THE SOLUTION OF SOME EQUATIONS OCCURRING IN POPULATION GENETICS

BY J. B. S. HALDANE AND S. D. JAYAKAR
Genetics and Biometry Laboratory, Government of Orissa, Bhubaneswar-3, Orissa, India

INTRODUCTION

If a population has sharply divided generations and a fixed breeding system, is subject to selection of constant intensity, and is so large that we can use deterministic hypotheses, a mathematical treatment of selection requires the solution of one or more recurrence, or finite difference, equations. In the case of inbreeding without selection, these equations are linear. If selection occurs, they are nonlinear.

These equations may be of the second or higher order. For example, if autosomal recessives are eliminated in the male but not the female sex, and mating is at random, while the frequencies of a recessive gene are x_n and y_n in the female and male sexes respectively of generation n,

$$\left.\begin{aligned} x_{n+1} &= \tfrac{1}{2}(x_n + y_n) \\ y_{n+1} &= \frac{x_n + y_n - 2x_n y_n}{2(1 - x_n y_n)}\,. \end{aligned}\right\} \tag{1.1}$$

These simultaneous equations are each of the first order, and on eliminating y_n and y_{n+1}, yield the equation

$$\left.\begin{aligned} x_{n+2} &= \frac{2x_{n+1} - x_n(1 + x_{n+1})(2x_{n+1} - x_n)}{2[1 - x_n(2x_{n+1} - x_n)]} \\ \text{or } x_{n+2} - x_{n+1} &= \frac{(x_{n+1} - 1)(2x_{n+1} - x_n)x_n}{2[1 - x_n(2x_{n+1} - x_n)]} \\ \text{or } \Delta x_{n+1} &= \frac{(x_n + \Delta x_n - 1)(x_n + 2\Delta x_n)x_n}{2(1 - x_n^2 - 2x_n \Delta x_n)}. \end{aligned}\right\} \tag{1.2}$$

These three equations, all equivalent, are a recurrence equation and two difference equations of the second order. For they involve frequencies in three successive generations. We shall not deal with them in this article. On the other hand if a fraction k of recessives is eliminated in both sexes in a random mating population, and q_n is the frequency of recessive gametes in generation n, then

$$\left.\begin{aligned} q_{n+1} &= \frac{q_n - kq_n^2}{1 - kq_n^2} \\ \text{or } q_{n+1} - q_n &= \frac{-kq_n^2(1 - q_n)}{1 - kq_n^2}. \end{aligned}\right\} \tag{1.3}$$

These are an equivalent recurrence and difference equation of the first order, the former of the second degree. That is to say q_{n+1} is a rational function of q_n containing no terms higher than its square. We desire to get an expression which will enable us to calculate q_n given q_0 and n, or n given q_0 and q_n with speed and accuracy even when n is large. Since n must be a positive integer, it will matter very little if the error in the calculated value of n is as large as 0·1. In the particular case (1.3) one can express n as a series in ascending powers of k which is pretty accurate over the whole range of q_n from 1 to zero. This cannot be done where several constants are involved. But when q_n is near zero we can get an expression for n in ascending powers of q_n, and when q_n is near unity we can get a similar expression in powers of $(1-q_n)$. The methods of doing so are quite general, and will be described.

Consider the general equation

$$z_{n+1}=f(z_n) \tag{1.4}$$

where f is a one-valued analytic function. A mathematical treatment, with rigorous proofs of certain theorems, and references, are given by Picard (1928) and Valiron (1954). If

$$f(Z)-Z=0 \tag{1.5}$$

then if $z_n=Z$, $z_{n+1}=Z$, so Z is said to be a *fixed point* of the iteration (1.4). We shall assume that $f(z)$ is a real function, but even so some or all the roots of (1.5) may be complex, and (1.4) can only be fully discussed if z_n is a complex variable.

If $f(z)$ is a rational function whose numerator is of order n_1 and denominator of order n_2, the number of roots of (1.5) is n_1 or n_2+1, whichever is larger. But two or more roots of (1.5) may be equal. Thus the fixed points of (1.3) are 0, 0, and 1.

If Z is any finite root of (1.5), and $z_n=Z+x_n$, then $x_{n+1}=f(Z+x_n)-f(Z)$

$$=x_n f'(Z)+\frac{1}{2!}x_n^2 f''(Z)+\frac{1}{3!}x_n^3 f'''(Z)+\cdots. \tag{1.6}$$

If a value of Z is infinite, as when $z_{n+1}=az_n^2+bz_n$, we put $x_n=\dfrac{1}{z_n+c}$. With the French authors, we shall denote $f'(Z)$ by s. The behaviour of x_n, and therefore z_n, in the neighbourhood of a fixed point, depends on the value of s. It is easily seen that if $|s|<1$, and $|x_n|$ is sufficiently small, $|x_{n+1}|$ is still smaller, and x_n approaches zero as n increases. Such a point is called an *attractive point*. If $s=0$, the approach is very rapid, and we shall call it a *highly attractive point*. If s is positive, x_n does not change its sign after a certain value of n, if s is negative the signs alternate. If $|s|>1$ successive values of $|x_n|$ increase at least below some value of n, and Z is said to be a *repulsive point*. If $|s|=1$, Valiron describes Z as an *indifferent point*. We think this is misleading. In fact such a point is attractive in certain directions in the complex plane and repulsive in others. For example in the case of (1.3), if k is positive, zero is attractive along the real positive axis. That is to say if q_n is positive and less than unity, q_{n+1} is smaller. But it is repulsive along the real negative axis. For if q_n is negative, q_{n+1} is also negative with a larger absolute value. Since negative values of q_n have no biological meaning, it is, for practical purposes, an attractive point. A point where $s=1$ corresponds to two or more equal roots of (1·5), and we shall call it a *confluent point.* If $s=-1$, we have only to put $y_n=(-1)^n x_n$, and y_n has a confluent point. Similarly if $s=\omega$, where ω is a complex root of unity with $\omega^k=1$, we can put $x_{kn}=y_n$. The case when $|s|=1$, but is not a complex root of unity, e.g. $s=\frac{1}{5}(4+3\sqrt{-1})$, is intractable, but has no relevance to genetics. It is easy to show that s is invariant under any homographic or antihomographic transformation of z_n.

If most of the cases of (1.4) which have yet occurred in genetics, $f(z)$ is a rational function of degree 1 or 2. We shall consider a few cases when $f(z)$ is a function of a more complicated kind. The solution when it is of the first degree is trivial. The most general expression is

$$x_{n+1}=\frac{s\,x_n}{1+ax_n}. \tag{1.7}$$

Hence $x_{n+1}^{-1}=s^{-1}(x_n^{-1}+a)$, or

$$x_{n+1}^{-1}+\frac{a}{1-s}=s^{-1}\left(x_n^{-1}+\frac{a}{1-s}\right).$$

Hence $$x_n^{-1}+\frac{a}{1-s}=s^{-n}\left(x_0^{-1}+\frac{a}{1-s}\right),$$

$$\left.\begin{aligned} x_n&=\frac{s^n(1-s)\,x_0}{1-s+a(1-s^n)x_0} \\ \text{or } n\log s&=\log\left(x_0^{-1}+\frac{a}{1-s}\right)-\log\left(x_n^{-1}+\frac{a}{1-s}\right). \end{aligned}\right\} \tag{1.8}$$

These are the appropriate forms when $|s|<1$, i.e. if zero is an attractive point. It will be seen that we have found a simple function of x_n whose values form a geometric series. We cannot do this exactly for recurrences of higher degree, but we can approximate to it. (1.8) has a second fixed point, $X=a^{-1}(s-1)$. If however $s=1$ the two fixed points are confluent, and $x_{n+1}^{-1}=x_n^{-1}+a$, so that the values of x_n form a harmonic series, and

$$\left.\begin{aligned} x_n&=\frac{x_0}{1+nax_0} \\ n&=a^{-1}(x_n^{-1}-x_0^{-1}). \end{aligned}\right\} \tag{1.9}$$

If a is positive, zero is attractive when x_n is positive, repulsive when it is negative.

When $f(z)$ in (1.4) is a rational function, n is an automorphic function of z_n or x_n of a type first described by Poincaré. Such functions have not been tabulated, and so far as we know have had no applications in physics. They can be represented by infinite series in certain regions of the z plane. Before showing how to deal with them, we shall say a little about the general theory of second order iterations. The most general difference equation with finite fixed points may be written

$$\Delta z_n = \frac{-(z_n - a_1)\,(z_n - a_2)\,(z_n - a_3)}{(z_n - b_1)\,(z_n - b_2)}, \tag{1.10}$$

whence $$z_{n+1} = \frac{(a_1 + a_2 + a_3 - b_1 - b_2)z_n^2 - (a_2a_3 + a_3a_1 + a_1a_2 - b_1b_2)z_n + a_1a_2a_3}{(z_n - b_1)\,(z_n - b_2)}, \tag{1.11}$$

which is the most general rational quadratic recurrence equation. For given b_1 and b_2 we can choose a_1, a_2, and a_3, so that the coefficients of the numerator assume any assigned values. a_1, a_2, and a_3 are the fixed points. If one or more is infinite, one or both of b_1 and b_2 must also be infinite. If $x_n = z_n - a_1$ then

$$\Delta x_n = \frac{-x_n\,(a_2 - a_1 - x_n)\,(a_3 - a_1 - x_n)}{(b_1 - a_1 - x_n)\,(b_2 - a_1 - x_n)}$$

$$x_{n+1} = \frac{x_n\,[b_1b_2 - a_2a_3 + a_1(a_2 + a_3 - b_1 - b_2) + (b_1 + b_2 - a_2 - a_3)x_n]}{(a_1 - b_1)(a_1 - b_2) + (b_1 + b_2 - 2a_1)x_n + x_n^2}.$$

Thus $s_1 = 1 - \dfrac{(a_1 - a_2)\,(a_1 - a_3)}{(a_1 - b_1)\,(a_1 - b_2)}$

which becomes unity if $a_1 = a_2$ or a_3 or both, that is to say if two or three fixed points coincide. Similarly one can show that the condition for a_1 to be a highly attractive point is that $(a_1 - b_1)\,(a_1 - b_2) = (a_1 - a_2)\,(a_1 - a_3)$. The condition that a_1 should be attractive is that $\dfrac{(a_1 - a_2)\,(a_1 - a_3)}{(a_1 - b_1)\,(a_1 - b_2)}$ should be positive but less than 2.

Standard Forms

The general recurrence equation of the second degree contains five arbitrary constants. This number is reduced to four on transferring the origin to a fixed point. Even so the solution is cumbrous. But by a further transformation we can reduce the number to two, one, or even zero. Further, in some cases, we can greatly simplify the final solution.

The most important group of cases has an ordinary attractive or repulsive point with s between 0 and 1 for an attractive point (stable equilibrium), exceeding unity for a repulsive point (unstable equilibrium). The exact value of unity is only reached if some condition is exactly fulfilled, for example if the loss or gain of fitness of a homozygote is completely recessive. This can very rarely be the case, even when dominance is so complete that we do not know whether the heterozygote is or is not fitter than the homozygous dominant. However s must sometimes be so near to unity that it is best taken to be exactly unity. We begin therefore with the general case.

Let $$x_{n+1} = \frac{sx_n(1 + Ax_n)}{1 + Bx_n + Cx_n^2}. \tag{2.1}$$

Here $|s|$ is not unity or zero, and A^2-AB+C is not zero, or $(1+Ax_n)$ would be a factor of the denominator. It is convenient to proceed to the standard form in two steps.

$$x_{n+1}^{-1}=\frac{x_n^{-2}+Bx_n^{-1}+C}{s(x_n^{-1}+A)}$$

$$=s^{-1}\left(x_n^{-1}+B-A+\frac{A^2-AB+C}{x_n^{-1}+A}\right).$$

If $\theta=\dfrac{A-B}{1-s}$, and $y_n^{-1}=x_n^{-1}-\theta$, $y_n=\dfrac{x_n}{1-\theta x_n}$, $x_n=\dfrac{y_n}{1+\theta y_n}$,
then

$$y_{n+1}^{-1}=s^{-1}\left(y_n^{-1}+\frac{A^2-AB+C}{y_n^{-1}+A+\theta}\right)$$

or

$$y_{n+1}=s\,y_n\left[1+\frac{(A^2-AB+C)y_n^2}{1+(A+\theta)y_n}\right]^{-1}.$$

If A^2-AB+C is negative, let $u_n=(-A^2+AB-C)^{\frac{1}{2}}y_n$, $a=-(A+\theta)(-A^2+AB-C)^{-\frac{1}{2}}$.

$$\text{Then } u_{n+1}=s\left(1-\frac{u_n^2}{1-au_n}\right)^{-1}. \tag{2.2}$$

This is the principal standard form. However if A^2-AB+C is positive we put $u_n=(A^2-AB+C)^{\frac{1}{2}}y_n$, and $a=-(A+\theta)(A^2-AB+C)^{-\frac{1}{2}}$.

$$\text{Then } u_{n+1}=su_n\left(1+\frac{u_n^2}{1-au_n}\right)^{-1}. \tag{2.3}$$

If $B=(2-s)A$, then $A+\theta=0$, and

$$u_{n+1}=su_n\ (1\pm u_n^2)^{-1} \tag{2.4}$$

the sign being that of A^2-AB+C. We can use the same method for any other function referred to one of its zeros, and obtain a standard form of the type $u_{n+1}=su_n\ (1\pm u_n^2+au_n^3+bu_n^4+\ -\ -\ -)^{-1}$.

When $s=1$, with two confluent roots, we have

$$x_{n+1}=\frac{x_n(1+Ax_n)}{1+Bx_n+Cx_n^2} \tag{2.5}$$

or $x_{n+1}^{-1}=x_n^{-1}+B-A+\dfrac{A^2-AB+C}{x_n^{-1}+A}$.

Zero is attractive for positive x_n if $B>A$. If so let $u_n=\dfrac{(B-A)x_n}{1+Ax_n}$, or

$$x_n^{-1}=(B-A)\ u_n^{-1}-A. \quad \text{Then } u_{n+1}=u_n(1+u_n-au_n^2)^{-1} \tag{2.6}$$

where $a=(A^2-AB+C)\ (A-B)^{-2}$.

If $A > B$, zero is repulsive for positive x_n, and the equilibrium is unstable. In this case we put $u_n = \frac{(A-B)x_n}{1+Ax_n}$, whence

$$u_{n+1} = u_n(1-u_n+au_n^2)^{-1}. \tag{2.7}$$

We may treat any function giving $s=1$ in the same way, and obtain the form $u_{n+1}=u_n(1 \pm u_n - au_n^2 - bu_n^3 - cu_n^4 - - -)^{-1}$. Another standard form is often more useful than (2.7) for a second degree equation with confluent roots. In equation (2.5) let us transfer our origin to the root $(A-B)\ C^{-1}$.

Let $y_n = A-B-C\,x_n$.

Then $y_{n+1}^{-1} = \dfrac{(A^2-AB+C)\ y_n^{-2}+(B-2A)\ y_n^{-1}+1}{(AB-B^2+C)\ y_n^{-1}-B}$.

Let $z_n = C^{-1}(A-B)\ (AB-B^2+C)\ y_n^{-1} - BC^{-1}(A-B)$.

Then $z_{n+1} = z_n + a(z_n-1)^2 z_n^{-1}$ (2.8)

where $a = \dfrac{B-A}{AB-B^2+C}$.

This transformation is clearly inapplicable if $AB-B^2+C=0$ and is useless when a is numerically large. The fixed points of (2.8) are 1, 1 and ∞ and it can be seen to be quite general. For Δz_n must have a square term since there are two confluent roots, and must be of order z_n, since ∞ is an ordinary attractive point.

If $B=A$, all three points are confluent at zero. Here

$$x_{n+1} = \frac{x_n(1+Ax_n)}{1+Ax_n+Cx_n^2}, \tag{2.9}$$

$$x_{n+1}^{-1} = x_n^{-1} + \frac{C}{x_n^{-1}+A}.$$

Zero is attractive for both positive and negative values of x_n if C is positive. If so let $x_n^{-1} = C^{\frac{1}{2}}u_n^{-1} - A$, or $x_n = \dfrac{u_n}{C^{\frac{1}{2}}-Au_n}$ or $u_n = \dfrac{C^{\frac{1}{2}}x_n}{1+Ax_n}$.

Then $u_{n+1} = u_n(1+u_n^2)^{-1}$. (2.10)

If C is negative the sign of u_n^2 must be reversed.

At a highly attractive point we have

$$x_n = \frac{x_n^2}{A+Bx_n+Cx_n^2} \tag{2.11}$$

or $x_{n+1}^{-1} = Ax_n^{-2} + Bx_n^{-1} + C$.

Let $Ax_n^{-1}=u_n^{-1}-\frac{1}{2}B$, or $x_n=\dfrac{Au_n}{1-\frac{1}{2}Bu_n}$, $u_n=\dfrac{x_n}{A+\frac{1}{2}Bx_n}$.

$$\text{Then } u_{n+1}^{-1}=u_n^{-2}+AC+\tfrac{1}{4}B(2-B)$$

$$\text{or if } a=\tfrac{1}{4}B(B-2)-AC,$$

$$u_{n+1}=u_n^2(1-au_n^2)^{-1}. \tag{2.12}$$

Similarly for highly attractive points of higher order, we can reduce $x_{n+1}=x_n^k$ $(A+Bx_n+Cx_n^2+\cdots)^{-1}$

$$\text{to } u_{n+1}=u_n^k(1-au_n^2-bu_n^3\cdots)^{-1}.$$

SOLUTION AT AN ORDINARY FIXED POINT

We have to solve the equation

$$u_{n+1}=su_n\left(1-\frac{u_n^2}{1-au_n}\right)^{-1}. \tag{2.2}$$

Using a method due to Abel (1881) let us put

$$(n+C)\ln s=\ln u_n+\sum_{r=2}^{\infty} b_r u_n^r$$

where the constant C depends on the value of u_0, while the constants b_r are independent of it. Then

$$(n+1+C)\ln s=\ln u_{n+1}+\sum_{r=2}^{\infty} b_r u_{n+1}^r;$$

on subtraction we find

$$\ln s\equiv\ln\left(\frac{u_{n+1}}{u_n}\right)+\sum_{r=2}^{\infty} b_r(u_{n+1}^r-u_n^r)$$

$$\equiv\ln s-\ln\left(\frac{u_n^2}{1-au_n}\right)-\sum_{r=2}^{\infty} b_r u_n^r\left[1-s^r\left(1-\frac{u_n^2}{1-au_n}\right)^{-r}\right].$$

This is an identity, in which we may determine the values of b_r by equating the coefficients of powers of u_n to zero. It is convenient to put

$$\frac{s^r}{1-s^r}=h_r, \text{ so that } 1+h_r=\frac{1}{1-s^r}.$$

We have $u_n^2+au_n^3+(a^2+\frac{1}{2})u_n^4+(a^3+a)\,u_n^5+\cdots$

$$\equiv b_2u_n^2[1-s^2-2s^2u_n^2-2as^2u_n^3\cdots]+b_3u_n^3[1-s^3-3s^3u_n^2\cdots]$$
$$+b_4u_n^4[1-s^4\cdots]+\cdots.$$

Clearly $b_2=\dfrac{1}{1-s^2}=1+h_2$, $b_3=\dfrac{a}{1-s_3}=a(1+h_3)$, $(1-s^4)\ b_4=2s^2b_2+a^2+\frac{1}{2}=a^2+\frac{1}{2}+2h_2$,

whence $b_4=(1+h_4)\ (a^2+\frac{1}{2}+2h_2)$ etc.

Thus we have

$$(n+C)\ \ln s = \ln u_n + (1+h_2)u_n^2 + (1+h_3)au_n^3 + (1+h_4)(a^2+\tfrac{1}{2}+2h_2)u_n^4 +$$
$$(1+h_5)a(a^2+1+2h_2+3h_3)u_n^5 + (1+h_6)[a^4+(\tfrac{3}{2}+2h_2+3h_3+4h_4)a^2+\tfrac{1}{3}+3h_2+2h_4+8h_2h_4]$$
$$u_n^6 + (1+h_7)a[a^4+(2+2h_2+3h_3+4h_4+5h_5)a^2+(1+6h_2+6h_3+2h_4+5h_5+8h_2h_4+$$
$$10h_2h_5+15h_3h_5]\ u_n^7 + \cdots. \qquad (3.1)$$

If a is large such a form as

$$(n+C)\ \ln s = \ln u_n + \sum_{r=2}^{\infty} (1+h_r)a^{r-2}u_n^r + (1+h_4)(\tfrac{1}{2}+2h_2)u_n^4 +$$
$$(1+h_5)(1+2h_2+3h_3)au_n^5 + \cdots$$

may be convenient.

When s is less than unity this may be written

$$n+C = \frac{\log_{10} u_n^{-1}}{\log_{10} s^{-1}} - \frac{3{\cdot}04\ u_n^2}{7\log_{10} s^{-1}}\ [1+h_2+(1+h_3)au_n+\cdots].$$

Here $\frac{3{\cdot}04}{7} = {\cdot}4342857$ is used as an approximation to ${\cdot}4342949$, or $\log_{10}e$. u_n will generally be less than $\frac{1}{2}$, and au_n less than unity. If so the terms in the bracket will amount to less than unity, and may be neglected if we merely wish to calculate n to the nearest integer. However when $s>1$ the series will diverge when $u_n > s^{-1}$, and will not be very accurate when su_n exceeds $\frac{1}{2}$. The other fixed points are given by $u^2+a(1-s)u-1+s=0$; they are $u = \frac{1}{2}[\pm\{(1-s)(4+a^2-a^2s)\}^{\frac{1}{2}} - a(1-s)]$. The series certainly diverges when $|u_n|$ exceeds the modulus of the smaller of these, or exceeds $(s^{-1}-1)^{\frac{1}{2}}$ if they are complex. The series (3.1) is often quite sufficient for computation. However it can be transformed as follows.

$$s^{n+C} = u_n \exp\ [b_2u_n^2 + b_3u_n^3 + b_4u_n^4 + \cdots]$$
$$= u_n + b_2u_n^3 + b_3u_n^4 + (b_4+\tfrac{1}{2}b_2^2)u_n^5 + (b_5+b_2b_3)u_n^6 + \cdots$$
$$= u_n + (1+h_2)u_n^3 + (1+h_3)au_n^4 + [(1+h_4)a^2+1+3h_2+\tfrac{1}{2}h_4+\tfrac{1}{2}h_2^2+2h_2h_4]u_n^5 + \cdots. \qquad (3.2)$$

If $s^{n+C} = t$,

$$u_n = t - (1+h_2)t^3 - (1+h_3)at^4 - [(1+h_4)a^2 - 2 - 2h_2 + \tfrac{1}{2}h_4 - \tfrac{5}{2}h_2^2 + 2h_2h_4]t^5 - \cdots. \qquad (3.3)$$

This series may be used for calculation, and can of course be extended, but (3.1) is generally sufficient. Examples are given later.

Next consider

$$u_{n+1} = s\,u_n\left(1+\frac{u_n^2}{1-au_n}\right)^{-1}. \qquad (2.3)$$

Let $u_n = i\, w_n$, or $w_n = -i\, u_n$. Then

$$w_{n+1} = s\, w_n \left(1 - \frac{w_n^2}{1 - i\, a w_n}\right).$$

So from (3.1)

$$(n+C') \ln s = \ln w_n + \Sigma b_r' w_n^r,$$

where b'_r is derived from b_r by substituting ia for a. Hence

$$(n+C) \ln s = \ln u_n + \Sigma b'_r (-i\, u_n)^r$$

$$= \ln u_n - (1+h_2) u_n^2 - (1+h_3) a u_n^3 + (1+h_4)(-a^2 + \tfrac{1}{2} + 2h_2) u_n^4 + (1+h_5) a (-a^2 + 1 + 2h_2 + 3h_3)\, u_n^5 + \cdots. \qquad (3.4)$$

This is derived from (3.1) by changing the sign of the first term in the coefficient of each power of u_n, conserving that of the second, changing that of the third, and so on. One can write down expressions corresponding to (3.2) and (3.3).

To solve

$$u_{n+1} = s\, u_n (1 + u_n^2)^{-1} \qquad (2.4)$$

we can put $u_{n+1}^2 = s^2 u_n^2 (1 + 2u_n^2 + u_n^4)^{-1}$,
which is an equation of type (2.6) in u_n^2, or use Abel's method directly. By the latter method we find

$$(n+C) \ln s = \ln u_n - (1+h_2) u_n^2 + (1+h_4)(\tfrac{1}{2} + 2h_2) u_n^4 - (1+h_6)(\tfrac{1}{3} + 3h_2 + 2h_4 + 8h_2h_4) u_n^6 + (1+h_8)(\tfrac{1}{4} + 4h_2 + 5h_4 + 2h_6 + 20h_2h_4 + 18h_2h_6 + 12h_4h_6 + 48h_2h_4h_6)\, u_n^8. \qquad (3.5)$$

In the general case we can always, by a homographic transformation of z_n, derive the iteration

$$u_{n+1} = s\, u_n (1 \pm u_n^2 - a_3 u_n^3 - a_4 u_n^4 - a_5 u_n^5 - \cdots).$$

Taking the negative sign for the ambiguity, Abel's method gives

$$(n+C) \ln s = \ln u_n + (1+h_2) u_n^2 + (1+h_3) a_3 u_n^3 + (1+h_4)(a_4 + \tfrac{1}{2} + 2h_2) u_n^4 + (1+h_5)[a_5 + (1 + 2h_2 + 3h_3)\, a_3]\, u_n^5 + \cdots \qquad (3.6)$$

which can be transformed like (3.1).

It may be remarked that (3.1) and similar equations may be written as differential equations provided C is assumed constant. (3.1) becomes

$$\ln s \frac{dn}{du} = u^{-1} + 2(1+h_2) u + 3(1+h_3) a u^2 + 2(1+h_4)(2a^2 + 1 + 4h_2) u^3 + \cdots.$$

We have thus got rid of the awkward logarithmic term.

Equations sometimes arise which can readily be reduced to the form

$$x_{n+1} = s\, x_n (1 - x_n). \qquad (3.7)$$

Putting $x_n = \frac{(1-s)\,u_n}{1-s-u_n}$, or $u_n = \frac{(1-s)\,x_n}{1-s+x_n}$, we have

$$u_{n+1} = s\,u_n\left(1 - \frac{u_n^2}{1-au_n}\right)^{-1}, \tag{2.2}$$

where $a = \frac{2-s}{1-s} = 2 + h_1$.

So from (3.1), $(n+C)\ln s = \ln u_n + (1+h_2)\,u_n^2 + (2+h_1)\,(1+h_3)\,u_n^3 + \cdots$. (3.8)

Solution at a Confluent Point

We have to solve

$$u_{n+1} = u_n\,(1 + u_n - au_n^2)^{-1}. \tag{2.6}$$

Let $n+C = u_n^{-1} - a\ln u_n + (a+\frac{1}{2})\ln(1-au_n) + a\sum_{r=2}^{\infty} b_r u_n^r$.

Then by Abel's method,

$$1 \equiv 1 - au_n + a\ln(1+u_n-au_n^2) + (a+\tfrac{1}{2})[\ln(1+u_n) - \ln(1+u_n-au_n^2)]$$
$$- a\sum_{r=2}^{\infty} b_r u_n^r[1-(1+u_n-au_n^2)^{-r}]$$

or

$$\sum_{r=2}^{\infty} b_r u_n^r[1-(1+u_n-au_n^2)^{-r}] \equiv -u_n + \ln(1+u_n) - \frac{1}{2a}\ln\left(1 - \frac{au_n^2}{1+u_n}\right).$$

The values of b_r are obtained by equating powers of coefficients of u_n. So

$$n+C = u_n^{-1} - a\ln u_n + (a+\tfrac{1}{2})\ln(1-au_n) -$$
$$\frac{au_n^2}{12}\left[1 - \frac{au_n}{3} + \frac{(5a^2+10a+2)}{20}u_n^2 - \cdots\right]. \tag{4.1}$$

The terms in the infinite series are often negligible. The second logarithmic term arises as follows. The fixed points of (2.6) are 0, 0, and a^{-1}. If $x_n = 1 - au_n$, $x_{n+1} = \frac{x_n(a+1-x_n)}{a+x_n-x_n^2}$. This is an equation of type (2.1) with $s = \frac{a+1}{a}$.

By (3.1) its solution for small x_n is

$$(n+C')\ln\frac{a+1}{a} = \ln x_n + O(x_n)$$

or $n+C' = \left(a + \frac{1}{2} - \frac{1}{12a} + \frac{1}{24a^2} - \cdots\right)\ln(1-au_n)$.

The use of the terms $a+\frac{1}{2}$ reduces the coefficient of u_n in (4.1) to zero, and considerably simplifies succeeding terms.

The inversion of (4.1) gives, writing n' for $n+C$,

$$u_n = n'^{-1} + a\, n'^{-2}\ln n' + (a^2 \ln n' + 2a^2 + \tfrac{1}{2})\, n'^{-3}\ln n' + \cdots \tag{4.2}$$

which is too cumbrous for serious computation.

To solve $u_{n+1} = u_n(1+u_n^2)^{-1}$ (2.9)

we put $2u_n^2 = w_n$, so that

$$w_{n+1} = w_n(1 + w_n + \tfrac{1}{4}w_n^2)^{-1}.$$

This is equation (2.6) with $a = -\frac{1}{4}$. The solution is therefore

$$n + C' = \frac{1}{2u_n^2} + \tfrac{1}{2}\ln u_n + \tfrac{1}{4}\ln(1 - \tfrac{1}{2}u_n^2) + \frac{u_n^4}{12}\left(1 + \frac{u_n^2}{6} - \frac{3u_n^4}{80} + \cdots\right). \tag{4.3}$$

In general if $u_{n+1} = u_n(1 + u_n - au_n^2 - bu_n^3 - \cdots)^{-1}$, the reciprocal and logarithmic terms will be as in (4.1) though the series will be different, and if $\Delta u_n = u_n^k + au_n^{k+1} + \cdots$, the solution will be of the form $n + C = \dfrac{1}{ku_n^k} + \cdots$.

SOLUTION AT A HIGHLY ATTRACTIVE POINT

If $u_{n+1} = u_n^2(1 - au_n^2)^{-1}$ (2.12)

let $2^{n+C} = -\ln u_n + \sum_{r=1}^{\infty} b_r u_n^{2r}$,

so $2^{n+1+C} = -\ln u_{n+1} + \sum_{r=1}^{\infty} b_r u_{n+1}^{2r}$.

Hence $-\ln u_{n+1} + \sum_{r=1}^{\infty} b_r u_{n+1}^{2r} + 2\ln u_n - 2\sum_{r=1}^{\infty} b_r u_n^{2r} \equiv 0$, or

$$\sum_{r=1}^{\infty} b_r u_n^{2r}[2 - u_n^{2r}(1 - au_n^2)^{-2r}] \equiv \ln(1 - au_n^2) = -(au_n^2 + \tfrac{1}{2}a^2u_n^4 + \tfrac{1}{3}a^3u_n^6 + \cdots).$$

On equating coefficients we find $b_1 = -\frac{1}{2}a$, etc., so

$$2^{n+C} = -\ln u_n - \tfrac{1}{2}au_n^2 - \tfrac{1}{4}a(a+1)u_n^4 - \tfrac{1}{6}a^2(a+3)u_n^6 - \tfrac{1}{8}a(a^3 + 6a^2 + a + 1)u_n^8 - \cdots$$

or $n + C = (\log 2)^{-1}\log[-\ln u_n - \frac{1}{2}au_n^2 - \frac{1}{4}a(a+1)u_n^4 + \cdots]$. (5.1)

u_n approaches zero very rapidly, and the first term in the power series is usually sufficient. The series may be readily inverted, and if

$A = e^{2^{C}}$, and $t = A^{-2^{n}}$,

$$u_n = t - \tfrac{1}{2}at^3 + \tfrac{1}{8}a(3a-2)t^5 - \tfrac{1}{16}a^2(5a-6)t^7 - \cdots. \tag{5.2}$$

The same method may be used when the iterated function is of higher order, or transcendental. If $u_{n+1}=u_n^2\left(1-\sum_{r=2}^{\infty} a_r u_n^r\right)^{-1}$, then it is easily shown that

$$2^{n+C}=-\ln u_n-\tfrac{1}{2}a_2u_n^2-\tfrac{1}{2}a_3u_n^3-\tfrac{1}{4}\left(a_2^2+a_3+2a_4\right)u_n^4- \cdots \quad (5.3)$$

which can be inverted if desired.

A General Solution for Singly Confluent Quadratic Iterations

We have seen that the general equation (2.5) can, except in one special case, be transformed to

$$x_{n+1}=x_n+a(x_n-1)^2x_n^{-1}. \quad (2.8)$$

Haldane (1932a, 1932b) showed that n could be expanded in ascending powers of a. In genetical applications a is never less than -1, but may assume fairly large positive values. If it is large and positive or close to -1, x_n changes quickly with n and can be calculated over much of its range. The series for n converges slowly or not at all when $|a|$ is large, but is quite satisfactory in the neighbourhood of the fixed points 1 (confluent) and ∞ (attractive). We need only consider values of $x_n>1$ in genetical applications. Haldane's series can be obtained simply by Abel's method as follows.

$$\text{If } x_{n+1}=x_n+ay, \quad (6.1)$$

where y is a known function of x_n, regular in the region considered,

let $y_r=\left(\frac{d}{dx_n}\right)^r y$. Let $an=\sum_{r=1}^{\infty}\frac{a^{r-1}}{r!}\int_{x_0}^{x_n} f_r(x)dx$,

where the functions $f_r(x)$ are to be determined. Provided $f_r(x)$ can be expanded in a Taylor's series,

$$\int_{x_n}^{x_{n+1}} f_r(x)\,dx=\sum_{i=1}^{\infty}\frac{(ay)^i}{i!}\left(\frac{d}{dx}\right)^{i-1} f_r(x_n).$$

So $\sum_{r=1}^{\infty}\left[\frac{a^{r-1}}{r!}\sum_{i=1}^{\infty}\left(\frac{d}{dx}\right)^{i-1} f_r(x_n)\right]-a\equiv 0.$

We can determine the values of $f_r(x_n)$ by equating the coefficients of powers of a to zero. The coefficient of a is

$y f_1(x_n)-1=0$, so $f_1(x_n)=y^{-1}$.

The coefficient of a^{m-1}, multiplied by $m!\,y^{-1}$, if $m>2$, is

$$\sum_{r=1}^{m-1}\binom{m}{r} y^{m-r-1}\left(\frac{d}{dx}\right)^{m-r-1} f_r(x)=0.$$

This is a recurrence equation for $f_r(x)$. For example if $m=4$, $4f_3(x)+6yf'_2(x)+4y^2f''_1(x)=0$.

Hence, provided the series converges uniformly,

$$n=\int_{x_0}^{x_n}\left[\sum_{r=1}^{\infty}\frac{a^{r-2}}{r!}\;f_r(x)\right]dx,$$

where

$$\left.\begin{aligned}
&f_1(x)=y^{-1},\\
&f_2(x)=y^{-1}y_1,\\
&f_3(x)=-\tfrac{1}{2}(y^{-1}y_1{}^2+y_1),\\
&f_4(x)=y^{-1}y_1{}^3+2y_1y_2,\\
&f_5(x)=\tfrac{1}{6}\,(-19y^{-1}y_1{}^4-59y_1{}^2y_2-yy_2{}^2+2yy_1y_3+y^2y_4),\\
&\quad\text{etc.}
\end{aligned}\right\}\qquad(6.2)$$

It is easily shown, by putting $y=x$, that the the leading terms are the expansion of

$$\frac{y_1}{y\ln\,(1+ay_1)}\,.$$

In the case here considered, $y=(x-1)^2\;x^{-1}, y_1=1-x^{-2}, y_r=(-)^r r!\;x^{-r-1}$, if $r>1$, so

$$\left.\begin{aligned}
&f_1(x)=x(x-1)^{-2},\\
&f_2(x)=x^{-1}(x-1)^{-1}(x+1),\\
&f_3(x)=-\tfrac{1}{2}(x^{-1}+2x^{-2}+3x^{-3}),\\
&f_4(x)=x^{-5}(x^2-1)\;(x^2+2x+5)=x^{-1}+2x^{-2}+4x^{-3}-2x^{-4}-5x^{-5},\\
&f_5(x)=-\tfrac{1}{6}x^{-7}\;(x-1)^2\;(19x^4+76x^3+220x^2+360x+105)\\
&\qquad=-\tfrac{1}{6}(19x^{-1}+38x^{-2}+87x^{-3}-\overset{3}{4}x^{-4}-395x^{-5}+150x^{-6}+105x^{-7}).
\end{aligned}\right\}\qquad(6.3)$$

The values of x_n thus change as if they were changing continuously with

$$\frac{dn}{dx}=a^{-1}f_1\,(x)+\frac{1}{2!}f_2(x)+\frac{a}{3!}f_3\;(x)+\frac{a^2}{4!}f_4(x)+\;-\;-\;-\;.$$

On integrating, and putting $z=x_n{}^{-1}$, we find

$$an+C=\ln(1-z)-\ln z-z(1-z)^{-1}+\tfrac{1}{2}a[2\ln(1-z)-\ln z]+\tfrac{1}{12}a^2(\ln z+2z+\tfrac{3}{2}z^2)$$
$$-\tfrac{1}{24}a^3(\ln z+2z+2z^2-\tfrac{2}{3}z^3-\tfrac{5}{4}z^4)+\tfrac{1}{720}a^4(19\ln z+38z+\tfrac{67}{2}z^2-\tfrac{4}{3}z^3-\tfrac{395}{4}z^4+30z^5$$
$$+\tfrac{35}{2}z^6)+\;-\;-\;-\;.$$

The coefficients of $\ln z+2z$ are the terms of the expansion of $-a\;[\ln(1+a)]^{-1}$. On taking these out, we find

$$an+C=(1+a)\ln(1-z)-z(1-z)^{-1}-\frac{a\ln z}{\ln(1+a)}+\left[2+a\;-\frac{2a}{\ln\,(1+a)}\right]\;z$$

$$+\frac{a^2z^2}{8}\left[1-\frac{a}{36}(24-8z-15z^2)+\frac{a^2}{1080}(522-16z-1185z^2+360z^3+210z^4)-O(a^3)\right].\quad(6.4)$$

The first four terms of this series give surprising accuracy even when $a\gg1$, when the reciprocal logarithmic series diverges. The fourth term which only reaches

$+0{\cdot}1146z$ when $a=1$, and $+0{\cdot}3183z$ when $a=-0{\cdot}9$, can be neglected for numerically small values of a.

To give an idea of the accuracy for numerically large values of a, let $a=1$, $z_0=0{\cdot}5$, $z_1=0{\cdot}4$. Then the first four terms of (6.4) give $C=-1{\cdot}32669$, $C+1=-0{\cdot}32055$, giving a difference of 1·00614, instead of unity. Similarly $z_0={\cdot}01$ gives exactly 1 (to the fifth place of decimals). $z_0=0{\cdot}99$ gives 1·00018.

The fit is thus extremely good when z_n or $1-z_n$ are small, but when z_n is about $\frac{1}{2}$, the error is 0·6%. If z_n is the frequency of a recessive gene, over 5,000 individuals would have to be counted to reduce the standard error of the estimate of z_n to this value. So the leading terms of (6·4) are likely to be sufficient for many years to come. When $a=-\frac{1}{2}$ the errors are of the same order of magnitude, as pointed out by Haldane (1932b) who however did not use the fourth term of (6.4). Thus even when dominants are half as fit or twice as fit as recessives, the approximation is excellent, and as the error is roughly proportional to a^2, it is much better for moderate intensities of selection.

If we write $n=F(a, z)$, n is an automorphic function of z. It has a denumerable infinity of poles corresponding to the periodic points of order c where $x_{n+c}=x_n=X_c$. X_c can be expressed as a function of a, and a as a function of X_c. Thus $F(a, z)$ has a denumerable infinity of poles in the a plane, and we conjecture that it is an automorphic function of a. For example if Z_c and A_c are values giving $z_{n+c}=z_n$, then

$$Z_1=0,\ 1,\ 1,\ Z_2=1\pm\sqrt{-a^{-1}},$$

$$A_1=0,\ A_2=\frac{1}{2z}\,[3z-1\pm\sqrt{z^2-6z+1}].$$

Selection of Constant Intensity, with Random Mating

We assume that a pair of autosomal allelomorphs is segregating normally, that mating is at random, and the relative fitnesses of the three genotypes are constant. If then the nth generation is formed from gametes p_n **A**$+q_n$ **a** $(p_n+q_n=1)$, the parents of the next generation are in the ratios

$$(1-k)\ p_n^2\ \mathbf{AA} : 2\ p_nq_n\ \mathbf{Aa} : (1-l)\ q_n^2\ \mathbf{aa}.$$

Here k and l may have any values not exceeding unity. They are most simply thought of as measures of differential mortality, but the same form is reached if they measure differential fertility. It follows that

$$\left.\begin{aligned} q_{n+1}&=\frac{q_n(1-q_n)}{1-kp_n^2-lq_n^2}, \\ \text{or } \Delta q_n&=\frac{p_n\,q_n\,(kp_n-lq_n)}{1-kp_n^2-lq_n^2}. \end{aligned}\right\} \qquad (7.1)$$

If $k+l-kl=0$, or $(1-k)(1-l)=1$, that is to say the fitnesses of the three genotypes are in geometric progression, this simplifies to the first order recurrence equation

$$q_{n+1}=\frac{(1-l)q_n}{1-lq_n}.$$

The fixed points are 0 and 1, and by (1.7),

$$\left.\begin{aligned} q_n &= \frac{(1-l)^n q_0}{1-q_0+(1-l)^n q_0} \\ &= \frac{q_0}{q_0+(1-k)^n(1-q_0)}. \end{aligned}\right\} \quad (7.2)$$

The gene ratio $\frac{p_n}{q_n}$ increases or diminishes in a geometric progression.

We also obtain a first order recurrence if $k=1$, or $l=1$, that is to say if a homozygote is effectively lethal. If $l=1$, (7.1) becomes

$$q_{n+1}=q_n[1-k+(1+k)q_n]^{-1}.$$

There are equilibria when $q_n=0$ or $q_n=Q=k(1+k)^{-1}$, the latter being stable and meaningful if and only if k is positive. From (1.8) and (1.9) the solutions are:

$$\left.\begin{array}{rcll} & q_n=\tfrac{1}{2} & & (k=1), \\ (n+C)\log(1-k) &=& \log(q_n-Q)-\log q_n & (0<k<1,\ q_n>Q), \\ (n+C)\log(1-k) &=& \log(Q-q_n)-\log q_n & (0<k<1,\ q_n<Q), \\ n+C &=& q_n^{-1} & (k=0), \\ (n+C)\log(1-k) &=& \log(q_n-Q)-\log q_n & (0>k>-1), \\ (n+C)\log 2 &=& -\log q_n & (k=-1), \\ (n+C)\log(1-k) &=& \log(Q-q_n)-\log q_n & (k<-1). \end{array}\right\} \quad (7.3)$$

Decimal logarithms may be used. The equations may of course also be written in such forms as

$$q_n=Q[1-(1-k)^{n+C}]^{-1},$$
$$q_n=(n+C)^{-1}$$

and so on.

In general the fixed points of (7.1) are 0, 1, and $k(k+l)^{-1}$. The latter is only biologically relevant if k and $k+l$ have the same sign. There are 9 cases.

(1) $k>0$, $l<0$. $q=0$ is an unstable, $q=1$ a stable equilibrium, both being ordinary fixed points, 0 repulsive and 1 attractive.

(2) $k<0$, $l>0$. As above with p substituted for q.

(3) $k=0$, $l<0$. $q=0$ is a confluent point, effectively repulsive, but with slow withdrawal; $q=1$ is an attractive point, giving a stable equilibrium.

(4) $k<0$, $l=0$. As above with p substituted for q.

(5) $k=0$, $l>0$. $q=0$ is a confluent point, effectively attractive, but with slow approach; $q=1$ is an ordinary repulsive point.

(6) $k>0$, $l=0$. As above with p substituted for q.

(7) $k<0$, $l<0$. There is a repulsive point giving an unstable equilibrium at $q=k\ (k+l)^{-1}$ and attractive points giving stable equilibrium at $q=0$ and 1.

(8) $k>0$, $l>0$. There is an attractive point giving a stable equilibrium at $q=k(k+l)^{-1}$, and repulsive points giving unstable equilibria at $q=0$ and 1.

(9) $k=l=0$. Every value of q gives a neutral equilibrium.

In fact the last case is negligible, though selection may be less important than mutation or drift. Also (3), (4), (5), and (6) are ideal cases, probably never realised in practice, though we often do not know whether the heterozygote is slightly fitter or slightly less fit than one homozygote. They must therefore be considered. It will be seen that (1) and (2), (3) and (4), (5) and (6) are equivalent. So (1), (3), (5), (7), and (8) demand investigation.

An approximate solution can be given if $|k|$ and $|l|$ are both small, when (7.1) can be treated as a differential equation

$$\frac{dq}{dn}=q(1-q)\ [k-(k+l)q]$$

$$\text{whence } n=\int \left[\frac{1}{kq}-\frac{1}{l(1-q)}+\frac{(k+l)^2}{kl\{k-(k+l)q\}}\right]dq$$

or

$$\left.\begin{aligned} n+C&=k^{-1}\ln q_n+l^{-1}\ln(1-q_n)+(k+l)\ (kl)^{-1}\ln[(k+l)\ q_n-k] \\ &\qquad\left(q_n>\frac{k}{k+l}\right) \\ &=k^{-1}\ln q_n+l^{-1}\ln(1-q_n)-(k+l)\ (kl)^{-1}\ln[k-(k+l)\ q_n] \\ &\qquad\left(q_n<\frac{k}{k+l}\right)\end{aligned}\right\} \quad (7.4)$$

with the well-known simplifications when $k=0$, or $l=0$. We shall see that these are approximations to the values found when k and l are not small.

There is a further simplification when $k=0$, or $l=0$, that is to say when dominance is complete. If $l=0$,

$$q_{n+1}=\frac{q_n}{1-kp_n^{\,2}},\ \text{or if } z_n=q_n^{-1},\ \ z_{n+1}=z_n-k\ (z_n-1)^2\ z_n^{-1}. \quad (7.5)$$

We have shown how to solve this equation, which is our (2.8) with $a\doteq-k$. In general we need expressions for:—

(*a*) The relation of n and q_n near zero when this is an ordinary attractive or repulsive point.

(*b*) The same relation when zero is a confluent point.

(*c*) The same relation when $Q=k\ (k+l)^{-1}$ is an ordinary attractive or repulsive point between zero and unity.

In the neighbourhood of zero we have, from (7.1)

$$q_{n+1}=\frac{q_n(1-q_n)}{1-k+2kq_n-(k+l)q_n^2}.$$

Thus in (2.1), $s=(1-k)^{-1}$, $A=-l$, $B=2k(1-k)^{-1}$, $C=-(k+l)(1-k)^{-1}$.

Hence $\theta=k^{-1}(2k+l-kl)$, $A+\theta=k^{-1}(2k+l-2kl)$,

$$A^2-AB+C=-(1-k)^{-1}(1-l)(k+l-kl).$$

If this is negative, we put

$$u_n=[(1-k)^{-1}(1-l)(k+l-kl)]^{\frac{1}{2}}q_n[1-k^{-1}(k+l-2kl)q_n]^{-1},$$
$$a=-k^{-1}(2k+l-2kl)[(1-k)^{-1}(1-l)(k+l-kl)]^{-1}, \tag{7.6}$$

whence $u_{n+1}=su_n\left(1-\frac{u_n^2}{1-au_n}\right)^{-1}$. (2.2)

Thus if $k=-0{\cdot}1$, $l=+0{\cdot}2$,

$$u_n=\frac{\sqrt{\frac{6}{11}}}{2{\cdot}5+3{\cdot}5\,q_n}=\frac{0{\cdot}2954196\,q_n}{1+1{\cdot}4\,q_n}$$

$$\text{or } q_n=\frac{u_n}{\cdot2954196-1{\cdot}4u_n},$$

$$s=\tfrac{10}{11},\; a=\sqrt{\tfrac{11}{6}}=1{\cdot}354004,$$

$$h_2=\frac{100}{21},\; h_3=\frac{1000}{331},\; h_4=\frac{10{,}000}{4{,}641},\; h_5=\frac{100{,}000}{61{,}051},\text{ etc.}$$

$$\text{and } n+C=-24{\cdot}15885\log_{10}u_n-10{\cdot}49184\,u_n^2\left[\frac{121}{21}+\frac{1331}{331}au_n+\frac{14641}{4641}\left(\frac{11}{6}+\frac{1}{2}+\frac{200}{21}\right)u_n^2+\cdots\right]$$

$$=-24{\cdot}15885\log_{10}u_n-60{\cdot}4530\,u_n^2(1+{\cdot}96006\,u_n+6{\cdot}4919\,u_n^2+\cdots).$$

If $u_0={\cdot}01$,

$$C=48{\cdot}3177-{\cdot}0060453\,(1+{\cdot}00960+{\cdot}0006492+\cdots)=48{\cdot}3116.$$

Now if $q_n=\frac{1}{2}$, $u_n=0{\cdot}086888$, so

$$n+48{\cdot}3116=25{\cdot}6335-0{\cdot}46915\,(1+{\cdot}083518+{\cdot}04908+\cdots)=25{\cdot}1396$$

or $n=-23{\cdot}1720$. The error is of the order of the last term in the series, or 0·02. If this were considered unacceptable, further terms could be taken, but in fact it is probably sufficient to determine n to the nearest unit. If $u_0={\cdot}01$, $q_0={\cdot}0355341$. So 23·2 generations would be needed to reduce the gene frequency q_n from 50% to 3·553%. Clearly for positive values of n, only the logarithmic term need be used.

If $k+l-kl$ is negative, A^2-AB+C is positive, so

$$u_n=[-(1-k)^{-1}(1-l)(k+l-kl)]^{\frac{1}{2}}q_n[1-k^{-1}(2k+l-kl)q_n]^{-1}$$

$$a=-k^{-1}(2k+l-2kl)[-(1-k)^{-1}(1-l)(k+l-kl)]^{\frac{1}{2}}$$

$$u_{n+1}=su_n\left(1+\frac{u_n^2}{1-au_n}\right)^{-1}. \tag{2.3}$$

The arithmetical work is quite similar.

When $k=0$, zero is a confluent point, and we can use equation (7.1). In the neighbourhood of $Q=\frac{k}{k+l}$, let $q_n=Q+x_n$, $P=1-Q$. Then (7.1) becomes

$$\Delta x_n=\Delta q_n=\frac{-(k+l)\,x_n\,(P-x_n)\,(Q+x_n)}{1-k(P-x_n)^2-l(Q+x_n)^2}$$

$$=\frac{-x_n\,[kl+(l^2-k^2)x_n-(k+l)^2x_n^2]}{k+l-kl-(k+l)^2\,x_n^2},$$

$$\text{or } x_{n+1}=\frac{x_n\,[k+l-2kl+(k^2-l^2)\,x_n]}{k+l-kl-(k+l)^2\,x_n^2}. \tag{7.7}$$

So in (2.1), $s=\frac{k+l-2kl}{k+l-kl}$, $A=\frac{k^2-l^2}{k+l-2kl}$, $B=0$, $C=\frac{-(k+l)^2}{k+l-kl}$.

$$\theta=\frac{(k^2-l^2)\,(k+l-kl)}{kl\,(k+l-2kl)}.$$

$$A^2-AB+C=-(1-k)\,(1-l)\,(k+l)^4\,(k+l-2kl)^{-2}\,(k+l-kl)^{-1},$$

which has the opposite sign to $k+l-kl$, and is negative if the equilibrium is stable. In the stable case let

$$u_n=\left[\frac{(1-k)\,(1-l)}{k+l-kl}\right]^{\frac{1}{2}}\frac{kl\,(k+l)^2x_n}{kl\,(k+l-2kl)+(l^2-k^2)\,(k+l-kl)x_n}$$

$$a=\frac{l-k}{kl}\left[\frac{k+l-kl}{(1-k)\,(1-l)}\right]^{\frac{1}{2}},$$

$$u_{n+1}=su_n\left(1-\frac{u_n^2}{1-au_n}\right)^{-1}.$$

For example if $k=\cdot 2$, $l=\cdot 1$,

$$Q=\frac{2}{3}\,,\ x_n=q_n-\frac{2}{3}\,,\ s=\frac{13}{14}\,.$$

$$u_n=\frac{27x_n}{\sqrt{14}\,(13-21x_n)}\quad=\frac{7\cdot 21605x_n}{13-21x_n}$$

$$a=-\frac{5}{6}\sqrt{14}=-3\cdot 11805$$

$$u_{n+1}=su_n\left(1-\frac{u_n{}^2}{1-au_n}\right)^{-1}.$$

We note that x_n may be positive or negative. If it is negative we put $v_n=-u_n$ to avoid complex logarithms.

SELECTION WITH CONSTANT INTENSITIES AND INBREEDING

Consider selection as in the last section, but with a constant mean coefficient of inbreeding f. If then the gametic frequencies are p_n **A** $+q_n$ **a** as before, the parents of the next generation are in the ratios

$$(1-k)p_n\,(p_n+f\,q_n)\ \mathbf{AA}\ :\ 2(1-f)p_nq_n\ \mathbf{Aa}\ :\ (1-l)q_n(q_n+f\,p_n)\ \mathbf{aa}.$$

Here $1\geqslant f\geqslant 0$. In human populations f rarely exceeds 0·02, but it may be much larger in plants and animals. We find without difficulty

$$q_{n+1}=\frac{q_n[1-l\,(q_n+f\,p_n)]}{1-k\,p_n(p_n+f\,q_n)-l\,q_n(q_n+f\,p_n)}$$

$$=\frac{q_n[1-fl-(1-f)lq_n]}{1-k+(2k-fk-fl)q_n-(1-f)(k+l)q_n{}^2}\,, \tag{8.1}$$

$$\triangle q_n=\frac{p_nq_n[\,(k-fl)p_n\,-\,(l-fk)q_n]}{1-k\,p_n{}^2-f\,(k+l)p_nq_n-lq_n{}^2}\,. \tag{8.2}$$

If $k+l-(1+f)\,kl=0$, or $(1-k)\,(1-l)+fkl=1$, (8.1) becomes

$$q_{n+1}=\frac{(1-l-fl)q_n}{1-(1+f)\,lq_n}\,, \tag{8.3}$$

which is of form (1.7).

We also obtain a first order recurrence if $k=1$, or $l=1$, that is to say a homozygote is lethal. If $l=1$, (8·1) becomes

$$q_{n+1}=(1-f)q_n[1-k+(1-f)(1+k)q_n]^{-1}.$$

If $Q=(k-f)(1-f)^{-1}(1+k)^{-1}$, $q_n=Q$ is the stable equilibrium if $k>f$. Otherwise $q_n=0$ is the stable equilibirum. By (1·8) and (1·9) the solutions are:

$$\left.\begin{array}{lll} q_n = \tfrac{1}{2} & & (k=1), \\ (n+C)\log\left(\dfrac{1-k}{1-f}\right) = \log(q_n-Q)-\log q_n & & (1>k>f,\ q_n>Q), \\ (n+C)\log\left(\dfrac{1-k}{1-f}\right) = \log(Q-q_n)-\log q_n & & (1>k>f,\ q_n<Q), \\ n+C = (1+k)^{-1}q_n^{-1} & & (k=f), \\ (n+C)\log\left(\dfrac{1-k}{1-f}\right) = \log(q_n-Q)-\log q_n & & (f>k>-1), \\ (n+C)\log\left(\dfrac{2}{1-f}\right) = -\log q_n & & (k=-1), \\ (n+C)\log\left(\dfrac{1-k}{1-f}\right) = \log(Q-q_n)-\log q_n & & (k<-1). \end{array}\right\} \quad (8\cdot 4)$$

As before the equations may be written in several forms. They could be used, for example, in the study of sickling in man, where one homozygote is nearly lethal, and mutation is generally thought to be unimportant.

In the general case the fixed points are $q_n=0$, $q_n=1$, and $q_n=Q=\dfrac{k-fl}{(1-f)(k+l)}$, the latter representing a biologically relevant equilibrium if $k-fl$ and $(1-f)(k+l)$ have the same sign. As before, there are 9 possible cases.

(1) $k-fl>0$, $l-fk<0$. $q=0$ is unstable, $q=1$ stable.
(2) $k-fl<0$, $l-fk>0$. $q=0$ is stable, $q=1$ unstable.
(3) $k=fl$, $l-fk<0$. $q=0$ is unstable but left slowly, $q=1$ stable.
(4) $k-fl<0$, $l=fk$. $q=0$ is stable, $q=1$ unstable but left slowly.
(5) $k=fl$, $l-fk>0$. $q=0$ is stable but approached slowly, $q=1$ is unstable.
(6) $k-fl>0$, $l=fk$. $q=0$ is unstable, $q=1$ stable but slowly approached.
(7) $k-fl<0$, $l-fk<0$. $q=0$ and $q=1$ are stable, $q=Q$ unstable.
(8) $k-fl>0$, $l-fk>0$. $q=0$ and $q=1$ are unstable, $q=Q$ stable.
(9) $k=fl$, $l=fk$. There is a neutral equilibrium at any point. This is only possible if $k=l=0$, or $k=l$, $f=1$.

At $q=0$ when $k\neq fl$, we have

$$s=\frac{1-fl}{1-k},\quad A=\frac{-(1-f)l}{1-fl},\quad B=\frac{2k-f(k+l)}{1-k},\quad C=\frac{-(1-f)(k+l)}{1-k}.$$

So $\theta=(k-fl)^{-1}(1-fl)^{-1}[2k+l-kl-f(k+2l+kl)+f^2l(k+l)]$;

$$A^2-AB+C=-(1-f)(1-l)(1-fl)^{-2}(1-k)^{-1}(k+l-kl-fkl)$$

and so on. It is generally more convenient to compute s, A, B, and C, etc. numerically.

$$\text{If } k = fl,\; q_{n+1} = \frac{q_n[1-fl-(1-f)l\,q_n]}{1-fl+f\,(1-f)l\,q_n-(1-f^2)l\,q_n^2}. \tag{8.5}$$

$$\text{So } A = \frac{-(1-f)l}{1-fl},\; B = \frac{f\,(1-f)l}{1-fl},\; C = \frac{-(1-f^2)l}{1-fl}.$$

$$\text{So if } u_n = \frac{(1-f^2)\,lq_n}{1-fl\;-(1-f)\,lq_n}, \tag{8.6}$$

$$u_{n+1} = u_n(1+u_n-au_n^2)^{-1},$$

where $a = -(1+f)^{-2}\, l^{-1}(1-l)$.

Or we may use the form (2.8).

If there is an equilibrium at $Q = \dfrac{k-fl}{(1-f)\;(k+l)}$, let $q_n = Q+x_n$.

Then

$$x_{n+1} = \frac{x_n[k+l-2kl-f\,(k+l-k^2-l^2)+(1-f)\;(k^2-l^2)x_n]}{k+l-kl-f\,kl-f\,(k^2-l^2)x_n-(1-f\,)\;(k+l)^2\,x_n^2}.$$

$$\text{Thus } s = \frac{k+l-2kl-f\,(k+l-k^2-l^2)}{k+l-kl-f\,kl} = 1 - \frac{kl+f\,(k+l-k^2-kl-l^2)}{k+l-kl-f\,kl},$$

$$A = \frac{(1-f\,)\;(k^2-l^2)}{k+l-2kl-f\,(k+l-k^2-l^2)},\; B = \frac{-f\,(k^2-l^2)}{k+l-kl-f\,kl},$$

$$C = \frac{-(1-f\,)\;(k+l)^2}{k+l-kl-f\,kl}.$$

Again, numerical calculation is usually most convenient.

PARTIAL SELF-FERTILIZATION WITH CULLING OF RECESSIVES

Sen (1961) has discussed this case. Suppose the nth generation consists of $(1-x_n)$ **AA**, x_n **Aa**, a fraction c is mated at random, while $1-c$ is self-fertilized, then the next generation consists of

$$\left[\left(1-c\right)\;(1-\tfrac{3}{4}x_n)+c\left(1-\frac{x_n}{2}\right)^2\right]\textbf{AA} + [(1-c)\tfrac{1}{2}x_n+c\,x_n(1-\tfrac{1}{2}x_n)]\;\textbf{Aa}$$

$$+\left[(1-c)\;\tfrac{1}{4}x_n+c\,\frac{x_n^2}{4}\right]\textbf{aa}.$$

Hence $x_{n+1}=\dfrac{\frac{1}{2}(1+c)x_n-\frac{1}{2}cx_n^2}{1-\frac{1}{4}(1-c)x_n-\frac{1}{4}c\,x_n^2}$. (9.1)

$x_n=0$ is a stable equilibrium,

$$s=\tfrac{1}{2}(1+c),\ A=-\frac{c}{1+c},\quad B=-\tfrac{1}{4}(1-c),\quad C=-\tfrac{1}{4}c.$$

$$\theta=\frac{1-4c-c^2}{2(1-c^2)},\ A^2-AB+C=-\frac{c\,(1-c)}{2(1+c)^2},$$ which is never positive.

So if $u_n=\dfrac{[2c(1-c)]^{\frac{1}{2}}\,(1-c)x_n}{2(1-c^2)-(1-4c-c^2)\,x_n}$,

$$a=\frac{-(1-6c+c^2)}{1-c}\left[\frac{2}{c(1-c)}\right]^{\frac{1}{2}}$$

$$=-2(2-4s+s^2)\,[(1-s)^3\,(2s-1)]^{-\frac{1}{2}},$$ and $s=\frac{1}{2}(1+c)$,

$$u_{n+1}=su_n\left(1-\frac{u_n^2}{1-au_n}\right)^{-1}. \qquad (2.2)$$

The frequency of recessives culled in generation $n+1$ is $\frac{1}{4}x_n\,(1-c+c\,x_n)$. Equation (3.2) should not be used until u_n^2 is sufficiently small.

In what follows we shall give some examples of the treatment of iterations of higher order than quadratic, or of transcendental functions.

Competition by Hybridization

Consider two species A and B whose hybrids are inviable or sterile. Let their relative fitnesses be as $1+k:1-k$. Let p_n be the frequency of A and q_n that of B in generation n, the sex ratios being the same in the two species. If encounters are at random the A females will encounter A and B males in proportion to their frequencies. Suppose that the probability that an encounter with a B male will lead to fertilization is λ times the corresponding probability for an encounter with an A male. Then the progeny of A females will be $\dfrac{p_n}{p_n+\lambda q_n}$ A, and $\dfrac{\lambda q_n}{p_n+\lambda q_n}$ hybrids. Let μ be a similar parameter for B females. Both λ and μ are probably usually much less than unity. Then it is easily seen that

$$q_{n+1}=\frac{(1-k)\,q_n^2}{q_n+\mu p_n}\left[\frac{(1+k)\,p_n^2}{p_n+\lambda q_n}+\frac{(1-k)\,q_n^2}{q+\mu p_n}\right]^{-1}$$

$$=\frac{(1-k)\,q_n^2\,(p_n+\lambda q_n)}{(1+k)\,p_n^2\,(\mu p_n+q_n)+(1-k)\,q_n^2\,(p_n+\lambda q_n)},$$

$$\Delta q_n = \frac{p_n q_n\,[(1-k)\,\lambda\, q_n^2 - 2k\, p_n q_n - (1+k)\,\mu\, p_n^2]}{(1+k)p_n^2\,(\mu p_n + q_n) + (1-k)\, q_n^2\,(p_n + \lambda q_n)}\,. \tag{9.1}$$

There are therefore four fixed points when $q=0$, $q=1$, or $(1-k)\,\lambda\, q^2 - 2k\, pq - (1+k)\,\mu\, p^2 = 0$. $q=0$ and $q=1$ are highly attractive, and represent very stable equilibria. Of the other two fixed points one gives a negative value of p or q, the fourth gives

$$Q = \frac{k - \mu\,(1+k) + (k^2 + \lambda\mu - \lambda\mu\, k^2)^{\frac{1}{2}}}{2k + \lambda\,(1-k) - \mu\,(1+k)} \tag{9.2}$$

which is a repulsive point with $s = 1 + \dfrac{2\,(k^2 + \lambda\mu - \lambda\mu\, k^2)^{\frac{1}{2}}}{1 + (k^2 + \lambda\mu - \lambda\mu k^2)^{\frac{1}{2}}}$,

representing an unstable equilibrium.

Since $\dfrac{P}{Q} = \dfrac{(k^2 + \lambda\mu - \lambda\mu\, k^2)^{\frac{1}{2}} - k}{\mu(1+k)}$

$$= \frac{\lambda\,(1-k)}{2k}\left[1 - \frac{\lambda\mu(1-k^2)}{4k^2} + \frac{\mu^2\lambda^2(1-k^2)^2}{8k^4} - \cdots\right]$$

we see that Q increases with μ and decreases with λ. So natural selection must tend to increase discrimination by females and decrease discrimination by males in each species.

Consider the dynamics when q_n is small. (9.1) may be written

$$q_{n+1} = \frac{(1-k)\; q_n^2\,[1-(1-\lambda)\; q_n]}{(1+k)\mu + (1+k)\,(1-3\mu)q_n - (1+3k-3\mu-3\mu k)q_n^2 + (2k+\lambda-\lambda k-\mu-\mu k)q_n^3}\,.$$

If $u_n = \dfrac{(1+k)\;\mu q_n}{(1-k)\,[1+\frac{1}{2}\,(1+k)\,(1-2\mu+\lambda\mu)q_n]}$, then

$$u_{n+1} = u_n^2\left[1 - a u_n^2 - \frac{b u_n^3}{1 - c u_n}\right]. \tag{9.3}$$

THE PROBABILITY OF SURVIVAL OF A MUTANT GENE

Suppose that a mutant gene occurs in a very large population of constant number. Suppose one member of the population carries a mutant gene, let p_r be the probability that this gene will be found in members of the next generation, and let $f(t) = \sum_{r=0}^{\infty} p_r\, t^r$ be the generating function of this probability distribution. If q_n be the probability that this gene has disappeared after n generations, Fisher (1930) showed that $q_{n+1} = f(q_n)$.

Fisher considered the case when $f(t) = e^{-1} \sum_{r=0}^{\infty} \frac{t^r}{r!} = e^{t-1}$,
that is to say the distribution is a Poisson distribution with mean unity. This implies that the gene is neutral. $q_{n+1} = e^{q_n} - 1.$ (10.1)

This has two confluent fixed points at $q=1$, and an infinity of repulsive points $\lambda(\cot\lambda + i)$, where $\lambda \operatorname{cosec} \lambda = e^{\lambda\cot\lambda - 1}$. Only the real confluent point is of biological interest. Let $z_n = 1 - q_n$. Then z_n is the probability that the mutant will still be present in the population after n generations. $z_0 = 1$, and $z_{n+1} = 1 - e^{-z_n}$ (10.2)

$$= z_n\left(\tfrac{1}{2}z_n \coth \tfrac{1}{2}z_n + \tfrac{1}{2}z_n\right)^{-1}$$

$$= z_n\left(1 + \frac{1}{2}z_n + \frac{1}{12}z_n^2 - \frac{1}{720}z_n^4 + \cdots\right)^{-1},$$

or if $u_n = \frac{1}{2}z_n$

$$u_{n+1} = u_n\left(1 + u_n + \frac{1}{3}u_n^2 - \frac{1}{45}u_n^4 + \frac{2}{945}u_n^6 - \frac{1}{4725}u_n^8 + \cdots\right)^{-1} = u_n(u_n \coth u_n + u_n)^{-1}.$$

Let $n + C = u_n^{-1} - \frac{1}{3}\ln u_n + \frac{1}{6}\ln\left(1 - \frac{1}{3}u_n\right) + \sum_{r=2}^{\infty} b_r u_n^r.$

Then $1 \equiv \coth u_n - u_n^{-1} + 1 + \frac{1}{3}\ln(u_n \coth u_n + u_n) + \frac{1}{6}\ln\left[1 - \frac{u_n}{3(u_n \coth u_n + u_n)}\right]$

$$-\frac{1}{6}\ln\left(1 - \frac{1}{3}u_n\right) + \sum_{r=2}^{\infty} b_r(u_{n+1}^r - u_n^r).$$

$$\sum_{r=2}^{\infty} b_r u_n^r\left[1 - (u_n \coth u_n + u_n)^{-r}\right] \equiv \coth u_n - u_n^{-1} + \frac{1}{6}\Big[\ln(u_n \coth u_n + u_n)$$
$$+ \ln\left(u_n \coth u_n + \frac{2}{3}u_n\right) - \ln\left(1 - \frac{1}{3}u_n\right)\Big].$$

Hence, or from Fisher's formula, we find

$$n + C = u_n^{-1} - \frac{1}{3}\ln u_n + \frac{1}{6}\ln\left(1 - \frac{1}{3}u_n\right) + \frac{u_n^2}{12}\left(\frac{1}{5} - \frac{u_n}{27} - \frac{5u_n^2}{42} - \frac{1051u_n^3}{47{,}250} + \frac{53u_n^4}{164{,}025}\right.$$
$$\left. + \frac{14{,}533u_n^5}{9{,}185{,}400} - \cdots\right).$$

$$n + C' = \frac{2}{z_n} - \frac{1}{3}\ln z_n + \frac{1}{6}\ln\left(1 - \frac{z_n}{6}\right) + \frac{z_n^2}{48}\left(\frac{1}{5} - \frac{z_n}{54} - \frac{5z_n^2}{168} - \frac{1{,}051z_n^3}{378{,}000} + \frac{53z_n^4}{2{,}624{,}400}\right.$$

$$\left. + \frac{1{,}453z_n^5}{293{,}932{,}800} - \cdots\right). \quad (10.3)$$

The coefficients are a little simpler than those of Fisher's series, but Fisher has made all the relevant calculations, including $C' = +\cdot 50732430$. The series terms are negligible in practice.

Next suppose that the mutant is not neutral, but confers a relative fitness $1+k$ on heterozygotes. Then if z_n is the probability of surviving for n generations,

$$z_{n+1} = 1 - e^{-(1+k)z_n}. \tag{10.4}$$

There is a fixed point at $z = 0$, and another real fixed point, besides an infinity of complex repulsive points.

If k is negative, then z_n tends to zero, which is an attractive point, while there is a repulsive and biologically irrelevant negative fixed point. Let $1+k = s < 1$.

Then $z_{n+1} = sz_n(1 - \frac{1}{2}sz_n + \frac{1}{6}s^2z_n^2 - \frac{1}{24}s^3z_n^3 + \cdots)$.

Putting $y_n = \dfrac{z_n}{1 - \dfrac{sz_n}{2(1-s)}}$, we have

$$y_{n+1} = s\,y_n\,[1 + \tfrac{1}{6}s^2y_n^2 - \tfrac{1}{24}s^3(3-s)y_n^3 + \cdots].$$

Hence $(n+C)\ln s = \ln y_n + \dfrac{s^2y_n^2}{6(1-s^2)} + O(y_n^3)$,

or $(n+C')\ln s = \ln z_n - \ln(2 - 2s - sz_n) - \dfrac{s^2z_n^2}{6(1-s^2)} + O(z_n^3)$.

Further terms can easily be calculated.

If k is positive, there is an attractive fixed point at a positive value Z, while zero is repulsive, with $s = 1+k$. To find Z, we revert the series

$$(1+k)Z = -\ln(1-Z) = Z + \tfrac{1}{2}Z^2 + \tfrac{1}{3}Z^3 + \cdots$$

whence $Z = 2k(1+k)^{-\frac{4}{3}} + \dfrac{2k^4}{405}(4 - \tfrac{31}{3}k + \cdots)$. (10.5)

This is the probability that a single mutant will survive indefinitely. Even if k is as large as $\frac{1}{2}$, the first term has an error under $\cdot 001$. To find the rate at which the probability approximates to Z, we put $z_n = Z - x_n$. Then

$$x_{n+1} = (1-Z)\,[e^{(1+k)x_n} - 1]$$
$$= (1+k)(1-Z)\,x_n\,[1 - \tfrac{1}{2}(1+k)x_n + \tfrac{1}{6}(1+k)^2x_n^2 - \cdots], \tag{10.6}$$
$$s = (1+k)(1-Z) = 1 + k - 2k(1+k)^{-\frac{1}{3}} = 1 - k + \tfrac{2}{3}k^2 - \tfrac{4}{9}k^3 + \cdots.$$

If $(1+k)x_n = y_n$ we have $y_{n+1} = s\,y_n(1 - \frac{1}{2}y_n + \frac{1}{6}y_n^2 - \frac{1}{24}y_n^3 + \cdots)$, Abel's method gives

$$(n+C)\ \ln s = \ln y_n - \ln(2-2s-y_n) + \frac{y_n^2}{12(1-s^2)} + \frac{y_n^3}{24(1-s)(1-s^3)} + \cdots$$

or $$(n+C')\ \ln s = \ln x_n - \ln[2-2s-(1+k)x_n] + \frac{(1+k)^2 x_n^2}{12(1-s^2)} + O(x_n^3). \qquad (10.7)$$

Discussion

Having shown how to solve one set of non-linear recurrence relations as exactly as is wished, it is perhaps worth enumerating some things which remain to be done.

1. Solution of equations of the types considered when one or more of the constants is replaced by a random variable. For example if $q_{n+1} = \frac{q_n(1-k\,q_n)}{1-k\,q_n^2}$, and k is a random variable never exceeding unity, whose distribution has given cumulants, it would be desirable to find an expression for the distribution of q_n in terms of q_0, and the above cumulants.

2. Solution of the above equations when the population is large, but a gene is rare enough to have a finite probability of extinction. Fisher (1930) opened this problem.

3. Solution when the population is finite. Here it is often desirable to consider mutation. Wright's analysis requires development when selection is intense.

4. Solution when a parameter such as k varies in a simple manner with time. We hope to solve this problem. In this case mutation must be considered.

5. Solution when generations overlap. Here the difference, or recurrence equations, are replaced by non-linear integral equations. Norton (1928) and Haldane (1927) opened up this field, but it is difficult because the ages of mates are correlated.

6. Solution of sets of simultaneous equations such as (1.1). Here there is more than one arbitrary constant like C; and when we express q_n as a function of n, its coefficient may tend to zero quicker than any negative power of n, while being non-negligible for several generations. Such equations generally arise when genes at several loci are considered.

7. Solution of equations combining two or more of these complications, for example those arising with finite populations and overlapping generations.

8. Tests for the truth of various hypotheses, e.g. that dominance is complete, mating at random, and the relative fitness of recessives constant, given a series of population samples.

9. Methods of estimating parameters giving the intensity of selection. This has only been done in the very simplest cases.

10. Study of n as a function of a in equation (6.4). We do not know whether n is a function of a type so far studied.

11. Investigation of the convergence of the series derived in this paper.

12. Tabulation of solutions of equations (2.2) and (2.3) and perhaps others, for different values of s and a.

Even if some of these tasks involve the use of electronic computers, the programming of such computations requires great skill.

SUMMARY

When q_n is a parameter of a population in generation n, q_{n+1} is often a simple function of q_n. If so, n is an automorphic function of q_n. A simple transformation of q_n permits the expression of n as an infinite series which often converges quickly, and allows numerical calculation. It is sometimes possible to obtain a very close approximation to the value of n in terms of logarithms. Examples are given.

REFERENCES

ABEL, H. (1881).* *Memoires Posthumeuses; Oeuvres.* Vol. II. Christiania.

FISHER, R. A. (1930). The distribution of gene ratio for rare mutations. *Proc. Roy. Soc. Edin.*, **50**, 205-220.

HALDANE, J. B. S. (1927). A mathematical theory of natural and artificial selection. Part. IV. *Proc. Camb. Phil. Soc.*, **23**, 607-615.

HALDANE, J. B. S. (1932a). A mathematical theory of natural and artificial selection. Part VI. *Proc. Camb. Phil. Soc.*, **28**, 244-248.

HALDANE, J. B. S. (1932b). On the non-linear difference equation $\Delta x_n = k\phi(x_n)$. *Proc. Camb. Phil. Soc.*, **28**, 234-243.

NORTON, H. T. J. (1928). Natural selection and Mendelian variation. *Proc. Lond. Math. Soc.*, **28**, 1-45.

PICARD, E. (1928). Leçons sur quelques équations fonctionelles avec des applications à divers problèmes d'analyse et de physique mathématique. *Cahiers Scientifiques*, Fascicule III. Paris

SEN, S. N. (1961). Complete selection with partial self fertilization. *J. Genet.*, **57**, 339-344.

VALIRON, G. (1954). *Fonctions Analytiques.* Presses Universitaires de France, Paris.

*Not read by authors; references taken from Picard (1928).