A LOST CHAPTER IN THE PRE-HISTORY OF ALGEBRAIC ANALYSIS

S. C. COUTINHO

Abstract. In the early 1930s W. O. Kermack and W. H. McCrea published three papers in which they attempted to prove a result of E. T. Whittaker on the solution of differential equations. In modern parlance, their key idea consisted in using quantized contact transformations over an algebra of differential operators. Although their papers do not seem to have had any impact, either then or at any later time, the same ideas were independently developed in the 1960-80s in the framework of the theory of modules over rings of microdifferential operators. In this paper we describe the results of Kermack and McCrea and discuss possible reasons why such promising papers had no impact on the mathematics of the 20th century.

Contents

Date: April 5, 2010.
2000 *Mathematics Subject Classification.* Primary: ; Secondary:
The author was partially supported by grants from CNPq and Pronex (commutative algebra and algebraic geometry.

1. Introduction

Among the ideas that made a deep impact on mathematics in the 1980s and 1990s we can certainly count the twin theories of differential and microdifferential operators, jointly known as algebraic analysis. Their impact extended far and wide, and spread through such areas as partial differential equations, representations of algebraic groups, noncommutative ring theory, and combinatorics. Nor has their influence waned, as a search of MathSciNet clearly shows.

The genesis of both theories is usually traced to the 1960s, when modules over rings of differential operators were first systematically studied by M. Sato in Japan and, independently, by J. Bernstein in the Soviet Union. These ideas were then developed systematically by Sato, Bernstein, M. Kashiwara, Z. Mebhkout and many other mathematicians. For more details see §9.2. However, 30 years earlier some of the ideas that we now associate with algebraic analysis had already been introduced in four papers of W. O. Kermack and W. H. McCrea. Published between 1931 and 1933, the papers were motivated by a 'research lecture' given by Sir Edmund Whittaker in Edinburgh. In this lecture, whose content appeared in [43], Whittaker proposed a theorem on the transformation of definite integrals which he used to solve various differential equations. It turned out, however, that Whittaker's proof of this theorem was, in his own words, 'defective'.

Whittaker's audience included the chemist W. O. Kermack and the physicist W. H. McCrea. In McCrea's words [12, p. 421]

> Kermack saw immediately that Whittaker's ideas required in the first place an algebra of operators of a novel sort. Within a day or two he sketched his thoughts to me and we proceeded together to develop them in four papers published soon afterwards.

This 'novel algebra of operators' turned out to be analogous to the algebra generated by the operators position and momentum in quantum mechanics, which had been introduced by Dirac in 1926; see [13]. This is a noncommutative algebra, for which they introduced 'contact transformations' like the ones used in Hamiltonian dynamics [41, p. 288ff]. These transformations allowed them to prove Whittaker's theorem.

In this paper we recount the history of the work on differential operators by Kermack and McCrea and relate it to modern work on rings of differential operators. Section 2 contains short biographies of the three main characters of our story: Whittaker, Kermack and McCrea. These do not cover in detail the whole life of each of these scientists, but only take the story up to the time when the work discussed here was done. In section 3 we review the basics of contact transformations, both the 1930s approach as expounded in Whittaker's *Analytical Dynamics* [41], and its modern presentation. The analysis of the work of Whittaker, Kermack and McCrea on differential equations begins in section 4, with a description of Whittaker's seminar in Edinburgh, where his results on integration were first presented to an audience that included both Kermack and McCrea. The next sections contain a presentation of six different papers. The first, by Whittaker, states his integration theorem and gives several examples; this is followed by two substantial papers, jointly written by Kermack and McCrea, that give their proof of Whittaker's theorem, based on noncommutative algebra. Of the last three papers, two develop some collateral themes, while the very last one (written by McCrea alone) is an exposition of their method aimed at a more general audience. In the final section we analyse some of the weak points of the theory introduced by Kermack and McCrea, explain how they anticipated some key ideas of modern algebraic analysis, and discuss the possibility of this being a case of missed opportunity, in the terminology of [15]. This section ends with a general conclusion which summarizes the main points of the article. To simplify cross referencing, equations will be numbered sequentially throughout the paper, even when they are part of a quotation from a primary source.

2. Dramatis personæ

In this section we give biographical data on the three main characters of our story, beginning with the oldest and best known of the three.

2.1. **E. T. Whittaker.**

Edmund Taylor Whittaker was born on 24 October 1873 in Southport. Whittaker was not very strong as a child, and his mother took care of his early education. It was only when he was eleven years old that he entered Manchester Grammar School, where he was on the classical side, with more than three-fifths of his time taken up with Latin and Greek. While the study was purely linguistic, he did very well; however, his lack of interest in poetry and drama led to a falling-off in the upper forms. In order to escape that he elected to specialize in mathematics.

In December, 1891 Whittaker obtained a scholarship to Trinity College, Cambridge, and in 1895 graduated as second wrangler, being beaten to the first place by T. J. l'A. Bromwich. The next year he was elected a Fellow of Trinity and was put on the staff. Although Whittaker's early interests where in applied mathematics, the work for which he was awarded the First Smith Prize in 1897 dealt with uniform functions, a theme of pure mathematics.

Among the courses he taught in Cambridge was the one on complex function theory which would later become the basis for Whittaker's first book *A Course of Modern Analysis*; [42]. Although the book is often called 'Whittaker and Watson', the first edition, published in 1902, was a solo effort; see [3]. Watson's collaboration began with the second edition, which appeared in 1915, and it is thanks to him that this edition included topics like Riemann integration and the zeta function.

Beside Watson himself, many well-known mathematicians attended Whittaker's Cambridge lectures, among them G. H. Hardy, J. E. Littlewood and H. W. Turnbull, as well as the astronomers J. H. Jeans and A. S. Eddington.

Whittaker remained in Cambridge until 1906, when he moved to Dublin to take the post of Royal Astronomer for Ireland. At the same time he was appointed professor of astronomy at the University of Dublin. By this time he had already been elected a fellow of the Royal Society. Despite the title of his chair, Whittaker recognized that he was more of a mathematician than an astronomer, his talents being in the line of developing the sort of mathematics that would be useful to astronomers; [3]. In 1912, Whittaker moved to the University of Edinburgh where he had been elected professor of mathematics and he stayed there for the next 35 years, by which time he had established a research school and founded (in 1914) what may have been the first university mathematical laboratory.

Whittaker was awarded many prizes, among them the Sylvester Medal (1931) and the Copley Medal (1956) by the Royal Society and he served as president of both the Edinburgh Mathematical Society (1939-1944) and the London Mathematical Society (1928-1929). He died on 24 March 1956. For more details see [31].

2.2. **W. O. Kermack.** Of the three main characters in our story, Kermack is surely the least known among mathematicians; and this is not at all surprising, for he was an experimental chemist.

William Ogilvy Kermack was born on 26 April 1898 in Kirriemuir, Scotland; the birthplace of J. M. Barrie. Actually, Barrie's first novels are based on stories that his mother told him about Kirriemuir, which is called 'Trums' in the novels.

On the death of his mother in 1904, Kermack was brought up by his father's sister. At the age of five he entered *Webster's Seminary*, a local school that provided primary and secondary education. At this school he was introduced to mathematics, mainly by the headmaster Thomas Pullar, who took him as far as coordinate geometry and conics, as well as topics of mixed mathematics like elementary dynamics and hydrostatics.

Kermack entered Aberdeen University in 1914 where he graduated M.A. four years later, with First Class Honours in Mathematics and Natural Philosophy, and B.Sc. with special Distinction in Mathematics, Natural Philosophy and Chemistry, having been awarded several prizes. He was taught mathematics at Aberdeen by J. H. Grace, who did a great deal to stimulate his interest in it.

After a few years at the Dyson Perrins Laboratory in Oxford, Kermack was appointed in charge of the Chemical Section of the Royal College of Physicians Laboratory in Edinburgh. This was in 1921. Three years later, tragedy struck. While he was alone at the laboratory, the preparation on which he was working exploded, driving caustic substances into his eyes and rendering him totally blind for life. Although he was only 26 at the time, Kermack was not put down by his blindness. Instead, he re-organised his life to meet his new needs, in which he was helped by friends and colleagues; with the Department of Scientific and Industrial Research and the Carnegie Trust providing funds for a special assistant.

Kermack would remain in Edinburgh until 1949, when he moved to Aberdeen to take the Macleaod-Smith Chair of Biological Chemistry, from which he retired in 1968. He died suddenly, two years later, while working at his desk. Most of Kermack's work is in the area of biochemistry. However, he kept his interest in mathematics and was an active member of the Edinburgh Mathematical Society during his stay in that city. Beside the work on differential equations that we will discuss here, Kermack's mathematical papers dealt also with Riemannian geometry, relativistic cosmology and statistics. Kermack was elected to the Royal Society in 1944. His obituary in the Biographical Memoirs [12] provides an extensive bibliography, as well as a detailed account of his chemical work.

2.3. **W. H. McCrea.** Although he is best known as an astronomer, Sir William Hunter McCrea held positions in mathematics departments for many years. Born in Dublin in 1904, McCrea was brought up in Derbyshire, after his family moved to England when he was two years old. He attended Chesterfield Grammar School and won an Entrance Scholarship to Trinity College Cambridge in 1923. While working for the Mathematical Tripos, he found time to obtain a first class B.Sc. with honours in mathematics from London University. McCrea graduated from Cambridge in 1926, and immediately began to work for a Ph.D. there under the supervision of R. H. Fowler. In 1928 he won a Rouse Ball Traveling Studentship that took him to Göttingen for the year 1928-29.

Back from Göttingen, McCrea took his first teaching position in 1930 as Lecturer in Mathematics at the University of Edinburgh. He spent two years in Edinburgh, during which the work that we report in this paper was undertaken. In 1932 he moved to Imperial College, London, as Reader in Mathematics, where he stayed until 1936, then moving to a Chair of Mathematics at Queen's University in Belfast. During the war McCrea worked at the the Operational Research Group, after which he took over as Professor and Head at the Mathematics Department, Royal Holloway College. It was only in 1966 that he had his first post in Astronomy, as founding Research Professor in the Astronomy Centre of the recently opened University of Sussex.

McCrea published some 280 scientific papers and six books. Although many of his early papers are on topics of physics and mathematics, his interests gradually focused on the application of theoretical physics to astronomy. Indeed, he is probably best known for his work on stellar atmospheres, relativity and cosmology.

McCrea was elected a Fellow of the Royal Astronomical Society in 1929, served as President between 1961 and 1963 and was awarded its Gold Medal in 1976. He was elected to a Fellowship of the Royal Society of London in 1952.

3. Contact transformations

As we saw at the introduction, the work of Kermack and McCrea on the application of rings of differential operators was prompted by Whittaker's idea of using contact transformations in order to solve differential equations. Actually, Whittaker's interest in these transformations predates this paper by many years, as we shall see in this section, where we also introduce the basic facts of the theory of contact transformations.

3.1. **Whittaker's 'Analytical Dynamics'.** As we saw in §2.1, Whittaker lectured in Cambridge between 1896 and 1906. At this time he wrote several books, the best known of which is probably his very first book *Modern Analysis*, which has already been mentioned in §2.1. Another book written at this time, and published in 1904, had the long title *A treatise on the analytical dynamics of particles and rigid bodies, with an introduction to the problem of three bodies*; nowadays it is known simply as *Whittaker's Analytical Dynamics*. Although mainly based on the notes of his Cambridge lectures, it also included material from the *Report on the progress of the solution of the problem of three bodies*, that he had prepared for the British Association for the Advancement of Science in 1899. At the time it appeared, *Analytical Dynamics* was, according to [32],

> the first systematic account to be given in English of general dynamical theory, i.e. the superbly beautiful theory which springs from Hamilton's equations and which has turned out to be of such fundamental importance for the development of quantum mechanics.

In its more than 450 pages, *Analytical Dynamics* ranges from the more elementary concepts of kinematics and dynamics, through specific problems that are completely soluble, to Hamiltonian dynamics and celestial mechanics. The chapters that will concern us here are those that deal with Hamiltonian mechanics, specially what he calls the 'transformation theory of dynamics'. This chapter played a decisive rôle in the history of modern physics by inspiring Dirac's approach to quantum mechanics; see [28, p. 17].

3.2. **Contact transformations.** Let us begin by reviewing the fundamental concepts of Hamiltonian mechanics in the terminology of Whittaker's *Analytical Dynamics*; see [43, §109, pp. 263-65].

Consider a conservative time independent (holonomic) dynamical system with coordinate functions $q_1, \dots, q_n$ and let $L(q_1, \dots, q_n, \dot{q_1}, \dots, \dot{q_n})$ be its kinetic potential. Writing the equations of motion in Lagrangian form we obtain,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q_r}}\right) - \frac{\partial L}{\partial q_r} = 0 \text{ for } \quad r = 1, \dots n.$$

If we write

$$p_r = \frac{\partial L}{\partial \dot{q_r}} \text{ then } \quad \dot{p_r} = \frac{\partial L}{\partial q_r}$$

where, as usual, the dot over a letter denotes the time derivative of the corresponding function. Using δ for the increment of a function with respect to a small change of its argument, we have

$$\delta L = \sum_{r=1}^{n}\left(\frac{\partial L}{\partial q_r}\delta q_r + \frac{\partial L}{\partial \dot{q_r}}\delta \dot{q_r}\right)$$

which, can be rewritten in the form

$$\delta L = \delta \sum_{r=1}^{n} p_r \dot{q_r} + \sum_{r=1}^{n} (\dot{p_r}\delta q_r - \dot{q_r}\delta p_r);$$

so that

$$\delta(L - \sum_{r=1}^{n} p_r \dot{q_r}) = \sum_{r=1}^{n} (\dot{p_r}\delta q_r - \dot{q_r}\delta p_r). \tag{3.1}$$

Denoting by H the quantity $L - \sum_{r=1}^{n} p_r \dot{q_r}$, and equating terms on δq_r and δp_r on both sides of (3.1) we find that

$$\frac{dq_r}{dt} = \frac{\partial H}{\partial p_r} \text{ and } \quad \frac{dp_r}{dt} = -\frac{\partial H}{\partial q_r}, \text{ for } \quad r = 1, \dots, n. \tag{3.2}$$

In present day terminology, H is the Hamiltonian function and (3.2) the Hamiltonian equations of the given system. Thus, in this framework, the equations of the system are described in terms of $2n$ variables, the n position variables $q_1, \dots, q_n$ and the n conjugate momentum variables $p_1, \dots, p_n$.

Suppose now that we have two sets of $2n$ variables, denoted by

$$(q_1, \dots, q_n, p_1, \dots, p_n) \text{ and } (Q_1, \dots, Q_n, Q_1, \dots, Q_n),$$

the second of which is defined in terms of the first. In other words, each Q and P is a function of the qs and ps. The change of variables from the capitalised variables to the noncapitalised ones defined by the formulae for the Qs and Ps is a *contact transformation* if the differential form

$$\sum_{r=1}^{n} P_r dQ_r - p_r dq_r,$$

when written in terms of the qs and ps is the differential of a function

$$W = W(q_1, \dots, q_n, p_1, \dots, p_n).$$

Take, for instance,

$$Q = (2q)^{1/2} e^k \cos(p) \text{ and } P = (2q)^{1/2} e^{-k} \sin(p); \tag{3.3}$$

see [43, Example 1, p. 293]. Since

$$PdQ = (-(2q)^{-1/2} \cos(p))dq - ((2q)^{1/2} \sin(p))dp$$

we have that $PdQ - pdq = dW$ for $W = q \sin(p) \cos(p) - qp$. Hence the formulae (3.3) define a contact transformation. This article of *Analytical Dynamics* ends with an argument purporting to prove that if the sets of variables q and Q are independent, then there exists a function $W_1(q_1, \dots, q_n, Q_1, \dots, Q_n)$, such that

$$P_r = \frac{\partial W_1}{\partial Q_r} \text{ and } p_r = -\frac{\partial W_1}{\partial q_r}. \tag{3.4}$$

The function W_1 is called a *generating function* of the corresponding contact transformation. For a more accurate discussion of these functions see [2, p. 258ff].

In §131 of *Analytical Dynamics* the conditions under which the Ps and Qs as functions of $q_1, \dots, q_n, p_1, \dots, p_n$ define a contact transformation are expressed in terms of the Poisson bracket. If u and v are functions of these same variables, then their *Poisson bracket* is

$$\sum_{r=1}^{n} \left(\frac{\partial u}{\partial q_r} \frac{\partial v}{\partial p_r} - \frac{\partial u}{\partial p_r} \frac{\partial v}{\partial q_r} \right).$$

Although Whittaker writes (u, v) for this bracket, we will use the modern notation $\{u, v\}$ instead. The main theorem of §131 is stated as follows:

> Now let $(Q_1, \dots, Q_n, P_1, \dots, P_n)$ denote $2n$ functions of $2n$ variables $(q_1, \dots, q_n, p_1, \dots, p_n)$; we shew that the conditions which must be satisfied in order that the transformation from one set of variables to the other may be a contact-transformation may be written in the form
>
> $$\{P_i, P_j\} = 0, \quad \{Q_i, Q_j\} = 0 \qquad (i, j = 1, 2, \dots, n),$$
> $$\{Q_i, P_j\} = 0 \qquad (i, j = 1, 2, \dots, n; i \neq j),$$
> $$\{Q_i, P_i\} = 1 \qquad (i, j = 1, 2, \dots, n)$$

Although Whittaker goes a lot deeper in his investigation of transformation theory, this is essentially all we need from *Analytical Dynamics*. However, before we move on, it is convenient to present this material under more modern garb, since it will reappear in this form when we come to relate the work of Kermack and McCrea to modern algebraic analysis.

3.3. **Symplectic geometry.** Although its roots can be found in the work of Lagrange, Poisson, Hamilton and Jacobi, symplectic geometry only took hold as the proper geometrical framework for conservative dynamical systems in the 20th century.

Let $q_1, \ldots, q_n, p_1, \ldots p_n$ be coordinates in a real $2n$-dimensional space $\mathbb{R}^{2n}$. The standard *symplectic structure* of such a space is defined by the closed 2-form

$$\Omega = \sum_{r=1}^{n} dq_r \wedge dp_r = d(\sum_{r=1}^{n} q_r \wedge dp_r),$$

where $\wedge$ denotes the exterior product of differential forms. Let $Q_1, \ldots, Q_n$ and $P_1, \ldots, P_n$ be another set of coordinates of $\mathbb{R}^{2n}$ and assume that $T : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is a map that transforms the P, Q coordinates into the p, q coordinates. In this terminology, T is a *contact transformation (or canonical)* if

$$T^*(\sum_{r=1}^{n} dq_r \wedge dp_r) = \sum_{r=1}^{n} dQ_r \wedge dP_r. \tag{3.5}$$

Nowadays these transformations are more often called *symplectomorphisms.* It follows from (3.5) that

$$T^*(\sum_{r=1}^{n} q_r \wedge dp_r) - \sum_{r=1}^{n} Q_r \wedge dP_r$$

is a closed 1-form, so it must be written as the differential of some function W, just as in Whittaker's definition. We can also easily define the Poisson bracket of two functions f and g on $\mathbb{R}^{2^n}$ by the formula

$$\{f, g\} = \Omega(df, dg).$$

4. Differential equations

We are lucky in having Sir William McCrea's personal testimony on his work with Kermack, which appeared at the end of [12].

4.1. **The 'research lectures'.** As part of his effort to create a research school at Edinburgh, Whittaker gave 'research lectures' on two afternoons a week on themes that interested him at the moment. On his obituary of Whittaker [31, p. 239], McCrea writes

> The most famous of the department's activities were his 'research lectures' to staff, postgraduate students and visitors. He gave them twice a week in the middle of the afternoon, throughout the academic year. Either he discussed his own current work or he gave his own development of topics of current interest in mathematics. One marvels at the mathematical power that enabled him always, year after year, to have material for these lectures—he never repeated the same ones–just as though he had nothing else to think about,

> when actually he was inundated with other duties. He had everything fully written out in his notes, which he kept tidy by a method of his own depending upon much use of scissors and paste. As he delivered the lecture, he wrote the whole of it on the blackboard as fast as he spoke. Apparently all his lectures to students were given in this style.

In his obituary of Whittaker for the Edinburgh Mathematical Society [32], D. Martin adds a few more details to McCrea's description

> For instance, when Einstein first produced a unified field theory the lectures dealt with that theory while a course on spinors followed the publication by Cartan of his important book on the subject. Whittaker's ability to absorb and digest so much fresh mathematical work and to lecture on it term after term was truly remarkable. For it must be emphasized that his lecture notes were not just verbatim copies of the essential parts of the published papers on the subject, but contained what was almost a redevelopment of the subject by himself.

And, after commenting, as does McCrea, on the fact that Whittaker wrote his whole talk on the blackboard, Martin adds

> At the end of a lecture he may have been physically tired but he was certainly mentally exhilarated. The research room in the Mathematical Institute is a small homely lecture room with a fire-place at the back, and a few minutes before the end of Whittaker's lecture someone in the back row would put the kettle on the fire so that by four o'clock tea would be ready. Whittaker would then relax in his armchair by the fire, serenely happy with his colleagues and visitors around him, while the animated discussion which arose would cover anything from the lecture just given to academic affairs or religion. In this atmosphere Whittaker was in a most exhilarating form and the inspiration seemed to flow from him. Indeed, as far as research work is concerned, it is rather for his power of inspiring others than for his own work that he will long be remembered.

It was at one of these lectures that he introduced the generalisation of Laplace's transform that led Kermack and McCrea to their work on differential operators.

4.2. **Edinburgh, autumn of 1930.** We now turn to McCrea's description of his joint work with Kermack, which originated in one of Whittaker's 'research lectures' at Edinburgh. According to him, Kermack often came to these lectures, even though they were aimed mainly at mathematicians. Here is what he says at the end of Kermack's obituary [12, pp. 420-422]:

> In one of these lectures in the autumn of 1930, Whittaker propounded a remarkable theorem on the solution of differential equations by definite integrals. Roughly stated, it dealt with the type of relation

$$\psi(q) = \int \chi(q, Q)\phi(Q)dQ \tag{4.1}$$

> where ψ and ϕ are solutions of linear differential equations related through a contact transformation, χ is a solution of a pair of partial differential equations derived from the same contact transformation, and the path of integration has to be suitably chosen. The theorem may be regarded as a greatly generalised Laplace transformation, and it covers all the well-known integral relations in the theory of special functions. It was a brilliant conjecture of Whittaker's, but for his part he contented himself with a number of interesting illustrations. He did not attempt to formulate a general statement nor to give a proof.

McCrea's comments should be compared with Whittaker's. In the paper [43] where he announces this 'brilliant conjecture', Whittaker says:

> My own original way of establishing the theorem was defective as a proof, and has been superseded by a proof devised by Dr Kermack and Dr M'Crea [sic], which is given in the paper following this.

So it seems that Whittaker originally had a proof, which turned out to be wrong; and that's why the result he presented ended up as no more than a conjecture.

McCrea goes on to explain how the collaboration with Kermack came about:

> Kermack saw immediately that Whittaker's ideas required in the first place an algebra of operators of a novel sort. Within a day or two he sketched out his thoughts to me and we proceeded together to develop them in four papers published soon afterwards.

The papers are [24], [25], [26] and [27], the first two of which we discuss in detail in sections 6 and 7. Next, McCrea explains the contents of each one of the papers:

> The first requirement of a theory was to show how to derive the equations for the function χ in a manner that is unambiguous and that makes them compatible. This was the main point of the first paper. We proceeded to give definite rules for passing from the differential equation for ϕ to that for ψ, and vice versa. We then showed how to generalise the equations for χ. If (4.1) be regarded as a transformation from ϕ to ψ, we also obtained some results on successive transformations of this sort.
>
> In our second paper we began by giving a number of theorems in the algebra of a pair of operators p, q that commute according to $qp - pq = 1$ and otherwise obey the usual laws of algebra. Probably most of these are now familiar in the literature, but at the time they were new and somewhat surprising. Kermack foresaw the forms of some of the most complicated results, doing all the working in his head; he had an altogether exceptional sense of algebraic form, in addition to his penetrating sense of mathematical significance. Functions $\rho(q, p), \sigma(q, p)$ such that $\rho\sigma - \sigma\rho = 1$ we called *conjugate* and the transformation from q, p to ρ, σ we called *canonical*. We were then able to reformulate the work on differential equations in terms of these concepts.

It should perhaps be recalled that by this time Kermack had been blind for 5 years. At the very beginning of his comments, McCrea says

> I understand that when [Kermack] was in hospital after the accident, as soon as he knew he would not see again, he got friends and nurses to read fairly elementary mathematics to him, so that he could train himself to take mathematics in this way and then to think about it for himself. It was incredible to find how much he could do without being able to put anything down on paper.

Kermack's feats are indeed quite remarkable, as anyone who has had to compute with noncommuting differential operators will surely appreciate. As to [26] and [27], here is what McCrea had to say about them:

> In the third paper we showed that in Whittaker's original theorem it is possible to replace the definite integral by a certain differential operator. Perhaps, not unnaturally, this could be thought of as a generalisation of Maclaurin's theorem. We showed that it gave as a particular case R. A. Fisher's rule for the transformation of the moment-generating function in probability theory.
>
> The fourth paper is one with the most general mathematical appeal. It is, of course, well known that the eigen-solutions of an equation of hypergeometric type form a set of functions orthogonal to the eigen-solutions of the adjoint differential equation. It is also well known that such functions satisfy a difference equation in the integral parameter of the set. We were able to show, at any rate formally, that the differential equation and the difference equation form a compatible or 'conjugate' pair in our sense, as do the adjoint pair, and that the two sets of solutions form *normalized* orthogonal sets. This seems to illustrate the unity and coherence that our work was bringing to the theory of special functions. We had plenty of ideas for further developments when circumstances brought our mathematical collaboration to an end.

Actually, there is a fifth paper [30] on the subject, this time by McCrea alone. It was published in the Mathematical Gazette and is of a more expository nature. We will come back to it in §8.3.

5. On the solution of differential equations by definite integrals

In this section we begin to analyse the contents of the papers whose genesis has been described in section 4, beginning with the one by Whittaker which gave rise to the whole sequence. The purpose of Whittaker's paper is clearly stated at its introduction [43, p. 189]:

> The object of the present paper is to communicate a general theorem regarding the solution of both ordinary and partial differential equations by means of families of definite integrals.

As he points out, the motivation is the fact that 'in many cases the solutions of a linear differential equation can be expressed as definite integrals'. Although Whittaker gives his results for any number of variables, we will state them only for two pairs of conjugate variables because this is the only case explicitly described by Kermack and McCrea in their papers. However, as they put it in a footnote to [24, p. 205]

> It is clear [...] that [the theory] may be generalised at once to include any number of pairs.

Consider any contact transformation from a pair of conjugate variables (q, p) to another pair (Q, P), defined by a function $W(q, Q)$. In other words,

$$P = \frac{\partial W}{\partial Q} \quad \text{and} \quad Q = -\frac{\partial W}{\partial q}; \tag{5.1}$$

see (3.4). Denoting by $Q(q, p)$ and $P(q, p)$ the variables when expressed in terms of p and q, the differential equations

$$Q(q, \frac{\partial}{\partial q})\chi = t\chi \quad \text{and} \quad P(q, \frac{\partial}{\partial q})\chi = -\frac{\partial \chi}{\partial t}, \tag{5.2}$$

are compatible and possess a common solution $\chi(q, t)$. Note that the operators that define the equations were obtained substituting p by $\partial/\partial q$ in the expressions of $Q(q, p)$ and $P(q, p)$. Whittaker does not go into any details on how this is to be done, and this is one of the weakest points in his approach. In the next article we will see how Kermack and McCrea proposed to get around this problem using the algebra of differential operators.

Whittaker's General Theorem. *Suppose that we wish to determine a function ψ which satisfies a differential equation*

$$F(q, \frac{\partial}{\partial q})\psi = 0. \tag{5.3}$$

Suppose that, on replacing q and p by their values in terms of P and Q in $F(q, p)$ we obtain $G(q, p)$. Then,

$$\psi(q) = \int \chi(q, t)\phi(t)dt$$

where ϕ is the solution of

$$G(t, \frac{\partial}{\partial t})\phi = 0.$$

Immediately after stating the theorem, Whittaker proceeds to show that it is an extension of the well-known Laplace transform. To see that, choose $W = -qQ$. Then,

$$P = \frac{\partial W}{\partial Q} = -q \quad \text{and} \quad p = -\frac{\partial W}{\partial q} = Q.$$

Replacing p by $\partial/\partial q$ in these equations, we find that the equations in (5.2) become,

$$\frac{\partial \chi}{\partial q} = -t\chi \quad \text{and} \quad -q\chi = -\frac{\partial \chi}{\partial t}.$$

The solution is readily seen to be $\chi = e^{qt}$ where we are, as Whittaker adds in brackets, 'neglecting now, as always, an arbitrary multiplicative constant'. Before we can state the conclusion of the General Theorem for this example, we must compute the operator G. Luckily, in this case all we have to do is replace p by Q and q by $-P$ in F, which gives $G(Q, P) = F(-P, Q)$. Thus, it follows from the General Theorem that if ψ is a solution of a differential equation (5.3), then

$$\psi(q) = \int e^{qt}\phi(t)dt$$

where ϕ is the solution of

$$F(-\frac{\partial}{\partial t}, t)\phi = 0;$$

which shows that Whittaker's theorem is a generalisation of the Laplace Transform, as claimed.

The remainder of the paper, comprising five more sections besides the first three discussed here, contains a a number of more sophisticated examples related to the hypergeometric function (§§4 to 7) and the Bessel function (§8).

6. On Professor Whittaker's solution of differential equations by definite integrals: Part I

As pointed out before, one of the problems with Whittaker's method is, in the words of Kermack and McCrea, how to pass

> from a contact transformation in its *algebraic* form to these partial differential equations, in a manner which is unambiguous and which makes them compatible.

The emphasis is theirs. The introduction of such a method is, according to themselves, the 'first object' of their paper. From now on we refer to this paper as Part I, while its sequel, discussed in the next section, will be called Part II.

6.1. **The method.** Following their approach, we begin with a contact transformation derived from a function $W(q, Q)$. Thus, by §3.2, the transformation is given by

$$P - \frac{\partial W}{\partial Q} = 0 \tag{6.1}$$

$$p + \frac{\partial W}{\partial q} = 0 \tag{6.2}$$

Next take p and P to be the operators $\partial/\partial Q$ and $-\partial/\partial q$, respectively. The simultaneous equations that define the function $\chi = \chi(Q, q)$ will then be

$$\left[P - \frac{\partial W}{\partial Q}\right]\chi = 0 \tag{6.3}$$

$$\left[p + \frac{\partial W}{\partial q}\right]\chi = 0; \tag{6.4}$$

in other words,

$$\frac{\partial \chi}{\partial Q} - \frac{\partial W}{\partial Q}\chi = 0 \tag{6.5}$$

$$\frac{\partial \chi}{\partial q} - \frac{\partial W}{\partial q}\chi = 0; \tag{6.6}$$

which clearly have $\chi = e^W$ for a solution. Then comes the clincher:

> Now this χ will satisfy any equations derived from [(6.5), (6.6)] by multiplying them by any functions of q or Q, or by differentiating them with respect to q or Q, or by any combination of such processes.

Thus, given operators $R(q, -\partial/\partial q)$ and $S(q, -\partial/\partial q)$, we have that any function that satisfies (6.5), (6.6) also satisfies the two equations

$$R\left(q, -\frac{\partial}{\partial q}\right)\left[\frac{\partial}{\partial Q} - \frac{\partial W}{\partial Q}\right]\chi = 0$$
$$S\left(q, -\frac{\partial}{\partial q}\right)\left[\frac{\partial}{\partial q} - \frac{\partial W}{\partial q}\right]\chi = 0.$$

Actually, for the method to work we have to be more careful. We need to know that χ satisfies equations of the form

$$Q\left(q, -\frac{\partial}{\partial q}\right)\chi - Q\chi = 0 \text{ and} \tag{6.7}$$
$$P\left(q, -\frac{\partial}{\partial q}\right)\chi - \frac{\partial \chi}{\partial Q} = 0, \tag{6.8}$$

assuming that they have been determined by the procedure of multiplication by an operator described above. Their next step is to simplify the above equations by writing them in the form

$$[Q(q,p) - Q]\chi = 0 \tag{6.9}$$
$$[P(q,p) - P]\chi = 0, \tag{6.10}$$

where we should not forget that $P = \partial/\partial Q$ and $p = -\partial/\partial q$. Then comes the following rather cryptic comment:

> It is clear that the passage from [(6.3), (6.4) to (6.9), (6.10)] by the stated processes is exactly equivalent to the algebraic solution of [(6.3), (6.4)] for Q, P in terms of q, p provided that the algebraic operations employed are:
> (i) Pre-multiplication by q, Q, p or P or any function of these.
> (ii) Permutation of the variables according to the equations
>
> $$\begin{aligned} qQ - Qq &= 0, & qP - Pq &= 0, \\ Qp - pQ &= 0, & pP - Pq &= 0, \\ qp - pq &= 1, & QP - PQ &= -1. \end{aligned} \tag{6.11}$$
>
> (iii) Equal quantities may be added or subtracted but not multiplied or divided.
>
> These are to replace the ordinary rules of algebra for the present set of variables.
>
> Since q, Q commute it does not matter in what order the terms in $W(q, Q)$ and its derivatives are written.

In modern parlance these are the equations that define the second Weyl algebra, which is the algebra of differential operators of the polynomial ring $\mathbb{C}[q, Q]$. Denoting by $\widehat{f}$ the operator of $\mathbb{C}[q, Q]$ that corresponds to multiplication by the polynomial $f = f(q, Q)$, it is easy to show that this algebra is generated by $\widehat{q}$, $\widehat{Q}$ and the partial differential operators $P = \partial/\partial Q$ and $p = -\partial/\partial q$; see [10, p. 8-10] for example. A straightforward computation shows that these operators satisfy the rules reproduced above.

6.2. **An example.** The paragraphs of Part I quoted above give a sketch of the procedure that Kermack and McCrea proposed to use to compute the operators in (6.9), (6.10). For the sake of clarity, we will perform the required computations in a particular example, as we go through the various stages of the procedure as described in [24, p. 207]. In the example, which is taken from [24, §3.12], $W = q^2Q/4$, so equations (6.1), (6.2) become

$$P - \frac{q^2}{4} = 0 \tag{6.12}$$

$$p + \frac{qQ}{2} = 0 \tag{6.13}$$

The description of the procedure begins as follows:

> We can solve equation [(6.2)], which involves only p, q and Q, for the variable Q according to the rules and obtain
>
> $$Q(q,p) - Q = 0 \tag{6.14}$$
>
> giving equation [(6.9)] at once.

Applying this to the example, we have to solve (6.13) with respect to Q, which gives

$$-2\frac{p}{q} - Q = 0; \text{ so that } \quad Q(q,p) = -2\frac{p}{q}.$$

Note that this is *not* quite as straightforward as it looks because the 'fraction' p/q is not well-defined. Since p and q do not commute, $pq^{-1} \neq q^{-1}p$. Indeed, it follows from (9) that

$$pq^{-1} - q^{-1}p = q^{-2}. \tag{6.15}$$

However, taking into account the meaning of p and q as operators, the most natural choice is $Q(q,p) = -2q^{-1}p$, which is the operator chosen in [24, §3.12, equation (19)]. Thus, in this example, equation (6.9) is given by

$$[-2q^{-1}p - Q]\chi = 0.$$

The description continues as follows:

> From [(6.14)] we obtain, for example,
>
> $$Q(q,p)^2 - Q^2 = 0, \tag{6.16}$$
>
> where we must obey the rules laid down in forming
>
> $$Q(q,p)^2 = Q(q,p) \times Q(q,p)[.]$$
>
> Proceeding in this way we may write in general, if $F(Q)$ is any function of Q,
>
> $$F\{Q(q,p)\} - F(Q) = 0.$$

Let us compute $Q(q,p)^2$ for $Q(q,p) = -2q^{-1}p$. Using (6.15) we find that

$$(-q^{-1}p)(-q^{-1}p) = q^{-1}(p(q^{-1})p = q^{-1}(-q^{-1}p + q^{-2})p.$$

Hence,

$$Q(q,p)^2 = -4q^{-2}p^2 - q^{-3}p.$$

Of course any *polynomial* function of p and q can then be written in the desired form. This is the method they apply to compute (6.10):

> Consequently, if we write [(6.1)] in such a way that in each term of $\partial W(q,Q)/\partial Q$ the part depending on Q comes last, we have merely to substitute for Q its value already found in [(6.14)]. This gives equation [(6.10)], and we have now derived [(6.9), (6.10)] in such a manner that χ is still a solution.

Since, in our example,

$$\frac{\partial W}{\partial Q} = \frac{q^2}{4},$$

which does not depend on Q, we obtain

$$P = \frac{q^2}{4}$$

without any need for further calculations. This example is so simple that it obscures the many difficulties that one may have to overcome in order to apply the procedure. However, Kermack and McCrea are quite clear about these difficulties:

> The processes employed may be called *pre-multiplication* and *post-substitution*. It is evident from the operational form of the equations that we may substitute only for a quantity which immediately precedes the χ. In practical cases we may always check the algebra by verifying that χ remains a solution of any equation obtained. These rules never permit of any ambiguity of interpretation. The only difficulty is that sometimes they do not yield a solution in finite terms of [(6.2)] for Q. This is the case, for example, if one requires the solution of a general quadratic equation. There seems to be nothing corresponding to 'completion of the square' of ordinary algebra, but we may assume there exists a formal solution in series.

Indeed, as we will see later in §9.1 the fact that the authors assume that such 'formal solutions' always exist is one of the weakest points of the paper even when judged by the standards of the 1930s.

6.3. **Definite integrals.** Immediately after the comment quoted above, Kermack and McCrea turn to the analysis of Whittaker's approach to the solution of definite integrals. Let us assume that

$$\psi(q) = \int \chi(q,t)\phi(t)dt \tag{6.17}$$

and that ϕ satisfies a given linear differential equation; say

$$G\left(t, \frac{d}{dt}\right)\phi(t) = 0; \tag{6.18}$$

where $G(Q,P)$ can be written as a sum of monomials of the form P^nQ^m. They begin by studying the action of one of these monomials on $\psi(q)$. Writing both P and Q in terms of q and $p = -\partial/\partial q$, we have

$$P^nQ^m\psi = \int P^n\left(q, -\frac{\partial}{\partial q}\right)Q^m\left(q, -\frac{\partial}{\partial q}\right)\chi(q,t)\phi(t)dt,$$

which equals

$$\int P^n\left(q, -\frac{\partial}{\partial q}\right)t^m\chi(q,t)\phi(t)dt = \int \frac{\partial}{\partial t}P^{n-1}\left(q, -\frac{\partial}{\partial q}\right)t^m\chi(q,t)\phi(t)dt,$$

by the two equations of (6.7). On integrating by parts once, we obtain

$$P^{n-1}\chi \cdot t^m\phi - \int P^{n-1}\left(q, -\frac{\partial}{\partial q}\right)\frac{d}{dt}\left(t^m\chi(q,t)\phi(t)\right)dt;$$

and if we keep doing this we end up with an expression of the form

$$\left[P^{n-1}\chi \cdot t^m\phi - P^{n-2}\chi\frac{d}{dt}(t^m\phi) + \cdots + (-1)^{n-1}\chi\left(\frac{d}{dt}\right)^{n-1}(t^m\phi)\right] + (-1)^n\int \chi\left(\frac{d}{dt}\right)^n(t^m\phi)dt. \quad (6.19)$$

Since these calculations apply to each one of the monomials of G we get that

$$\int \chi(q,t)G\left(t, \frac{d}{dt}\right)\phi(t)dt = \mathbb{I} + G\left(Q\left(q, -\frac{\partial}{\partial q}\right), -P\left(q, -\frac{\partial}{\partial q}\right)\right)\psi(q),$$

where $\mathbb{I}$ is a sum of terms like the one that appears in square brackets in (6.19). Choosing the path of integration C such that $\mathbb{I}$ vanishes on C and taking (6.18) into account,

$$G\left(Q\left(q, -\frac{\partial}{\partial q}\right), -P\left(q, -\frac{\partial}{\partial q}\right)\right)\psi(q) = 0,$$

which is actually the converse of Whittaker's original theorem. But, as the authors point out in §3.2 of their paper, the original statement can be proved in essentially the same way. The result is summarised on page 208 of Part 1 as follows:

> If the function ϕ satisfies a given differential equation $G(Q,P)\phi = 0$, then the definite integral [(6.17)], taken along a suitable path, will satisfy the differential equation obtained by substituting for Q, P in $G(Q,P)$ the values of $Q(q,p)$ and $-P(q,p)$ derived according to the given rules. The order of the factors in terms like P^nQ^m must be preserved.

They give a nice application of this result to the integral

$$\psi(q) = \int_C \exp\left(\frac{q^2t}{4}\right)(1+t)^{-\frac{n}{2}-1}(1-t)^{\frac{1}{2}(n-1)}dt;$$

with C chosen as the 'contour that encircles $t = -1$ and begins and ends at $-\infty$'. Taking $\chi = \exp(q^2t/4)$, it follows that $W = q^2t/4$. Thus, in this case, (5.1) gives $P = q^2/4$ and $p = -qQ/2$; which when solved for Q, P, give $Q = -2q^{-1}p$ and $P = q^2/4$. Now, our choice of χ requires that $\phi(t) = (1+t)^{-\frac{n}{2}-1}(1-t)^{\frac{1}{2}(n-1)}$, which is easily shown to satisfy the differential equation

$$(t^2-1)\frac{d\phi}{dt} - (n + \frac{1}{2} - \frac{3}{2}t)\phi = 0;$$

which corresponds to the differential operator

$$G(Q,P) = (Q^2-1)P - (n + \frac{1}{2} - \frac{3}{2}Q).$$

Substituting the expressions for Q, P in terms of q, p obtained above in this operator, we find that ψ must be a zero of

$$\left(4q^{-1}\frac{d}{dq}\cdot q^{-1}\frac{d}{dq} - 1\right)\frac{q^2}{4} - (n + \frac{1}{2} - 3q^{-1}\frac{d}{dq}).$$

Note that $(q^{-1}d/dq)(q^{-1}d/dq)$ must be computed according to the rules of the algebra established in §6.1. Performing this rather messy calculation, we may write the equation for ψ in the form

$$\frac{d^2\phi}{dt^2} - (n + \frac{1}{2} - \frac{1}{4}q^2)\psi = 0.$$

A look at some standard reference like Whittaker and Watson [42, §16.5, p. 347] shows that this is the *Weber equation*, which received this name because it appears in a paper [40] of H. Weber on the solution of a partial differential equation that arises when one studies certain physical problems (like heat propagation) on cylindrical surfaces. In the language of Whittaker and Watson, the calculations we performed above amount to proving that ψ is a constant multiple of the *parabolic cylinder function* D_n. In other words, we have given an integral representation for D_n. See also [42, Exercise 11, p. 353], where the constant of proportionality is given explicitly.

6.4. **Generalised equations.** It turns out that to subsume all the examples in Whittaker's paper [43] under this scheme it is necessary to consider what Kermack and McCrea call *generalised equations* for the function χ. Let

$$\begin{aligned} P\chi &= h(p,q)\chi \\ Q\chi &= k(p,q)\chi, \end{aligned} \tag{6.20}$$

be a pair of compatible differential equations, and let $\widetilde{\omega}$ be a function of q and p. Recall that we are assuming that $P = \partial/\partial Q$ and $p = -\partial/\partial q$. Note that, with this interpretation, $\widetilde{\omega}$ is actually a *differential operator* on q, p. Multiplying both sides of (6.20) by $\widetilde{\omega}^{-1}$, we have

$$\begin{aligned} \widetilde{\omega}^{-1}P\chi &= \widetilde{\omega}^{-1}h(p,q)\chi \\ \widetilde{\omega}^{-1}Q\chi &= \widetilde{\omega}^{-1}k(p,q)\chi. \end{aligned}$$

Since, according to the relations satisfied by q, p, Q, P, the operator $\widetilde{\omega}$ commutes with both Q and P, these last equations become

$$\begin{aligned} P\widetilde{\omega}^{-1}\chi &= \widetilde{\omega}^{-1}h(p,q)\chi \\ Q\widetilde{\omega}^{-1}\chi &= \widetilde{\omega}^{-1}k(p,q)\chi. \end{aligned}$$

Writing $\chi^* = \widetilde{\omega}^{-1}\chi$ and taking into account that $\widetilde{\omega}^{-1}\widetilde{\omega} = 1$, we end up with a new pair of compatible equation, namely

$$\begin{aligned} P\chi^* &= \widetilde{\omega}^{-1}h(p,q)\widetilde{\omega}\chi^* \\ Q\chi^* &= \widetilde{\omega}^{-1}k(p,q)\widetilde{\omega}\chi^*. \end{aligned}$$

More generally, Kermack and McCrea point out that if the relations

$$\begin{aligned} H(Q,P)\chi &= h(p,q)\chi \\ K(Q,P)\chi &= k(p,q)\chi, \end{aligned}$$

yield a compatible pair of equations with the solution χ then, given a 'function' Π of Q, P, the relations

$$\begin{aligned} \Pi^{-1}H(Q,P)\Pi\chi^* &= \widetilde{\omega}^{-1}h(p,q)\widetilde{\omega}\chi^* \\ \Pi^{-1}K(Q,P)\Pi\chi^* &= \widetilde{\omega}^{-1}k(p,q)\widetilde{\omega}\chi^*. \end{aligned}$$

yield a new compatible pair of compatible equations whose solution is given by $\chi^* = \widetilde{\omega}^{-1}\Pi^{-1}\chi$. For example, by taking $\Pi = Q$, we can transform the system

$$\begin{aligned} P\chi &= qQ^{-1}\chi \\ p\chi &= -\log Q\chi, \end{aligned}$$

into

$$\begin{aligned} q\chi^* &= PQ\chi^* \\ p\chi^* &= -\log Q\chi^*, \end{aligned} \tag{6.21}$$

where $\chi^* = Q^{-1}\chi$. Moreover, since Q^q is a solution of the original system, it follows that $\chi^* = Q^{q-1}$ satisfies (6.21).

In the remainder of their paper, Kermack and McCrea discuss the case whereby the generating function of the contact transformation has P and p as its independent variables (§5 · 1) and the composition of transformations, which they call *successive transformations* (§6 · 1), giving several examples along the way. This last topic leads to a method that can be used to evaluate definite integrals based on the elimination of operators between two given equations (§§6 · 2 and 6 · 3).

7. On Professor Whittaker's solution of differential equations by definite integrals: Part II

Subtitled *Applications of the methods of non-commutative algebra*, *Part II* continues the search for solutions of ever more general differential equations that began in *Part I*. Indeed, as the authors state at the very beginning of the paper:

> It will be shown in the present communication that the necessary and sufficient condition that [the system (6.7)] should be compatible is that
>
> $$Q(q,p)P(q,p) - P(q,p)Q(q,p) = 1, \tag{7.1}$$
>
> regarded as an equation in the non-commutative variables q, p which themselves satisfy the condition
>
> $$qp - pq = 1. \tag{7.2}$$
>
> We shall call functions satisfying this condition *conjugate functions.* From this point of view the method employed by Professor Whittaker in his original paper, involving the use of a contact transformation, was really a particular method of generating conjugate functions. This powerful method may be supplemented and extended by the other methods developed in the following pages.

However, in order to do this they found it necessary to dig deeper into the underlying non-commutative algebra

> In working out this theory it has been found necessary to develop somewhat the algebra of non-commutative variables obeying the law [(7.2)].

The paper is divided into three parts, which we call II.A, II.B, and II.C, that we discuss in some detail below. All the proofs in Part II are formal calculations with expressions in q and p, taking into account only the commutation relation $qp - pq = 1$.

7.1. **Non-commutative algebra.** In Part II.A the authors investigate some properties of the algebra of non-commutative operators that are required for the applications in subsequent parts. They begin by formally defining the derivative of an operator with respect to q and p, denoted by $\delta/\delta q$ and $\delta/\delta p$ respectively, and by proving that the rules they introduce correspond to commuting with the conjugate operator. Thus, if $K(q,p)$ is an element in the non-commutative algebra generated by q and p, then

$$\frac{\delta K(q,p)}{\delta q} = K(q,p)p - pK(q,p);$$

as shown in §3, Theorem I. As Kermack and McCrea themselves point out, this had already been observed by Dirac; see [13] and [14]. The uniqueness of this derivative is proved in Theorem II. Defining *integration* as the 'process inverse to differentiation', it is shown in Theorem III that every function has an integral with respect to q and that this integral is unique up to a function of q. Of course, a similar result holds for integrals relative to p.

Following present terminology we have chosen to refer to $K(q,p)$ above as an element of the non-commutative algebra generated by q and p, but Kermack and McCrea call it a 'function of q and p'. Indeed, although they call attention to some of the problems encountered in calculating with these 'functions', like using the correct relations when swapping sides, they are never clear about the hypotheses these 'functions' have to satisfy for the theory to make sense mathematically. As they put it:

> All mixed derivatives are independent of the order of differentiation, provided the necessary differentiability conditions are satisfied. It would, however, be difficult to formulate conditions of convergence or continuity or differentiability in the present variables. In this paper, we leave aside such considerations and assume that the functions with which we deal are such that we may legitimately perform the required operations upon them; [25, p. 221].

This is essentially Dirac's point of view in [14].

The fact is that it would have been easy to give a formal justification for these results if one could restrict oneself to polynomial functions in q and p. This, however, is not enough for the intended applications. Indeed, most of Part II.A of the paper is concerned with functions defined by series, specially the exponential function. In order to do this, it is necessary first to find a binomial formula for $(p+q)^n$. As a preparation they prove in Theorem IV that if the derivatives of $L(q,p)$ with respect to q and p coincide, then $L(q,p)$ is a function of $q+p$. Theorem V gives the familiarly looking formula

$$\frac{\delta(q+p)^n}{\delta q} = \frac{\delta(q+p)^n}{\delta p} = n(q+p)^{n-1}.$$

An induction is then used in Theorem VI to show that

$$\frac{(p+q)^n}{n!} = \sum_{j=0}^{\infty} \left(\frac{1}{2}\right)^j \frac{1}{j!} \frac{(p_1+q)^{n-2j}}{(n-2j)!}; \tag{7.3}$$

where

$$(p_1+q)^n = \sum_{j=0}^{n} \binom{n}{j} p^{n-j} q^j$$

is the 'usual' binomial formula, written so that p *always* precedes q. Note that (7.3) is a finite sum, since only finitely many of its terms are non-zero.

Turning now to §6, we find the definition and properties of the exponential function in q and p. Take e^q, for instance. Since this is a function of q only, it may be defined with the help of the usual infinite series. The same, of course, holds for e^p. However, problems related to the non-commutativity of q and p take over when we try to find what e^{p+q} should be. The solution is given in Theorems VII and VIII, where it is shown that

$$e^{p+q} = e^p e^q e^{1/2} = e^{-1/2} e^q e^p.$$

The remainder of Part II.A (§8) consists of a long and drawn out formal calculation that leads to a formula for $(q + \phi(p))^n$ in terms of the exponential function of a series, where ϕ is 'any function of p'.

7.2. **The solution of differential equations.** In Part II.B we come to the applications of non-commutative algebra to differential equations. The section begins with two 'preliminary results', the first of which (Theorem X) states that every element of the algebra in q, p is conjugate to q. In other words, given any $\rho(q,p)$, there exists $\widetilde{\omega}(q,p)$ such that

$$\rho(q,p) = \widetilde{\omega}(q,p) q \widetilde{\omega}(q,p)^{-1}. \tag{7.4}$$

Their 'proof' of this result is as follows. First, (7.4) is equivalent to

$$\rho(q,p)\widetilde{\omega}(q,p) = \widetilde{\omega}(q,p)q,$$

while,

$$\widetilde{\omega}(q,p)q = q\widetilde{\omega}(q,p) - \frac{\delta\widetilde{\omega}(q,p)}{\delta p};$$

so that

$$\frac{\delta\widetilde{\omega}(q,p)}{\delta p} - (q-\rho)\widetilde{\omega}(q,p) = 0,$$

which they consider as a differential equation to be solved by the method of undetermined coefficients. In order to do this both $\widetilde{\omega}(q,p)$ and $q - \rho$ are written as power series on p whose coefficients are functions on q. This leads to a recursive system that the authors claim can be formally solved.

Now let $\rho(q,p)$ be any element of the algebra in q, p, and let $\widetilde{\omega}(q,p)$ be chosen so that (7.4) is satisfied. Then, if

$$\sigma(q,p) = \widetilde{\omega}(q,p)p\widetilde{\omega}(q,p)^{-1}$$

it follows that $\rho\sigma - \sigma\rho = 1$; and ρ and σ are said to be *conjugate functions*. This is Theorem XI, which states that for every 'function' $\rho(q,p)$ there is a 'function' $\sigma(q,p)$ such that ρ and σ are conjugate. The converse of this result is given in Theorem XII, where it is shown that every pair of conjugate functions can be obtained by conjugating q and p with an appropriate 'function' $\widetilde{\omega}(q,p)$. Their proof of this result makes use of the exponential defined in the Part II.A.

It is only with Theorem XIII (§10) that we properly turn to the applications of these results to differential equations. More precisely, the authors prove that a pair of equations of the form

$$\begin{aligned} Q(q,p)\chi &= Q\chi \\ P(q,p)\chi &= P\chi \end{aligned} \tag{7.5}$$

is compatible if and only if $Q(q,p)$ and $P(q,p)$ are conjugate as functions of the operators q and $p = -\partial/\partial q$. Writing the commutator $Q(q,p)P(q,p)-P(q,p)Q(q,p)$ as $[Q(q,p), P(q,p)]$, we have from (7.5) that

$$[P(q,p), Q(q,p)]\chi = [Q,P]\chi = \chi;$$

where the last equality follows by taking $P = \partial/\partial Q$ as in Part I. Next, the authors claim that if $[P(q,p), Q(q,p)] \neq 1$ then the previous equation implies that $\chi(q,Q)$ may be written as a product,

$$\chi(q,Q) = \lambda(q)\mu(Q),$$

where $([Q(q,p), P(q,p)] - 1)\lambda = 0$. This, however, is not compatible with the original system, thus proving that $[P(q,p), Q(q,p)] = 1$. Conversely, if $P(q,p)$ and $Q(q,p)$ are conjugate, then there exists a function $\widetilde{\omega}$ such that

$$Q = \widetilde{\omega}(q,p)q\widetilde{\omega}(q,p)^{-1} \quad \text{and} \quad P = \widetilde{\omega}^{-1}(q,p)p\widetilde{\omega}(q,p). \tag{7.6}$$

In this case, (7.5) has a solution given by $\widetilde{\omega}(q,p)e^{-qQ}$; so the equations are indeed compatible.

Theorem XIII is used to give a unified approach to the two methods of constructing contact transformations used in Part I. Recall that, in the first method, the contact transformation was constructed from a generating function $W(q,Q)$; while, in the second, $P(q,p)$ and $Q(q,p)$ were given as in (7.6). In terms of the integral (6.17), the kernel χ (which they call a nucleus) equals e^W in the first method and $\widetilde{\omega}e^{-qQ}$ in the second. The equivalence of these two methods is made explicit in the following theorem.

Theorem XIV. *To any given $\widetilde{\omega}$-function there corresponds a χ-function given by $\chi(q,Q) = \widetilde{\omega}(q,p)e^{-qQ} = e^{-qQ}\widetilde{\omega}(q,Q)$, and to any χ-function there corresponds a $\widetilde{\omega}$-function given by $\widetilde{\omega}(q,Q) = e^{qQ}\chi(q,Q)$.*

The next result (Theorem XV) proposes a method for constructing a family of contact transformations. It is followed by two theorems of a very general nature. Consider a differential equation

$$g(q,p)\psi = 0, \quad \text{where} \quad p = -\partial/\partial q.$$

Theorem XVI: the solution of this equation is given by

$$\psi(q) = \int \widetilde{\omega}(q,p)e^{-qt}dt \tag{7.7}$$

where $\widetilde{\omega}$ is defined by $g(q,p) = -\widetilde{\omega}q\widetilde{\omega}^{-1}$;

Theorem XVII: if $G(Q,P)\phi = 0$ is any other linear differential equation, then there exists a function χ such that

$$\psi(q) = \int \chi(q,p)\phi(Q)dQ, \tag{7.8}$$

for a suitable path of integration.

Thus, while Theorem XVI gives a general formula for the solution of any linear differential equation, Theorem XVII shows that the solutions of any two such equations are connected by an integral formula.

The proofs of both theorems follow by combining previous results from Parts I and II. Thus, by Theorem X, there exists $\widetilde{\omega}$ such that $g(q,p) = -\widetilde{\omega}q\widetilde{\omega}^{-1}$, and by Theorem XIV we have an explicit formula for the corresponding χ-function. Let

$\phi = \phi(t)$ be a function that satisfies $d\phi/dt = 0$. Hence, $P\phi(q) = 0$ and by Part I, the function ψ defined in (7.7) is a solution of $g(q,p)\psi = 0$; which proves Theorem XVI.

In order to prove Theorem XVII we must first introduce the *adjoint* $\widetilde{G}$ of G as the operator obtained by 'reversing the order of each term and changing the sign of P', [25, p. 234]. Next, we determine (Theorem X) $\widetilde{\omega}(q,p)$ and $\Pi(q,p)$ such that

$$g(q,p) = -\widetilde{\omega}q\widetilde{\omega}^{-1} \text{ and } \widetilde{G}(q,p) = -\Pi q\Pi^{-1}.$$

Thus, $\chi = \widetilde{\omega}\Pi e^{-qQ}$ is a solution of the system

$$-\widetilde{\omega}q\widetilde{\omega}^{-1}\chi = \Pi P\Pi^{-1}\chi \tag{7.9}$$

$$\widetilde{\omega}p\widetilde{\omega}^{-1}\chi = \Pi Q\Pi^{-1}\chi \tag{7.10}$$

which, being compatible, define a contact transformation from q,p to Q,P. Thus, by (7.9), $g(q,p) = \widetilde{G}(Q,P)$. But this is exactly the condition required to apply the results of Part I in order to conclude that if ϕ satisfies $G(Q,P)\phi = 0$ then the function ψ defined by (7.8) satisfies $g(q,p)\psi = 0$.

Part II.B ends with an integral formula for the function $\widetilde{\omega}$ that satisfies $Q(q,p) = \widetilde{\omega}p\widetilde{\omega}^{-1}$. Suppose that the function χ corresponding to the given contact transformation is known, and let $\theta(Q,t)$ be the function defined by

$$e^{-qQ} = \int \theta(Q,t)\chi(q,t)dt.$$

Since $\chi = \widetilde{\omega}e^{-qt}$, then

$$\widetilde{\omega}^{-1}e^{-qQ} = \int \theta(Q,t)e^{-qt}dt,$$

is the desired formula.

7.3. **The identical transformation and infinitesimal transformations.** This short part collects a miscellany of results. Of course the identical transformation of the title is the one given by $Q = q$ and $P = p$. Applying Whittaker's Theorem to this transformation it is possible to prove the well-known property of Dirac's δ-function; namely,

$$\psi(q) = \int \psi(t)\delta(t-q)dt.$$

An infinitesimal transformation is a deformation of the identical transformation by a function depending on a parameter ϵ whose square vanishes. The authors prove that, for any $\theta(q,p)$ the formulae

$$Q = q - \epsilon\delta\theta/\delta p \text{ and } P = p + \epsilon\delta\theta/\delta p$$

define a contact transformation. Articles 15 to 17 discuss some identities, including a generalisation of Taylor's Theorem to non-commutative variables. Finally, in §18 the authors state that every algebra whose generators q,r satisfy $qr - rq = F(r,q)$ 'can be reduced to the algebra for which' $qp - pq = 1$. This has to be interpreted with great care because it is not true that every algebra with a relation of this sort is isomorphic to the Weyl algebra; indeed, we need only take $qr - rq = 0$ to see that. The argument given in the paper depends on infinite series in q,p which, as we shall see, can cause all sorts of problems; see §9.1 for more details.

8. Three more papers

In this section we discuss two shorter papers of Kermack and McCrea that explore and extend some of the themes discussed in Parts I and II, and a more elementary exposition of these methods written by McCrea for the *Mathematical Gazette.* The notation we have been using will remain in force throughout this section.

8.1. **An operational method.** The main result of this paper is a formula for ψ in terms ϕ that, unlike (7.7), does not involve integration. In keeping with the policy of Parts I and II we consider only the 2-dimensional case of the result. This is essentially what the authors do for, although the statement is given in full generality in [26, p. 177], they prove it only in dimension two.

Theorem 8.1. *Consider a contact transformation from the variables q, p to the variables Q, P derived from the function $W(q, P)$, so that*

$$Q = \frac{\partial W}{\partial P} \quad \textit{and} \quad p = \frac{\partial W}{\partial q}$$

and suppose that the variables Q, P when expressed in terms of q, p are denoted by $\mathbf{Q}(q, p), \mathbf{P}(q, p)$. Suppose that $\phi(Q)$ satisfies the differential equation $G(Q, P)\phi = 0$, where $P = \partial/\partial Q$. Let $F(q, p)\psi = 0$ be the equation obtained by substituting $\mathbf{Q}(q, p)$ for Q and $-\mathbf{P}(q, p)$ for P, where $p = -\partial/\partial q$. Then

$$\psi(q) = e^{W(q, -\partial/\partial Q)}\phi(Q).$$

Kermack and McCrea give two proofs of this result, both of which take as their starting point the integral expression for ψ in terms of ϕ given in (6.17) and formally deduce from it the required differential expression; see [26, p. 177]. Neither of these proofs is complete, a fact that the authors themselves point out; adding that:

> There is much more that requires investigation about the nature of the functions for which the result holds and the paths of integration used to obtain it, and about the degree of generality when obtained. We do not attempt this investigation, but pass on to applications of the general theorem.

See [26, p. 176]. After a brief discussion of how Theorem 8.1 can be generalised in the spirit of Part II, the authors turn to what they call *special cases.* By taking $W = qP$, they argue in §6 that the theorem can be seen as a generalisation of Taylor's Theorem, which they write in the form

$$\psi(x + h) = e^{hd/dx}\psi(x).$$

In §7 they use their method to give a new proof of a result by R. A. Fisher [16, §10, p. 226ff] on the moment generating function; while §8 is concerned with obtaining a differential equation satisfied by a function series. This in turn leads to a method that allows them 'formally to reduce the solution of any [linear] differential equation [...] to a simple quadrature, together with the solution of a problem in non-commutative algebra'; see [26, §9, p. 184]. The 'problem in non-commutative algebra' is that of finding the 'roots' p of the equation $P = F(q, p) - 1$, where F is a polynomial expression in $q, p = -\partial/\partial q$. Generating functions are the theme of §10, where it is shown that their method can be used to derive differential and difference equations for the functions of q that constitute the coefficients of the power series in t of a generating function $e^{U(q,t)}$.

From §12 (infinitesimal transformations) and throughout the whole of the final section (called 'examples and applications') Kermack and McCrea apply these results to various choices of W, deriving from some of them new proofs of well-known results about special functions such as Bessel's function (§14), Hermite polynomials (§15) and Laguerre polynomials (§16).

8.2. **On compatible differential equations.** The authors explain the aim of this paper as follows:

> In some previous work [i.e. the three papers we have already discussed] we have defined conjugate partial differential equations for a function of two variables. In the present paper we begin by examining more fully the condition for their compatibility and that of their adjoint equations. We then extend the work to the case where, instead of partial derivatives with respect to one of the variables, we have *differences* with respect to that variable. This is the case applicable to discrete eigenvalues. It is well known that the eigensolutions of a linear differential equation involving a parameter are orthogonal to those of its adjoint equation. This parameter can be treated as the second variable in the present work, and a difference equation with respect to its eigenvalues can be written down for the eigenfunctions. What we now show is that the eigensolutions of this differential and difference equation, and those of the adjoint pair of equations, form *normalized* orthogonal sets.

See [27, p. 81]. This is followed by the usual disclaimer:

> We do not examine fully the general conditions under which the latter theorem is true. Our purpose is rather to show how it can be used to predict the existence and form of orthogonality relations, the strict validity of which may require independent investigation.

As the statement above indicates, §2 begins with a study of the system

$$(Q(q,p) - Q)\chi = 0 \tag{8.1}$$

$$(P(q,p) - P)\chi = 0 \tag{8.2}$$

which, by Theorem XIII of Part II (see §7), is compatible if and only if

$$Q(q,p)P(q,p) - P(q,p)Q(q,p) = 1. \tag{8.3}$$

This is followed by one of the main results of the paper.

Theorem II. *If the condition [8.3] is satisfied, and if $Q(q,p)$ is an integral function of p of degree n, then there are in general n, and only n, common solutions $\chi(q,Q)$ of [8.2], linearly independent in both q and Q.*

The proof they proposed is rather clever and deserves to be reproduced here. Following the authors, we give it only for $n = 3$. Since $Q(q,p)$ has degree $n = 3$ in $p = -\partial/\partial q$, it follows that equation (8.2) is a linear differential equation of the third order. As such, it has a basis of three elements, which we denote by χ_1, χ_2 and χ_3. These are functions of q and Q, and every other solution of (8.1) can be written as a linear combination of χ_1, χ_2 and χ_3 with coefficients that are functions of Q. All the computations that follow take place in the algebra generated by q, p, Q, P, subject to the rules stated in §6.1. Multiplying $(Q(q,p) - Q)\chi_j = 0$ respectively by $P(q,p)$ and P, and using the commutation relations, we get the two equations

$$(Q(q,p)P(q,p) - QP(q,p) - 1)\chi_j = 0$$
$$(Q(q,p)P - QP - 1)\chi_j = 0;$$

which, when added up and factored, give rise to

$$(Q(q,p) - Q)(P(q,p) - P)\chi_j = 0.$$

Thus $(P(q,p) - P)\chi_j$ is a solution of (8.1) and, as such, may be written in the form

$$(P(q,p) - P)\chi_j = a_{j1}\chi_1 + a_{j2}\chi_2 + a_{j3}\chi_3.$$

Thus, if $\chi = b_1\chi_1 + b_2\chi_2 + b_3\chi_3$, we obtain from the previous equations that

$$(P(q,p) - P)\chi = \sum_{j=1}^{3}(a_{1j}b_1 + a_{2j}b_2 + a_{3j}b_3 - b_j')\chi_j, \tag{8.4}$$

where $b_j' = db_j/dQ$. Thus χ is also a solution of (8.2) if and only if the bs satisfy the system of differential equations given by the vanishing of the coefficients of χ_1, χ_2 and χ_3 in (8.4). By eliminating two of the bs, say b_1 and b_2, we can produce a linear differential equation of the third order in the remaining b_3. This equation is then solved, and each one of its three solutions will give rise to a triple b_1, b_2, b_3 that is a solution of the required system, thus completing the proof of the theorem.

Adjoint equations (already defined in §7.2 above) are introduced in section 2, where it is also proved that if a pair of equations is compatible then so are their adjoints. Difference equations appear for the first time in section 3 as an equality of the form $(S(q,p) - e^P)\chi = 0$, where e^P is the operator formally defined defined by $e^P f(Q) = f(Q+1)$. Two theorems are then stated. According to the first (Theorem IV) the system

$$(Q(q,p) - Q)\chi = 0 \tag{8.5}$$
$$(S(q,p) - e^P)\chi = 0, \tag{8.6}$$

is compatible if and only if

$$Q(q,p)S(q,p) - S(q,p)Q(q,p) = S(q,p),$$

when $qp - pq = 1$; whilst the second (Theorem V) is the analogue of Theorem II above for this system of equations. Finally, it is pointed out that the adjoint of the operator e^P is e^{-P}; so the adjoints of equations (8.5) are

$$(Q^*(q,p) - Q^*)\chi^* = 0 \tag{8.7}$$
$$(S^*(q,p) - e^{-P})\chi^* = 0, \tag{8.8}$$

The orthogonality of χ and χ^*, mentioned in the introduction to the paper (see quotation above), can be formally stated as

$$\int_a^b \chi(q,Q)\chi^*(q,Q')dq = k\delta_{QQ'},$$

where $\delta_{QQ'}$ is Kronecker's delta symbol. These orthogonality relations are proved in Theorem VII of section 5 by an integration by parts argument. A comment is required concerning the limits of integration a and b. It is assumed that

> the pair of solutions χ and χ^* can be made to satisfy 'suitable boundary conditions at end-points $q = a, q = b$ (say), by restricting Q to assume integral eigenvalues.

See [27, p. 87]. What these 'suitable boundary conditions' are, depends on the equations, some of which are discussed in sections 9 to 12. These include Legendre polynomials (section 10), parabolic cylinder functions (section 11) and Bessel functions (section 12). This leaves out sections 6 to 8, which treat of some possible generalisations of the previously discussed methods. Article 8, on continuous eigenvalues, deserves to be quoted in full:

> A similar theory can be written down *formally* for the case of continuous eigenvalues, with Dirac's δ-function replacing the Kronecker δ-symbol. But the satisfaction of the boundary conditions appears to be more difficult, and we have not succeeded in constructing examples for which it holds.

See [27, p. 89].

8.3. **Operational proofs of some identities.** This is the shortest of the four papers. It contains a brief explanation, by McCrea alone, of the method developed in the previous papers with some applications to special functions (Bessel and Legendre). We quote the first paragraph in full for the opinion it contains on the merits of the method and the status of the theory of special functions in the 1930s:

> The interest of pure mathematicians is nowadays shifting from the study of special functions to general function theory. In fact the study of special functions might ultimately be forsaken, as for example in geometry the study of the triangle is now virtually a closed chapter, were it not for one important consideration. That is, that these functions must always be in constant use in mathematical physics. Hence their properties tend to become questions of practical application rather than of theoretical interest. It follows that there is a demand for some method, if only a formal one, which will derive the required properties in an expeditious manner, without always calling for the services of the professional pure mathematician. The purpose of this note is to show how a beginning might be made by using a calculus of operators. We illustrate the method by obtaining some identities for functions of hypergeometric type.

See [30, p. 43]. This point of view is reinforced at the conclusion, where it is said that

> The point illustrated, however, is that [the identities] can all be *discovered* by the single method of transforming operators.

See [30, p. 45].

9. Discussion

As far as I can determine these papers had essentially no impact on the mathematics of the 20th century. In this section we seek to answer two questions. The first is why this might be so; the second is whether this constitutes a case of a *missed opportunity*, as defined by Freeman Dyson in [15].

9.1. **On rigour, and its absence.** As the quotations given in §§7.1, 8.1 and 8.2 indicate, Kermack and McCrea knew very well that theirs was a formal method, that would require a lot of work to become acceptable by the standards of rigour of the 1930s. We begin with a problem that had already been explicitly discussed in a paper published at about the same time, the convergence of a series in q and $p = -\partial/\partial q$.

Recall that such series are necessary as long as we consider generalised equations of the type introduced in §6.4, which are defined by conjugating with an operator $\widetilde{\omega}(q,p)$. Since all elements in the algebra generated by q, p are implicitly taken to be sums of monomials in q and p, we can only apply the results of Part I if $\widetilde{\omega}(q,p)^{-1}$ can be so written. However, such a sum would very often be infinite; a series, in other words.

The convergence of series in the algebra generated by q, p subject to $pq - qp = 1$ was discussed by D. E. Littlewood in [29, pp. 217–219]. As Littlewood points out, there is no difficulty in defining what it should mean to say that a sum of monomials in q, p is convergent. However, as he puts it:

> If two such series were multiplied, a question of convergence would arise, as indeed it would in the multiplication of two infinite matrices.

See [29, p. 217]. Indeed, it follows from $pq - qp = 1$ that

$$p^n q^n = pq(pq+1)\cdots(pq+n-1).$$

Thus, for example, if we try to write the product $(\sum_{i\geq 0} a_i p^i)(\sum_{j\geq 0} b_i q^i)$ as a series in the monomials $q^k p^\ell$, we must deal with the sum

$$\sum_{i\geq 0} a_i b_i$$

which need not be finite even when $\sum_{i\geq 0} a_i p^i$ and $\sum_{j\geq 0} b_i q^i$ are convergent. Littlewood, however, adds that

> Without making any attempt to obtain best possible conditions, we give a sufficient condition that multiplication should be possible according to the usual rules. If, in the series of the form $\sum a_{m,n} p^m x^n$ for existent terms, $m - n$ is bounded above (alternatively below) products are always convergent, and in the product $m - n$ is still bounded above (below).

Of course x denotes what we have been calling q, and an 'existent term' is one whose coefficient is nonzero.

Actually, as Littlewood also points out, although it is always possible to rewrite a convergent series $\sum a_{m,n} p^n q^m$ into the form $\sum b_{m,n} q^m p^n$, the latter need not be convergent, so that 'it would appear to be necessary to restrict our infinite series to some specific form'; [29, p. 217]. The way out he proposes is to allow only infinite series of the form $\sum b_{m,n} q^m p^n$. Unfortunately, by doing that we introduce zero divisors; for example,

$$\left(1 - qp + \frac{q^2p^2}{2!} - \frac{q^3p^3}{3!} + \cdots\right) q = 0,$$

as one readily checks; [29, p. 218]. Hence:

> Instead of extending our original algebra thus by means of infinite series, we can extend it by defining new quantities which are both right-hand and left-hand inverses of polynomial expressions in x [q in our notation] and p.

See [29, p. 219]. This approach is satisfactory from Littlewood's point of view because his aim is to construct a division ring that contains the ring generated by q, p subject to $pq - qp = 1$. We, however, require more, because our elements must act on functions in some way, and it is not clear what the action of such an inverse should be. For more details on Littlewood's paper, see [11]. In §9.3 we will see how these problems were dealt with in the context of modern algebraic analysis.

Reading these papers in chronological order, one gets the impression that the authors became so deeply aware of the obstacles to formalizing their results that they began to suspect that such a formalization might not be possible. Thus, in Part I they seem rather optimistic, writing that:

> The only difficulty is that sometimes they do not yield a solution in finite terms of [(6.2)] for Q. [...] but we may assume there exists a formal solution in series.

By the time they wrote Part II, the difficulties are spelt out more clearly:

> It would, however, be difficult to formulate conditions of convergence or continuity or differentiability in the present variables.

But are immediately swept under the carpet:

> we leave aside such considerations [of convergence] and assume that the functions with which we deal are such that we may legitimately perform the required operations upon them.

In [26] this has become the less optimistic:

> There is much more that requires investigation about the nature of the functions for which the result holds and the paths of integration used to obtain it, and about the degree of generality when obtained.

And, once again, they make clear that such investigation will not be carried out. A similar observation is made in [27]. By the time McCrea wrote [30] one almost senses a feeling of despondency. To begin with, even the study of special functions seems to be at a low ebb, for

> [t]he interest of pure mathematicians is nowadays shifting from the study of special functions to general function theory.

However, since such functions are still important in physics,

> [i]t follows that there is a demand for some method, if only a formal one, which will derive the required properties in an expeditious manner[...].

Thus,

> [t]he purpose of this note is to show how a *beginning* [my emphasis] might be made by using a calculus of operators.

As the conclusion quoted above makes clear, what is proposed is no more than a method for the *discovery* of these identities, that can then be proved rigorously by pure mathematicians.

One final point concerns the local character of the results. Modern algebraic analysis makes use of sheaf theory to prove results over general manifolds, but

Kermack and McCrea do not even say whether they mean their results to be taken locally or globally. This, however, is not a true fault of theirs, for such considerations would have to wait for the second half of the 20th century. At that time, results were implicitly taken to hold only in the neighbourhood of a point, whenever that may be necessary. Of course there were conflicts. After all, some of the definite integrals are computed along paths that go to infinity; as we saw in §6.3. However, these problems were dealt with case by case.

9.2. **Algebraic analysis.** Although the papers of Kermack and McCrea do not seem to have had any influence on the mathematics of the 20th century (see §9.1), many of their results entered mainstream mathematics as part of what became known as *algebraic analysis*. This is an umbrella term for the study of (holomorphic) linear differential equations by algebro-geometric methods: rings, modules, sheaves and cohomology; see [19] and [20]. More concretely, one can think of algebraic analysis as the study of modules over rings of differential operators. Such a module can be associated to both, a system of linear differential equations and the space in which we seek its solutions. In this language the solutions of a given differential equation are homomorphisms from the module that represents the differential equation to the one where the solutions are sought; see [10] or [7].

Algebraic analysis is a rather young subject, whose history is yet to be written. Luckily, we need little more than a time line of its main achievements up to the early 1990s, for our aim is merely to show in what way the work of Kermack and McCrea can be formalized within this theory.

Algebraic geometrical methods entered the theory of linear differential equations in the late 1950s and early 1960s. In the case of constant coefficients, a differential operator can be considered as a polynomial in the $\partial/\partial x$s. Thus a system of linear differential equations with constant coefficients gives rise to an algebraic variety, that can be used to tease out the properties of the differential system. This is done, for example, in [18, §3.1, p. 211] and [33, §4, p. 96]. At about the same time M. Sato, then a postgraduate student and high school teacher in Japan, was developing the theory of hyperfunctions:

> during the summer of 1957 I tried to prepare something that I could show him [his advisor, Shōkichi Iyanaga], and that was hyperfunctions. I worked out hyperfunction series and outlined the theory for several variables—though the complete theory was finished later, since it required a generalisation of cohomology theory. In December of that year, I went to see Professor Iyanaga, after an interruption of some years, and told him about it.

See [1, p. 211]. Sato's algebro-geometric approach to the theory of linear differential equations was first systematized in the 1970 Master thesis of his student M. Kashiwara. This often quoted work, whose original was handwritten in pencil, was only published in 1995; see [21]. Kashiwara and T. Kawai, also a student of Sato, helped him to develop and write his ideas on microdifferential analysis, which can be roughly described as analysis done in the cotangent bundle. This is the origin of *Microfunctions and pseudo-differential equations* [34] published in 1973. It is in this context that the work of Kermack and McCrea can be formalized as we explain in §9.3. For more details on the work of Sato and Kashiwara see [1], [35], [36] and [37]. In a parallel development, rings of differential operators with polynomial (rather than holomorphic) coefficients were also being studied by I. N. Bernstein in the USSR; see [5] and [6].

Although originally developed in the cradle of partial differential equations, the scope of algebraic analysis widened to include representation theory. Indeed, in the 1980s it provided one of the key ingredients in the solution of the *the Kazhdan-Lusztig conjecture* in the representation theory of Lie algebras; see [4] and [9]. In this context one of the key results of the theory is the *Riemann-Hilbert correspondence*, which is one of the latest developments in a subject that goes back to the work of Gauss and Riemann on the hypergeometric function and includes such classics as Fuch's work on equations with regular singularities and Hilbert's twenty first problem; see [17, chapters I and II] and [8].

9.3. **Kermack-McCrea and algebraic analysis.** In this article we give a rough sketch of the local theory of microdifferential operators and explain how the results in the papers of Kermack and McCrea can be formalized in the context of this theory. As usual, we consider only the one dimensional case. In this section we will follow the modern convention of setting $p = \partial/\partial q$, rather than $-\partial/\partial q$ as Kermack and McCrea do.

Let q, p be coordinates in the cotangent bundle $T^*\mathbb{C} \cong \mathbb{C} \times \mathbb{C}$, and consider the open set of those points $T^*\mathbb{C}$ for which $p \neq 0$. We will work at the neighbourhood of a point in this open set which, without loss of generality, we can assume to be $z = (0, 1)$. A function $f = f(q, p)$, holomorphic in a neighbourhood of z, is *p-homogeneous* of order s if $f(q, p) = g(q)p^s$, where $g(q)$ is holomorphic in a neighbourhood of z. An infinite sum $F = \sum_s f_s(q, p)$ of p-homogeneous functions is a *microdifferential operator* in a neighbourhood of z if

(1) there exists a neighbourhood U of z such that $f_s(q, p)$ is holomorphic and p-homogeneous of order s in U;
(2) there exists an integer t such that $f_s = 0$ for all $s > t$;
(3) there exist constants A and K such that

$$|f_s|_U = \sup\{|f_s(q, p)| \mid (q, p) \in U\} \leq A(|s|!)K^{|s|}, \text{ for all } s.$$

The set $\mathcal{E}$ of all these operators can be made into a ring. The addition is defined as usual, while the multiplication of $F = \sum_s f_s$ and $G = \sum_s g_s$ is given by

$$F \circ G = \sum_{k,s,t} \frac{1}{k!} \frac{\partial^k f_s}{\partial p^k} \frac{\partial^k g_t}{\partial q^k}.$$

A generalisation of these definitions to several variables and a proof that if F and G belong to $\mathcal{E}$, then so does $F \circ G$ can be found in [7, p. 136].

For this ring to play its intended rôle, the commutation relations of §6.1 must hold among its generators. However, by definition,

$$p \circ f(q, p) = f(q, p)p + \frac{\partial f}{\partial q} \text{ while } f(q, p) \circ p = f(q, p)p$$

from which the relation $p \circ q - q \circ p = 1$ immediately follows. Since, under this definition, p plays the part of the operator $\partial/\partial q$, from now on we use ζ to denote the coordinate of $T^*\mathbb{C}$ conjugate to q.

Given an element $F = \sum_s f_s \in \mathcal{E}$, it follows from (2) that there exists a largest integer m such that $f_m \neq 0$ but $f_s = 0$ for all $s > m$. This integer is called the *order* of F, and we write $\sigma(F) = f_m(q, \zeta)$, a polynomial in the commutative variables q

and ζ. Using the order we can define a filtration in the ring $\mathcal{E}$. It can now be proved that if $\sigma(F)(0,1) \neq 0$ then F is invertible in $\mathcal{E}$. Recall that inverses were required in Part II to prove that every 'function' of q,p was conjugated to q; see (7.4). A similar result holds in $\mathcal{E}$:

> if $\sigma(F) = q$, there exists an invertible element $A \in \mathcal{E}$ such that $A \circ F \circ A^{-1} = q$.

Let us now turn to contact transformations. First of all, if $F = \sum_s f_s$ and $G = \sum_s g_s$ have orders, respectively, m and n, then the highest order term in $F \circ G - G \circ F$ is

$$\frac{\partial f_m}{\partial p}\frac{\partial g_n}{\partial q} - \frac{\partial g_n}{\partial p}\frac{\partial f_m}{\partial q},$$

which is equal to the Poisson bracket $\{f_m, g_n\} = \{\sigma(F), \sigma(G)\}$. Suppose now that Φ is an isomorphism of $\mathcal{E}$ that *preserves operator order*; in other words, if $F \in \mathcal{E}$ has order m, than $\Phi(F)$ has order at most m. If $\Phi(q) = Q(q,p)$ and $\Phi(p) = P(q,p)$, then

$$1 = \Phi(p \circ q - q \circ p) = \Phi(p) \circ \Phi(q) - \Phi(q) \circ \Phi(p) = P \circ Q - Q \circ P.$$

Hence,

$$1 = \sigma(P \circ Q - Q \circ P) = \{\sigma(P), \sigma(Q)\}.$$

This implies that the map of $\mathbb{C}^2$ defined by $\phi(q) = \sigma(Q)$ and $\phi(p) = \sigma(Q)$ is a contact transformation, as explained in §3.2. The converse also holds.

Theorem 9.1. *Let f and g be holomorphic functions in q, ζ which are homogeneous in ζ and satisfy $\{f, g\} = 1$. There exists an order preserving isomorphism Φ of $\mathcal{E}$ such that*

$$\sigma(\Phi(q)) = g \text{ and } \sigma(\Phi(p)) = f.$$

Suppose now that Φ is an order preserving isomorphism of $\mathcal{E}$ defined by

$$\Phi(F) = A \circ F \circ A^{-1},$$

for some invertible $A \in \mathcal{E}$. Then, using the multiplicativity of σ,

$$\sigma(\Phi(F)) = \sigma(A \circ F \circ A^{-1}) = \sigma(A)\sigma(F)\sigma(A^{-1}) = \sigma(F),$$

since the symbols are commutative polynomials; see [7, p. 138]. In other words,

> if an order preserving isomorphism Φ is defined by conjugation with an invertible element of $\mathcal{E}$ then $\sigma(\Phi(F)) = \sigma(F)$.

Remarkably, the converse of this result is also true.

Theorem 9.2. *If Φ is an order preserving isomorphism of $\mathcal{E}$ for which*

$$\sigma(\Phi(q)) = \sigma(q) = q$$

then there exists an invertible element $A \in \mathcal{E}$ such that $\Phi(F) = A \circ F \circ A^{-1}$.

Finally, let us turn to Whittaker's theorem on integrals. In order even to state the theorem we ought first to explain how a microdifferential operator acts on a holomorphic function. That, however, is neither very natural, nor very useful for our purposes. Indeed, instead of taking a holomorphic function for the kernel of the integral, the modern version of the theorem takes a microfunction defined in terms of Dirac's δ. For an elementary definition of microfunctions see [10, p. 48ff]. Since $p = -\partial/\partial q$ is surjective as an operator on the space of microfunctions, the integration operator can be defined as its inverse. The result we require is the following:

> Let $W = W(q, p)$ be a function and let $P = P(q, p)$ be a microdifferential operator. Under certain technical hypothesis on W there exists a microdifferential operator $Q(q, p)$ such that
> $$\int P(q,p)\delta(W(q,t))u(t)dt = \int \delta(W(q,t))Q(q,p)u(t)dt.$$
> Conversely, if P is given, a Q can be constructed so that the same equation holds. Furthermore, P and Q are operators of the same order.

For a detailed statement and proof of this result see [22, p. 221]. Thus, if we assume that $u(t)$ satisfies $Q(q, p)u(t) = 0$, then

$$0 = \int P(q,p)\delta(W(q,t))u(t)dt = P(q,p)\left(\int \delta(W(q,t))u(t)dt\right)$$

which is a modern analogue of *Whittaker's General Theorem*; see §4.2.

9.4. **A case of missed opportunity?** In his well-known paper [15], Freeman Dyson introduced the idea of *missed opportunity*, which he defines as

> occasions on which mathematicians and physicists lost chances of making discoveries by neglecting to talk to each other.

A quick glance might suggest that the work of Kermack and McCrea fits this definition fairly well, if by mathematician we understand 'pure mathematician'; but, is this really so? More precisely, had a pure mathematician of the 1930s looked carefully into these papers of Kermack and McCrea, could he or she have developed the necessary machinery and thus discovered the foundations of algebraic analysis thirty years before anyone began to apply modern algebra to linear differential equations? It seems to me that so far as this question has an answer, it must be no. Indeed, it is difficult to imagine algebraic analysis being developed without the aid of cohomology and sheaves, a technology that had not been developed at that time; see [19] and [20].

On the other hand, consider the history of the δ-function. It had been used by Heaviside, and was independently introduced by Dirac in 1927, but had to wait for L. Schwartz's work on distributions in the 1940s before it was accepted as a *bona fide* mathematical concept. That, however, did not stop physicists from using it, nor mathematicians from trying to provide an adequate formalization. The same is true of Feynman integrals, which to this day have not been formalized in complete generality. Nothing like that happened to the work of Kermack and McCrea. Their anticipation of some of the ideas of algebraic analysis was totally forgotten and the developments that led to an adequate formalization of what they had in mind was not influenced by their work.

The difference surely lies in the many applications of both the δ-function and Feynman's integral. Physicists realized they were too useful to be put on hold until they could be rigorously explained to the mathematicians' content. Nothing similar happened to the work of Kermack and McCrea. To a pure mathematician of those times, it must have seemed a rather strange sort of thing to study, with its quaint mixture of analysis and noncommutative algebra. Moreover, most of the work on noncommutative rings before the 1960s was concerned with structure results that would have very little to say about the algebra generated by p and q under the relation $qp - pq = 1$. On top of this, let us not forget that those were the times when algebraic structures were most often studied for their own sake, while the study of microdifferential operators makes very little sense outside the framework of searching for generalised solutions of partial differential equations. Finally, their work was presented as a method to handle properties of special functions, a subject that, as McCrea himself [30] points out, seemed to have exhausted itself at the time.

Summing up, although this may seem at first to be a case of missed opportunity, a deeper look suggests that it would have been very unlikely for anyone at the time to see much of an opportunity in it. There is one final question that we may ask, although no answer can be found in the evidence available: had the papers by Kermack and McCrea been brought to the attention of the early developers of algebraic analysis, would they have influenced them in any way?

9.5. **Conclusion.** In a 'research lecture' that took place in the autumn of 1930 at the University of Edinburgh, E. T. Whittaker presented a new method he had devised for computing certain definite integrals. The method was based on the use of contact transformations, a subject dear to Whittaker, who wrote its first presentation in English in a chapter of his book *Analytical Dynamics*. However, Whittaker's proof of his theorem proved to be faulty, and this prompted two of his auditors to attempt to provide a correct proof. The team consisted of a chemist (Kermack) and a physicist (McCrea), both of whom worked in Edinburgh at the time. Their solution made use of the newly invented algebra of quantum mechanics, which turns out to be the noncommutative algebra generated by the operators that describe position and momentum in quantum dynamics. As part of their attempt they uncovered a number of concepts and results that are now part of algebraic analysis: the study of linear partial differential equations from the algebro-geometric point of view. Among these results, one finds several tools of the theory of microdifferential operators, such as:

- the existence of *quantized contact transformations* associated to a given contact (canonical) transformation;
- the fact that some of these transformations can be given as conjugation by appropriate invertible operators;
- the relation between the solutions of two differential equations that are connected by a quantized contact transformation.

These results are presented in the papers [24] and [25], which were written in what was by then a very unrigorous language. No hypotheses are made concerning the convergence of any of the series that come up in the paper, a problem that is compounded by the difficulties inherent in defining convergent series of noncommuting operators. Although the authors make clear that they are aware of at least some of these problems, they always assume that their results hold as long as the required hypotheses are satisfied. That such hypotheses actually exist one has to take on trust, for they are never investigated.

Even though Kermack and McCrea anticipated a few of the ideas that make up the present theory of rings of microdifferential operators, there is no evidence that their work had any influence in their contemporaries, or at any later time. Remarkable as their work is, it does not seem legitimate to argue that it is a case of missed opportunity in the sense used by Dyson in [15] for, at the time the papers were published, there was no one who could have taken that opportunity. The mathematical technology required to carry out the development of the theory would only be created more than twenty years later. Moreover, the work was presented in the context of the theory of special functions, at a moment when the focus in pure mathematics was shifting to general function theory. This meant that no other potential applications of these methods were explored, so the papers disappeared from view leaving hardly a trace. To complicate matters, algebraists were beginning their study of structures in the 1930s, and this mixture of noncommutative algebra and analysis, with a touch of Hamiltonian mechanics, did not fit the mathematical panorama that the leading mathematicians of those decades were composing. So we seem to have, in these early applications of noncommutative algebra to linear differential equations, a case of non-mathematicians stumbling upon very important results, at a time when these results could not have been derived by standard rigorous methods, even at a great cost of time and effort. That caused them to be forgotten, until they were rediscovered, in a different context, more than twenty years later.

References

[1] E. Andronikof, *Interview with Michio Sato*, Notices of the AMS, **54** (2007), 208–222.

[2] V. I. Arnold, *Mathematical methods of classical mechanics*, second edition, Springer (1989)

[3] J. Barrow-Green, *Whittaker and Watson's 'Modern Analysis'*, European Mathematical Society Newsletter, **45** (2002), 14–15.

[4] A. Beilinson and I. N. Bernstein, *Localisation des G-modules*, C. R. Acad. Sci. Paris 292 (1981), 15–18.

[5] I. N. Bernstein, *Modules over the ring of differential operators: a study of the fundamental solutions of equations with constant coefficients*, Funz. Analysis, Akademia Nauk CCCR **5**(2) (1971), 1–16.

[6] I. N. Bernstein, *The analytic continuation of generalised functions with respect to a parameter*, Funz. Analysis, Akademia Nauk CCCR **6**(4) (1972), 29–40.

[7] J.-E. Björk, *Rings of differential operators*, North-Holland (1979).

[8] A. Borel et al., *Algebraic D-modules*, Academic Press (1987).

[9] J.-L. Brylinski and M. Kashiwara, *Kazhdan-Lusztig conjecture and holonomic systems*, Inv. Math. 64 (1981), 387–410.

[10] S. C. Coutinho, *A primer of algebraic D-modules* , Cambridge University Press, London Mathematical Society Student Texts 33, Cambridge (1995).

[11] S. C. Coutinho, *Quotient rings of noncommutative rings in the first half of the 20th century*, Arch. Hist. Exact Sci. **58**(3) (2004), 255–281.

[12] J. N. Davidson, *William Ogilvy Kermack. 1898-1970*, Biographical Memoirs of Fellows of the Royal Society, **17**, (Nov., 1971), 399–429.

[13] P. A. M. Dirac, *The fundamental equations of quantum mechanics*, Proc. Royal Soc. A **109** (1926), 642–653; also in [39, 307-320].

[14] P. A. M. Dirac, *Quantum mechanics and a preliminary investigation of the Hydrogen atom*, Proc. Royal Soc. A **110** (1926), 561–569; also in [39, p. 417-427].

[15] F. J. Dyson, *Missed oppotunities*, Bull. Amer. Math. Soc., **78** (1972), 635–652.

[16] R. A. Fischer, *Moments and Product Moments of Sampling Distributions*, Proc. London Math. Soc., Series 2, **30** (1929), 199–238.

[17] J. J. Gray, *Linear differential equations and group theory: from Riemann to Poincaré*, second edition, Birkhäuser (2000).

[18] Hormander, L., *On the theory of general partial differential operators*, Acta Math., **94** (1955), 161–248.

[19] C. Houzel, *Les débuts de la théorie des faisceaux*, in [23, p. 7-22].

[20] C. Houzel, *Histoire de la théorie des faisceaux*, in Matériaux pour l'histoire des mathématiques au XXe siécle–Actes du colloque à la mémoire de Jean Dieudonné (Nice 1996) Séminaires et Congrés **3** (1998), 101–119.
[21] M. Kashiwara, *Algebraic study of systems of partial differential equations (Master's thesis, Tokyo University, December 1970*, translated by A. D'Agnolo and J.-P. Schneiders, Mémoires de la Société Mathématiques de France, **63** (1995).
[22] M. Kashiwara, T. Kawai and T. Kimura, *Foundations of algebraic analysis*, translated by G. Kato, Princeton University Press, Princeton (1986).
[23] M. Kashiwara and P. Schapira, *Sheaves on Manifolds*, Springer (1990).
[24] W. O. Kermack and W. H. McCrea, *On Professor Whittaker's solution of differential equations by definite integrals: Part I*, Proc. Edinb. Math. Soc. **2** (1931), 205–219.
[25] W. O. Kermack and W. H. McCrea, *On Professor Whittaker's solution of differential equations by definite integrals: Part II*, Proc. Edinb. Math. Soc. **2** (1931), 220–239.
[26] W. O. Kermack and W. H. McCrea, *An operational method for the solution of linear partial differential equations*, Proc. R. Soc. Edinb. **51** (1931), 176–189.
[27] W. O. Kermack and W. H. McCrea, *On compatible differential equations and the orthogonal properties of their solutions*, Quart. J. Math. Oxford Series, **4** (1933), 81–92.
[28] H. Kragh, *Dirac: A Scientific Biography*, Cambridge University Press (1990).
[29] D. E. Littlewood, *On the classification of algebras*, Proc. London Math. Soc.(2) **35** (1931), 200-240.
[30] W. H. McCrea, *Operational proofs of some identities*, The Math. Gazette, **17** (1933), 43–45.
[31] W. H. McCrea, *Edmund Taylor Whittaker*, Journal of the London Mat. Soc., **32** (1957), 234–256.
[32] D. Martin, *Sir Edmund Taylor Whittaker, F.R.S.*, Proc. Edinburgh Math. Soc. (2) **11** (1958), 1-10.
[33] S. Matsuura, *On general systems of partial differential operators with constant coefficients*, J. Math. Soc. Japan, **13** (1961), 94–103.
[34] M. Sato, T. Kawai and M. Kashiwara, *Microfunctions and pseudo-differential equations*, in Hyperfunctions and pseudo-differential equations, Proc. Conf., Katata, 1971, 265529, Lecture Notes in Math., **287**, Springer (1973).
[35] P. Schapira, *Mikio Sato, un visionnaire des mathématiques*, Gazette des Mathmaticiens, **97** (2003), 23–28.
[36] P. Schapira, *Mikio Sato, a visionary of mathematics*, Notices AMS, **54** (2007), 243–245.
[37] P. Schapira, *Masaki Kashiwara and algebraic analysis*, in Algebraic Analysis and Around: In Honor of Professor Masaki Kashiwara's 60th Birthday (Advanced Studies in Pure Mathematics), World Scientific (2009).
[38] G. Temple, *Edmund Taylor Whittaker*, Biographical Memoirs of Fellows of The Royal Society **2** (1956), 299–325.
[39] B.L. van der Waerden, *Sources of quantum mechanics*, Dover Publications (1967).
[40] H.F. Weber, *Ueber die Integration der partiellen Differentialgleichung* $\partial^2 u/\partial x^2 + \partial^2 u/\partial y^2 + k^2 u = 0$, Math. Ann. **1** (1869), 1-36.
[41] E. T. Whittaker, *A treatise on the analytical dynamics of particles and rigid bodies*, fourth edition, Cambridge University Press, Cambridge, (1999).
[42] G. N. Watson and E. T. Whittaker, *A course of modern analysis*, Cambridge University Press, Fourth Edition (1963).
[43] E. T. Whittaker, *On the solution of differential equations by definite integrals*, Proc. Edinb. Math. Soc. **2** (1931), 189–204.

Departamento de Ciência da Computação, Instituto de Matemática, Universidade Federal do Rio de Janeiro, P.O. Box 68530, 21945-970 Rio de Janeiro, RJ, Brazil.

Programa de Engenharia de Sistemas e Computação, COPPE, UFRJ, PO Box 68511, 21941-972, Rio de Janeiro, RJ, Brazil.
E-mail address: `collier@impa.br`