive and negative terms for even and odd integral values of t respectively (the time unit now being a subdivision of the original time unit) with no appropriate interpolation between these integral values.

One further point calls for comment. In the process of idealization of the problem discussed, it has been assumed that $p(a)$ and $m(a)$ are independent of time, and the conclusions reached must be construed in the light of this assumption. In itself this would hardly call for comment, as it is a matter of common understanding. But the question does arise whether the assumption itself is free from implied internal contradictions.

In a recent publication, P. K. Whelpton[13] has drawn attention to the fact that in times of rapid changes in the birth rate, the assumption of age specific fertilities being held constant at the values observed in a given calendar year may imply that some of the women had more than one *first* child, a logical impossibility.

The data used in the present numerical example are derived from a period of relatively undisturbed birth rate (1920), and do not involve any such conflict. But, in the light of Whelpton's contribution one may ask the broader question whether the computation of an intrinsic rate of natural increase and related parameters based on age specific fertility as observed in one calendar year retain any practical value at all.

In answering this question, two considerations will be weighed. First, that *ordinarily* the rates computed in the usual way differ but little from those obtained by taking into account order of birth as in Whelpton's procedure. Secondly, that the computation using over-all values of $m(a)$ for all orders of birth combined is a relatively simple matter based on data commonly available; whereas the more complete treatment of the problem taking into account order of birth is considerably more complicated and often not possible at all for lack of detailed data.

REFERENCES

[1] E. C. Rhodes, "Population mathematics," *Roy. Stat. Soc. Jour.*, (1940), pp. 229, 232.

[2] S. D. Wicksell, "Bidrag till den formella befolkningsteorien," *Statsokonomisk Tidsskrift*, Heft 1–2 (1934), p. 1.

[13] See [11]. Another refinement recently introduced into the measurement of reproductivity is to take into account duration of marriage. See Colin Clark and R. E. Dyne, *Economic Record* (Australia), June 1946, p. 23.

[3] E. J. GUMBEL, "Eine Darstellung statistischer Reihen durch Euler," *Jahresber. der Deutschen Mathematiker Vereinigung*, Vol. 25, p. 257.

[4] W. FELLER, "On the integral equation of renewal theory," *Annals of Math. Stat.*, Vol. 12 (1941), p. 243.

[5] A. J. LOTKA, "Contribution to the theory of self-renewing aggregates with special reference to industrial replacement," *Annals of Math. Stat.*, Vol. 10 (1939), p. 11.

[6] A. J. LOTKA, "The progeny of a population element," *Am. Jour. of Hygiene*, Vol. 8 (1928), p. 892.

[7] A. J. LOTKA, "The progeny of an entire population," *Annals of Math. Stat.*, Vol. 13 (1942), pp. 117–118.

[8] W. DOBBERNACK AND G. TIETZ, *Twelfth International Congress of Actuaries*, Vol. 4 (1940), p. 233.

[9] P. H. LESLIE, "On the use of matrices in certain population mathematics," *Biometrika*, Vol. 33 (1945), pp. 202 and 206.

[10] E. G. LEWIS, "On the generation and growth of a population," *Sankhya*, Vol. 6 (1942), Part I, p. 93 et seq.

[11] P. K. WHELPTON, "Reproduction rates adjusted for age, parity, fecundity, and marriage, *Am. Stat. Assn. Jour.*, Vol. 41 (1946), p. 501.