JIA 71 (1943) 0228-0258

ON THE APPLICATION OF QUANTUM MECHANICS TO MORTALITY TABLES

By R. D. ANDERSON, F.I.A.
Assistant Actuary of the Commercial Union Assurance Co., Ltd.

If you have had your attention directed to the novelties in thought in your own lifetime, you will have observed that almost all really new ideas have a certain aspect of foolishness when they are first produced.
Prof. A. N. WHITEHEAD, *Science and the Modern World.*

PART 1. INTRODUCTION

1. Quantum Mechanics is a portentous name; the alternative—Wave Mechanics—is almost as bad. The mathematics are formidable, the literature large and growing rapidly, and the subject-matter dealt with is the behaviour of physical things, such as electrons, protons, atoms, and so on. Why, then, should actuaries as such take any interest in the subject?

Because it is an application of the statistical theory of probability, and one of the objects of the Institute, for which it obtained its charter, is 'the extension and improvement of the data and methods of the science which has its origin in the application of the doctrine of probabilities to the affairs of life...'.

We, as well as physicists (or ought I to say scientists?), deal with probabilities because we also are concerned with the unpredictable behaviour of individuals. The formulae of Quantum Mechanics may not be of any use to us, but the ideas and the methods of their application are, because they may help us in our problems.

2. 'When a physicist predicts the result of an atomic experiment', says Mr Gurney,* 'his prediction is often embodied in a distribution curve. His recording instruments are not usually sensitive to individual atoms, and in any case, the conditions of experiment cannot be made sufficiently precise to enable him to predict a single definite value for the quantity being measured. Often the best he can do then is to express the expected result by

* *Elementary Quantum Mechanics*, R. W. Gurney, M.A., Ph.D.

means of a curve showing the number of particles for which the measured value will be, say, between q and $q+dq$. To such a curve it will be convenient to give a name; we will call it the pattern of the predicted results, characteristic of the particular conditions and apparatus used.'

We are acquainted with such curves and call them 'frequency curves'. But the curves of the physicist are not the 'Pearson group'; generally they are wave curves, such as, for example, $A \cos 2\pi \frac{x}{\lambda} + B \sin 2\pi \frac{x}{\lambda}$, where A, B and λ are constants. This expression is the measure of the probability that a particle is at a point x in a potential box within a certain critical boundary. Outside this boundary the formula becomes $Ce^{-kx} + De^{kx}$, where C, D and k are constants. At the critical boundary the curves have to fit on to each other. The latter curve is just two Gompertz mortality curves (when x is positive, D is zero; when x is negative, C is zero) back to back with the above curve in the middle. The curves are 'wave' curves. Hence the name 'Wave Mechanics'. But the waves are not waves of anything material. They are waves of probability.

In order to make use of familiar terms, I have called the curves Gompertz curves, but of course the physicists did not borrow these curves from actuarial science. Neither do they base their use of them entirely on experience. 'Of course', Mr Gurney says, 'we intend to predict the patterns from pure theory, not to obtain them from measurements.' The method, therefore, is not empirical. The pure deductive approach is, I think, best found in Sir A. Eddington's *Relativity Theory of Protons and Electrons*. What the physicists have done for physical statistics we ought to be able to do for vital statistics.

3. The quantum physicist does not deal with absolute frequencies but with the relative frequencies obtained by dividing by the total number of observations. Thus the area of any strip between the curve and the base becomes the probability that the result of any particular observation has a value between certain limits. The curve is then said to be 'normalized'.*

Suppose the thing observed is a function of several variables,

* Nothing to do with the 'normal' curve of error.

say three, x, y, z, the measure of the probability that it is at the point x, y, z is $\psi(x, y, z) = a_1\psi_1 + a_2\psi_2 + a_3\psi_3 + a_4\psi_4 + \ldots$, where a_1, a_2, a_3, a_4, etc., are the probabilities that the observable is in a certain state and the ψ's are the states or frequency distributions. Thus the probability required is an analysed composite like a Fourier Analysis.

Eddington expresses it as follows:* 'The method of wave mechanics is to analyse the whole probability of the system, which must be unity, into the probabilities p_a, p_b, p_c, ... of a set of elementary states a, b, c, ... Then if $q_a(x_\mu, t)$ is the probability of the configuration x_μ at time t in the state a, the whole probability of the configuration x_μ at time t is

$$p_a q_a(x_\mu, t) + p_b q_b(x_\mu, t) + p_c q_c(x_\mu, t) + \ldots .\text{'}$$

Stated in this way the method may be thought to be, one might say, may be accused of being, inverse probability. In a sense, all probabilities obtained by induction from observed facts are inverse probabilities.† In this sense, which it shares with all other methods of obtaining relative frequencies from statistics, it is inverse probability. But it does not pretend to discover causes. It sets out to describe what is observed and to analyse it, and claims no more knowledge than what is obtained by observation and thought. The probabilities p_a, p_b, p_c, ..., are not fixed but can vary according to our state of knowledge. They are described as a 'probability fluid'. When all the probability flows into one term, so that that term becomes 1 and the rest 0, the observation relates to a thing that has been observed in the state whose probability has become 1. When the observation is over, the thing may pass into another state, and there are new values of the p's (or a's) representing the probabilities that the observable is in one of the different possible states. This avoids the assumption, which did so much to bring inverse probability into disrepute, that the probabilities of the different states are all equal.

4. At this point it is necessary to mention a curious feature of the theory, because it is a prominent feature, yet one which I have felt compelled to discard in applying the ideas to our own statistics.

* *Relativity Theory of Protons and Electrons*, p. 115.

† Cf. *The relation between probability and statistics*, Dr W. F. Sheppard, *T.F.A.* Vol. XII, p. 38.

The probability functions of the different values of the thing observed (the ψ's) and the probability functions that the observation relates to the respective states (the a's) are not the probabilities themselves but their square roots or moduli.

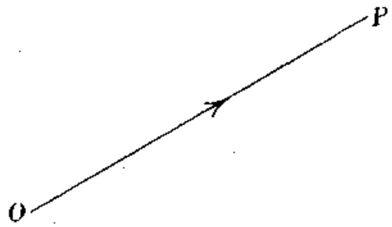
This is because physicists deal with motion in space, so that the conception of distance enters into their observations, while they use vectors for expressing their formulae.

Vectors do not involve any assumption as to how a distance is measured. The metric is introduced later as a separate step. Using them may be compared in some ways to working in decimals instead of in £ *s. d.* or in rupees, annas, pies.

It is impossible to follow the mathematics of quantum mechanics without having at least an elementary knowledge of vector analysis. I have therefore devoted the next part to an excursus on this subject, which can be omitted by those already familiar with the theory of vectors. I am not, of course, professing to supply a text-book on vector analysis, and I must refer those who require formal proofs of the theorems and a complete account of the subject to the standard works on it.

PART 2. VECTOR ANALYSIS

5. A vector, geometrically, is a directed line. It may be regarded as an instruction to perform an operation, viz. the vector OP is



an instruction to go from O to P. (The name is derived from the Latin—*veho*, I carry.)

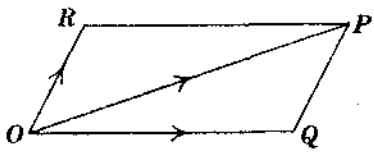
A vector has length, but that only represents a comparative measure of quantity and does not imply a measure of distance. One could, for example, represent a number of balls of different colours by a vector.* Choose three unit vectors at right angles, and call

* Cf. *Space, Time, Matter*, Hermann Weyl, p. 23.

them R, B, Y, and let them represent the three colours red, blue, yellow. Then $3R+5B+2Y$ will represent 3 red, 5 blue and 2 yellow balls. All different combinations of balls of those colours would be then represented by lines radiating from the origin. A rotation of the vector to all possible positions, keeping its length constant, would run through all reallotments of the same number of balls, 10, to the different colours. Only those positions would be possible that gave an integral number of balls of each colour.

The principal characteristic of vectors, as will become plain when one considers their multiplication, is their direction. A vector is independent of position, the vector not being localized in any definite line. Like vectors are vectors with the same direction.

Vectors are accordingly *compounded* or *added* by the parallelogram rule. If the vectors OR and OQ are added their sum is OP.



For if one is told to go from O to R and then go in the direction of OQ as far as is equal to OQ one reaches P. It is plain from this that OP, OQ and OR are not measures of length in the ordinary sense which is based on the assumption that distances are compounded by the rule given by the Theorem of Pythagoras (better known to old-stagers as Euclid, Bk. 1, Prop. 47).

6. Let i, j be unit vectors along OQ and OR respectively. (A unit vector is a vector whose square is 1, or -1, as may be defined.)

Then OQ can be expressed as xi and OR as yj, where the coefficients x and y are the lengths of OQ and OR measured in those directions and are called the co-ordinates.

Then
$$OP=xi+yj.$$

This expression retains the same form whatever the directions of OQ and OR, which we may now regard as axes along which the vector OP is resolved, and whether the angle between OQ and OR is a right angle or not, i.e. whether the axes are orthogonal or not, and it can obviously be extended to any number of dimensions. Thus the expression of a vector in terms of co-ordinates is extremely simple, being just a linear equation of as many terms as there are dimensions.

The most important case of resolutions of vectors is that in which the axes are orthogonal.

If α and β are the angles OP makes with OQ and OR respectively, and r is the length of OP, called the scalar part of the vector, and $\alpha+\beta=\frac{1}{2}\pi$,

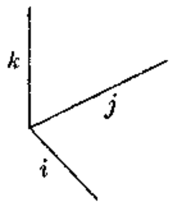
$$x=r\cos\alpha;\quad y=r\cos\beta=r\sin\alpha,$$

and

$$\text{OP}=r\cos\alpha . i+r\sin\alpha . j,$$

and

$$x^2+y^2=r^2.$$



7. Hamilton's idea in introducing vectors was to find a method of multiplication that was not continued addition. His quaternions were based on orthogonal unit vectors i, j, k defined so that

$$i^2=j^2=k^2=-1, \tag{i}$$

$$ij=k;\quad jk=i;\quad ki=j. \tag{ii}$$

Thus

$$ji=-k \quad\text{and}\quad ij+ji=0.$$

These vectors, therefore, anti-commute, that is, the sign is changed if the order of multiplication is reversed.

In the elementary vector theory now in use, which departed from Hamilton's quaternions, much to Tait's indignation (vide introduction to *Elementary Vector Analysis*, by Weatherburn), the unit vectors are defined by the relations

$$i^2=j^2=k^2=1, \tag{i}$$

$$i . j=j . k=k . i=0, \tag{ii}$$

$$i\times j=k;\quad j\times k=i;\quad k\times i=j, \tag{iii}$$

the product $i . j$ being called the scalar or dot product and $i\times j$ the vector or cross product. (This notation for products is not universal; on the contrary, there are almost as many notations as writers.)

Eddington when writing on quantum mechanics began apparently by adopting the relations (for four dimensions)

$$\text{E}_\mu^2=1\quad(\mu=1, 2, 3, 4)$$

corresponding to the $i^2=j^2=k^2=1$, but later changed to

$$\text{E}_\mu^2=-1;\quad \text{E}_\mu\text{E}_\nu=-\text{E}_\nu\text{E}_\mu\quad(\mu, \nu=1, 2, 3, 4)$$

(*Relativity Theory of Protons and Electrons*, Introduction and p. 21).

He thus appears to have returned to Hamilton's original definitions, which seem better suited to the representation of physical phenomena from a theoretical point of view.

I adopt Hamilton's definitions.

8. Let us consider the result of multiplying the vector

$$\mathrm{OP} = r\cos\alpha . i + r\sin\alpha . j$$

by $e^{k\gamma}$, where k is another square root of -1 and $e^{k\gamma}$ as in De Moivre's theorem is defined by the exponential series, so that

$$e^{k\gamma} = \cos\gamma + k\sin\gamma. \tag{8·1}$$

Thus

$$e^{k\gamma}(r\cos\alpha . i + r\sin\alpha . j)$$

$$= r\cos\gamma\cos\alpha . i + r\sin\gamma\sin\alpha . kj + r\sin\alpha\cos\gamma . j + r\sin\gamma\cos\alpha . ki$$

$$= r\cos\gamma\cos\alpha . i - r\sin\gamma\sin\alpha . i + r\sin\alpha\cos\gamma . j + r\sin\gamma\cos\alpha . j$$

$$= r\cos(\alpha+\gamma) . i + r\sin(\alpha+\gamma) . j. \tag{8·2}$$

Thus the vector OP has been rotated through an angle γ.

It will be observed that in order to rotate a vector in the plane i, j we have to call in the aid of a vector orthogonal to both of them.

This can be understood by transferring our ideas to one of the familiar 'balls in a bag' problems of probability.

Suppose we have 10 white and 5 black balls in a bag. On the usual assumptions the probability of drawing a white ball is $\frac{2}{3}$.

Now we can represent the constitution of the bag as $\frac{2}{3}\mathrm{W} + \frac{1}{3}\mathrm{B}$, where W and B stand for orthogonal co-ordinates, whiteness and blackness, two things which have each of them no element of the other, which is the fundamental idea of orthogonality, in space or elsewhere.

We can, provided we do not disturb the probability of $\frac{2}{3}$ for the whole bag, regard the balls as divided into two lots and the probability of drawing a white ball from each lot as compounded of a probability of selecting that lot and then drawing a white ball from it, e.g. such an arrangement might be 6W 1B in one lot, and 4W 4B in the other. Then, provided the probabilities of selecting the lots were $\frac{7}{15}$ and $\frac{8}{15}$ respectively, the probability of drawing a white ball would still be $\frac{2}{3}$, being

$$\tfrac{7}{15}\times\tfrac{6}{7}+\tfrac{8}{15}\times\tfrac{1}{2}=\tfrac{2}{3}.$$

All such arrangements are permissible, and correspond to choosing co-ordinates that are not orthogonal and do not represent pure states. Changes from one to another are relativistic changes. The one arrangement that has two pure states is the arrangement $10W+5B$, or, when normalized, $\frac{2}{3}W+\frac{1}{3}B$. An arrangement having one pure state would be, say, $4W+(6W+5B)$ or $\frac{4}{15}W+\frac{11}{15}M$ where M stands for mixed, giving a probability of $\frac{4}{15}+\frac{11}{15}\times\frac{6}{11}=\frac{2}{3}$.

Now how can this probability of $\frac{2}{3}$ be altered? Not by any conceivable rearrangement of the mental division into two groups *within the bag*, i.e. not by any relativistic change, which is only an alteration of description, but only by changing the colour of some of the balls.

If we divide the balls mentally into the pure states, i.e. resolve them orthogonally, the probability that, after one or other of these states has been picked at random, a white ball will be drawn will be 1 or 0.

Let us regard the probability, then, as made up of the probabilities of performing the operations of selecting a state from the pure states, drawing a ball from it and replacing the ball. On two-thirds of the occasions we draw from the white state. If now we draw a ball, find it to be white and instead of replacing it, substitute a black ball, we perform an operation of the same kind and representable by a co-ordinate but orthogonal to the others, and we bring about a change in the direction of the vector, i.e. we rotate it.

9. Let us continue consideration of the process of keeping the vector fixed and changing the co-ordinates, i.e. leaving the probability of $\frac{2}{3}$ unaltered and rearranging the distribution of balls into two lots.

If we actually divide the balls into pure white and pure black, putting them in separate bags, we have $\frac{2}{3}W+\frac{1}{3}B$.

If now we draw a ball from the 'white' bag and put it in with the black we do not alter the original probability of $\frac{2}{3}$ nor do we alter the co-ordinate W, but we alter the proportions $\frac{2}{3}$ and $\frac{1}{3}$ and we alter the co-ordinate B. Thus we rotate one co-ordinate. As we continue the process we gradually rotate the co-ordinate B until it coincides with the vector and the co-efficient of W disappears and we are left with the balls all mixed together again. In rotating from the position B to the vector, the co-ordinate rotated passes through in reverse order all the positions that the vector would occupy if it were rotated from its original position to B.

Transference of balls then may be regarded as causing a rotation in a vector co-ordinate representing either a mixed state of different coloured balls, or a probability of drawing a ball of a certain colour. Such a representation is imaginary, the angles are imaginary, and the rotational force is a real transference operating in time and producing the real function $e^{-\delta t}$ (for a uniform rate) instead of the imaginary function $e^{k\gamma}$. Correspondingly, in real space, time enters as an imaginary co-ordinate, it, in the theory of relativity.

10. Returning to real vectors, with imaginary functions, suppose we rotate the vector through an angle π. We get

$$e^{k\pi} = \cos\pi + k\sin\pi = -1. \qquad (10\cdot1)$$

For the vector we get

$$\begin{aligned} & r\cos(\alpha+\pi).i + r\sin(\alpha+\pi).j \\ & \qquad = -r\cos\alpha.i - r\sin\alpha.j. \end{aligned} \qquad (10\cdot2)$$

The vector is reversed in direction. If now we require the *distance* to the end of the vector to be the same in both directions we must square the vector, obtaining

$$\begin{aligned} (r\cos\alpha.i + r\sin\alpha.j)^2 &= (-r\cos\alpha.i - r\sin\alpha.j)^2 \\ &= r^2(-\cos^2\alpha - \sin^2\alpha) \\ &= -r^2, \end{aligned} \qquad (10\cdot3)$$

in both cases.

That is, the measure of distance, to be the same in both directions and make space isotropic, is the square of the vector.* Hence the Theorem of Pythagoras. Our ordinary convention as to distance, which takes the positive square root of the square of the modulus of the vector, was no doubt adopted as the result of assuming that lengths in the same straight line would be added. It is now so ingrained in us that it is difficult to realize that it is a convention.†

* Cf. the fact that the index in the normal curve of error is $-x^2/c^2$ implying the assumption that positive and negative errors are equally likely.

† 'No deduction of a really geometrical kind can be legitimately based on statements of which any particular conception of distance forms a part; such statements are equivalent only to statements in regard to the behaviour of particular measuring instruments, which must rest on physical hypotheses' (H.F. Baker, *Principles of Geometry*, Vol. II, p. 186).

It is therefore understandable that the probabilities of the location of an object in space must be compounded with the squares of the vectors and when the vectors themselves are used the square roots of the probabilities are the coefficients of the vectors.

When we deal, however, with probabilities which *are* vectors, the coefficients representing the probabilities of the existence of different states are themselves the actual probabilities and not their square roots.

11. Since $j=k.i$, the vectors

$$r\cos\alpha.i+r\sin\alpha.j \quad \text{and} \quad r\cos\beta.i+r\sin\beta.j$$

may be written $re^{k\alpha}.i$ and $re^{k\beta}.i.$

In view of what has been said above as to rotation it will be seen that this form for the vector expresses it as a rotation of the fundamental vector i.

The i must be written after the exponential because i, j, k do not commute. It is, therefore, necessary to have a convention as to the meaning of 'multiplying one vector by another'. Since we are thinking of operations, the convention is that the multiplier comes first and the multiplicand second. Thus in the above we multiply i by $e^{k\alpha}$ and by $e^{k\beta}$. The symbol r is not a vector but a scalar, or constant, and commutes with all other symbols.

Let us now multiply the two vectors together, multiplying $re^{k\alpha}.i$ by $re^{k\beta}.i$. We have

$$r^2.e^{k\beta}.i.e^{k\alpha}.i.$$

Now

$$\begin{aligned} ie^{k\alpha} &= i(\cos\alpha+k\sin\alpha) \\ &= i\cos\alpha+ik\sin\alpha \\ &= i\cos\alpha-ki\sin\alpha. \end{aligned}$$

Multiply i by this,

$$\begin{aligned} i.e^{k\alpha}.i &= (i\cos\alpha-ki\sin\alpha)\,i \\ &= i^2\cos\alpha-k.i.i\sin\alpha \\ &= -\cos\alpha+k\sin\alpha \\ &= -e^{-k\alpha}. \end{aligned}$$

So that the product, in the above order, is

$$-r^2e^{k(\beta-\alpha)}. \qquad (11\cdot1)$$

If, therefore, $\alpha=\beta$, or the difference between them is a multiple of π, the square root of the square of the modulus is unaltered and the distance measured in any direction is the same. For a difference of $-\frac{1}{2}\pi$ we get $-r^2[\cos(-\frac{1}{2}\pi)+k\sin(-\frac{1}{2}\pi)]$. The first part of this is 0 and is the dot product of the ordinary theory. The second part is the cross product and is equal to r^2k. The condition for orthogonality of the vectors is that the dot product $=0$. When $\alpha=\beta$ or the difference is a multiple of π, the cross product is 0. This is the condition for parallelism.

Ordinary vector theory takes the two parts of the product separately and changes the sign of the square of the vector.

This explains the signs in the formula for an interval in a Euclidean continuum in Einstein's theory,

$$ds^2=dt^2-dx_1^2-dx_2^2-dx_3^2. \tag{11·2}$$

12. Let us deal a little more formally with the process of changing the co-ordinate system.

An n-dimensional manifold requires n independent vectors to represent it; $n+1$ vectors would be linearly dependent. Thus, in the case of a surface we require two fundamental vectors, say e_1 and e_2, and choosing any point O as origin a vector OP in the surface will be expressed by the equation

$$\mathrm{OP}=a_{11}e_1+a_{12}e_2. \tag{12·1}$$

The double suffixes in the coefficients a_{11}, a_{12} are not necessary but they are convenient as will appear later.

The same vector OP can be represented by a linear equation of two other fundamental vectors, thus

$$\mathrm{OP}=a'_{11}e'_1+a'_{12}e'_2. \tag{12·2}$$

These fundamental vectors e'_1 and e'_2 being themselves vectors in the continuum are linear functions of the old fundamental vectors e_1, e_2:

$$\left.\begin{aligned}e'_1&=b_{11}e_1+b_{12}e_2,\\ e'_2&=b_{21}e_1+b_{22}e_2.\end{aligned}\right\} \tag{12·3}$$

Substituting these values for e'_1 and e'_2 in (12·2) we have

$$\mathrm{OP}=(a'_{11}b_{11}+a'_{12}b_{21})\,e_1+(a'_{11}b_{12}+a'_{12}b_{22})\,e_2. \tag{12·4}$$

Hence, equating coefficients of e_1 and e_2,

$$\left.\begin{aligned}a_{11}&=a'_{11}b_{11}+a'_{12}b_{21},\\ a_{12}&=a'_{11}b_{12}+a'_{12}b_{22}.\end{aligned}\right\} \tag{12·5}$$

These relations are independent of the fundamental vectors. They are written in the form

$$a'_{11} \quad a'_{12} \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} = a_{11} \quad a_{12}. \tag{12·6}$$

The expression between the lines is called a matrix. It is written in the form of a determinant but it has no 'value', though one may have the 'determinant of the matrix', when the value is obtained in the usual way.

13. Matrices, when the elements b_{nm} are given numerical values, form a higher order of numbers. The matrix corresponding to the number 1 is the matrix that has

$$\left. \begin{aligned} b_{nm} &= 1, \quad \text{when} \quad n = m. \\ &= 0, \quad \text{when} \quad n \neq m. \end{aligned} \right\} \tag{13·1}$$

There is an indefinite number of matrices corresponding to the number 0. Such a matrix is described as 'singular'. A singular matrix has no reciprocal. The test for singularity is that the determinant of the matrix is 0.

It will be seen from the way in which a matrix was obtained above that the rule for multiplication of a vector by a matrix is 'row by column'. The same rule applies when two matrices are multiplied together. Thus if c_{rs} are the elements of the product matrix and a_{rs}, b_{rs} the elements of the matrices that are being multiplied together, the rule is that

$$c_{rs} = a_{r1} b_{1s} + a_{r2} b_{2s} + a_{r3} b_{3s} + \dots. \tag{13·2}$$

The order of multiplication of matrices is important because matrices are not necessarily commutative. In other respects they obey the laws of ordinary algebra, viz. association and distribution. To add matrices together one adds the elements in corresponding positions:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} = \begin{vmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{vmatrix}. \tag{13·3}$$

14. Suppose we have the vector $10W + 5B$ representing 10 white and 5 black balls. Apply the matrix $\begin{vmatrix} \frac{2}{5} & \frac{3}{5} \\ 0 & 1 \end{vmatrix}$ which will change the co-ordinate system.

For the first co-ordinate we get $10 \times \frac{2}{5} + 5 \times 0 = 4$. For the second we get $10 \times \frac{3}{5} + 5 \times 1 = 11$. This alters the arrangement to 4 white balls and 11 mixed, or when normalized so as to give a probability instead of an absolute measurement, $\frac{4}{15}$ and $\frac{11}{15}$.

This matrix could be expressed as the sum of the two matrices

$$\begin{vmatrix} \frac{2}{5} & 0 \\ 0 & \frac{2}{5} \end{vmatrix} \quad \text{and} \quad \begin{vmatrix} 0 & \frac{3}{5} \\ 0 & \frac{3}{5} \end{vmatrix}$$

which may be written with the numerical multipliers outside, thus:

$$\frac{2}{5}\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \quad \text{and} \quad \frac{3}{5}\begin{vmatrix} 0 & 1 \\ 0 & 1 \end{vmatrix}.$$

Applying each of these separately, we get for the co-ordinates 4 and 2, and 0 and 9, which by addition of vectors gives 4 and 11 for the co-ordinates of the compounded vector. It will be noticed that the former of these two matrices is $\frac{2}{5}$ of the 'unit' matrix or matrix corresponding to the number 1 and that it produces the same probability for drawing a white ball, viz. $\frac{2}{3}$, as the original 10 white and 5 black collection, while the latter matrix is a singular matrix (determinant $= 0$), and it produces a 'pure' black selection of balls, namely 9 black.

Since the effect of applying a matrix to the vector is to change the co-ordinate system, it is clear that applying another matrix to the result changes the co-ordinate system again and that, correspondingly, the matrix that is the product of two or more matrices changes the co-ordinate system to the system that would be obtained by applying the latter matrices successively in the same order. The order is material because matrices do not necessarily commute.

15. The matrix obtained above 'followed' the vector and would be called a post-factor. In a similar manner one can obtain a matrix for multiplying a following vector in order to change the co-ordinate system, and the matrix would then be a pre-factor. Thus:

Let the vector OP be $e_1 a_{11} + e_2 a_{21}$ in one co-ordinate system and $e'_1 a'_{11} + e'_2 a'_{21}$ in another.

Let
$$\left.\begin{aligned} e_1 &= e'_1 b_{11} + e'_2 b_{21}, \\ e_2 &= e'_1 b_{12} + e'_2 b_{22}. \end{aligned}\right\} \qquad (15 \cdot 1)$$

Then, substituting for e_1 and e_2 and equating coefficients of e'_1 and e'_2,

$$b_{11}a_{11}+b_{12}a_{21}=a'_{11},$$
$$b_{21}a_{11}+b_{22}a_{21}=a'_{21},$$

written in the form

$$\begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} \begin{matrix} a_{11} \\ a_{21} \end{matrix} = \begin{matrix} a'_{11} \\ a'_{21} \end{matrix} \qquad (15\cdot2)$$

The vector is now written downwards instead of across because of the rule of multiplication, row by column, and the vector now constitutes a single-column matrix instead of a single-row matrix. The suffixes indicate this. This is one of the reasons why I have used double suffixes for the vector.

Vectors have been described, as above, as single-row or single-column matrices. It would be equally correct to say that an n-fold matrix is a set of n vectors. This is easy to see if one works in polar co-ordinates.

Omitting the scalar part of a vector, which is irrelevant, the vector is, say, $\cos\theta . e_1 + \sin\theta . e_2$. Similarly, the new co-ordinate system expressed in terms of the old is

$$e'_1 = \cos\phi . e_1 + \sin\phi . e_2,$$

$$e'_2 = \cos\psi . e_1 + \sin\psi . e_2.$$

The $\cos\theta$ and $\sin\theta$, then, are the a's of the vector and the $\cos\phi$, $\sin\phi$, $\cos\psi$, $\sin\psi$, the b's of the matrix, showing that the a's and b's are essentially of the same kind. Double suffixes are essential for the elements of a matrix, which is a kind of double entry table, and hence it is desirable to use double suffixes for the co-ordinates of a vector when dealing with changes of co-ordinate systems.

PART 3. QUANTUM MECHANICS

16. We are rather inclined, I think, to talk of selection as if it applied only to lives entering into assurance or buying annuities. It was only tentatively, and, as it were, apologetically, that the idea of class-selection was put forward a few years ago.

But statistics are one form of selection after another. We select when we decide to form one mortality table for men and another for women. We select when we make restrictions of the data with regard to race, with regard to climate, with regard to occupation, with regard to kind of policy. When we classify the lives by age we select.

The physicist does the same; he selects his thing to be observed, and he selects the conditions under which he observes it.

What is the object of all this selection? It is to obtain homogeneity, or purity, because by reaching purity one reaches constancy. A mixture would produce differences in results, when the observation was repeated, that arose merely from variations in the proportions of the pure ingredients.

The idea of Quantum Analysis then is to express the observations in terms of pure or select states. Such states will be represented by orthogonal co-ordinates, because it is only lines that are at right angles that contain no element of each other's direction. Intermediate states or mixtures, if they exist, will be represented by lines not at right angles, that is, by rotation of one of the vectors.

The method therefore proceeds to study changes in co-ordinate systems with the aim of finding ways of selecting orthogonal co-ordinates that will represent select states and rotated vectors that will represent non-select states.

The operation of changing the co-ordinate system has been explained in Part 2. It leads to a double entry table, like a correlation table, called a matrix.

The matrix is at the same time a 'Probability Operator' and an 'Observable', because by changing the co-ordinate system it alters the probabilities that the thing observed is in different states, and in doing so it represents the thing one is observing, which produces different observational results of an experiment according to the state it is in.

Real matrices are time-like: imaginary matrices space-like.*

The sum total of the probabilities of the different states must be unity, because the thing observed must be in some one of the possible states. That is, the formulae must satisfy the principle of the 'Conservation of Probability'.

In order to find pure states, the matrix must be split up. There is no sealed pattern for doing this; it depends on the nature of the problem. But there are certain conditions that the parts, or

* *Relativity Theory of Protons and Electrons*, Eddington, p. 96.

selective operators, must satisfy if they are to produce pure states.

(1) They must satisfy the principle of the conservation of probability. Thus they are 'exhaustive'.

(2) They must when repeated give the same result, representing the fact that when you select twice according to the same rule, you ought to have the same result, and that if you apply parts of a matrix to make your co-ordinates orthogonal, you cannot make them more orthogonal by performing the operation again on your already orthogonal co-ordinates. This characteristic is called 'idempotency', and the selective operators are 'idempotent'.

(3) The product of any two of them must be 0, representing the fact that if you select for a certain characteristic and then apply selection again for an inconsistent characteristic you get nothing, and the fact that you can find no direction that goes simultaneously entirely in two orthogonal directions. Thus the operators are 'orthogonal'.

A set of operators which satisfy these conditions is called a spectral set.

The pure states that are studied under any one system of these functions are mutually exclusive. States which can exist simultaneously would be represented by mutually commuting observables or matrices.

A general state will then be represented by a matrix that gives a rotation between the pure states. This state will be composed of proportions of the pure states between which it lies. Thus any mixed state is regarded as obtained by the superposition of pure states. Conversely it is possible to regard a pure state as obtainable by superposition of mixed states. The essence of it is that there is a number, n, of distinct directions, and any direction can be regarded as composed of proportions of n arbitrarily chosen co-ordinate axes. This is called the Principle of Superposition.

17. When the result of a particular observation made on a system in a particular state is with certainty one particular number, a say (instead of being one of two or more numbers according to a probability law), the operator or matrix representing the observable, say α, and the vector representing the state, say ϕ_α, are connected by the equation $\phi_\alpha \alpha = \phi_\alpha a$. In this equation a is called an eigen-value or expectation value and ϕ an eigen-ϕ.

The literature of quantum mechanics is so spattered with eigen-values and eigen-ϕ's that the most elementary description of the theory can hardly avoid mentioning them. I have, however, in the comparatively simple application of the ideas of the theory to the mortality table, avoided the use of these expressions as unnecessary. It may be mentioned, though, that each of the elements of a matrix is an eigen-value. A life under observation will be represented by the matrix, and if we examine him and find that he is a first-class life, the probability that he is a first-class life becomes 1 and the force of mortality that relates to him becomes $\mu_{[x]}$.

18. In giving this brief account of quantum mechanics, I have not dealt with the principle of indeterminacy, nor differential operators and many other aspects of the theory, because I could not in the space of a paper give any more than a description of such parts as I at present want to use for the application of the method to the mortality table.

I must also make it clear that I have described what the theory means to me and that I stand to be corrected by those who have produced it. The authors sometimes do not condescend low enough to explain to this obtuse person the meaning of what they are doing. In fact, they make such remarks as 'the main object of physical science is not the provision of pictures, but is the formulation of laws governing phenomena and the application of those laws to the discovery of new phenomena. If a picture exists, so much the better; but whether a picture exists or not is a matter of only secondary importance.'*

As a picture of the philosophical outlook of the theory I think that a parable by Eddington is unequalled.† It is rather long to quote in full; briefly it is as follows:

'An ichthyologist is exploring the ocean. He casts a net into the water and brings up a fishy assortment. Surveying his catch, he proceeds to systematize what it reveals. He arrives at two generalizations:

'(1) No sea-creature is less than two inches long.

'(2) All sea-creatures have gills.

'They are both true of his catch and he assumes tentatively that they will remain true however often he repeats it.

* *Quantum Mechanics*, P.A.M. Dirac, p. 6.

† *The Philosophy of Physical Science*, Eddington, p. 26.

'An onlooker objects that the first generalization is wrong. There are plenty of sea-creatures under two inches long, only your net is not adapted to catch them.

"Anything uncatchable by my net", says the ichthyologist, "is ipso facto outside the scope of ichthyological knowledge, and is not part of the kingdom of fishes which has been defined as the theme of ichthyological knowledge. In short, what my net can't catch isn't fish."'

Another onlooker makes a different suggestion:

'I realize that you are right in refusing our friend's hypothesis of uncatchable fish, which cannot be verified by any tests you and I would consider valid. By keeping to your own method of study, you have reached a generalization of the highest importance—to fishmongers, who will not be interested in generalizations about uncatchable fish. Since these generalizations are so important, I would like to help you. You arrived at your generalization in the traditional way by examining the fish. May I point out that you could have arrived more easily at the same generalization by examining the net and the method of using it?'

Our methods of collecting the data for mortality tables and tabulating them are an actuarial net. The general shape of a mortality table is due to the nature of those methods and the way we use them.

PART 4. APPLICATION TO THE MORTALITY TABLE

19. Science proceeds by three stages:

(1) Selection for the purpose of observation.

(2) Comparison for the purpose of classification.

(3) Measurement or comparison in respect of the extent of the possession of homogeneous characteristics.

The first requires 'an' object.

The second requires 'two' objects.

The third requires 'four' objects, two to give a standard of measurement, and two to give something to be measured.

We may therefore expect the 'law of mortality' to require four values of μ to express it.

20. In the case of mortality statistics most of the variables for which we 'select' data are discontinuous. Sex, for example, and

class of policy and to a large extent, race. Unless there is continuous variation which can be linked with a variation of time or place, it is necessary to form separate mortality tables. If it were possible to find data in sufficient quantity to form tables showing all stages of mixture of black and white so as to vary continuously for colour we could form a three-dimensional mortality table showing mortality according to age, colour, and time since selection. In practice we are restricted to two continuous variables. Hence the mortality table, in so far as it is homogeneous, must be representable in two dimensions and by two pure states, and in so far as such a representation departs from the facts, there is evidence of lack of homogeneity. The legendary professor who, when told that his theory did not agree with the facts, replied 'So much the worse for the facts' was quite right!

21. As shown in my paper on 'Select mortality tables' we can split l_{x+t} lives into two parts and express μ in the form

$$\mu_{x+t} = \frac{l^{(1)}_{x+t}}{l_{x+t}} \mu^{(1)}_{x+t} + \frac{l^{(2)}_{x+t}}{l_{x+t}} \mu^{(2)}_{x+t}. \tag{21·1}$$

We can associate one of these forces, say $\mu^{(1)}$, with time since selection, and the other with age, but since both are measured by time we can divide through by dx or dt and write

$$\mu = a_{11}\mu_1 + a_{12}\mu_2. \tag{21·2}$$

(The μ_2 in this formula is what I called $\mu^{(3)}$ in the previous paper. I have altered its symbol in order to conform to the notation of quantum mechanics and for the sake of symmetry.)

We can divide the l_{x+t} lives in some other way, $l'^{(1)}$ and $l'^{(2)}$. It is evident that the $l'^{(1)}$ and $l'^{(2)}$ are then related to the $l^{(1)}$ and $l^{(2)}$ by a linear relationship of the form

$$\left.\begin{aligned} l'^{(1)} &= b_{11} l^{(1)} + b_{12} l^{(2)}, \\ l'^{(2)} &= b_{21} l^{(1)} + b_{22} l^{(2)}. \end{aligned}\right\} \tag{21·3}$$

Hence the μ''s are similarly related to the undashed μ's.

The new values of μ_1 and μ_2 are related then to the old values in the same way as vectors, and since μ obeys the other laws of algebra, association and distribution, as do vectors, μ is a vector. Such differences in treatment as arise between the application of quantum mechanics to the force of mortality and to physical phenomena such as atoms, electrons and so on, are due to the fact that we are dealing with a continuum that is two-dimensional in time, with no spatial dimension, whereas the continuum dealt with in physics is four-dimensional, 3+1, three spatial and one temporal. Our problem is like that of the balls of two colours in Part 3. Those of physics relate principally to distances, and related measurements such as velocities.

22. Consider the following matrices:

$$\underset{A}{\begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix}} \quad \underset{B}{\begin{vmatrix} 0 & -1 \\ 0 & 1 \end{vmatrix}} \quad \underset{I}{\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}} \quad \underset{C}{\begin{vmatrix} 1 & -1 \\ 0 & 0 \end{vmatrix}} \quad \underset{D}{\begin{vmatrix} 0 & 1 \\ 0 & 1 \end{vmatrix}} \qquad (22{\cdot}1)$$

These matrices satisfy the following conditions:

$$A+B=I, \quad C+D=I, \qquad (a)$$

$$A^2=A;\ B^2=B, \quad C^2=C;\ D^2=D, \qquad (b)$$

$$AB=0;\ BA=0, \quad CD=0;\ DC=0. \qquad (c)$$

I is the unit matrix.

Thus the matrices A, B, and C, D, satisfy respectively the conditions necessary for a spectral set, viz.: they are exhaustive by (*a*), idempotent by (*b*) and orthogonal by (*c*). They therefore form two spectral sets.

Applying A as a pre-factor to the vector $a_{11}\mu_1 + a_{12}\mu_2$ in which, by the conservation of probability, $a_{11}+a_{12}=1$, so that $a_{12}=1-a_{11}$, we have

$$\begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix} \begin{vmatrix} a_{11} \\ 1-a_{11} \end{vmatrix} = \begin{matrix} 1 \\ 0 \end{matrix} \qquad (22{\cdot}2)$$

Thus the vector becomes μ_1.

Applying B to the vector as a pre-factor, we have

$$\begin{vmatrix} 0 & -1 \\ 0 & 1 \end{vmatrix} \begin{vmatrix} a_{11} \\ 1-a_{11} \end{vmatrix} = \begin{matrix} -(1-a_{11}) \\ (1-a_{11}) \end{matrix} \qquad (22{\cdot}3)$$

Thus the vector becomes

$$(1-a_{11})(\mu_2-\mu_1).$$

In the formula $\mu=a_{11}\mu_1+(1-a_{11})\mu_2$, $1-a_{11}$ is identically $(\mu-\mu_1)/(\mu_2-\mu_1)$, so that the result is to produce $\mu-\mu_1$, which is what I called $\mu^{(2)}$ in my paper on Select Tables. Thus the operator B has an effect corresponding to the 'force of transfer' conceived as operating as an independent force on the select lives.

Let us obtain a mixed state by taking a part only of B with A, say $A+k_1 B$, where k_1 is less than 1. We then have

$$\begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix} + k_1 \begin{vmatrix} 0 & -1 \\ 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 1-k_1 \\ 0 & k_1 \end{vmatrix}. \qquad (22{\cdot}4)$$

Applying this to the vector, we have:

$$\begin{vmatrix} 1 & 1-k_1 \\ 0 & k_1 \end{vmatrix} \begin{matrix} a_{11} \\ 1-a_{11} \end{matrix} = \begin{matrix} a_{11}+(1-k_1)(1-a_{11}) \\ k_1(1-a_{11}) \end{matrix} \qquad (22{\cdot}5)$$

This gives $[1-k_1\ (1-a_{11})]\ \mu_1+k_1\ (1-a_{11})\ \mu_2$. A repetition of the operations, taking k_2 of B, will give

$$(A+k_2 B)\ (A+k_1 B)$$

which by virtue of the rules above mentioned is equal to $A+k_2 k_1 B$, and generally by performing the operations n times we get

$$A+k_n k_{n-1} k_{n-2} \dots k_1 B, \qquad (22{\cdot}6)$$

where $k_n k_{n-1} k_{n-2} \dots k$ is a continued product of terms each less than 1. In the limit we get the pure state A. Thus we have worked backwards from the mixed state $A+B$ which produces μ to the pure state A which produces $\mu_{[x]}$.

If the k's were all equal we should have $A+k^n B$ giving

$$[1-k^n\ (1-a_{11})]\ \mu_1+k^n\ (1-a_{11})\ \mu_2. \qquad (22{\cdot}7)$$

In this formula n is not continuous, it is a discrete number of operations. The function k^n only has values for integral values of n; in between n and $n+1$ it remains k^n. Thus it applies to such a process as selecting balls of one colour out of a mixture.

By assuming that the operations, though separate and complete, are performed at regular intervals of time, we bridge over the gap between the idea of discrete operations and the idea of a rate over a period of time.

Suppose the operation is performed n times a year. In t years it will be performed nt times and give the result $A+k^{nt}B$. Both k and n are unknown; all that is known is k^n or an empirical constant v. We may therefore make $A+k^{nt}B=A+v^t B$ and, if we are using fairly large numbers, assume that v^t is continuous.

We then get

$$[1-v^t\ (1-a_{11})]\ \mu_1+v^t\ (1-a_{11})\ \mu_2. \qquad (22{\cdot}8)$$

For $t=0$ this reduces to the original vector μ, where $1-a_{11}=\dfrac{\mu-\mu_1}{\mu_2-\mu_1}$. Identically,

$$\mu_x=\left(\frac{\mu_2-\mu_x}{\mu_2-\mu_{[x-t]+t}}\right)\mu_{[x-t]+t}+\left(\frac{\mu_x-\mu_{[x-t]+t}}{\mu_2-\mu_{[x-t]+t}}\right)\mu_2. \qquad (22\cdot 9)$$

Equating coefficients of μ_2,

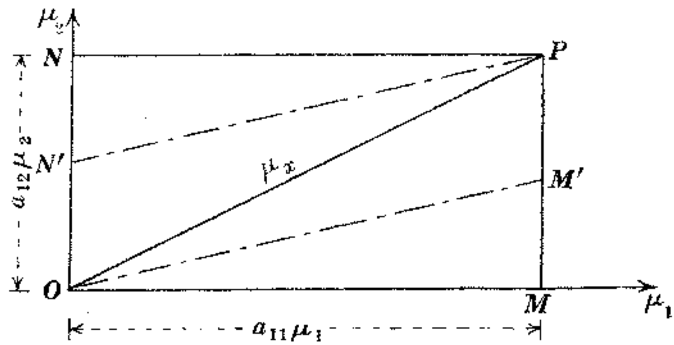
$$\frac{\mu-\mu_1}{\mu_2-\mu_1}\,v^t \quad \text{or} \quad \frac{\mu_x-\mu_{[x]}}{\mu_2-\mu_{[x]}}\,v^t=\frac{\mu_x-\mu_{[x-t]+t}}{\mu_2-\mu_{[x-t]+t}},$$

whence

$$\mu_{[x-t]+t}=\frac{\mu_x-\dfrac{\mu_x-\mu_{[x]}}{\mu_2-\mu_{[x]}}\,.\,v^t\,.\,\mu_2}{1-\dfrac{\mu_x-\mu_{[x]}}{\mu_2-\mu_{[x]}}\,v^t}, \qquad (22\cdot 10)$$

which is formula (3) of my paper on Select Tables, allowing for the renaming of $\mu^{(3)}$.

23. The following is a diagrammatic representation of the process. In the diagram μ_x is represented by the vector OP, which



is resolved along the orthogonal vectors μ_1, μ_2. The resolved parts are $a_{11}\mu_1$ and $a_{12}\mu_2$.

If we reduce the part resolved along μ_2 by multiplying it a number of times by a factor less than 1, N moves back gradually to a new position, say N′, such that $ON'=v^t.ON$.

Then in order to complete the parallelogram, OP being fixed, OM must rotate forwards to OM′, μ_1 becoming $\mu_{[x-t]+t}$, the angle of rotation forwards corresponding to the extent of the movement backwards of N.

The substitution of v^t for the continued product $k_n \ldots k_1$ is equivalent to the assumption that the rotation of μ_1 is uniform in time. Actually, as the paper on Select Tables showed, the rate of rotation should increase slightly with time.

24. Applying D to the vector as a post-factor we have

$$a_{11} \quad 1-a_{11} \begin{vmatrix} 0 & 1 \\ 0 & 1 \end{vmatrix} = 0 \quad 1. \tag{24·1}$$

Thus the vector becomes μ_2, the maximum force of mortality reached at age ω.

Applying C as a post-factor we have

$$a_{11} \quad 1-a_{11} \begin{vmatrix} 1 & -1 \\ 0 & 0 \end{vmatrix} = a_{11} \quad -a_{11}. \tag{24·2}$$

Thus the vector becomes

$$\begin{aligned} &-a_{11}(\mu_2-\mu_1) \\ &= -(\mu_2-\mu). \end{aligned} \tag{24·3}$$

Forming a superimposed state by adding to D a part of C, say $\phi(y)$ C, where y is measured from ω, so that the origin of the ultimate table is placed at the extremity of life and not at birth, we have

$$\begin{vmatrix} 0 & 1 \\ 0 & 1 \end{vmatrix} + \phi(y) \begin{vmatrix} 1 & -1 \\ 0 & 0 \end{vmatrix} = \begin{vmatrix} \phi(y) & 1-\phi(y) \\ 0 & 1 \end{vmatrix} \tag{24·4}$$

Just as the matrix applied previously changed the co-ordinate system by rotating μ_1, and keeping μ_2 fixed, so this matrix rotates μ_2 and keeps μ_1 fixed.

Applying it we have

$$a_{11} \quad 1-a_{11} \begin{vmatrix} \phi(y) & 1-\phi(y) \\ 0 & 1 \end{vmatrix} = \phi(y)\, a_{11} \quad 1-\phi(y)\, a_{11},$$

so that, if μ_α is an arbitrary μ,

$$\mu_\alpha = \left(\frac{\mu_\omega - \mu_\alpha}{\mu_\omega - \mu_{[\alpha]}}\phi(y)\right)\mu_{[\alpha]} + \left(1 - \frac{\mu_\omega - \mu_\alpha}{\mu_\omega - \mu_{[\alpha]}}\phi(y)\right)\mu_{\omega-y}. \tag{24·5}$$

Identically,

$$\mu_\alpha = \left(\frac{\mu_{\omega-y} - \mu_\alpha}{\mu_{\omega-y} - \mu_{[\alpha]}}\right)\mu_{[\alpha]} + \left(\frac{\mu_\alpha - \mu_{[\alpha]}}{\mu_{\omega-y} - \mu_{[\alpha]}}\right)\mu_{\omega-y}. \tag{24·6}$$

Equating coefficients of $\mu_{[\alpha]}$

$$\frac{\mu_{\omega-y}-\mu_\alpha}{\mu_{\omega-y}-\mu_{[\alpha]}}=\frac{\mu_\omega-\mu_\alpha}{\mu_\omega-\mu_{[\alpha]}}.\phi(y),$$

whence
$$\mu_{\omega-y}=\frac{\mu_\alpha-\dfrac{\mu_\omega-\mu_\alpha}{\mu_\omega-\mu_{[\alpha]}}\mu_{[\alpha]}\phi(y)}{1-\dfrac{\mu_\omega-\mu_\alpha}{\mu_\omega-\mu_{[\alpha]}}.\phi(y)}. \qquad (24\cdot7)$$

If the rotation were uniform in time, $\phi(y)$ would be of the form v^y where v was constant. An assumption that was satisfactory for the comparatively short range of the select period would probably not be satisfactory for the range of the ultimate table and it is necessary to investigate $\phi(y)$.

It may be remarked, however, that if we put $\phi(y)=v^y$, change the origin to birth, and consolidate the constants we obtain as an approximation

$$\mu_x \doteq \frac{A-DC^x}{1-EC^x} \quad (C>1), \qquad (24\cdot8)$$

or still further approximating

$$\mu_x \doteq A+BC^x. \qquad (24\cdot9)$$

The fact that the true origin is at the age where the maximum rate of mortality is reached explains, I think, why it has been found that Makeham's formula fits mortality tables better at old ages than at younger ages.

25. In investigating $\phi(y)$ there are the following unknowns:

(1) The value of the maximum force μ_z, or μ_ω.

(2) The age at which μ_ω is attained.

(3) The value of the (minimum) ultimate force, μ_α, which theoretically is never reached.

(4) The value of the (minimum) select force $\mu_{[\alpha]}$.

Expressing $\phi(y)$ in terms of the μ's we have

$$\phi(y)=\frac{\mu_{\omega-y}-\mu_\alpha}{\mu_{\omega-y}-\mu_{[\alpha]}}.\frac{\mu_\omega-\mu_{[\alpha]}}{\mu_\omega-\mu_\alpha}. \qquad (25\cdot1)$$

The factor $\dfrac{\mu_\omega-\mu_{[\alpha]}}{\mu_\omega-\mu_\alpha}$ is a constant, so that if we take logs and difference we shall eliminate it.

$$\Delta \log \phi(y)=\Delta \{\log(\mu_{\omega-y}-\mu_\alpha)-\log(\mu_{\omega-y}-\mu_{[\alpha]})\}.$$

If $\phi(y)$ were e^{-ky}, k constant, $\Delta \log \phi(y)$ would be constant.

From a study of the A 1924–29 Select Tables it appears reasonable to take the maximum value of μ as ·56714, which is reached at about age 101. For the purpose of an investigation of $\phi(y)$ in relation to those tables the origin was therefore placed at age 101.

When beginning the work of investigating $\phi(y)$ I used the values of μ_{15} and $\mu_{[15]}$ from the A 1924–29 Tables for μ_α and $\mu_{[\alpha]}$. The ultimate rate at birth is unknown to us and so is μ_1 for that age. I regard the descending rates of mortality in early life as due to an excess number of non-select lives existing at birth, possibly due to the stress of birth, or possibly to the infant sharing the mother's life up to that point, and therefore also her rates of mortality until birth. After birth, the non-select lives tend to die off quickly and the mortality reaches the unstable equilibrium of the ultimate table at the end of the period that selection takes to wear off, about 10 years. The A 1924–29 rates of mortality at age 10, however, are, I think, not trustworthy as homogeneous with the rest of the table, and it seemed likely that by age 15 the minimum values of μ would have been reached in practice. I began, then, with $\mu_\alpha = \cdot 00201$, $\mu_{[\alpha]} = \cdot 00115$, and tabulated $\Delta \log \phi(y)$ and $\Delta \log [-\Delta \log \phi(y)]$. On coming to a conclusion as to the form of $\phi(y)$ I proceeded to regraduate the table and was led to new values of μ_α and $\mu_{[\alpha]}$. It would be tedious to detail all the work done, and in describing the process of the investigation of $\phi(y)$ and the fixing of the values of the constants I shall assume a knowledge of the final values which were the result of an iterative process.

The values of μ_α and $\mu_{[\alpha]}$ finally adopted, then, were

$$\mu_\alpha = \cdot 002156,$$

$$\mu_{[\alpha]} = \cdot 001330.$$

Table 1 shows the calculation of

$$\Delta \log \phi(y) \quad \text{and} \quad \Delta \log [-\Delta \log \phi(y)]$$

for ages 30 to 40, using ·00216 for μ_{15} since μ is tabulated to five figures only, and in order to use five-figure logarithms. The complete table stops at age 28 because the succession of uniform values of μ in the official Table from that age backwards makes

$$\Delta \log \phi(y) = 0$$

for a series of ages.

Table 1

(1) Age	(2) $\lambda\,(\mu_x-\cdot00216)$	(3) $\lambda\,(\mu_x-\cdot00133)$	(4) (2)–(3) = $\lambda\phi\,(y)$	(5) $-\Delta$ (4)	(6) λ (5)	(7) Δ (6)
40	$\bar{3}\cdot20683$	$\bar{3}\cdot38739$	$\bar{1}\cdot81944$	·02514	$\bar{2}\cdot40037$	·10301
39	$\bar{3}\cdot13672$	$\bar{3}\cdot34242$	$\bar{1}\cdot79430$	·03187	$\bar{2}\cdot50338$	·04230
38	$\bar{3}\cdot05690$	$\bar{3}\cdot29447$	$\bar{1}\cdot76243$	·03513	$\bar{2}\cdot54568$	·07798
37	$\bar{4}\cdot97772$	$\bar{3}\cdot25042$	$\bar{1}\cdot72730$	·04204	$\bar{2}\cdot62366$	·07765
36	$\bar{4}\cdot89209$	$\bar{3}\cdot20683$	$\bar{1}\cdot68526$	·05027	$\bar{2}\cdot70131$	$-$·00829
35	$\bar{4}\cdot79934$	$\bar{3}\cdot16435$	$\bar{1}\cdot63499$	·04932	$\bar{2}\cdot69302$	·08025
34	$\bar{4}\cdot71600$	$\bar{3}\cdot13033$	$\bar{1}\cdot58567$	·05933	$\bar{2}\cdot77327$	·08581
33	$\bar{4}\cdot62325$	$\bar{3}\cdot09691$	$\bar{1}\cdot52634$	·07229	$\bar{2}\cdot85908$	$-$·05236
32	$\bar{4}\cdot51851$	$\bar{3}\cdot06446$	$\bar{1}\cdot45405$	·06408	$\bar{2}\cdot80672$	$-$·07796
31	$\bar{4}\cdot43136$	$\bar{3}\cdot04139$	$\bar{1}\cdot38997$	·05355	$\bar{2}\cdot72876$	$-$·23419
30	$\bar{4}\cdot36173$	$\bar{3}\cdot02531$	$\bar{1}\cdot33642$	·03123	$\bar{2}\cdot49457$	·04999

It will be found that $-\Delta\log\phi\,(y)$ shows a practically continuous increase but $\Delta\log\,[-\Delta\log\phi\,(y)]$, in the last column of the table, does not show any marked progression and apart from the ends of the table can be satisfactorily represented by a constant. It may be seen from (24·5) that $\phi\,(0)=1$, and it appears therefore that $\phi\,(y)$ can empirically be represented satisfactorily by the function $e^{-ks_{\overline{y}|}}$ where k is a constant, or, by a suitable modification of k for purposes of calculation, by $10^{-ks_{\overline{y}|}}$.

The appropriate rate at which to calculate $s_{\overline{y}|}$ was found as follows:

Assuming values of μ_α and $\mu_{[\alpha]}$, and writing $1/h$ for the second factor in (25·1),

$$-ks_{\overline{y}|}=\log\,(\mu_{101-y}-\mu_\alpha)-\log\,(\mu_{101-y}-\mu_{[\alpha]})-\log h,$$

$$\therefore\ \sum_{y=26}^{y=45}-ks_{\overline{y}|}=\sum_{y=26}^{y=45}\log\,(\mu_{101-y}-\mu_\alpha)-\log\,(\mu_{101-y}-\mu_{[\alpha]})-\log h$$

$$=-\frac{k}{i}\,[(1+i)^{26}s_{\overline{20}|}-20].$$

Similarly for 46–65.

The ratio gives $[(1+i)^{46}s_{\overline{20}|}-20]/[(1+i)^{26}s_{\overline{20}|}-20]$, from which i can be found.

The values of i and k finally adopted were

$$i=\cdot1162,\quad k=\cdot000026494. \tag{25·2}$$

26. Having calculated $\phi(y)$, we find μ_α and $\mu_{[\alpha]}$ by simultaneous equations from the exposed and deaths using first and second summations.

These should practically reproduce the assumed values but do not do so. If, then, using new values of μ_α and $\mu_{[\alpha]}$ we again find i and k and, hence, new $\phi(y)$, we find new values of μ_α and $\mu_{[\alpha]}$. I have not been able to find a completely self-consistent set of values of i, k, μ_α and $\mu_{[\alpha]}$. This may mean either that the A 1924–29 data are not homogeneous or that the form of $\phi(y)$ adopted is not quite correct. A further point materially affecting the values is that I aimed at reproducing the actual deaths as adjusted for the errors set out in *J.I.A.* Vol. LXVIII, p. 83, so that the graduated curve of μ should be on the whole above that of the official table.

The formula finally adopted for the regraduation is

$$\mu_{101-y}=\frac{\mu_\alpha-\dfrac{\mu_{101}-\mu_\alpha}{\mu_{101}-\mu_{[\alpha]}}\,.\,\mu_{[\alpha]}\,.\,10^{-\frac{\cdot000026494}{\cdot1162}(1\cdot1162^y-1)}}{1-\dfrac{\mu_{101}-\mu_\alpha}{\mu_{101}-\mu_{[\alpha]}}\,.\,10^{-\frac{\cdot000026494}{\cdot1162}(1\cdot1162^y-1)}}, \qquad (26\cdot1)$$

or

$$\mu_{101-y}=\mu_\alpha+(\mu_\alpha-\mu_{[\alpha]})\,.\,\frac{\dfrac{\mu_{101}-\mu_\alpha}{\mu_{101}-\mu_{[\alpha]}}\,.\,10^c\,.\,10^{-c(1\cdot1162)^y}}{1-\dfrac{\mu_{101}-\mu_\alpha}{\mu_{101}-\mu_{[\alpha]}}\,.\,10^c\,.\,10^{-c(1\cdot1162)^y}}, \qquad (26\cdot2)$$

where

$$c=\cdot000228,$$

$$\mu_\alpha=\cdot002156,$$

$$\mu_{[\alpha]}=\cdot001330,$$

$$\mu_{101}=\cdot56714.$$

Tables 2 and 3 show the resulting values of μ_x compared with those of the official table, and the expected and actual deaths.

The regraduation given may not be the best that can be produced by this method and is given rather as an illustration. I therefore refrain from commenting on it in detail.

Table 2. Values of 100,000μ_x

(*a*) *By the formula;* (*b*) *From the official Table A* 1924–29

Age	(*a*)	(*b*)	(*a*)−(*b*)	Age	(*a*)	(*b*)	(*a*)−(*b*)
15	216	201	+15	60	1877	1894	−17
16	216	211	+5	61	2071	2094	−23
17	216	220	−4	62	2285	2308	−23
18	216	228	−12	63	2525	2541	−16
19	217	233	−16	64	2787	2795	−8
20	217	235	−18	65	3082	3084	−2
21	218	235	−17	66	3409	3405	+4
22	219	235	−16	67	3766	3779	−13
23	221	235	−14	68	4167	4202	−35
24	223	235	−12	69	4609	4676	−67
25	226	235	−9	70	5091	5195	−104
26	229	235	−6	71	5632	5760	−128
27	232	235	−3	72	6219	6369	−150
28	236	235	+1	73	6871	7023	−152
29	241	237	+4	74	7580	7724	−144
30	247	239	+8	75	8351	8478	−127
31	254	243	+11	76	9220	9292	−72
32	261	249	+12	77	10158	10171	−13
33	270	258	+12	78	11160	11120	+40
34	279	268	+11	79	12276	12141	+135
35	290	279	+11	80	13445	13240	+205
36	303	294	+9	81	14751	14418	+333
37	317	311	+6	82	16194	15671	+523
38	332	330	+2	83	17646	17000	+646
39	350	353	−3	84	19234	18409	+825
40	370	377	−7	85	20847	19896	+951
41	391	402	−11	86	22664	21463	+1201
42	416	427	−11	87	24681	23112	+1569
43	444	453	−9	88	26625	24846	+1779
44	474	481	−7	89	28743	26670	+2073
45	509	512	−3	90	30764	28588	+2176
46	547	546	+1	91	32919	30604	+2315
47	590	584	+6	92	35149	32725	+2424
48	637	629	+8	93	37704	34955	+2749
49	690	678	+12	94	40264	37302	+2962
50	750	736	+14	95	42752	39774	+2978
51	816	800	+16	96	44830	42380	+2450
52	889	871	+18	97	47061	45132	+1929
53	972	951	+21	98	49009	48042	+967
54	1063	1040	+23	99	51763	51124	+639
55	1165	1141	+24	100	54128	54396	−268
56	1279	1256	+23	101	56714	57875	−1161
57	1405	1387	+18				
58	1547	1538	+9				
59	1703	1707	−4				

Table 3. Comparison of the actual and expectant deaths

Ages	Expected	Actual	E − A	Σ (E − A)
15½–19½	61	63	− 2	− 2
20½–24½	621	668	− 47	− 49
25½–29½	1539	1505	+ 34	− 15
30½–34½	2420	2326	+ 94	+ 79
35½–39½	3770	3727	+ 43	+ 122
40½–44½	5493	5737	− 244	− 122
45½–49½	7832	7633	+ 199	+ 77
50½–54½	10135	9978	+ 157	+ 234
55½–59½	11858	11856	+ 2	+ 236
60½–64½	12316	12594	− 278	− 42
65½–69½	13613	13624	− 11	− 53
70½–74½	15262	15692	− 430	− 483
75½–79½	14204	14245	− 41	− 524
80½–84½	9448	9238	+ 210	− 314
85½–89½	4163	3941	+ 222	− 92
90½–94½	1072	1007	+ 65	− 27
95½–99½	149	122	+ 27	—
100½–	13	9	+ 4	+ 4
	113969	113965	+ 1057 − 1053	

Note. The actual deaths have been adjusted to include 559 deaths originally omitted as set out in *J.I.A.* Vol. LXVIII, p. 83.

CONCLUSION

I have been looking for a Law of Mortality, which some think equivalent to chasing a will-o'-the-wisp. Whether they are right or not depends, I think, on what one understands by a 'Law of Mortality'. It will not be a law that fits any and every mortality experience exactly. It will be, rather, a law of the essential characteristics of a mortality table that is constructed in the way in which we construct mortality tables. Roughly it may be said that the shape of the earth is spherical. This does not quite fit such facts as the flattening of the earth at the poles, and the existence of mountains and valleys, but it is substantially correct, ignores irrelevant detail, and puts the earth in a general class of heavenly bodies, all described approximately by the same formula. This is the kind of law to look for, and when we find it, we shall find that, as in the case of many other laws of nature, it is there because we have put it there. So perhaps, after all, they are right who think there is no law of mortality; it is a law of the way in which we think about mortality rates.

I must admit that I do not think that I have found out all that there is to find out about this law. For example, there is the relationship, if any, between $\psi(t)$ and $\phi(y)$, but I do not feel that at present I can venture to add anything more to what is already a long, and, superficially, a difficult paper. So here is the law as it appears to me:

$$\mu = \left\{\begin{vmatrix}1 & 0\\ 1 & 0\end{vmatrix} + \psi(t)\begin{vmatrix}0 & 0\\ -1 & 1\end{vmatrix}\right\}\begin{vmatrix}a\\ 1-a\end{vmatrix}\left\{\begin{vmatrix}1 & 0\\ 1 & 0\end{vmatrix} + \phi(y)\begin{vmatrix}0 & 0\\ -1 & 1\end{vmatrix}\right\}\mu$$

where $\Delta \log[-\Delta \log \psi(t)]$ and $\Delta \log[-\Delta \log \phi(y)]$ are constants.

Written in this way, the representatives of the select portion run across the paper, as do the select rates, and the representatives of the ultimate portion run up and down the paper, as do the ultimate rates. The general symmetry of the formula then becomes obvious.

BIBLIOGRAPHY

Title	Author	Publishers
	Philosophical	
New Pathways in Science	Sir A. Eddington	C.U.P.
The Philosophy of Physical Science	Sir A. Eddington	C.U.P.
The New Background of Science	Sir James Jeans	C.U.P.
The Physical Significance of the Quantum Theory	F. A. Lindemann	Clarendon Press
	Pure Mathematics	
Vector Methods	D. E. Rutherford	Oliver and Boyd
Elementary Vector Analysis	C. E. Weatherburn	G. Bell and Sons, Ltd.
Advanced Vector Analysis	C. E. Weatherburn	G. Bell and Sons, Ltd.
Introduction to Quaternions	Kelland and Tait	Macmillan and Co.
Determinants and Matrices	A. C. Aitken	Oliver and Boyd
Principles of Geometry, Vol. 1	H. F. Baker	C.U.P.
	Applied Mathematics	
Space, Time, Matter	Hermann Weyl	Methuen and Co., Ltd.
Elementary Quantum Mechanics	R. W. Gurney	C.U.P.
Relativity Theory of Protons and Electrons	Sir A. Eddington	C.U.P.
The Principles of Quantum Mechanics	P. A. M. Dirac	Clarendon Press

Remarks. The above works are not the only books in their respective fields, and not necessarily the best for a student. The basis of recommendation is personal acquaintance. None of the books is in a foreign language.

It is not suggested that the whole book need be studied in each case. For example, Kelland and Tait's *Quaternions* is mentioned for the sake of the introduction and ch. III, §§ 15–21 on rotations, and Baker's *Geometry* for pp. 62 et seq. on matrices. In some of the books classified as Pure Mathematics, the examples and exercises are nearly all based on Physics, and the books on Quantum Mechanics are very largely written with direct application to physical problems. The student who is interested will have to pick out what he thinks of use to him.