THE CONTRIBUTIONS OF E. T. WHITTAKER TO ALGEBRA AND NUMERICAL ANALYSIS

by A. C. AITKEN

1. WHITTAKER'S contributions to algebra are not numerous and are confined to special problems ; in one or two cases they are rediscoveries, of the kind that add a certain illumination to the original. His outlook on algebra was that of his time ; one might characterise it by referring to the spirit and content of Perron's *Algebra*, as contrasted with the *Moderne Algebra* of van der Waerden. Bred in the Cayley-Sylvester tradition of matrix algebra and the invariant theory of forms, he was expert also in the manipulation of continued fractions and determinants. His lectures to honours classes in Edinburgh always included a course on matrices, vector analysis and invariants, historical reference being made to Grassmann, Cayley and Sylvester ; it is to such a course that the writer owes his first acquaintance with matrix algebra. In later years, from about 1925, Whittaker acquired an adequate knowledge, though never a marked taste, for those parts of modern abstract algebra, in particular the representation of groups and algebras, that serve a purpose in mathematical physics ; his courses of research lectures, which he kept up until a late period in his tenure of the chair, bore witness of this.

2. His few papers on algebra concern matrices, determinants and continued fractions, and are ancillary to more general researches into the theory of functions satisfying certain differential equations. His chief paper on continued fractions (2) follows in this way an earlier one (1) in which he had used continued fractions to represent the confluent hypergeometric function, and had pointed out the advantages in respect of convergence and region of validity. The introductory example of the paper of 1916 is a good one, namely the C.F.

$$\frac{1}{2x-} \quad \frac{1}{2x-} \quad \frac{1}{2x-} \quad \cdots$$

as compared with the power series

$$\frac{1}{2x} + \frac{1}{2!2^2x^3} + \frac{1}{3!2^3x^5} + \frac{1}{4!2^4x^7} + \cdots,$$

both being representations of the function $y = x - (x^2-1)^{\frac{1}{2}}$. Whereas the series converges only outside the unit circle in the complex plane, the C.F. converges everywhere except along the segment of the real axis given by $-1 < x < 1$.

The manipulative disadvantage of continued fractions (this is the chief deterrent to their more general use) is that they do not lend themselves to addition, multiplication or differentiation—fundamental processes. Whittaker's aim was in part to remedy this, in respect of differentiation at least. His first result is that the C.F.

$$\frac{1}{b_0+x-} \quad \frac{a_1}{b_1+x-} \quad \frac{a_2}{b_2+x-} \quad \cdots \quad \frac{a_n}{b_n+x}$$

admits an expansion in powers of x^{-1}, generalising the corresponding expansion of $1/(b+x)$; it is in fact

$$\frac{1}{x} - \frac{b^{(1)}}{x^2} + \frac{b^{(2)}}{x^3} - \frac{b^{(3)}}{x^4} + \ldots,$$

where $b^{(p)}$ is the leading element in the pth power of the continuant matrix

$$B = \begin{bmatrix} b_0 & c_1 \\ d_1 & b_1 & c_2 \\ & d_2 & b_2 & \cdot \\ & & \cdot & \cdot & \cdot \\ & & & \cdot & \cdot & \cdot \\ & & \cdot & & \cdot & \cdot & \cdot \\ & & & & & \cdot & b_{n-1} & c_n \\ & & & & & & d_n & b_n \end{bmatrix},$$

c_i and d_i being any numbers such that $c_i d_i = a_i$. (As is pointed out, a natural choice would be to take all $c_i = 1$, or to ensure symmetry of B by making $c_i = d_i = a_i^{\frac{1}{2}}$.) The theorem becomes more interesting when expressed (as at p. 246) in matrix notation. It is then that the C.F. in question is the leading element in the reciprocal matrix $(B+xI)^{-1}$, connexion being thus established with the theorem that a continued fraction is the quotient of two determinants, the numerator being a leading diagonal cofactor in the denominator. The rest of the paper is concerned with the relation of the result to Stieltjes' expansion of a continued fraction in a power series.

The differentiation of the C.F., let us say S, follows from the above result; for $-dS/dx$ must be the leading element in $(B+xI)^{-2}$. The rest of the paper is devoted to obtaining an explicit form $-dS/dx$ as the quotient of two determinants, the numerator of order $2n$, the denominator of order $2n+2$. In the notation of partitioned matrices we may write this quotient as

$$\begin{vmatrix} O & B_0+xI \\ B_0+xI & -I \end{vmatrix} \div \begin{vmatrix} O & B+xI \\ B+xI & -I \end{vmatrix},$$

where B_0 is B with the first row and column removed, in fact the submatrix complementary to the leading element b_0 of B, and O denotes a block of zero elements.

3. In another paper (3) Whittaker considers "the latent roots of compound determinants and Brill's determinants". The results, established directly from the associated linear transformations, were already known. Indeed the three types of matrix considered, the mth compound of A, the mth Schläflian and the mth direct or Kronecker power, were the types of "invariant matrix" explicitly mentioned by Schur in his celebrated thesis of 1901 and generalised by him. The nature of their latent roots, with respect to those of A, is the subject of theorems due to Radós (1891) and Franklin (1894).

To the same field of compound determinants Whittaker makes an original

contribution (**6**), the novel feature being a demonstration that certain types of such determinants are expressible as other determinants having ordinary monomial elements. A clear account of this paper, with instructive examples, is to be found in Muir, *History of Determinants*, 1900-1920 (Blackie, 1930), pp. 241-242. In his class lectures Whittaker was accustomed to remark that there were really only three main theorems in determinants, namely the expansion theorem of Laplace, the multiplication theorem and Jacobi's theorem on minors of the adjugate determinant. In this paper the Laplace expansion is the chief instrument. We may quote the deduction with which the paper ends, generalising a theorem of Bazin :

The determinant whose rows are the first k rows of the adjugate of A and and the last $n-k$ rows of the adjugate of B is equal to

$$| B |^{n-k-1} \quad \begin{vmatrix} (A_1, B_1) & (A_1, B_2) & \cdots & (A_1, B_k) \\ (A_2, B_1) & (A_2, B_2) & \cdots & (A_2, B_k) \\ \cdots & \cdots & \cdots & \cdots \\ (A_k, B_1) & (A_k, B_2) & \cdots & (A_k, B_k) \end{vmatrix},$$

where A and B are of order n and (A_r, B_s) denotes the determinant obtained by replacing the rth row of A by the sth row of B.

4. It is convenient at this point to refer to a paper (**5**) belonging in part to algebra, in part (and this is the chief intention) to numerical analysis. It develops a series formula for the root, smallest in modulus, of the equation

$$0 = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots.$$

If we denote, for example,

$$\begin{vmatrix} a_2 & a_3 & a_4 \\ a_1 & a_2 & a_3 \\ a_0 & a_1 & a_2 \end{vmatrix} \quad \text{by} \quad | \, a_2 \, a_2 \, a_2 \, |,$$

and similarly in general, the series is

$$-\left(\frac{a_0}{a_1} + \frac{a_0^2 a_2}{a_1 \, | \, a_1 a_1 \, |} + \frac{a_0^3 \, | \, a_2 a_2 \, |}{| \, a_1 a_1 \, || \, a_1 a_1 a_1 \, |} + \dots \right).$$

Truncated at the nth term, it is actually a type of expansion due in the first place to Schweins (1825) but easily established, as here, by Jacobi's theorem on minors of the adjugate. It is equivalent in this case to the quotient

$$- \, | \, a_1 a_1 a_1 \dots a_1 \, |_n \div | \, a_1 a_1 a_1 \dots a_1 \, |_{n-1}$$

of two determinants of orders n and $n-1$. It is perhaps easier to modify the notation and to consider the problem as that of determining the root, greatest in modulus, of the equation

$$0 = f(x) = a_0 x^n - a_1 x^{n-1} + a_2 x^{n-2} - \dots + (-)^n a_n, \quad a_0 = 1.$$

The a_k are then elementary symmetric functions of the roots of $f(x) = 0$, while

$$\{f(x)\}^{-1} = x^{-n}(h_0 + h_1 x^{-1} + h_2 x^{-2} + \dots), \quad h_0 = 1,$$

E.M.S.—C

generates the complete homogeneous symmetric functions h_k. Then from Wronski's relations

$$0 = h_1 a_0 - h_0 a_1, \ 0 = h_2 a_0 - h_1 a_1 + h_0 a_2, \ \ldots,$$

used in Whittaker's proof, it follows that

$$h_k = |\, a_1 a_1 a_1 \ldots a_1 \,|_k,$$

whence the Whittaker series is equivalent to $\lim h_n/h_{n-1}$. The use of quotients of determinants of type $|\, a_k a_k \ldots a_k \,|$ to approximate to the numerically greatest root, or to the product of the r numerically greatest roots, goes back (Muir, *History of Determinants*, 1900-1920, p. 272) to E. Fürstenau. Indeed, since the full Wronski relation

$$h_n a_0 - h_{n-1} a_1 + h_{n-2} a_2 - \ldots + (-)^n h_0 a_n = 0$$

may be regarded as a linear difference equation in h_x for which $f(x) = 0$ is the characteristic equation, it goes back further, to the method of D. Bernoulli [(9), p. 98]. Since the terms of such a Bernoullian sequence, and the determinants derived from it, satisfy multiplicative relations well adapted to modern calculating machines capable of automatic multiplication and division Whittaker's paper has been followed up by others, and methods based on Bernoullian sequences have had a reassessment.

5. Whittaker's other contributions to numerical analysis are principally concerned with the numerical solution of integral equations and with the problem of the graduation of observational data, that is, with the representation of a set of values $u(x)$, affected by errors of observation, by presumptive values $y(x)$ smoothed according to some criterion.

Whittaker's work on the numerical solution of integral equations (4) is set out in a selected version in pp. 376-381 of (9). The types of integral equation considered are

$$\int_0^x \phi(s)\kappa(x-s)ds = f(x)$$

and

$$\phi(x) + \int_0^x \phi(x)\kappa(x-s)ds = f(x).$$

The first is solved for the case where the nucleus $\kappa(x)$ becomes infinite at $x = 0$ in such a way that

$$x^p \kappa(x) < M, \quad 0 < p < 1,$$

and $\kappa(x)$ is then presumed to be representable with sufficient accuracy by

$$x^{-p}(a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n), \quad a_0 \neq 0.$$

Whittaker then shows, generalising a result due to Abel, that

$$\phi(x) = \frac{\sin p\pi}{\pi} \int_0^x f'(s)L(x-s)ds$$

where

$$L(x) = \frac{x^{p-1}}{a_0} + \frac{a^{n-p}}{F'(a)}\gamma_p(ax) + \frac{\beta^{n-p}}{F'(\beta)}\gamma_p(\beta x) + \ldots + \frac{\nu^{n-p}}{F'(\nu)}\gamma_p(\nu x),$$

where $a, \beta, \ldots, \nu$ are the roots of the algebraic equation

$$0 = a_0 x^n - (p-1)a_1 x^{n-1} + (p-1)(p-2)a_2 x^{n-2} - \ldots + (-)^n(p-1)(p-2)\ldots$$
$$(p-n)a_n = 0,$$

and where $\gamma_p(x)$ denotes the Incomplete Gamma Function

$$\gamma_p(x) = e^x \int_0^x s^{p-1}e^{-s}ds.$$

The details of computation are still rather formidable; a procedure of interpolation to express $\kappa(x)$, the solution of an algebraic equation of degree n, and finally the consultation of tables of the Incomplete Gamma Function.

The second integral equation requires a treatment perhaps even more elaborate. Of the several methods given by Whittaker in his memoir, he regarded that set out at p. 378 of (9) as most serviceable. The nucleus $\kappa(x)$ is first represented as a sum of m exponential functions

$$\kappa(x) = Pe^{px} + Qe^{qx} + \ldots + Ve^{vx},$$

the $2m$ constants being determined by Prony's method [(9), pp. 369-370]. This requires the solution of a set of simultaneous linear equations, and then the setting and solving of a further set of normal equations. The solution of the integral equation is then

$$\phi(x) = f(x) - \int_0^x K(x-s)f(s)ds,$$

where

$$K(x) = -\frac{(a-p)(a-q)\ldots(a-v)}{(a-\beta)(a-\gamma)\ldots(a-\nu)} e^{ax} - \ldots - \frac{(\nu-p)(\nu-q)\ldots(\nu-v)}{(\nu-a)(\nu-\beta)\ldots(\nu-\mu)} e^{vx},$$

and where $a, \beta, \ldots, \nu$ are the roots of the algebraic equation

$$\frac{P}{x-p} + \frac{Q}{x-q} + \ldots + \frac{V}{x-v} + 1 = 0.$$

In his memoir Whittaker suggests and develops alternative methods, which he has not included in (9). The most promising would seem to be one which expands $\kappa(x)$ as a Taylor series

$$\kappa(x) = \kappa_0 + \kappa_1 x + \kappa_2 \frac{x^2}{2!} + \kappa_3 \frac{x^3}{3!} + \ldots,$$

the solution being then given by

$$\phi(x) = f(x) - \int_0^x K(x-s)f(s)ds,$$

where

$$K(x) = \kappa_0 - \begin{vmatrix} \kappa_0 & 1 \\ \kappa_1 & \kappa_0 \end{vmatrix} x + \begin{vmatrix} \kappa_0 & 1 & 0 \\ \kappa_1 & \kappa_0 & 1 \\ \kappa_2 & \kappa_1 & \kappa_0 \end{vmatrix} \frac{x^2}{2!} - \ldots \quad .$$

In fact

$$K(x) = \frac{1}{x}\left(1 - \frac{1}{1+\kappa_0 x + \kappa_1 x^2 + \kappa_2 x^3 + \ldots}\right).$$

So far as the writer is aware, these methods of Whittaker have not been tried out in the respect of adapting them to modern machines, whether electrically operated or electronic. It would seem that the last one, depending as it does on reciprocating a series, could be so adapted.

6. Another field in which Whittaker did original work is that of the smoothing or "graduation" of functional values corresponding to equally spaced abscissæ and given by observations affected by random error. He was led to consider this problem through his friendship and mathematical association with the distinguished actuary, the late G. J. Lidstone, F.I.A., F.F.A., LL.D. In actuarial theory and practice the problem of graduating tables of mortality has a history and literature of its own. "Summation formulæ" (so called because the operations of linear combination were carried out by repeatedly summing the data in consecutive runs of so many at a time, as by threes, or fives, or sevens) such as the Spencer 21-term or the Karup-King 19-term formulæ were in regular use and had been found to give a high empirical efficiency, as judged by the smoothness of a graduated set and its fidelity to the ungraduated set. There was also the method of least squares, where in particular the fitting of a polynomial to a long series of data had been considered in great detail by W. F. Sheppard and others. Yet a reasonably satisfactory theory of graduation was still a desideratum. Whittaker sought a basis in the theory of inverse or inductive probability. The concepts of smoothness and of fidelity required to be objectively specified, with of course an unavoidable element of arbitrariness in the criteria adopted, and to be embedded in a formulation from which a unique solution might emerge. According to Whittaker, fidelity might be measured by the compound probability of deviations, supposed for simplicity to be of equal weight and independently and normally distributed; smoothness by the exponential of the negative sum of squared third differences of values. Let then

$$F = \Sigma(y_x - u_x)^2, \quad S = \Sigma(\Delta^3 y_x)^2.$$

The argument from inductive probability then dictated the minimising of $S + \epsilon F$, where ϵ was a constant, small in all practical cases, expressing the balance of smoothness against fidelity.

The element of arbitrariness is evident, especially in the choice of S, which could equally well have been based on any other order of differences. We may note also, for the benefit of any who may feel antipathetic to the argument from inverse probability, that precisely the same minimal formulation could have been derived from other postulates, namely the minimising of F subject to the condition that S assumes some *prescribed* value; not zero, for this would be equivalent to fitting a quadratic polynomial by least squares. Then ϵ would in effect be a Lagrange multiplier in the conditioned minimal problem. Either approach leads to a central difference equation of the 6th order,

$$\epsilon y_x - \delta^6 y_x = \epsilon u_x,$$

with six initial conditions in the form of vanishing third differences, three at each end of the range.

Whittaker developed two solutions of this problem, which is a heavy one when it is considered that the number of data may in practice be fifty or more. In the first [7] he proposed to expand the smoothed function y in powers of ϵ, supposed small. The first term in such an expansion would correspond to $\epsilon = 0$, and would thus be the quadratic polynomial as fitted by least squares. He showed how to obtain the second term. None the less the matter remained laborious and liable, in a long series of data, to accumulation of error. In a second paper (8) he introduced a " graduation function " G_x in the form of a series satisfying the difference equation with right-hand member equal to zero, and proposed to expand y in terms of this. The disadvantage was that G_x, even for fairly small values of ϵ, attained very large values as x increased. The writer may be permitted to say that Whittaker's work suggested to him that the reciprocal of the difference operator might be expanded as a Laurent series in operators, and that this led to the expression of y for each x as a convergent symmetrical linear combination of values of u.

Whittaker's method of graduation has received little attention in Britain, but much attention from actuaries in America, notably the late Robert Henderson, and more recently Mr Charles Spoerl and Mr Kingsland Camp. It could be programmed without difficulty on electronic machines. It is also possible that alternative methods of solution suitable to desk machines may even yet be developed.

7. A word, in conclusion, concerning the calculating facilities available in the early years of the mathematical laboratory which Whittaker instituted at the University of Edinburgh in 1913. To introduce numerical analysis in this way into an honours curriculum was quite an innovation at the time. In default of machines, which at that period were not very efficient and were rather noisy to operate, tables such as those of Crelle (products of integers up to 1000 by 1000), Barlow and Schrön were extensively used ; and the fitting of Fourier terms to a cycle of 12 or 24 observations was accomplished by a special form of computation modified and adapted from Runge. (There were machines in the department such as the Millionaire, the Archimedes, early forms of the Brunsviga and the Mercedes-Euklid, and others ; these were for the use of the staff and research students.) In 1924, with the co-operation of G. Robinson, who was then a member of the staff, the *Calculus of Observations* (9) was published. Perusal of this work will give an accurate picture of numerical analysis as understood and practised in 1925. There is some resemblance to the *Numerisches Rechnen* of Runge and König, but even now, thirty years afterwards, the freshness is preserved, and there is much to be found between those covers that appears nowhere else.

REFERENCES

(1) On the continued fractions which represent the functions of Hermite, *Proc. Edin. Math. Soc.*, **32** (1914), 65-74.

(2) On the theory of continued fractions, *Proc. Roy. Soc. Edin.*, **36** (1916), 243-255.

E.M.S.—C 2

(3) On the latent roots of compound determinants and Brill's determinants, *Proc. Edin. Math. Soc.*, **35** (1917), 2-9.

(4) On the numerical solution of integral equations, *Proc. Roy. Soc.*, **94** (1918), 367-383.

(5) A formula for the solution of algebraic and transcendental equations, *Proc. Edin. Math. Soc.*, **36** (1918), 103-106.

(6) On determinants whose elements are determinants, *Proc. Edin. Math. Soc.*, **36** (1918), 107-115.

(7) On a new method of graduation, *Proc. Edin. Math. Soc.*, **41** (1923), 63-75.

(8) On the theory of graduation, *Proc. Roy. Soc. Edin.*, **44** (1924), 77-83.

(9) *The Calculus of Observations : a Treatise on Numerical Mathematics* (with G. Robinson) (Blackie, 1924).

THE MATHEMATICAL INSTITUTE
THE UNIVERSITY
EDINBURGH