Arch. Hist. Exact Sci. 59 (2005) 381–436
Digital Object Identifier (DOI) 10.1007/s00407-004-0095-4

Frobenius, Cartan, and the Problem of Pfaff

THOMAS HAWKINS

Communicated by J. J. GRAY

Contents

1. Introduction

Judging by the nomenclature of present-day mathematics Georg Frobenius (1849–1917) made many contributions of lasting significance to mathematics. One thinks for example of the Frobenius substitution in the theory of numbers, Frobenius algebras, Frobenius groups, and Frobenius's complete integrability theorem in differential equations. With the notable exception of his ground-breaking work related to group characters and representations, however, relatively little has been done in the way of an analysis and assessment of Frobenius's mathematics within its historical context, that is, on the one hand an analysis of the considerations that motivated and informed it and on the other hand an assessment of its impact upon subsequent developments.[1] The following study is intended as a small contribution to such an undertaking.[2] It is focused exclusively

[1] Regarding Frobenius's work on group characters and representations, see, eg [17, 38, 39].

[2] I wish to express my gratitude to several individuals and institutions for the role they played in facilitating the final version of this paper: J–P. Serre, whose perceptive critical reading of preliminary versions induced me to revise the text in a variety of ways; S. Sternberg for illuminating discussions of Frobenius's Theorems 8.1 and 9.2; J. Gray for cajoling me into a better, more concrete explanation of now unfamiliar notions; the Department of Mathematics at Brigham Young

upon a paper written by Frobenius in 1876 [26] and published early the next year on what had become known as the problem of Pfaff, a problem in the area of differential equations that was of considerable interest during the 19th century.

Frobenius's paper marked a watershed in his mathematical career in that it marked the beginning of his activity in the theory and application of what would now be described as linear algebra, an area in which he made many important contributions. Most of his earlier work (1870–75) had been devoted to problems involving Weierstrassian analysis, especially as applied to linear differential equations in directions suggested by the work of Fuchs.[3] As we shall see, the linear algebra that Frobenius began to investigate was linear algebra as cultivated at the University of Berlin, where he received his mathematical education; and the disciplinary ideals that underlay Berlin-style linear algebra were a significant factor in motivating his work on the problem of Pfaff. I have already shown this to be the case regarding other work by Frobenius as well as the work of Wilhelm Killing on Lie algebras [40, 41, 42, 43, 46]. In addition to the novelty of how those same ideals inspired Frobenius to work on the problem of Pfaff and informed his approach, here we see as well a greater influence of ideals articulated by Kronecker.

The problem of Pfaff also inspired a paper by Élie Cartan (1869–1951) in 1899 [8] that represented a landmark in his own mathematical career, and more generally in the history of mathematics, because it resulted in his development of the exterior calculus of differential forms, which he went on to use as a tool for investigating systems of partial differential equations, the structure of continuous groups, and algebraic topology. Since Cartan wrote his paper in the light of Frobenius's contribution to the problem, one of the objectives of the following study is to assess the extent to which Frobenius influenced Cartan. A byproduct of this assessment is a history of Frobenius's complete integrability theorem – how it arose from the mathematics surrounding the problem of Pfaff, what Frobenius and Cartan contributed to its formulation, and why it bears Frobenius's name. This assessment also suggests a principal reason why the mathematics of Frobenius has had such a widespread impact upon present-day mathematics. In the concluding section I offer some tentative arguments to this effect.

2. Berlin-style linear algebra

Frobenius grew up in Berlin, and except for one semester spent at the University of Göttingen his entire university education (1867–1870) was at the University of Berlin, to which he retained a connection until called in 1875 to a professorship at the Polytechnicum in Zürich (now the Eidgenössische Technische Hochschule or ETH). During his student days, Berlin was the most vital center for mathematics within Germany. The mathematics faculty was small by today's standards, but it included Kummer, Weierstrass and Kronecker. Weierstrass, in particular, was a charismatic figure, and many students were drawn to Berlin by his presence. Frobenius was one of Weierstrass's many

University for their hospitality during the final stages of work on this paper; the NSF Science and Technology Studies program for partial financial support under grant SES-0312697.

[3] Much of this work is discussed thoroughly by Gray in [35].

doctoral students. His doctoral dissertation (1870) was on the representation of analytic functions by certain types of infinite series of functions.[4] Among Weierstrass's other doctoral students at this time were Georg Cantor, whose Weierstrass-style investigations on the representation of functions by trigonometric series eventually led him to his theory of sets and transfinite numbers, and Wilhelm Killing, whose application of Weierstrass's theory of elementary divisors to the study of pencils of quadric surfaces presaged his ground-breaking work on the classification of simple Lie algebras.

The theory of determinants and, as an application, Weierstrass's theory of elementary divisors was the backbone of Berlin-style linear algebra. Weierstrass had published the theory in a paper of 1868 [73]. The theory provides necessary and sufficient conditions that one family of bilinear forms $\lambda\Phi - \Psi$, where λ is complex, $\Phi = \sum_{i,j=1}^{n} a_{ij}x_i y_j$, and $\Psi = \sum_{i,j=1}^{n} b_{ij}x_i y_j$, can be transformed into another such family by means of (possibly different) nonsingular linear transformations of the x and y variables. The theory was developed under the assumption that Φ is nonsingular, i.e., $\det(a_{ij}) \neq 0$. Weierstrass showed that the necessary and sufficient conditions were that the two families have the same elementary divisors. The elementary divisors were defined in terms of the determinant $\det[(\lambda a_{ij} - b_{ij})]$ and its minor determinants,[5] and the proof that two families with the same elementary divisors could be transformed into one another utilized a canonical form for the matrix $(\lambda a_{ij} - b_{ij})$ that is essentially what has become known as the Jordan canonical form.[6]

In Berlin, Weierstrass's paper was viewed as demonstrating more than purely mathematical theorems. He had demonstrated that it was possible to provide a more systematic and rigorous approach to algebraic analysis than that practiced by most mathematicians in the 18th and early 19th centuries. Characteristic of that earlier practice was what I have called *generic reasoning*. It was a natural byproduct of the analytical revolution of the 17th century set in motion by the work of Viète and especially Descartes.[7] The great power of the method of analysis was that it involved reasoning with abstract symbols rather than with specific numbers or geometrical lines. Its power lay in its generality, which brought with it a tendency to regard the symbols involved as possessing "general" values, thereby drawing attention and concern away from potential difficulties that might arise by assigning certain specific values to such symbols. For example, viewed generically, the roots of the characteristic equation $p(\lambda) = \det[\lambda I - A]$ of a matrix A are distinct, and its determinant is nonzero. Sometimes an author would expressly acknowledge the generic nature of the proposition being set forth by the addition of a phrase such as "in general." Jacobi was a prime example of a mathematician who practiced generic reasoning, which of course could be quite ingenious and sophisticated from a purely formal point of view. Thus he gave determinant-based proofs that a quadratic form could be diagonalized by means of an orthogonal transformation [50] and also

[4] Frobenius's dissertation was in Latin, but he incorporated its results into a paper [24] written in German and published in 1871.

[5] See [46, p. 106] for the definition of elementary divisors and my paper [41] for study of the background to Weierstrass's paper and its influence on the history of linear algebra.

[6] See my paper [41, p. 138] for further details.

[7] See [46, pp. 108–111] for a fuller exposition of generic reasoning and its historical roots.

that a bilinear form could be diagonalized [52]. These results indicate the problem with generic proofs, for the former theorem turns out to be generally true – as Weierstrass first proved in 1858 [72] – whereas the latter is not – as Weierstrass's theory of elementary divisors shows.

Kronecker, who studied the transformation of $\lambda\Phi - \Psi$ in the formidable case in which Φ is singular, seems to have been the main spokesman for the Berlin attitude towards generic reasoning and related matters. In 1874 he wrote, with Weierstrass's 1868 paper expressly in mind [58, p. 405]:

> It is common – especially in algebraic questions – to encounter essentially new difficulties when one breaks away from those cases which are customarily designated as general. As soon as one penetrates beneath the surface of the so-called generality, which excludes every particularity, into the true generality, which comprises all singularities, the real difficulties of the investigation are usually first encountered; but, at the same time, also the wealth of new viewpoints and phenomena which are contained in its depths.

The really interesting – and challenging – mathematics thus involved going beyond the superficialities of the generic case. Weierstrass had showed through his paper on elementary divisors that problems could be systematically and rigorously resolved within a genuinely general context, thereby setting a precedent for his colleagues and students – including Frobenius, who was particularly receptive to the disciplinary ideal implicit in Weierstrass's paper [73] and explicit in Kronecker's above remarks. Indeed, in Sect. 6 we will see Frobenius quoting, in his paper on the problem of Pfaff, another passage from Kronecker's paper [58] involving a different, but related, disciplinary ideal.

As suggested by the above brief summary of Berlin-style linear algebra, the theory of determinants provided the basis for the linear algebra. Indeed, the theory of bilinear forms was classified as a part of the theory of determinants. The standard text on that theory for the Berliners was Richard Baltzer's book, *Theory and Application of Determinants*, which went through several editions beginning in 1857. The third edition [3] was published in 1870 – the year Frobenius obtained his doctorate, and there is little doubt that he had mastered its contents and absorbed its spirit. Although it was not the first treatise on the subject, in composing it Baltzer sought to present the material with Euclidean-style rigor. As he wrote in the preface to the first edition. "In order to unfold the theoretical kernel of the subject most clearly, I have dealt with the main properties of determinants and the algorithms based upon them in a synthetic, precisely articulated format as in the treatises of the ancients" [3, p. v]. In this endeavor he was encouraged initially by his friend Carl Borchardt, Weierstrass's close friend and the Berliner who during this period was editor of the principal journal of the Berlin school, *Journal für die reine und angewandte Mathematik* – or Crelle's *Journal* as it was often called in honor of its first editor. Subsequent editions of Baltzer's book were written in close contact with Kronecker, who provided him with new material drawn from his Berlin lectures. An example of such is Theorem 9.1 below.

3. Mathematical preliminaries & caveats

Before entering into the diverse mathematical treatments of the problem of Pfaff, it will be helpful to make some general comments on the nature of the mathematical reasoning in this period. The theory involves functions $f = f(x_1, \dots, x_n)$ of any number of variables whose properties are not explicitly specified. It is not even clear whether the variables are assumed to be real or are allowed to be complex, although it does seem that the latter possibility is the operative one, since occasionally telltale expressions such as $\log(-1)$ occur.[8] In general, mathematicians in this period tended to regard variables as complex rather than real [35, p. 29]. It is taken for granted that partial derivatives of these functions exist and frequent use is made of the equality of mixed partial derivatives, e.g., $\partial^2 f/\partial x_i \partial x_j = \partial^2 f/\partial x_j \partial x_i$. Also the inverse function theorem and the implicit function theorem are applied whenever needed. In the case of Frobenius, who had been trained in a school emphasizing mathematical rigor, these theorems are never applied without first showing that the requisite Jacobian determinant does not vanish, but even Frobenius never expressly points out that these theorems are local in nature. It seems likely to me that Frobenius regarded the functions under consideration as complex-analytic functions of complex variables $x_1, \dots, x_n$ but continued the tradition of not being explicit about such assumptions.[9] Whether or not he was fully aware of the local nature of his results is far less certain, but the reader should understand them as local results valid in the neighborhood of any point satisfying the specified conditions. It was not until the 20th century that the distinction between local and global results began to be taken seriously by mathematicians.[10]

In Sect. 2, I pointed out that the Berlin school stressed the importance of going beyond the generic case in dealing with algebraic matters. As applied, e.g., to a matrix $A = (a_{ij})$ this meant not thinking of the coefficients a_{ij} as symbols so that ("in general") A has full rank. As we shall see, as a student of that school, Frobenius was careful to base his reasoning upon the rank of A, which is not presumed to be maximal. In the problem of Pfaff, however, matrices arise whose coefficients are functions of $x = (x_1, \dots, x_n)$, so that $A = A(x)$ is likewise a function of these variables. Nonetheless, Frobenius spoke of *the* rank of $A(x)$ without any clarification, whereas (as he certainly realized) it can vary with x in the type of matrices that occur in the theory. For example, in the theory as developed by Frobenius, corresponding to the Pfaffian expression $\omega = 2x_1x_2dx_1 + 2x_1dx_2$ is the matrix of its associated bilinear covariant

$$A(x) = \begin{pmatrix} 0 & x_1 - 1 \\ -(x_1 - 1) & 0 \end{pmatrix},$$

which has rank 2 for points $x = (x_1, x_2)$ with $x_1 \neq 1$ but rank 0 at points with $x_1 = 1$. By *the* rank of these matrices Frobenius evidently meant their maximal rank, so that in the above example the rank of $A(x)$ is 2.

For any matrix $A(x)$, the points at which it has maximal rank r are the points of $\mathbb{C}^n$ that do not lie on the intersection of the hypersurfaces formed by the vanishing of the degree r minors of $A(x)$. Assuming, e.g., that all functions are complex-valued and analytic, which appears to be Frobenius's tacit assumption, the points of maximal rank

8 For example, in a key paper by Clebsch [14, pp. 210–12] discussed in Sect. 5.

9 This assumption is explicitly made in the 1900 treatise on Pfaff's problem by E. von Weber [71, Ch.2, §1], although the local nature of the results is glossed over.

10 How this occurred in the theory of Lie groups is depicted in [45]. See also "local vs. global viewpoint" in the index of [46].

form an open, dense subset G of $\mathbb{C}^n$. It is the points in G that Frobenius was tacitly considering. For purposes of reference I will follow Cartan's book of 1945 [12, p. 45] and refer to them as the *generic points* of $\mathbb{C}^n$. Of course, G depends upon the choice of $A(x)$. It seems reasonable to assume that Frobenius realized in effect that G is open, for all his analytical proofs implicitly depend upon the rank of $A(x)$ maintaining its maximal value in a neighborhood of a fixed point $x_0 \in G$. Of course, G can now be seen to form a manifold in the modern sense with charts being, say, open balls in $\mathbb{C}^n$ on which $A(x)$ has maximal rank.

In what follows I will present the deliberations of the various mathematicians involved more or less as they did, and so these preliminary remarks should be kept in mind. In the case of Frobenius's main analytical theorems, namely Theorems 8.1 and 9.2, besides stating them as he did I have given my interpretation of them as the local theorems that his reasoning implies. As proofs of local theorems Frobenius's reasoning is rigorous in the sense that the necessary details for a proof by present-day standards can be filled in. I do not think this is a coincidence. Based upon my continuing study of a variety of Frobenius's papers on diverse subjects, I would venture to say that Frobenius was very careful to to present his mathematics in a clear and precise manner. I believe that he himself could have filled in the omitted details in his proofs. However, since this was not the custom in the theory of partial differential and Pfaffian equations at the time, he omitted them, content to focus on the algebraic aspects of the problem, which were his primary interest.

One final caveat regarding notation: I have not been entirely consistent in my notation for vectors, including column (i.e., $n \times 1$) matrices and row matrices. Boldface is used only when it makes the mathematics easier to follow and thus only in a portion of Sect. 9.

4. The problem of Pfaff

The problem that became known as Pfaff's problem had its origins in the theory of first order partial differential equations, which as a general theory began with Lagrange.[11] Although other 18th century mathematicians such as Euler had studied various special types of first order partial differential equations, Lagrange was primarily responsible for initiating the general theory of such equations, which I will express with the notation

$$F(x_1, \ldots, x_m, z, p_1, \ldots, p_m) = 0, \quad p_i = \frac{\partial z}{\partial x_i}, \; i = 1, \ldots, m. \tag{4.1}$$

Here z is singled out as the dependent variable and the goal is to obtain a general (or complete) solution $z = \varphi(x_1, \ldots, x_m, C_1, \ldots, C_m)$, where the C_i are arbitrary constants. Here by "obtain a solution" I mean to show how such a solution can be obtained by means of solutions to one or more systems of ordinary differential equations. This is what it meant to integrate a partial differential equation throughout the 18th century

[11] The following introductory remarks are based on the more extensive discussion in Sect. 2 of my paper [44].

and up to the time period of Frobenius. As Lagrange put it, "the art of the calculus of partial derivatives is known to consist in nothing more than reducing this calculus to that of ordinary derivatives, and a partial differential equation is regarded as integrated when its integral depends on nothing more than that of one or more ordinary differential equations" [59, p. 625].

For linear first order equations, Lagrange showed how to do this for any number m of variables. He was, however, less successful in dealing with nonlinear equations and was able to integrate any first order partial differential equation (4.1) only when $m = 2$, i.e., in the case of two independent variables. The integration of nonlinear equations with $m > 2$ was first achieved in 1815 by Johann Friedrich Pfaff (1765–1825), a professor of mathematics at the University of Halle. Pfaff's bold and brilliant idea was to consider the more general problem of "integrating," in a sense to be discussed, a total differential equation

$$\omega = a_1(x)dx_1 + \cdots + a_n(x)dx_n = 0 \tag{4.2}$$

in any number of variables $x = (x_1, \ldots, x_n)$ [66]. The reason he sought to deal with the integration of total differential equations (in the sense explained below) was that he had discovered that the integration of any first order partial differential equation in m variables can be reduced to the integration of a total differential equation in $n = 2m$ variables.[12] Thus, by solving the more general problem of integrating (4.2), he obtained as a special case what had eluded Lagrange and won thereby the praise of Gauss, who described Pfaff's result as "a beautiful extension of the integral calculus" [32, p. 1026].

At the time Pfaff wrote his memoir there was no consensus as what it meant to integrate (4.2) even for $n = 3$. Pfaff observed [66, p. 6] that Euler had expressed the view that it only makes sense to speak of the integration of $\omega = 0$ when $M\omega$ is exact for some nonvanishing factor $M = M(x_1, x_2, x_3)$. This means that if $M\omega = d\Phi$, or equivalently that $\omega = Nd\Phi$ with $N = 1/M$. Then equation $\Phi(x_1, x_2, x_3) = C$ represents an integral of $\omega = 0$ in the sense that for x_1, x_2, x_3 so constrained $d\Phi = 0$ and so $\omega = Nd\Phi = 0$ for these x_1, x_2, x_3. In geometrical terms, the integral $\Phi(x_1, x_2, x_3) = C$ defines a surface with the property that all the vectors (dx_1, dx_2, dx_3) in the tangent plane to the surface at a point $P = (x_1, x_2, x_3)$ lying on it are perpendicular to the vector (a_1, a_2, a_3) evaluated at P, i.e., $a_1dx_1 + a_2dx_2 + a_3dx_3 = 0$.

Pfaff pointed out that Monge had disagreed with Euler and stated that two simultaneous equations $\Phi = C_1$, $\Psi = C_2$ could also be regarded as an integral of $\omega = 0$. That is, viewed geometrically the simultaneous equations $\Phi = C_1$ and $\Psi = C_2$ define a curve. The equation (4.2) stipulates that the tangent space at each point x of a solution manifold should consist of vectors $(dx_1, \ldots, dx_n)$ orthogonal to $(a_1, \ldots, a_n)$. If this is true of the tangent to the above curve, then, from Monge's viewpoint $\Phi = C_1$, $\Phi = C_2$ would constitute a solution to (4.2).

Monge's viewpoint, which Pfaff accepted, can be stated more analytically and more generally as follows. A system of k simultaneous equations

$$\Phi_i(x_1, \ldots, x_n) = C_i, \quad i = 1, \ldots, k \tag{4.3}$$

[12] A description of this reduction is given in [44, §2].

is an integral of $\omega = a_1 dx_1 + \cdots + a_n dx_n$ if: (1) the Φ_i are functionally independent in the sense that the $k \times n$ Jacobian matrix

$$\partial(\Phi_1, \ldots, \Phi_k)/\partial(x_1, \ldots, x_n)$$

has full rank k; (2) For the points satisfying the constraints imposed by (4.3) $\omega = 0$. That is, if (by the implicit function theorem) we express (4.3) in the form $x_i = \Psi(t_1, \ldots, t_d)$, $i = 1, \ldots n$, where $d = n - k$ and $t_1, \ldots, t_d$ denotes d of the variables $x_1, \ldots, x_n$, then setting

$$dx_i = \sum_{j=1}^{d} \frac{\partial \Psi_i}{\partial x_j} dt_j$$

in the expression for ω, makes $\omega = 0$. Of course, as indicated in Sect. 3, all this needs to be understood locally.

In general it is not at all clear how to determine an integral in this sense for a given equation $\omega = 0$. Pfaff's idea was that an integral of $\omega = 0$ is immediate when ω has a simple form. Consider, for example

$$\omega = dx_1 + x_2 dx_3 = 0. \tag{4.4}$$

Then it turns out that $\omega = 0$ does not have an integral in Euler's sense, i.e., $\omega \neq N d\Phi$.[13] On the other hand, the simultaneous equations $x_1 = C_1$, $x_3 = C_3$ define a curve on which $dx_1 = dx_3 = 0$ and so $\omega = 0$ there. Pfaff's idea was to show that by a suitable variable change a total differential ω could "in general" be put in a comparably simple form that made finding an integral a trivial matter.

The generic theorem implicit in Pfaff's memoir may be stated as follows using the sort of index notation introduced later by Jacobi.

Theorem 4.1 (Pfaff's theorem) *"In general" a change of variables $x_i = f_i(y_1, \ldots, y_n)$, $i = 1, \ldots, n$, exists such that ω as given by (4.2) transforms into an expression involving m differentials*

$$\omega = b_1(y)dy_1 + \cdots + b_m(y)dy_m, \tag{4.5}$$

where $m = n/2$ if n is even and $m = (n + 1)/2$ if n is odd.

Here it is tacitly understood that a bonafide variable change has a nonvanishing Jacobian, since the inversion of the variable change is necessary to produce an integral of $\omega = 0$. That is, if $y_i = \Phi_i(x_1, \ldots, x_n)$, $i = 1, \ldots, n$, denotes the inverse of the variable transformation posited in Pfaff's theorem, and if $C_1, \ldots, C_m$ are constants, then the m simultaneous equations $\Phi_i(x_1, \ldots, x_n) = C_i$, $i = 1, \cdots, m$, represent an integral of $\omega = 0$ because these equations imply that $dy_i = 0$ for $i = 1, \ldots, m$, and whence by (4.5) that $\omega = 0$ for the values of $(x_1, \ldots, x_n)$ satisfying these equations. This solution

[13] The necessary condition for $\omega = N d\Phi$ was given in an elegant form by Jacobi and is displayed below in (4.6). The Pfaffian equation (4.4) fails to satisfy this condition.

can be thought of geometrically as the integral manifold formed by the intersection of m hypersurfaces $\Phi_i = C_i$ and hence "in general" of dimension $d = n - m$. In the case of ordinary space ($n = 3$), $m = 2$ and so $d = 1$, i.e., the solution to the generic equation $\omega = 0$ in this case is a curve – the sort of solution envisioned by Monge.

Pfaff pointed out [66, p. 7] that exceptional cases existed for which the number m of terms in (4.5) could be less than $n/2$ or $(n+1)/2$, respectively. Indeed, as explained below, when $n = 3$ in (4.2) the case envisioned by Euler as the sole meaningful one, namely $\omega = N d\Phi$, means that a variable change, namely $y_1 = \Phi(x_1, x_2, x_3)$, $y_2 = x_2$, $y_3 = x_3$, exists so that $\omega = N(y_1, y_2, y_3) dy_1$ and thus $m = 1 < (3+1)/2$, whereas Monge had argued for the legitimacy of the case $m = 2 = (3+1)/2$, i.e., what turns out in Pfaff's Theorem to be the "general" case. Although Pfaff recognized exceptions to his theorem, he restricted his attention to the generic case stated therein. He worked out detailed, successive proofs for the generic cases of $n = 4, \ldots, 10$ variables in (4.2) and then stated the general generic theorem along with a brief proof-sketch reflecting the approach detailed in the worked out cases $n = 4, \ldots, 10$ [66, §16].

In an important and influential paper of 1827 [49, p. 28] Jacobi expressed the necessary condition that Euler's relation $\omega = N d\Phi$ hold in the following elegant form:

$$a_1 a_{23} + a_2 a_{31} + a_3 a_{12} = 0, \quad \text{where} \quad a_{ij} = \frac{\partial a_i}{\partial x_j} - \frac{\partial a_j}{\partial x_i}. \tag{4.6}$$

The above expressions a_{ij} were introduced by Jacobi with the notation (i, j) for a_{ij} and were defined for Pfaffian equations $\omega = 0$ in any number n of variables. Perhaps consideration of the a_{ij} was motivated initially by the fact that when ω is exact all $a_{ij} = 0$ necessarily. Then, more generally, (4.6) gives the necessary condition for Euler's condition to hold. Thus in general the a_{ij}, along with the a_i, seem to contain the information needed to decide about the nature of the integrals of $\omega = 0$.

The main object of Jacobi's paper [49] was not (4.6) but a general proof, using the elegant n-variable notation he had introduced into analysis, of Pfaff's Theorem in the case of $n = 2m$ variables – the case that was relevant to its application to partial differential equations. In Pfaff's proof-method the reduction from $n = 2m$ to $n - 1$ differential terms was attained by a variable change provided by the complete solution to a system of ordinary differential equations, and in Jacobi's rendition he introduced the coefficient system

$$A = (a_{ij}), \quad a_{ij} = \frac{\partial a_i}{\partial x_j} - \frac{\partial a_j}{\partial x_i}, \tag{4.7}$$

to write down and manipulate the system of differential equations. Since $a_{ji} = -a_{ij}$, the $n \times n$ matrix A is what is now called skew symmetric. In what is to follow I will refer to it as *Jacobi's skew symmetric matrix.*

With the skew symmetry of A evidently in mind, Jacobi remarked that the system (4.7) shows "great analogy" with the symmetric linear systems that had arisen in diverse analytical applications [49, p. 28],[14] and at first he apparently believed that skew

---

[14] Although Jacobi implied there were many such applications, the only one he explicitly mentioned was the method of least squares, presumably because of the symmetric normal equations $M^t M\mathbf{x} = M^t\mathbf{b}$ used to obtain a least squares solution to the linear system $M\mathbf{x} = \mathbf{b}$.

symmetric systems were new in applications of analysis.[15] In any case, he realized, as we shall see below, that the properties of skew symmetric matrices associated to determinants implied many many interesting algebraic relationships. Jacobi stressed that it was the introduction of the skew symmetric system (4.7) and its application to Pfaff's procedure that constituted the original contribution of his paper. Since he was proceeding on the generic level of Pfaff's Theorem he took it for granted that $\det A \neq 0$, and he realized that this was permissible because n was assumed even and that when n is odd the skew symmetry of A forces $\det A$ to vanish identically, i.e., no matter what values are given to the a_{ij}.[16] He did briefly discuss what happens when n is odd so that $\det A = 0$, and again it was on the generic level. Expressed in more modern terms, Jacobi's tacit assumption was that A has rank $n - 1$ when n is odd [49, p. 28].

Pfaff's method of integrating (4.1), namely by constructing the variable change leading to the normal form (4.5), reduced the integration of $\omega = 0$ to the integration of systems of ordinary differential equations, which was also the generally accepted goal in the theory of partial differential equations as he noted. The variable change of Theorem 4.1 was obtained by a succession of $n - m$ variable changes, the kth of which transformed ω from an expression involving $n - k$ differentials into one involving $n - k - 1$ differentials, where $k = 0, \ldots, n - (m + 1)$. Each such variable change required completely integrating a system of first order ordinary differential equations.

For $n \gg 2$ Pfaff's method required the complete solution of a large number of systems of ordinary differential equations, but it was not until after Cauchy and Jacobi had discovered direct methods of solving nonlinear partial differential equations (4.1) in any number of variables that the inefficiency of Pfaff's method was made evident by the new methods. After Jacobi published his method in 1837 the goal in the theory of first order partial differential equations became to devise methods for integrating them that were as efficient as possible, i.e., that reduced the number of systems of ordinary differential equations that needed to be considered, as well as their size and the number of solutions to such a system that is required by the method. In this connection Jacobi devised a new method that was published posthumously in 1862 and which spurred on the quest for more efficient methods in the late 1860s and early 1870s.[17]

Although the original motivation for Pfaff's theory – the integration of first order partial differential equations – had lost its special significance due to the later direct methods of Cauchy and Jacobi, during the 1860s and 1870s the interest in partial differential equations extended to Pfaffian equations as well, since the papers by Pfaff and Jacobi raised several questions. First of all, there was the matter of the admitted exceptions to Pfaff's generic theorem. Given a Pfaffian equation $\omega = 0$, how can one tell whether or not Pfaff's theorem applies, and, if it does not, what can one say of the integrals of $\omega = 0$? For example, what is the minimal number k of equations $f_i(x_1, \ldots, x_n) = C_i$,

[15] Thus at the conclusion of his paper he made a point of noting that after he had written it he discovered that Lagrange and Poisson had already introduced skew symmetric systems in their work on the calculus of variations.

[16] Since $A^t = -A$, $\det A = \det A^t = \det(-A) = (-1)^n \det A$. Thus when n is odd $\det A = 0$ follows.

[17] On Jacobi's two methods and their influence see Sects. 2–3 of [44].

$i = 1, \ldots, k$, that are needed to specify an integral? If we express this question in geometrical terms – something rarely done in print in this period – the question becomes: what is the maximal dimension $d = n - k$ of an integral manifold for $\omega = 0$? Secondly, there was the problem of determining more efficient ways to integrate a Pfaffian equation, in the generic case and also the nongeneric ones. These two questions constituted what was known as "the problem of Pfaff" in the period leading up to the work of Frobenius.

Nowadays the problem of Pfaff is characterized more specifically as the problem of determining for a given Pfaffian equation $\omega = 0$ in n variables the maximal dimension d of its integral manifolds [47, p. 1623] or equivalently, since $d = n-k$, the minimal value possible for k. Thus in the modern conception of the problem, the second "efficiency" question is ignored. As we shall see, Frobenius also ignored it, but, utilizing the work of his predecessors – above all that of Clebsch – and spurred on by the disciplinary ideals of the Berlin school, he definitively and rigorously resolved the first problem.[18] Frobenius showed that the integration of $\omega = 0$ depends upon the class number p of ω. This may be defined by $p = \text{rank } (A \quad a)$, where A is the Jacobi matrix of $\omega = \sum_{i=1}^{m} a_i(x)dx_i$ and a is the column matrix $(a_1, \ldots, a_n)^t$ (t denoting matrix transposition).[19]

By way of an introduction to Frobenius's results, let us consider the case $n = 5$, so $\omega = a_1(x)dx_1 + \cdots + a_5(x)dx_5$. Pfaff's Theorem gives the generic value of $k = m = 3$ differential terms in (4.5) and so $m = 3$ equations defining an integral of $\omega = 0$ and thus an integral manifold of dimension $d = 5-3 = 2$. Frobenius showed that a variable change $x_i = f_i(y_1, \ldots, y_5)$ exists[20] such that, in the z-variables one has, depending on the value of p:

$$p = 1 \Rightarrow \omega = dz_5, \quad p = 2 \Rightarrow \omega = z_2 dz_1, \quad p = 3 \Rightarrow \omega = dz_5 + z_3 dz_2,$$

$$p = 4 \Rightarrow \omega = z_3 dz_1 + z_4 dz_2, \quad p = 5 \Rightarrow \omega = dz_5 + z_3 dz_1 + z_4 dz_2.$$

Let $z_i = g_i(x_1, \ldots, x_5)$ denote the (local) inverse. Then in the case $p = 1$, $\omega = dg_5$ is exact with integral manifold $g_5(x_1, \ldots, x_5) = C$ of dimension $d = 5-1 = 4$. The case $p = 2$ covers the Pfaffian equations integrable in the sense of Euler since $\omega = N(x)d\Phi$ with $N(x) = g_2(x)$ and $\Phi = g_1(x)$. Here also $k = 1$ and $d = 4$. When $p = 3, k = 2$, and an integral manifold is defined by the simultaneous equations $g_2(x) = C_1, g_5(x) = C_2$ and so has dimension $d = 3$. Likewise when $p = 4$, $k = 2$ and the integral manifold $g_1(x) = C_1, g_2(x) = C_2$ has dimension $d = 3$. Finally, when $p = 5$, $k = m = 3$ and we are in the generic case of Pfaff's Theorem with integral manifold defined by three equations and so of dimension $d = 5 - 3 = 2$. Thus Pfaff's Theorem covers just one of the five possibilities distinguished by Frobenius's results as expressed below in Theorem 8.1 in the case of $n = 5$ variables.

[18] Rigorously, in the sense indicated in Sect. 3, which for the 19th century was indeed exceptionally rigorous.

[19] Frobenius actually defined p as the average of the ranks of two matrices A and $\widehat{A}$ (Sects 7–8), the second of which involves A and a, but realized the above simpler characterization of p as well (as explained in a footnote relating to Lie's work in Sect. 10).

[20] Strictly speaking, such a transformation exists in the neighborhood of any generic point as explained in the discussion of Frobenius's Theorem 8.1.

5. The contributions of Clebsch

After Jacobi, significant advances on Pfaff's problem were not made until the early 1860s, when two mathematicians, Clebsch and Natani, independently and almost simultaneously took up the matter.[21] Natani's paper appeared first, but it was the work of Clebsch that made the greatest impression upon Frobenius and so will occupy our attention here. Some aspects of Natani's work will be discussed briefly in Sect. 9.

Alfred Clebsch (1833–1873) had obtained his doctorate in mathematics in 1854 from Königsberg in the post-Jacobi era. His teachers were Franz Neumann, Richelot and Hesse (who had been Jacobi's student). At Königsberg Clebsch received a broad and thorough training in mathematical physics, which included on his part a detailed study of the publications of Euler and Jacobi. He was known personally to the mathematicians in Berlin, where Jacobi had ended his career, since during 1854–58 he had taught in various high schools there as well as (briefly) at the University. Apparently Borchardt, as editor of Crelle's *Journal*, asked him to edit the manuscript by Jacobi that contained his new method of integrating partial differential equations so that it might be published in his journal, where it did in fact appear in 1862 [53].[22]

The study of Jacobi's new method led Clebsch to ponder the possibility of its extension to the integration of a Pfaffian equation

$$\omega = a_1(x)dx_1 + \cdots + a_n(x)dx_n = 0. \tag{5.1}$$

Such an extension would greatly increase the efficiency of integrating a Pfaffian equation in the sense explained above. Clebsch satisfied himself that he could do this and in fact do it in complete generality – not just for the case of an even number of variables with a nonsingular Jacobi matrix $A = (a_{ij})$. Thus he wrote that, "The extension of this method [of Jacobi] to the problem of Pfaff in complete generality and in all possible cases is the subject of the following work" [14, p. 193].

Clebsch's entire treatment of the problem was based upon a distinction he did not adequately justify. Let m denote the minimum number of differential terms into which ω can be transformed by a variable change so that $\omega = F_1(x)df_1 + \cdots + F_m(x)df_m$ and $2m \leq n$. Consider the $2m \times n$ Jacobian matrix $\partial(f_1, \ldots F_m)/\partial(x_1, \ldots, x_n)$. Then (to use modern terms) either (I) this matrix has full rank or (II) it does not have full rank. In case (I) the $2m$ functions $f_1, \ldots, F_m$ are by definition *independent*. In the "indeterminate case" (II) they are not, but Clebsch went on to claim without adequate justification that it is always possible in this case to take $F_1 \equiv 1$ with the remaining functions independent [14, pp. 217–20]. For convenience of reference I will refer to the proposition implicit in Clebsch's remarks as Clebsch's Theorem. With a change of notation in case (II) it may be stated in the following form.

Theorem 5.1 (Clebsch's theorem) *Let m denote the minimal number of differential terms into which ω can be transformed. Then either (I) $2m$ independent functions $f_1, \ldots, f_m, F_1, \ldots, F_m$ exist such that*

---

[21] What follows does not represent a full account of work on the problem of Pfaff up to that of Frobenius. For more details see [22, 71].

[22] For further details about Clebsch's career, see the anonymous memorial essay [2] and [70, p. 7ff].

$$\omega = F_1(x)df_1 + \cdots + F_m(x)df_m \tag{5.2}$$

or (II) $2m + 1$ *independent functions* $f_0, f_1, \ldots, f_m, F_1, \ldots, F_m$ *exist such that*

$$\omega = df_0 + F_1(x)df_1 + \cdots + F_m(x)df_m. \tag{5.3}$$

Clebsch stressed that (I) and (II) represented two general and quite distinct classes into which Pfaffians ω are divided, and he was apparently the first to emphasize this distinction [14, p. 217]. He called them the *determinate* and the *indeterminate* cases, respectively, and pointed out that by starting from the assumption of Theorem 5.1, "one is spared the trouble of carrying out direct proofs that lead to very complicated algebraic considerations, which, to be sure, are of interest in their own right ..." [14, p. 194]. This led him to later refer to this approach as his *indirect method.*

Clebsch's Theorem is evidently equivalent to asserting that a change of variables $x = \varphi(z)$ is possible so that

$$\text{I}: \ \omega = z_{m+1}dz_1 + \cdots + z_{2m}dz_m$$

$$\text{II}: \ \omega = dz_0 + z_{m+1}dz_1 + \cdots + z_{2m}dz_m$$

For example, in case I, we can add $n-2m$ functions $g_1, \ldots, g_{n-2m}$ so that $z_1, \ldots, z_n \stackrel{\text{def}}{=} f_1, \ldots, F_m, g_1, \ldots, g_{n-2m}$ have nonvanishing Jacobian determinant and then apply the inverse function theorem to get the above-mentioned (local) variable change $x = \varphi(z)$. He believed that he could distinguish the determinate and indeterminate cases by means of the $n \times n$ Jacobi skew symmetric matrix $A = (a_{ij})$ associated to ω by (4.7). According to him, the determinate case (I) occurred when (in modern parlance) the rank of A is $2m$ [14, p. 208]. That is – as articulated in Clebsch's time – when all $k \times k$ minors of A with $k > 2m$ vanish, but some of the $2m \times 2m$ minors do not. In his own paper on the problem of Pfaff [26], Frobenius also used similar cumbersome language, but in a brief sequel submitted seven months later (June 1877), he introduced the notion of the rank of a matrix so that (by his definition of rank) when, e.g., the minors of the matrix A have the above property it is said to have rank $2m$ [29, p. 435]. For conciseness I will use Frobenius's now-familiar rank terminology in what follows, even though it is slightly anachronistic. As for the indeterminate case (II), Clebsch claimed that it corresponded to the case in which the rank of A is $2m + 1$ [14, p. 218], although this turns out to be impossible, since (as Frobenius was to prove) the rank of any skew symmetric matrix is always even.

In either case (I) or (II), if $z_i = f_i(x)$ for $i \leq m$, then the system of integrals $f_i(x) = C_i$ defines a solution manifold for $\omega = 0$. Taking the above theorem for granted, Clebsch turned to the real problem of interest to him, namely to determine the functions f_i by means of solutions to systems of first order ordinary differential equations – and in a way that was more efficient than Pfaff's original method. Clebsch's idea in case (I) was first to determine one of the functions f_i, e.g., f_m. Then the equation $f_m(x_1, \ldots, x_n) = C$ is used to eliminate, e.g., x_n, thereby diminishing both the number of variables and the number m of differential terms by one unit. This new, reduced case could then be handled in the same manner to get f_{m-1}, and so on until all m functions f_i were determined [14, p. 204].

In his indirect method the function f_m was a solution to a system of linear homogeneous partial differential equations, i.e., equations of the form (in Jacobi's notation)

$$A_i(f) = 0,\ i = 1, \ldots r,\ \text{where } A_i(f) = \sum_{j=1}^{n} \alpha_{ij}(x) \frac{\partial f}{\partial x_j}. \tag{5.4}$$

Whereas a single such equation $A(f) = 0$ was known already in the 18th century to be equivalent to a system of first order ordinary differential equations – so that a solution was known to exist – a simultaneous system need not have a solution. Such systems had already occurred in Jacobi's new method,[23] and Jacobi showed that when the r equations in (5.4) are linearly independent,[24] then $n - r$ independent solutions exist provided the system satisfies the integrability condition

$$A_j[A_k(f)] - A_k[A_j(f)] \equiv 0 \quad \text{for all} \quad j \neq k. \tag{5.5}$$

For systems satisfying this condition, Jacobi sketched a method for finding the solutions via integration of systems of ordinary differential equations. Clebsch argued that the system $A_i(f) = 0$ he had arrived at also satisfied Jacobi's integrability condition, and so a solution f_m could be obtained by Jacobi's method of integration, which Jacobi had already proved to be "efficient" for integrating nonlinear partial differential equations. However, the reasoning leading to these equations in the nongeneric cases was vague and sketchy and reflected an incorrect understanding of the algebraic implications of cases (I) and (II) of Theorem 5.1, implications that Frobenius first correctly determined (Theorems 7.2 and 8.1). Thus Clebsch's proof that the partial differential equations at which one arrives satisfy Jacobi's integrability condition is not rigorous.

In a paper of 1863 Clebsch himself expressed dissatisfaction with his indirect method because it assumed the forms (I) and (II) of Theorem 5.1 rather than revealing how they are obtained. "In the present essay I will therefore directly derive these defining equations of the problem of Pfaff" [15, p. 146]. One key point he failed to mention in his introductory remarks was that he had only succeeded in providing a direct derivation of these defining equations in the case in which n is even and $\det A \neq 0$, i.e., the generic case dealt with by Jacobi in 1827. Within that limited framework, however, his direct approach was indeed far more satisfying.

In Jacobi's new method the systems $A_i(f) = 0$ had been defined in terms of the Poisson bracket operation [46, p. 47]. Clebsch introduced two bracket operations as analogs of the Poisson bracket.[25] They may be defined as follows. Let $n = 2m$ so $\omega = \sum_{i=1}^{2m} a_i(x) dx_i$, and let $A = (a_{ij})$ denote the associated skew symmetric Jacobi matrix, which is now assumed to be nonsingular. Let $C_{ij}(A)$ denote the (i, j) cofactor of A, and for any functions φ and ψ of $x_1, \ldots, x_{2m}$ set

[23] See [44, p. 209] for details.

[24] By "linearly independent" I mean that the $r \times n$ matrix (α_{ij}) has full rank r.

[25] These operations were already introduced by Clebsch in his 1861 paper [14, p. 243]. There they are defined in terms of Pfaffians, but the equivalence with the above was also indicated by Clebsch [15, p. 148].

$$(\varphi) = \sum_{i,j=1}^{2m} \frac{C_{ij}(A)}{\det A} a_i \frac{\partial \varphi}{\partial x_j}, \quad [\varphi, \psi] = \sum_{i,j=1}^{2m} \frac{C_{ij}(A)}{\det A} \frac{\partial \varphi}{\partial x_i} \frac{\partial \psi}{\partial x_j}.$$

The bracket operation $[\varphi, \psi]$ is a generalization of the Jacobi bracket and, like it, satisfies

$$[\varphi, \psi] = -[\psi, \varphi] \text{ and } [[\varphi, \psi], \chi] + [[\psi, \chi], \varphi] + [[\chi, \varphi], \psi] = 0. \tag{5.6}$$

In addition, there is a new identity involving both bracket operations:

$$[(\varphi), \psi] - [(\psi), \varphi] = ([\varphi, \psi]) + [\varphi, \psi]). \tag{5.7}$$

The main theorem that results from the direct method may be stated as follows [15, p. 153].

Theorem 5.2 (Clebsch's direct method) *If $f_1, \dots, f_m$ are functionally independent and satisfy the relations*

$$(f_i) = 0 \quad \text{and} \quad [f_i, f_j] = 0 \quad \text{for all} \quad i, j, \text{ with } i \neq j \tag{5.8}$$

then functions $F_1, \dots, F_m$ exist so that $f_1, \dots, f_m, F_1, \dots, F_m$ are functionally independent and $\omega = F_1 df_1 + \cdots + F_m df_m$.

In view of this theorem, what remains is the determination of the functions f_i. Clebsch observed that these functions are solutions to the progression of systems of linear homogeneous partial differential equations implied by (5.8) [15, p. 168ff]. The equations (5.8) may be written as follows.

(1) $(f_1) = 0$
(2) $(f_2) = 0,\ [f_2, f_1] = 0$
(3) $(f_3) = 0,\ [f_3, f_1] = 0,\ [f_3, f_2] = 0$
$\dots \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$
(m) $(f_m) = 0,\ [f_m, f_1] = 0,\ \dots, [f_m, f_{m-1}] = 0.$

If we set $A(\varphi) = (\varphi)$ and $B_k(\varphi) = [\varphi, f_k]$, then (1) corresponds to the "system" $A(\varphi) = 0$ of which f_1 denotes a solution, and (2) corresponds to the system $A(\varphi) = 0$, $B_1(\varphi) = 0$ of which f_2 denotes a solution. In general, once $f_1, \dots, f_k$ have been determined by the systems corresponding to (1)–(k), corresponding to $(k + 1)$ is the system

$$A(\varphi) = 0, B_1(\varphi) = 0, \cdots, B_k(\varphi) = 0 \tag{5.9}$$

where f_{k+1} denotes a solution.

Of course it remains to show that these systems have solutions and, in addition, that, e.g., system (5.9) has at least $k + 1$ independent solutions so that f_{k+1} may be chosen to be independent of $f_1, \dots, f_k$. This system does not satisfy Jacobi's integrability condition (5.5), since, although the identities (5.11) imply that the $B_i(\varphi)$ satisfy Jacobi's integrability condition, the identity (5.7) implies that

$$B_i(A(\varphi)) - A(B_i(\varphi)) = B_i(\varphi). \tag{5.10}$$

Clebsch was nonetheless able to deal with these systems. In fact in 1866 he published a paper [16], which presented his solution in a more detailed and systematic fashion.

Given a system $A_i(f) = 0$ as in (5.4) and assumed to be linearly independent, the property in equation (5.10) is a special case of

$$A_i(A_j(f)) - A_j(A_i(f)) = \sum_{k=1}^{r} c_{ijk}(x)A_k(f), \ i, j = 1, \ldots, r, \tag{5.11}$$

for any function f. A system with this property Clebsch termed a complete system [16, p. 258], although I will follow Frobenius's terminology and reserve that term for an equivalent property. The equations in (5.11) will be referred to as Clebsch's integrability condition. It includes as a special case Jacobi's integrability condition.

Clebsch showed that given any linearly independent system $A_i(f) = 0$, $i = 1, \ldots, r$, in $n > r$ variables $x_1, \ldots, x_n$ that satisfies his integrability condition, it is always possible to determine r operators $B_i(f)$ that are linear combinations of the $A_i(f)$ and satisfy Jacobi's integrability condition. Since each system is a linear combination of the other, they have the same solutions, and so the existence of simultaneous solutions to the system $A_i(f) = 0$ follows provided the system $B_i(f) = 0$ possesses them [15, pp. 258–60]. For subsequent reference, I will summarize Clebsch's result as follows.

Theorem 5.3 (Jacobi–Clebsch) *A system of $r < n$ linearly independent first order partial differential equations $A_i(f) = 0$ in n variables is complete in the sense that there exist $n - r$ functionally independent solutions $f_1, \ldots, f_{n-r}$ to the system if and only if the integrability condition (5.11) is satisfied.*

This theorem may now be applied to the system (5.9), which consists of $k + 1$ equations in $n = 2m$ variables with $k < m$ and satisfies the integrability condition of the theorem by virtue of (5.10). Hence it has $2m - (k + 1) \geq m$ independent solutions, and so we may pick a solution f_{k+1} that is independent of the k independent functions $f_1, \ldots, f_k$ already determined.

Clebsch's new method was thus direct in the sense that it avoided reliance upon his earlier Theorem 5.1. Furthermore, the successive systems of partial differential equations that need to be integrated are precisely described and consequently the integrability conditions guaranteeing an adequate supply of solutions are seen to be satisfied – unlike the situation in the indirect method, where only the first system is written down. But the direct method was limited to the even-generic case $n = 2m$ and $\det A \neq 0$. Inspired by Clebsch's efforts, Frobenius sought to deal with the completely general problem of Pfaff by a direct method, i.e., one that did not start from Theorem 5.1. Indeed the challenge of the "very complicated algebraic considerations" predicted by Clebsch (in the nongeneric cases) when that theorem is avoided, evidently appealed to Frobenius, who sought to deal with the challenge Berlin-style by first seeking to determine the intrinsic algebraic grounds for Clebsch's Theorem 5.1.

6. Kronecker & the quest for intrinsic grounds

The opening paragraphs of Frobenius's 1877 paper on the problem of Pfaff make it clear that Clebsch's work had provided the principal source of motivation. Thus Frobenius wrote [26, pp. 249–50]:

> After the preliminary work by Jacobi ... the problem of Pfaff was made the subject of detailed investigations primarily by Messrs. Natani ... and Clebsch .... In his first work, Clebsch reduces the solution of the problem to the integration of many systems of homogeneous linear partial differential equations by means of an indirect method, which he himself later said was not suited for presenting the nature of the relevant equations in the right light. For this reason in the second work he attacked the problem in another, direct manner but only treated such differential equations ... [$\omega = 0$] ... for which the determinant of the magnitudes $a_{\alpha\beta}$ ... differs from zero.
> It seems desirable to me to deal with the more general case ... by means of a similar direct method, especially since from the cited works I cannot convince myself that the methods developed for integrating the Pfaffian differential equation in this case actually attain this goal .... Under the above-mentioned assumption[26] in the very first step towards the solution one arrives at a system of many homogeneous linear partial differential equations, rather than a single one. Such a system must satisfy certain integrability conditions if it is to have a nonconstant integral.... I fail to see, on the part of either author[27] a rigorous proof for the compatibility of the partial differential equations to be integrated in the case where the determinant $|a_{\alpha\beta}|$ vanishes.
> Clebsch distinguishes two cases in the problem of Pfaff, which he calls *determinate* and *indeterminate*.... However the criterion for distinguishing the two cases has not been correctly understood by Clebsch.... Were the distinction specified by Clebsch correct, the indeterminate case would never be able to occur.
> For the purposes of integration, the left side of a first order linear differential equation [$\omega = 0$] is reduced by Clebsch to a canonical form that is characterized by great formal simplicity. It was while seeking to derive the posited canonical form on intrinsic grounds (cf. Kronecker, Berl. Monatsberichte 1874, January ...) that I arrived at a new way of formulating the problem of Pfaff, which I now wish to explicate.

Frobenius's above-quoted words not only indicate the many ways in which Clebsch's work motivated his own, they also reveal how he hit upon the approach that he sets forth in his paper, namely by seeking to derive the canonical forms I–II of Clebsch's Theorem 5.1 on "intrinsic grounds" (*innere Gründen*) in the sense of Kronecker. To see what Kronecker had in mind and how it inspired Frobenius, let us turn to the work of Kronecker cited by Frobenius. It bears the title "On families of bilinear and quadratic forms" [58] and is also the source for the quotation given in Sect. 2. The quotation cited by Frobenius was provoked by a note in the Paris *Comptes Rendus* of 1873 on the same subject by Camille Jordan [54]. Jordan's note was motivated by the papers by Weierstrass and Kronecker on this subject starting in 1868 with Weierstrass's paper on elementary divisors. He found them difficult to follow and proposed his own approach. Jordan's note evidently piqued Kronecker, who overreacted with minor criticisms of

[26] Namely, that $A = (a_{\alpha\beta})$ does not have full rank.
[27] Meaning Clebsch in his papers [14, 15] and Natani in his paper [65].

Jordan's comments. For example, Kronecker criticized the way Jordan used the term "canonical form," and this caused him to launch into a discourse on canonical forms, which was no doubt dismissed by Jordan but made an impression upon the youthful Frobenius.

Kronecker wrote as follows [58, pp. 367–8]:

> In fact the expression "canonical form" ... has no generally accepted meaning and in and of itself represents a concept devoid of objective content. No doubt someone who is faced with the question of the simultaneous transformation of two bilinear forms may, as an initial vague goal of his efforts have in mind finding general and simple expressions to which both forms are to be simultaneously reduced. But a "problem" in the serious and rigorous meaning which justifiably attends the word in scientific discourse certainly may not refer to such a vague endeavor. In retrospect, after such general expressions have been found the designation of them as canonical forms may at best be motivated by their generality and simplicity. But if one does not wish to remain with the purely formal viewpoint, which frequently comes to the fore in recent algebra – certainly not to the benefit of true knowledge – then one must not neglect the justification of the posited canonical form on the basis of intrinsic grounds. Usually the so-called canonical or normal forms are merely determined by the tendency of the investigation and hence are only regarded as the means, not the goal of the investigation. In particular, this is always much in evidence when algebraic work is performed in the service of another mathematical discipline, from which it obtains its starting point and goal. But, of course, algebra itself can also supply sufficient inducement for positing canonical forms; and thus, e.g., in the two works by Mr. Weierstrass and myself cited by Mr. Jordan the motives leading to the introduction of certain normal forms are clearly and distinctly emphasized.

It is not difficult to see how Frobenius could view Kronecker's words as applying to Clebsch. Formal analysis had led Clebsch to posit the canonical forms (5.2) and (5.3) of cases (I) and (II) of Theorem 5.1, and Frobenius was heeding Kronecker's words that "one must not neglect the justification of the posited canonical form on the basis of intrinsic grounds." For Kronecker this meant turning to rigorous Berlin-style algebra, but of course a Pfaffian form $\omega = \sum_{i=1}^{n} a_i(x)dx_i$ was not entirely an algebraic object, since its coefficients are not constant but vary with $x = (x_1, \dots, x_n)$, and the transformations $x = \varphi(x')$ to which ω is subjected are not generally linear. For this sort of a situation, however, Frobenius had a paradigm conveniently at hand in the papers on the transformation of differential forms by Christoffel and Lipschitz, who had independently developed the mathematics hinted at in Riemann's 1854 lecture "On the Hypotheses at the Basis of Geometry," which had been published posthumously in 1868 [67]. As we shall see, using ideas gleaned from their work he confirmed Kronecker's above declaration that "algebra itself can also supply sufficient inducement for positing canonical forms."

In papers published back-to-back in Crelle's *Journal* in 1869, Christoffel [13] and Lipschitz [61] concerned themselves, among other things, with the problem of determining the conditions under which two nonsingular quadratic differential forms $\sum_{i,j=1}^{n} g_{ij}(x)dx_i dx_j$ and $\sum_{i,j=1}^{n} g'_{ij}(x')dx'_i dx'_j$ can be transformed into one another by means of general (presumably analytic) transformations $x = \varphi(x')$. Of particular interest was the question of when $\sum_{i,j=1}^{n} g_{ij}(x)dx_i dx_j$ could be transformed into a sum of squares $\sum_{i=1}^{n}(dx'_i)^2$ so as to define (when $n = 3$) Euclidean geometry.[28] For the discussion of Lipschitz's paper below it is helpful to note that if a transformation $x = \varphi(x')$ exists for which $\sum_{i,j=1}^{n} g_{ij}(x)dx_i dx_j = \sum_{i,j=1}^{n} c_{ij} dx_i dx_j$, where the c_{ij} are constants, then when (g_{ij}) is symmetric and positive definite as in Riemann's lecture, a further linear transformation may be made so that the original quadratic form becomes a sum of squares $\sum_{i=1}^{n}(dx_i'')^2$.

As we shall see, Frobenius extracted mathematical ideas from each author's paper, ideas which enabled him to formulate a path to the Kroneckerian intrinsic grounds of Clebsch's canonical forms. Let us consider first what he found in Lipschitz's paper. The approach of Lipschitz was somewhat more general than that of Christoffel in that he considered homogeneous functions $f(dx)$ of the differentials $dx = (dx_1, \dots, dx_n)$ of $x = (x_1, \dots, x_n)$ of any fixed degree k – the analytic analog of the homogeneous polynomials of algebra. With those polynomials in mind he suggested that two such functions $f(dx)$ and $f'(dx')$ of the same degree should be regarded as belonging to the same class if there existed a nonsingular variable transformation $x = \varphi(x')$ so that $f(dx) = f'(dx')$. Of particular interest was a class containing a function $f'(dx')$ with constant coefficients for the reason indicated above.

By way of illustrative example, Lipschitz considered the case $k = 1$ so that $f(dx) = a_1(x)dx_1 + \cdots + a_n(x)dx_n$ [61, pp. 72–3]. He used this example to explain his interest in what he regarded as the analytical counterpart to a covariant in the algebraic theory of invariants. As an example he gave the bilinear form in variables $dx = dx_1, \dots, dx_n$ and $\delta x = \delta x_1, \dots, \delta x_n$

$$\Omega = \sum_{i,j=1}^{n} \left(\frac{\partial a_i}{\partial x_j} - \frac{\partial a_j}{\partial x_i} \right) dx_i \delta x_j. \tag{6.1}$$

The coefficient matrix of Ω is of course the skew symmetric Jacobi matrix of the theory of Pfaff's problem, although Lipschitz made no allusion to that theory. However, in what follows it will be convenient to express (6.1) with the notation $\Omega = \sum_{i,j=1}^{n} a_{ij} dx_i \delta x_j$ with a_{ij} as in (6.1) defining the Jacobi matrix associated to $a_1, \dots, a_n$.

Using results from Lagrange's treatment of the calculus of variations in *Méchanique analytique*, he showed [61, pp. 75–7]:

Theorem 6.1 (Lipschitz) *Ω is a covariant of $\omega = f(dx)$ in the sense that if $f(dx) = f'(dx')$ under $x = \varphi(x')$ and the concomitant linear transformation of differentials $d\varphi : dx \to dx'$, then one has as well*

$$\Omega = \sum_{i,j=1}^{n} a_{ij} dx_i \delta x_j = \sum_{i,j=1}^{n} a'_{ij} dx'_i \delta x'_j = \Omega', \tag{6.2}$$

where also $d\varphi : \delta x \to \delta x'$.

[28] Keep in mind that there is at this time no sensitivity to a distinction between local and global results and that the actual mathematics is being done on a strictly local level.

Lipschitz emphasized that Ω gives the conditions that $f(dx) = \sum_{i=1}^{n} a_n(x)dx_i$ can be transformed into $\sum_{i=1}^{n} c_i dx_i'$, namely that the bilinear form vanish identically, i.e.,

$$a_{ij} = \frac{\partial a_i}{\partial x_j} - \frac{\partial a_j}{\partial x_i} \equiv 0. \tag{6.3}$$

As Lipschitz observed, these were the well-know conditions that the differential $\omega = \sum_{i=1}^{n} a_n(x)dx_i$ be exact, from which the above transformation property followed.[29] All this was by way of a preliminary to motivate the case of degree $k = 2$, namely quadratic differential forms ds^2, where the same process leads to a quadrilinear "covariant" form in four variables that defines what later became known as the Riemann curvature tensor, so that the vanishing of its coefficients gives the condition that ds^2 may be transformed into a sum of squares.

Of primary interest to Frobenius, however, was Lipschitz's result in the motivational case $k = 1$ – his Theorem 6.1. In keeping with the work on quadratic and bilinear forms within the Berlin school, as well as the above-mentioned work of Christoffel and Lipschitz, Frobenius focused on the question of when two Pfaffian expressions $\omega = \sum_{i=1}^{n} a_i(x)dx_i$ and $\omega' = \sum_{i=1}^{n} a_i'(x')dx_i'$ are analytically equivalent in the sense that a nonsingular transformation $x = \varphi(x')$ exists such that $\omega = \omega'$. His goal was to see if Clebsch's normal forms characterized these equivalence classes and, if so, to find the "intrinsic grounds" for this fact. What Lipschitz's Theorem 6.1 showed him was that the analytical equivalence of ω and ω' brought with it the analytical equivalence of their associated "bilinear covariants" Ω and Ω' – as Frobenius called them in keeping with Lipschitz's Theorem and terminology. Today the bilinear covariant Ω associated to ω is understood within the framework of the theory of differential forms initiated by Élie Cartan, where $\Omega = -d\omega$. In Sect. 11 the influence of Frobenius's work on Cartan's development of his theory of differential forms will be considered.

It was in seeking to use Lipschitz's theorem 6.1 that Frobenius drew inspiration from Christoffel's paper. To determine necessary and sufficient conditions for the analytical equivalence of two quadratic differential forms, Christoffel had determined purely algebraic conditions involving a quadrilinear form that were necessary for equivalence and he then asked whether they were sufficient for the analytical equivalence or whether additional analytical conditions needed to be imposed. He characterized this question as "the heart (*Kernpunkt*) of the entire transformation problem" [13, p. 60].

Applying this strategy to the problem at hand, Frobenius began by giving an algebraic proof of Lipschitz's Theorem 6.1 [26, p. 252–3] that was elegant, clear, and simple, and, in particular, did not rely on results from the calculus of variations. His proof makes it clear to present-day readers that Lipschitz's theorem follows from a simple calculation in the tangent space of the manifold of generic points where Ω and Ω' have maximal rank.

---

[29] If $f_1(x)$ is such that $\omega = df_1$ in accordance with exactness, then additional functions $f_2, \dots, f_n$ may be chosen so that $f_1, \dots, f_n$ are independent and so define locally a variable change $y_i = f_i(x)$, $i = 1, \dots, n$ for which $\omega = dy_1$ has constant coefficients.

Here is the gist of Frobenius's proof. With the relations between the x- and x'-coordinates given by[30]

$$x = \varphi(x'), \quad dx = d\varphi \cdot dx' \quad \text{or} \quad x' = \psi(x), \quad dx' = d\psi \cdot dx, \tag{6.4}$$

he considered x as an arbitrary function of two parameters, which I will denote by s, and t. Then the components u_i, v_i of the n-tuples

$$\begin{aligned} \partial x/\partial s &= (\partial x_1/\partial s, \ldots, \partial x_n/\partial s) = (u_1, \ldots, u_n) = u \\ \partial x/\partial t &= (\partial x_1/\partial t, \ldots, \partial x_n/\partial t) = (v_1, \ldots, v_n) = v \end{aligned}$$

are completely arbitrary and so are the vectors u and v. Since $x' = \psi(x)$ and $x = x(s,t)$ we have $x' = x'(s,t)$ and so we may analogously define the vectors $u' = \partial x'/\partial s = (u'_1, \ldots, u'_n)$ and $v' = \partial x'/\partial t = (v'_1, \ldots, v'_n)$. Then by (6.4) $u = d\psi \cdot u'$ and $v = d\psi \cdot v'$. The equivalence $\omega = \omega'$ implies in particular that, regarded as functions of s and t, $\omega(u) = \sum_i a_i u_i$ equals $\omega'(u') = \sum_i a'_i u'_i$ and likewise $\omega(v) = \omega'(v')$. Thus

$$\frac{\partial \omega(u)}{\partial t} - \frac{\partial \omega(v)}{\partial s} = \frac{\partial \omega'(u')}{\partial t} - \frac{\partial \omega'(s)}{\partial s}, \tag{6.5}$$

and when (6.5) is computed out using the definition (4.7) of the Jacobi coefficients a_{ij}, a'_{ij}, the result is that for any two vectors u and v, if $u' = d\psi \cdot u$ and $v' = d\psi \cdot v$, then

$$\Omega(u,v) \stackrel{\text{def}}{\equiv} \sum_{i,j} a_{ij} u_i v_j = \sum_{i,j} a'_{ij} u'_i v'_j \stackrel{\text{def}}{\equiv} \Omega'(u',v'),$$

which is the conclusion of Lipschitz's Theorem expressed in a notation that makes its meaning clearer. That is, for each fixed x, ω is seen to be a linear form, $\omega = \omega(u)$, and Ω and alternating bilinear form, $\Omega = \Omega(u,v)$ and likewise for ω' and Ω'.

Frobenius could readily see that a necessary consequence of the analytical equivalence of ω and ω' is the *algebraic equivalence* of the form-pairs (ω, Ω) and (ω', Ω') in the following sense: Fix x at x_0 and x' at $x'_0 = \varphi(x_0)$.[31] Then we have two form-pairs $(\omega, \Omega)_{x_0}$, $(\omega', \Omega')_{x'_0}$ with constant coefficients that are equivalent in the sense that $\omega_{x_0}(u) = \omega'_{x'_0}(u')$ and $\Omega_{x_0}(u,v) = \Omega'_{x'_0}(u',v')$ by means of a nonsingular linear transformation $u = Pu'$, $v = Pv'$, where $P = d\varphi_{x_0}$. The question that Frobenius posed to himself was whether the algebraic equivalence of the form-pairs was sufficient to guarantee the (local) analytical equivalence of ω and ω'. Thus it was first necessary to study the algebraic equivalence of form-pairs (w, W) under a nonsingular linear transformation $u = Pu'$, $v = Pv'$, where $w(u) = \sum_{i=1}^n a_i u_i$ is a linear form and $W(u,v) = \sum_{i,j=1}^n a_{ij} u_i v_j$ is an alternating bilinear form ($a_{ji} = -a_{ij}$) and all coefficients a_i, a_{ij} are constant. The hope would be that the algebraic analog of Clebsch's two canonical forms (5.2) and (5.3) of cases (I) and (II) of Theorem 5.1 would yield the distinct equivalence classes for (w, W). As we shall see in the following section,

[30] In what follows I have used a more compact notation than Frobenius, who expressed all transformations in terms of coordinate equations, e.g., $x_\alpha = \varphi_\alpha(x'_1, \ldots, x'_n)$ rather than $x = \varphi(x')$.

[31] As indicated in Sect. 3, strictly speaking x_0 and x'_0 should be generic points, i.e., points at which the bilinear forms Ω, Ω' have maximal rank.

Frobenius confirmed that this was the case. Indeed, his proof provided a paradigm which led him by analogy to a proof of Clebsch's Theorem 5.1 – including a correct way to algebraically distinguish the two cases – and thereby to the conclusion that the algebraic equivalence of (ω, Ω) and (ω', Ω') at generic points x_0 and x_0' (in the sense of Sect. 3) implies the analytical equivalence of ω and ω' at those points. In this manner he found what he perceived to be the true "intrinsic grounds" in Kronecker's sense for Clebsch's canonical forms.

Clebsch had shied away from a direct approach in the nongeneric case in favor of his indirect approach because "In this way, one is spared the trouble of carrying out direct proofs that lead to very complicated algebraic considerations, which, to be sure, are of interest in their own right . . . " [14, p. 194]. It was just these sorts of "complicated algebraic considerations" that attended nongeneric reasoning in linear algebra and that Weierstrass and Kronecker had shown could be successfully transformed into a satisfying theory, and the paradigm of their work clearly encouraged Frobenius to deal in the above-described manner with the theory of Pfaffian equations. Indeed, Frobenius realized that his friend Ludwig Stickelberger, a fellow student at Berlin and a colleague at the ETH in Zürich when Frobenius wrote [26], had already considered the simultaneous transformation of a bilinear or quadratic form together with one or more linear forms in his 1874 Berlin doctoral thesis [26, p. 264n].

7. The algebraic classification theorem

In discussing form-pairs (w, W) I will use more familiar matrix notation and write $W = u^t Av$, where A is skew symmetric ($A^t = -A$), and $w = a^t u$, with a, u and v here being regarded as $n \times 1$ column matrices, e.g., $a = (a_1, \ldots, a_n)^t$. Frobenius began to develop such notation himself shortly after his work on Pfaffian equations.[32]

The first algebraic question that Frobenius considered concerned the rank of A, i.e., the rank of a skew symmetric matrix or, as he called it, an alternating system [26, pp. 255–261]. I would guess that this was also one of the first questions related to Pfaffian equations he investigated, since the resultant answer plays a fundamental role in the ensuing theory. For ease of reference I will name it the even rank theorem.

Theorem 7.1 (Even rank theorem) *If A is skew symmetric then its rank r must be even.*

Following Frobenius, let us say that a principal minor[33] is one obtained from A by deleting the same numbered rows and columns, e.g., the $n - 3 \times n - 3$ minor obtained by deletion of rows 1,3,5 and columns 1,3,5 of A. It is easy to give examples of matrices of rank r for which all the principal minors of degree r vanish, but Frobenius showed that when A is symmetric or skew symmetric of rank r then there is always a principal minor of degree r which does not vanish. From this result Theorem 7.1 follows directly, since the matrix of a principal minor of A is also skew symmetric and as we saw in

[32] He did this in a paper of 1878 [27], which is discussed in historical context in my paper [41].
[33] Literally "principal determinant" (*Hauptdeterminant*) [26, p. 261].

Sect. 4, Jacobi had already observed that skew symmetric determinants of odd degree must vanish. Thus the degree r of the nonvanishing principal minor must be even.

With this theorem in mind, let us consider the two canonical forms of cases (I) and (II) of Clebsch's Theorem 5.1. In (I) $\omega = z_{m+1}dz_1 + \cdots + z_{2m}dz_m$, where the total number of variables is $n = 2m + q$, $q \geq 0$. In this case

$$\Omega = (dz_1\delta z_{m+1} - dz_{m+1}\delta z_1) + \cdots + (dz_m\delta z_{2m} - dz_{2m}\delta z_m)$$

The corresponding pair of algebraic forms (w, W) is obtained by setting the variables z_i equal to constants and the differentials dz_i, δz_i equal to variables u_i, v_i, respectively, to obtain $w = c_1u_1 + \cdots + c_mu_m$ and $W = (u_1v_{m+1} - u_{m+1}v_1) + \cdots + (u_mv_{2m} - u_{2m}v_m)$. Thus $W = u^tAv$, where in block matrix form

$$A = \begin{pmatrix} 0 & I_m & 0 \\ -I_m & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{7.1}$$

Here I_m denotes the $m \times m$ identity matrix. For case (II) where ω is given by $\omega = dz_0 + z_{m+1}dz_1 + \cdots + z_{2m}dz_m$, $n = 2m + q$, $q \geq 1$, the algebraic form-pair is $w = u_0 + c_1u_1 + \cdots c_mu_m$ and $W = u^tAv$ with

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & I_m & 0 \\ 0 & -I_m & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{7.2}$$

In both cases the matrix A has rank $2m$, so that contrary to Clebsch's claim the rank of the skew symmetric Jacobi matrix A does not distinguish the two cases *a priori*.

On the algebraic level considered here, where the link between ω and Ω no longer exists, one could say that this is because we are seeking an invariant of a form-pair $(w, W) = (a^tu, u^tAv)$ under linear transformations but only looking at A and ignoring a. In his Berlin doctoral thesis of 1874 [69, §2] Stickelberger had already introduced an appropriate invariant for a system consisting of a bilinear form and several linear forms. Stickelberger's thesis was well known to Frobenius, who along with Killing was one of the three appointed "adversaries" at Stickelberger's thesis defense. Consider, for example, $\Gamma = u^tCv$, $\gamma_1 = c^tu$, $\gamma_2 = d^tu$. From Weierstrass's paper on elementary divisors it was well-known that by virtue of a theorem on minor determinants due to Cauchy, the rank of the bilinear form is invariant under nonsingular linear transformations $u = P\bar{u}$, $v = Q\bar{v}$, i.e., if

$$u^tCv = \bar{u}^t(P^tCQ)\bar{v} \equiv \bar{u}^t\bar{C}\bar{v},$$

then rank $(\bar{C})$ = rank C. Since the linear forms transform by

$$c^tu = u^tc = \bar{u}^t(P^tc) \equiv \bar{c}^t\bar{u},$$

Stickelberger observed that if one introduces the bilinear form $\widehat{\Gamma}$ in $n+1$ variables with $n+1 \times n+1$ coefficient matrix

$$\widehat{C} = \begin{pmatrix} C & c \\ d^t & 0 \end{pmatrix},$$

which amalgamates the original bilinear and linear forms into a single bilinear form, then under the linear transformations $u = \widehat{P}\bar{u}$, $v = \widehat{Q}\bar{v}$, where

$$\widehat{P} = \begin{pmatrix} P & 0 \\ 0 & 1 \end{pmatrix}, \quad \widehat{Q} = \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix},$$

one has

$$\widehat{C} \to \widehat{P}^t \widehat{C} \widehat{Q} = \begin{pmatrix} \bar{C} & \bar{c} \\ \bar{d}^t & 0 \end{pmatrix}.$$

This shows that the rank of $\widehat{C}$ is an invariant of the system $(\Gamma, \gamma_1, \gamma_2)$.

To apply this to the pair $(w, W) = (a^t u, u^t A v)$ with A skew symmetric, Frobenius introduced the analogous alternating form $\widehat{W}$ with coefficient matrix

$$\widehat{A} = \begin{pmatrix} A & a \\ -a^t & 0 \end{pmatrix}. \tag{7.3}$$

Then it follows that the rank of $\widehat{A}$ is an invariant of the system (w, W). Since $\widehat{A}$ is skew symmetric it follows readily from the even rank theorem that either rank $\widehat{A}$ = rank A or rank $\widehat{A}$ = rank $A + 2$.

Going back to the pairs (w, W) obtained above corresponding to Clebsch's cases (I) and (II) with coefficients given in (7.1) and (7.2), respectively, it follows that rank $\widehat{A}$ = rank $A = 2m$ in case (I) but rank $\widehat{A}$ = rank $A + 2 = 2m + 2$ in case (II). For example when $m = 2$ and $n = 5$ so that $\widehat{A}$ is 6×6 we get in cases (I) and (II), respectively, matrices of ranks 4 and 6, namely:

$$\widehat{A} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & c_1 \\ 0 & -1 & 0 & 0 & 0 & c_2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -c_1 & -c_2 & 0 & 0 \end{pmatrix}, \quad \widehat{A} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & c_1 \\ 0 & 0 & -1 & 0 & 0 & c_2 \\ -1 & 0 & 0 & -c_1 & -c_2 & 0 \end{pmatrix}.$$

These examples suggest that Clebsch was at least partly right in thinking that the rank of a skew symmetric matrix would distinguish *a priori* the two cases, but the relevant matrix would seem to be $\widehat{A}$ and not A. Frobenius proved that this was in fact the case. His main result may be summarized as follows.

Theorem 7.2 (Algebraic classification theorem) *Let $w = a^t u$ be a linear form and $W = u^t A v$ an alternating bilinear form. Then $p =$ (rank A + rank $\widehat{A}$)$/2$ is an integer and is invariant with respect to nonsingular linear transformations of the pair (w, W), which is thus said to be of* **class p** . *When* rank $\widehat{A}$ = rank $A = 2m$, *$p = 2m$ is even; and when* rank $\widehat{A}$ = rank $A + 2 = 2m + 2$, *$p = 2m + 1$ is odd. If (w, W) is of class $p = 2m$, then there exists a nonsingular linear transformation $u = P\bar{u}$, $v = P\bar{v}$ such that $w = c_{m+1}\bar{u}_1 + \cdots + c_{2m}\bar{u}_m$, and $W = \bar{u}^t \bar{A} \bar{v}$ with $\bar{A}$ as in (7.1). When $p = 2m + 1$ a nonsingular linear transformation exists so that w takes the form $w = \bar{u}_0 + c_{m+1}\bar{u}_1 + \cdots + c_{2m}\bar{u}_m$ and $W = \bar{u}^t \bar{A} \bar{v}$ with $\bar{A}$ as in (7.2). Consequently, two form-pairs (w, W) and (w', W') are equivalent if and only if they are of the same class p.*

The developments leading up to this theorem as well as collateral results regarding Pfaffian determinants were presented in Sects. 6–11 of Frobenius's paper and totaled 21 pages. Turning next to the analytical theory of the equivalence of differential forms $\omega = \sum_{i=1}^{n} a_i(x)dx_i$, however, Frobenius explained that, in effect, his Theorem 7.2 would not serve as the mathematical foundation of the analytical theory but rather as a guide. In developing the analytical theory "I will rely on the developments of §§. 6–11 as little as possible, and utilize them more by analogy than as a foundation" [26, p. 309]. No doubt he took this approach to encourage analysts not enamored (as he was) by algebra to read the largely self-contained analytical part. Comparison of the two parts shows that the algebraic part, however, provided the blueprint for the analysis, for the reasoning closely parallels the line of reasoning leading to the algebraic classification theorem. Indeed, as we shall see in Sect. 10, it proved too algebraic for the tastes of many mathematicians primarily interested in the integration of differential equations.

The fact that Frobenius deemed the purely algebraic results sufficiently noteworthy in their own right to present them carefully worked out in Sects. 6–11 is indicative of his appreciation for Berlin-style linear algebra, and indeed the theory presented in Sects. 6–11 and culminating in Theorem 7.2 was the first but hardly the last instance of Frobenius's creative involvement with linear algebra. For example, undoubtedly inspired by the theory of Sects. 6–11, during 1878–1880 Frobenius published several highly important and influential memoirs on further new and important aspects of the theory of bilinear forms [27, 28, 30]. [34]

8. The analytic classification theorem

The main theorem in the analytical part of his paper is the result at which Frobenius arrived by developing the analog of the reasoning leading to the above algebraic classification theorem. Thus given a Pfaffian form $\omega = \sum_{i=1}^{n} a_i(x)dx_i$, he considered the form-pair (ω, Ω) determined by ω and its bilinear covariant

$$\Omega = \sum_{i,j=1}^{n} a_{ij}(x)u_i v_j, \quad a_{ij} = \frac{\partial a_i}{\partial x_j} - \frac{\partial a_j}{\partial x_i}. \tag{8.1}$$

Corresponding to this form-pair we have by analogy with (7.3) the augmented skew symmetric matrix

$$\widehat{A}(x) = \begin{pmatrix} A(x) & a(x) \\ -a^t(x) & 0 \end{pmatrix}, \quad A = (a_{ij}(x)). \tag{8.2}$$

Since the ranks of A and $\widehat{A}$ figure prominently in what is to follow, recall from Sect. 3 that by the rank of, e.g., $A = A(x)$ Frobenius apparently meant the maximal rank of $A(x)$. What Frobenius tacitly showed [26, p. 309] was that the maximal rank of $\widehat{A}(x)$ is also invariant with respect to nonsingular transformations $x = \varphi(x')$ at the generic points corresponding to $\widehat{A}(x)$. Hence $p = (1/2)[\text{rank } A(x) + \text{rank } \widehat{A}(x)]$ is also an invariant in this sense on the (open and dense) set of generic points x where both $A(x)$

[34] See [40, 41] for a discussion of these papers and their historical significance.

and $\widehat{A}(x)$ attain their maximal ranks, and p so defined can therefore be used to define the class of ω. The main goal of his paper was to establish the following analytical analog of the algebraic classification theorem, which should be interpreted as a theorem about the existence of local transformations at the above-mentioned generic points.

Theorem 8.1 (Analytical classification theorem) *Let ω be of class p. If $p = 2m$ there exists a transformation $x = \varphi(z)$ such that*

$$\text{(I)} \quad \omega = z_{m+1}dz_1 + \cdots + z_{2m}dz_m;$$

and if $p = 2m + 1$, then a transformation $x = \varphi(z)$ exists so that

$$\text{(II)} \quad \omega = dz_0 + z_{m+1}dz_1 + \cdots + z_{2m}dz_m.$$

Consequently ω and ω' are equivalent if and only if they are of the same class p.

This theorem implies Clebsch's Theorem 5.1 and, in addition, provides through the notion of the class of ω a correct algebraic criterion distinguishing cases (I) and (II). The theorem also shows that the algebraic equivalence of two form-pairs (ω, Ω) and (ω', Ω') is sufficient for the analytical equivalence of ω and ω'. That is, if x and x' are fixed generic points with respect to both ω and ω', and if they are algebraically equivalent for the fixed values x and x', then by the algebraic classification theorem, $(\omega, \Omega)_x$ and $(\omega', \Omega')_{x'}$ must be of the same class p. But this then means that ω and ω' are of the same class p and so by the analytical classification theorem each of ω and ω' can both be locally transformed into the same canonical form (I) or (II) and hence into each other, i.e., they are analytically equivalent.

The proofs of both the algebraic and analytical classification theorems are quite similar in most respects, with the analytical version evidently inspired by the algebraic version. Both are lengthy, but the analytical version is longer by virtue of a complication attending the analytical analog of one point in the proof. It is this complication that led Frobenius to formulate his integrability theorem for systems of Pfaffian equations, which is discussed in the next section. The remainder of this section is devoted to a brief summary of the complication.[35]

In proving the Algebraic Classification Theorem Frobenius at one point was faced with the following situation, which I describe in the more familiar vector space terms into which his reasoning readily translates. Given k linearly independent vectors $w_1, \ldots, w_k$ in $\mathbf{C}^n$, determine a vector $w \neq 0$ such that w is linearly independent of $w_1, \ldots, w_k$ and also w is perpendicular to a certain subspace $\mathcal{V}$, with basis vectors $v_1, \ldots, v_d$. Thus w needs to be picked so that

$$w \cdot v_i = 0, \quad i = 1, ..., d. \tag{8.3}$$

This means w is to be picked from $\mathcal{V}^{\perp}$, the orthogonal complement of $\mathcal{V}$. The problem is complicated by the fact that the vectors $w_1, \ldots, w_k$ also lie in $\mathcal{V}^{\perp}$ and w must be

[35] Frobenius's proof is spread out over pp. 309–331 of [26] and also draws on results from other parts of the paper.

linearly independent of them. However, it is known that the dimension d of $\mathcal{V}$ satisfies the inequality $d \leq n + 1 + k - 2m$, where m is a given fixed integer such that $k < m$. Since the dimension of $\mathcal{V}^{\perp}$ is $n - d \geq n - (n + 1 + k - 2m) = 2m - (k + 1) \geq m > k$, it is possible to pick w to be linearly independent of $w_1, \ldots, w_k$ as desired.

In the analytical version of this situation that arises en route to the Analytical Classification Theorem, k functions $f_1, \ldots, f_k$ of $x_1, \ldots, x_n$ are given that are functionally independent, i.e., their gradient vectors $\nabla f_1, \ldots, \nabla f_k$ are linearly independent in the sense that the $k \times n$ matrix with the ∇f_i as its rows has full rank k (in the neighborhood of the point under consideration). In lieu of the d vectors $v_1, \ldots, v_d$ of the algebraic proof, there are now d vector-valued functions $v_i = (b_{i1}(x), ..., b_{in}(x))$, where again $d \leq n + 1 + k - 2m$ and m is a given fixed integer such that $k < m$. The problem now is to determine a function f, so that $\nabla f_1, \ldots, \nabla f_k, \nabla f$ are linearly independent and in lieu of (8.3) f must satisfy

$$\nabla f \cdot v_i = 0, \text{ i.e., } B_i(f) \stackrel{\text{def}}{=} \sum_{j=1}^{n} b_{ij}(x) \frac{\partial f}{\partial x_j} = 0, \; i = 1, \ldots d. \tag{8.4}$$

By analogy with the algebraic proof, the situation is complicated by the fact that the functions $f_1, \ldots, f_k$ are also solutions to this system of partial differential equations. Thus for the desired f to exist the system (8.4) must have at least $k + 1$ functionally independent solutions. This, of course is the type of system considered by Jacobi and Clebsch and by the Jacobi–Clebsch Theorem 5.3, if the system satisfies Clebsch's integrability condition (5.11) for completeness – $[B_i, B_j] = \sum_{l=1}^{d} c_{ijk} B_l$ – it will have $n - d$ independent solutions, and since (as we already saw) $n - d \geq m > k$, the existence of the desired function f will then follow. The system, however, was not explicitly given (see [26, pp. 312-13]) – as was also the case with the systems in Clebsch's indirect method. It was consequently uncertain whether Clebsch's integrability condition was satisfied. As the quotation given at the beginning of Sect. 6 shows, Frobenius had criticized Clebsch for glossing over a similar lack of certainty in his indirect method. He was certainly not about to fall into the same trap now! But how to salvage the proof? To this end, he turned to a general duality between systems such as $B_i(f) = 0$ and systems of Pfaffian equations. As we shall see in the next section, this duality had come to light in reaction to the work of Clebsch and Natani, but Frobenius developed it in a more elegant and general form than his predecessors.

9. Frobenius's integrability theorem

In his paper of 1861 on Pfaff's problem [65], Leopold Natani, who was unfamiliar with the still unpublished new method of Jacobi that had inspired Clebsch, did not seek to determine the functions $f_1, \ldots, f_m$ in $\omega = \sum_{i=1}^{2m} a_i dx_i = \sum_{i=1}^{m} F_i df_i$ by means of partial differential equations. Instead he utilized successive systems of special Pfaffian equations. With the representation $\omega = F_1 df_1 + \cdots + F_m df_m$ as the goal, Natani first constructed a system of Pfaffian equations out of the coefficients a_i, a_{ij} that yielded f_1 as a solution. Then he constructed a second system using as well f_1 to obtain f_2, and so on.[36]

[36] For a clear exposition of the details of Natani's method, see [37].

The contrasting treatments of Pfaff's problem by Natani and Clebsch turned the attention of some mathematicians to the connections between: (1) systems of linear homogeneous partial differential equations, such as Clebsch's systems, and (2) systems of Pfaffian equations, such as Natani's. It turns out that associated to a system of type (1) is "dual" system of type (2) with the property that the independent solutions f to (1) are precisely the independent integrals $f = C$ to (2). This general duality was apparently not common knowledge in 1861, since, judging by Clebsch's remarks on Natani's work [15, p. 146n], he failed to realize that Natani's successive systems of Pfaffian equations were the duals in the above sense to the systems of partial differential equations in his direct method. This was pointed out by Hamburger in a paper of 1877 [37] to be discussed below.

Apparently the first mathematician to call attention to the general existence of a dual relation between systems of Pfaffian equations and systems of linear homogeneous partial differential equations was Adolph Mayer (1839–1908) in a paper of 1872 [62]. Consideration of Mayer's way of looking at and establishing this reciprocity is of interest by way of comparison with the approach to the matter taken by Frobenius. We will see that on the local level where both tacitly reasoned, Mayer's approach lacked the complete generality and strikingly modern algebraic elegance achieved by Frobenius.

Ever since the late 18th century mathematicians had realized that the integration of a single linear homogeneous partial differential equation

$$\xi_1 \frac{\partial z}{\partial x_1} + \cdots + \xi_n \frac{\partial z}{\partial x_n} = 0 \tag{9.1}$$

was equivalent to the integration of a system of first order ordinary differential equations, namely the system that – with x_n picked as independent variable – can be written as

$$\frac{dx_1}{dx_n} = \frac{\xi_1}{\xi_n}, \ \ldots\ , \frac{dx_{n-1}}{dx_n} = \frac{\xi_{n-1}}{\xi_n}. \tag{9.2}$$

Jacobi, for example, gave an elegant treatment of this equivalence in a paper of 1827 on partial differential equations [48].[37] Mayer began by noting the above-described equivalence between (9.1) and (9.2), which evidently inspired his observation that "in an entirely similar way" it is easy to establish a reciprocal connection between systems of linear homogeneous partial differential equations and Pfaffian systems, which "in particular cases ... has already been observed and utilized many times" [62, p. 448].

Mayer began with a system of m linearly independent partial differential equations $A_i(f) = 0$ in $n > m$ variables. Normally, the $A_i(f)$ would be written in general form $A_i(f) = \sum_{j=1}^{n} \alpha_{ij}(x)\partial f/\partial x_j$, but Mayer assumed they were written in the special form

$$A_i(f) = \frac{\partial f}{\partial x_i} + \sum_{k=m+1}^{n} a_{ik}(x) \frac{\partial f}{\partial x_k}, \quad i = 1, \ldots, m. \tag{9.3}$$

In other words, he assumed the $m \times n$ matrix $M(x) = (\alpha_{ij}(x))$ can be put in the reduced echelon form $(I_m \quad A)$. It was well known that any particular equation $A_i(f) = 0$ of

[37] For an exposition of the equivalence presented in the spirit of Jacobi see [44, p. 201ff.].

the system could be replaced by a suitable linear combination of the equations. In other words, elementary row operations may be performed on the matrix $M(x)$. In this manner for a *fixed* value of x, $M(x)$ can be transformed into its reduced echelon form, and then by permuting columns, which corresponds to reindexing the variables $x_1, \ldots, x_n$, the form $(I_m \quad A)$ can be obtained. However, Mayer apparently did not realize that this cannot be done analytically, i.e., for all x in a neighborhood of a fixed point x_0.

For example with $m = 2$ and $n = 3$ consider

$$M(x, y, z) = \begin{pmatrix} x & y & 1 \\ z & -1 + x + \frac{yz}{x} & 3 + \frac{z}{x} \end{pmatrix}$$

in the neighborhood of $(x, y, z) = (1, 1, 1)$. There $M(x, y, z)$ has full rank, which means that the corresponding partial differential equations are linearly independent as Mayer evidently assumed. The reduced echelon forms of $M(1, y, z)$ and of $M(x, y, z)$ with $x \neq 1$, for all (x, y, z) close to $(1, 1, 1)$ are, respectively,

$$\begin{pmatrix} 1 & y & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 & \frac{x-3y-1}{x(x-1)} \\ 0 & 1 & \frac{3}{x-1} \end{pmatrix},$$

so that it is impossible to bring $M(x, y, z)$ into the form $(I_2 \quad A)$ for all (x, y, z) in a neighborhood of $(1, 1, 1)$. Hence his treatment of duality on the customary local level was not completely general. That being said, let us consider his introduction of the dual to system (9.3).

Since simultaneous solutions to this system are precisely the solutions to

$$A_\lambda(f) = \lambda_1(x)A_1(f) + \cdots + \lambda_m A_m(f) = 0,$$

for any choice of functions $\lambda_i(x)$ the generic equation $A_\lambda(f) = 0$ is equivalent to the system. Now this is a single equation and so corresponds in the above-described manner to a system of ordinary differential equations, namely

$$(1) \quad \frac{dx_i}{dx_n} = \lambda_i, \; i = 1, \ldots, m, \qquad (2) \quad \frac{dx_k}{dx_n} = \sum_{i=1}^{m} \lambda_i a_{ik}$$

Substituting (1) in (2) and multiplying through by dx_n then yields the system of $n - m$ Pfaffian equations

$$dx_k = \sum_{i=1}^{m} a_{ik}(x_1, \ldots, x_n)dx_i, \quad k = m + 1, \ldots, n. \tag{9.4}$$

Thus (9.4) is the Pfaffian system which corresponds to (9.3).

Mayer defined a Pfaffian system in the form (9.4) and thus consisting of $n - m$ equations to be "completely integrable" (*unbeschränkt integrable*) if $n - m$ independent integrals $f_k(x_1, \ldots, x_n) = C_k$, $k = m + 1, \ldots, n$, exist in the sense that if this integral system is solved for the variables $x_{m+1}, \ldots, x_n$ to get the functions $x_k = \varphi_k(x_1, \ldots, x_{m-1}, C_m, \ldots, C_n)$, $k = m + 1, \ldots, n$, then (9.4) is satisfied identically if these functions are substituted, i.e., if x_k is replaced by $\varphi_k(x_1, \ldots, x_m, C_{m+1}, \ldots, C_n)$ and dx_k is replaced by $\sum_{i=1}^{m}(\partial\varphi_k/\partial x_i)dx_i$ for $k = m + 1, \ldots, n$.

It follows that if (9.4) is completely integrable in this sense and the above-mentioned substitutions are made, then by comparing the two sides of the resulting equation

$$\frac{\partial \varphi_k}{\partial x_i} = a_{ik}(x_1, \ldots, x_m, C_{m+1}, \ldots, C_n).$$

From this and the chain rule Mayer then obtained the condition

$$0 = \frac{\partial^2 \varphi_k}{\partial x_j \partial x_i} - \frac{\partial^2 \varphi_k}{\partial x_i \partial x_j} = \frac{\partial a_{ik}}{\partial x_j} - \frac{\partial a_{jk}}{\partial x_i} + \sum_{l=m+1}^{n} a_{jl} \frac{\partial a_{ik}}{\partial x_l} - a_{il} \frac{\partial a_{jk}}{\partial x_l},$$

which in Jacobi operator notation is

$$A_j(a_{ik}) - A_i(a_{jk}) = 0, \quad i, j = 1, \ldots, m, \; k = m+1, \ldots, n.$$

Since this condition must hold for all $x_1, \ldots, x_m$ and any values of $C_{m+1}, \ldots C_n$, it must hold identically in $x_1, \ldots, x_n$ and so is equivalent to the condition that $A_i(A_j(f)) - A_j(A_i(f)) = 0$ for all f and all $i, j = 1, \ldots, m-1$, which is the Jacobi integrability condition (5.5). This then is a necessary condition of the complete integrability of the Pfaffian system (9.4): the corresponding system (9.3) of partial differential equations must satisfy Jacobi's integrability condition. According to Mayer, implicit in his discussion of the integration of (9.4) was proof that this condition on (9.4) is also sufficient.

Mayer's interest in the above duality was motivated by his interest in the goal of integration efficiency mentioned in Sect. 4, which was of interest to many analysts at this time. His idea was to start with a system of partial differential equations satisfying Jacobi's integrability condition – the type of system involved in the integration of nonlinear partial differential equations in accordance with Jacobi's new method and its extensions – and then go over to the dual system of Pfaffian equations and integrate it to see if it yielded a more efficient method. He showed that it did, that the number of integrations needed could be reduced by almost 50% over what the latest theories offered.

Although Mayer had indeed established a reciprocity or duality between systems of linear homogeneous partial differential equations and systems of Pfaffian equations, he had done so by assuming the systems in a special form so that the duality could be obtained from the well-known equivalence of a single linear homogeneous partial differential equation and a system of ordinary differential equations. As we shall see, by taking a more algebraic and elegant approach and by virtue of a powerful new construct – the bilinear covariant – Frobenius, who never mentions Mayer's paper, was able to establish the reciprocity without assuming the systems in a special generic form. This enabled him to formulate a criterion for complete integrability that was directly applicable to any system of Pfaffian equations. Frobenius was no better than Mayer when it came to glossing over distinctions between local and global results, but his more elegant algebraic approach does seem to point the way to a more rigorous, modern interpretation and for that reason the integrability theorem at which he arrived (Theorem 9.2) does seem to represent an indication of the modern Frobenius integrability theorem. In presenting Frobenius's treatment, I will first present it as he did, glossing over any mention of the values of x for which the reasoning makes sense, and then I will indicate what his proof actually established.

Given the special form of the systems required for duality in Mayer's sense and the focus of his attention on efficiency matters, it is not clear that he himself realized that the systems of partial differential equations of Clebsch's direct method and the Pfaffian systems of Natani were in fact duals of one another. This was pointed out by Meyer Hamburger (1838–1903) in a paper [37] submitted for publication a few months after Frobenius had submitted his own paper on the problem of Pfaff. Although Hamburger's paper [37] is not mentioned by Frobenius, Hamburger's method for establishing the duality between Natani's systems of Pfaffian equations and Clebsch's systems of partial differential equations was probably known to him because Hamburger had presented it in the context of a different problem in a paper of 1876 [36, p. 252] in Crelle's *Journal* that Frobenius did cite in his paper. Hamburger's method is more algebraic than Mayer's and does not require that the systems be put in a special generic form. It may have encouraged Frobenius's own simpler algebraic method.

In accordance with the approach of Hamburger but especially that of Frobenius [26, §13], who unlike Hamburger presented everything with elegant simplicity and algebraic clarity, let us consider a system of r linearly independent Pfaffian equations

$$\omega_i = a_{i1}(x)dx_1 + \cdots + a_{in}(x)dx_n = 0,\ i = 1, \dots, r. \tag{9.5}$$

Frobenius regarded the ω_i as linear forms in $(dx_1, \dots, dx_n)$, and since linear forms may be identified with linear functionals or elements in the dual of the tangent space, his way of thinking translates readily into the modern view of differentials, as will be seen.

Every Pfaffian equation $\omega = \sum_{i=1}^n b_i(x)dx_i = 0$ that is a linear combination of the ω_i is said by Frobenius to belong to the system (9.5). As he observed, this is equivalent to saying that the n-tuple $b = (b_1, \dots, b_n)$ is a linear combination of the rows of $A = (a_{ij})$. He then defined $f(x) = C$ to be an integral of the system (9.5) if f is such that its differential df belongs to the system (9.5), or, equivalently, if $\nabla f = (\partial f/\partial x_1, \dots, \partial f/\partial x_n)$ is a linear combination of the rows of A. (This simple, algebraically oriented definition should be compared with the one given by Mayer.) There can exist at most r independent functions defining integrals of the system. When this is the case the system (9.5) is said to be *complete* [26, p. 286].

In his approach to duality, Hamburger tacitly assumed for convenience that since $A = (a_{ij})$ has rank r the $r \times r$ minor matrix A_r of A defined by the first r rows and columns of A has nonzero determinant. Suppose $f(x_1, \dots, x_n) = C$ in an integral of the Pfaffian system (9.5) so that ∇f is a linear combination of the the rows of A.[38] Then the $(r+1) \times n$ matrix

$$A^* = \begin{pmatrix} A \\ \nabla f \end{pmatrix} = \begin{pmatrix} a_1^{(1)} & \cdots & a_n^{(1)} \\ \vdots & \ddots & \vdots \\ a_1^{(r)} & \cdots & a_n^{(r)} \\ \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{pmatrix}$$

has rank r, i.e., every $(r+1) \times (r+1)$ minor matrix of A^* has vanishing determinant. These determinants set equal to 0 yield a system of linear homogeneous partial differential equations having f as solution. Not all of these equations are independent, however,

[38] Hamburger's presentation is not as clear as I am indicating. He never defined what it means for $f = C$ to be an integral, but concluded that the equations (9.5) imply that $df = 0$ and that this in turn implies the $n - r$ vanishing determinants indicated below.

and Hamburger singled out the $n-r$ systems that arise by forming minors using the first r columns of M^* plus one of the remaining $n-r$ columns. Hamburger's approach was thus more algebraic and general than Mayer's, and that is undoubtedly why Frobenius mentioned it. Hamburger, however, never justified that his equations were independent. Frobenius avoided the need for this by proceeding somewhat differently after mentioning Hamburger's approach.

Frobenius considered the linear homogeneous system of equations associated to the Pfaffian system (9.5), namely,

$$a_{i1}u_1 + \cdots + a_{in}u_n = 0, \ i = 1, \ldots, r, \tag{9.6}$$

or in more familiar notation $A\mathbf{u} = \mathbf{0}$, where $A = (a_{ij})$ is the $r \times n$ matrix of the coefficients and $\mathbf{u}$ is now used to denote the vector with u_i as components. It was well-known at this time that the system $A\mathbf{u} = \mathbf{0}$ – with A a matrix of constants and of rank r – has $n-r$ linearly independent solutions. For example, in the 1870 edition of Baltzer's text on determinants, which is often cited by Frobenius, there is a theorem due to Kronecker which gives a formula for the general solution to $A\mathbf{u} = \mathbf{0}$ [3, pp. 66–7].

Theorem 9.1 (Kronecker's theorem) *Given an $m \times n$ matrix $M = (a_{ij})$ of rank r, suppose for specificity that the minor determinant of M formed from its first r rows and columns is nonzero. Then the solutions to the homogeneous system $Mu = 0$ are given as follows. Consider the $(r+1) \times (r+1)$ matrix*

$$\begin{pmatrix} a_{11} & \cdots & a_{1r} & a_{1r+1}u_{r+1} + \cdots + a_{1n}u_n \\ \vdots & \ddots & \vdots & \vdots \\ a_{r1} & \cdots & a_{rr} & a_{rr+1}u_{r+1} + \cdots + a_{rn}u_n \\ * & \cdots & * & * \end{pmatrix},$$

where the last row can be anything. Then the general solution to $Mu = 0$ is $u^t = (C_1/C_{r+1}, \ldots, C_r/C_{r+1}, u_{r+1}, \ldots u_n)$, where $C_1, \ldots, C_{r+1}$ denote the cofactors along the last row.

The general solution thus involves $n-r$ "free variables" $u_{r+1}, \ldots, u_n$, which occur in the cofactors C_i, $i \neq r+1$, and by successively setting one of the u_i equal to 1 and the rest 0, we obtain $n-r$ solutions that can be shown to be linearly independent. Frobenius surely knew this theorem, but in the purely algebraic part of his paper, he presented a more elegant, perfectly general way of establishing the $n-r$ independent solutions, which seems to have originated with him.[39] It is of historical interest because it shows how he was compelled by a penchant for algebraic elegance and generality to a type of linear algebra with strikingly modern overtones despite the continued reliance upon determinants.

Frobenius proceeded as follows [26, p. 255ff]. Consider an $r \times n$ system of equations $A\mathbf{u} = \mathbf{0}$ with the coefficients of A assumed constant and $r = \text{rank } A$. Then we may pick

[39] Judging by his remarks some thirty years later [31, p. 349f.].

$n - r$ n-tuples $\mathbf{w}_k = (w_{k1}, \ldots, w_{kn})$ so that the $n \times n$ matrix D with rows consisting of the r rows $\mathbf{a}_i = (a_{i1}, \ldots, a_{in})$ of A followed by the $n - r$ rows $\mathbf{w}_k$, viz.

$$D = \begin{pmatrix} A \\ \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_{n-r} \end{pmatrix},$$

satisfies $\det D \neq 0$. Then let b_{kj} denote the cofactor of $\det D$ corresponding to the w_{kj} coefficient. If D is modified to D^* by replacing row $\mathbf{w}_k$ by row $\mathbf{a}_i$, then $\det D^* = 0$ since it has two rows equal. Expanding $\det D^*$ by cofactors along the changed row yields the relation

$$\sum_{j=1}^{n} a_{ij} b_{kj} = 0, \; i = 1, \ldots, r, \; k = 1, \ldots n - r. \tag{9.7}$$

I will express this relation in the more succinct form $\mathbf{a}_i \cdot \mathbf{b}_k = 0$ for any $i = 1, \ldots, r$ and any $k = 1, \ldots, n - r$, where $\mathbf{b}_k = (b_{k1}, \ldots, b_{kn})$. Using well-known determinant-theoretic results that go back to Jacobi, Frobenius easily concluded that the $(n - r) \times n$ matrix B with the $\mathbf{b}_k$ as its rows has full rank $n - r$ so that the $\mathbf{b}_k$ represent $n - r$ linearly independent solutions to the homogeneous system $A\mathbf{u} = \mathbf{0}$. He also showed that no more than $n - r$ independent solutions of $A\mathbf{u} = \mathbf{0}$ can exist. However, he did not stop with this result.

Frobenius defined the two coefficient systems $A = (a_{ij})$ and $B = (b_{ij})$ to be "associated" or "adjoined"; likewise the two systems of equations $A\mathbf{u} = \mathbf{0}$ and $B\mathbf{v} = \mathbf{0}$ are said to be adjoined. The reason for this terminology was the following immediate consequence of (9.7): $\mathbf{a} \cdot \mathbf{b} = 0$, where $\mathbf{a}$ is any linear combination of the rows of A and $\mathbf{b}$ is any linear combination of the rows of B. In other words, as we could now say it, the row space of A, Row A, and Row B are orthogonal. In fact, given the ranks of A and B, Frobenius realized that any $\mathbf{a}$ such that $\mathbf{a} \cdot \mathbf{b} = 0$ for all $\mathbf{b} \in$ Row B must belong to Row A and vice versa, i.e., he realized what would now be expressed by

$$\text{Row } A = [\text{Row } B]^{\perp} \quad \text{and} \quad \text{Row } B = [\text{Row } A]^{\perp}. \tag{9.8}$$

Furthermore, adjoinedness meant "the coefficients of the one system of equations are the solutions of the other" [26, p. 257], i.e., in modern terms

$$\text{Row } A = \text{Null } B \quad \text{and} \quad \text{Row } B = \text{Null } A, \tag{9.9}$$

where, e.g., Null A denotes the nullspace of A. This was an immediate consequence of the obvious fact, frequently used by Frobenius, that $\mathbf{u}$ is a solution to a system of equations $A\mathbf{u} = \mathbf{0}$ if and only if $\mathbf{u} \cdot \mathbf{a} = 0$ for any $\mathbf{a}$ that is a row of A or a linear combination thereof, i.e., what would now be expressed by

$$\text{Null } A = [\text{Row } A]^{\perp}. \tag{9.10}$$

Returning now from this excursion into Frobenius-style linear algebra, let us consider how he applied the key relation (9.8) as it relates to the matrix A of (9.6) to obtain the dual of the Pfaffian system (9.5). By his definition $f = C$ is a solution to the Pfaffian system (9.5) if and only if ∇f is a linear combination of the rows of $A = (a_{ij})$, and by (9.8) this is equivalent to saying that ∇f is perpendicular to the rows of B, i.e.,

$$X_k(f) \stackrel{\text{def}}{=} b_{k1}\frac{\partial f}{\partial x_1} + \cdots + b_{kn}\frac{\partial f}{\partial x_n} = 0, \quad k = 1, \ldots, n-r. \tag{9.11}$$

The above system is then defined to be the system of linear partial differential equations adjoined to the Pfaffian system (9.5) – the dual system as I will call it. Likewise, as Frobenius showed, if one starts with a system $X_k(f) = 0$, $k = 1, \ldots, n-r$, the above reasoning can be reversed to define the associated Pfaffian system $\omega_i = 0$, $i = 1, \ldots, r$.

Frobenius used the correspondence between the systems $X_\nu(f) = 0$ of (9.11) and $\omega_i = 0$ of (9.5) to translate Clebsch's integrability condition (5.11) for the system $X_\nu(f) = 0$ into an integrability condition for the system $\omega_i = 0$ [26, §14]. His reasoning utilized his version of (9.8)–(9.9). With this fact in mind, note the following implications of the duality correspondence $X_k(f) = 0 \leftrightarrow \omega_i = 0$. First of all, $\omega = \sum_{i=1}^n a_i(x)dx_i$ belongs to the system $\omega_i = 0$ if and only if $\mathbf{a} = (a_1, \ldots, a_n) \in \text{Row } A = [\text{Row } B]^\perp = \text{Null } B$. Secondly, $X(f) = \sum_{i=1}^n b_i(x)\partial f/\partial x_i = 0$ belongs to the system $X_k(f) = 0$, i.e., $X(f)$ is a linear combination of the $X_k(f)$, if and only if $\mathbf{b} = (b_1, \ldots, b_n)$ is a linear combination of the $\mathbf{b}_k$, i.e., $\mathbf{b} \in \text{Row } B = [\text{Null } B]^\perp$.

Let us now consider, along with Frobenius, Clebsch's integrability condition (5.11). It implies that if the equations

$$X(f) = \sum_{i=1}^n b_i(x)\partial f/\partial x_i = 0 \quad \text{and} \quad Y(f) = \sum_{i=1}^n c_i(x)\partial f/\partial x_i = 0$$

belong to the system $X_k(f) = 0$, then so does $X(Y(f)) - Y(X(f)) = 0$, where

$$X(Y(f)) - Y(X(f)) = \sum_{i=1}^n \sum_{j=1}^n \left(b_j\frac{\partial c_i}{\partial x_j} - c_j\frac{\partial b_i}{\partial x_j}\right)\frac{\partial f}{\partial x_i}. \tag{9.12}$$

In view of the above preliminary remarks, the integrability condition may be stated as follows. Let $\mathbf{b} = (b_1, \ldots, b_n)$ and $\mathbf{c} = (c_1, \ldots, c_n)$ and let $[\mathbf{b}, \mathbf{c}]$ denote the coefficient n-tuple of $X(Y(f)) - Y(X(f))$ as given in (9.12), i.e.,

$$[\mathbf{b}, \mathbf{c}]_i = \sum_{j=1}^n \left(b_j\frac{\partial c_i}{\partial x_j} - c_j\frac{\partial b_i}{\partial x_j}\right), \; i = 1, \ldots, n.$$

Then the integrability criterion is that if $\mathbf{b} \in [\text{Null } B]^\perp$ and $\mathbf{c} \in [\text{Null } B]^\perp$, then also $[\mathbf{b}, \mathbf{c}] \in [\text{Null } B]^\perp$. Since $\mathbf{a} \in \text{Null } B$, this says that if $\mathbf{b} \cdot \mathbf{a} = 0$ and $\mathbf{c} \cdot \mathbf{a} = 0$ for all $\mathbf{a} \in \text{Null } B$, then $[\mathbf{b}, \mathbf{c}] \cdot \mathbf{a} = 0$ for all $\mathbf{a} \in \text{Null } B$. In terms of the Pfaffian expression $\omega = \sum_{i=1}^n a_i(x)dx_i$ defined by $\mathbf{a}$, the criterion takes the following form. Let $\omega(\mathbf{b}) = \sum_{i=1}^n a_i(x)b_i = \mathbf{a} \cdot \mathbf{b}$. (In effect ω is being evaluated at the vector $\mathbf{b}$ in the tangent space at $x \in G$, which in a natural way is a complex-analytic manifold.) The above considerations may then be summarized in the following form.

$$\text{If } \omega(\mathbf{b}) = 0 \text{ and } \omega(\mathbf{c}) = 0 \text{ for all } \omega \text{ in (9.5), then } \omega([\mathbf{b}, \mathbf{c}]) = 0. \tag{9.13}$$

Now

$$\omega([\mathbf{b},\mathbf{c}]) = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(b_j\frac{\partial c_i}{\partial x_j} - c_j\frac{\partial b_i}{\partial x_j}\right)a_i. \tag{9.14}$$

Frobenius observed that he could rewrite this expression by utilizing the relations $\omega(\mathbf{b}) = 0$ and $\omega(\mathbf{c}) = 0$. For example, differentiating the first equation with respect to x_j yields

$$0 = \frac{\partial}{\partial x_j}\left(\sum_{i=1}^{n} a_i b_i\right) = \sum_{i=1}^{n}\frac{\partial a_i}{\partial x_j}b_i + \sum_{i=1}^{n} a_i\frac{\partial b_i}{\partial x_j},$$

which gives $\sum_{i=1}^{n} a_i\partial b_i/\partial x_j = -\sum_{i=1}^{n} b_i\partial a_i/\partial x_j$. Doing the same to $\omega(\mathbf{c}) = 0$ likewise shows that $\sum_{i=1}^{n} a_i\partial c_i/\partial x_j = -\sum_{i=1}^{n} c_i\partial a_i/\partial x_j$. If these expressions are substituted in (9.14), the result is that

$$\omega([\mathbf{b},\mathbf{c}]) = \sum_{i,j=1}^{n}\left(\frac{\partial a_i}{\partial x_j} - \frac{\partial a_j}{\partial x_i}\right)b_i c_j.$$

Thus $\omega([\mathbf{b},\mathbf{c}]) = \Omega(\mathbf{b},\mathbf{c})$, where Ω is the bilinear covariant associated to ω, and so (9.13) can be reformulated in terms of bilinear covariants. In this manner Frobenius showed that the Jacobi–Clebsch Theorem 5.3, which asserts that the system $X_k(f) = 0$ is complete if and only if it satisfies the Clebsch integrability condition (5.11), translates into the following completeness theorem for the system $\omega_i = 0$ [26, p. 290].

Theorem 9.2 (Frobenius integrability theorem) *Given a system of r linearly independent Pfaffian equations $\omega_i = \sum_{j=1}^{n} a_{ij}dx_j = 0$, $i = 1,\ldots,r$, it is complete if and only if the following integrability condition holds: whenever $\omega_i(\mathbf{b}) = 0$ and $\omega_i(\mathbf{c}) = 0$ for all i, it follows that $\Omega_i(\mathbf{b},\mathbf{c}) = 0$ for all i, where Ω_i is the bilinear covariant associated to ω_i.*

Frobenius's theorem has been stated more or less as he did, i.e., with all reference to the dependency of everything upon the x-variables suppressed as was customary at the time. However, the theorem and its attendant definitions, as well as Frobenius's direct proof (mentioned below) make sense only for $x \in G$, the set of generic points x where rank $\big(a_{ij}(x)\big) = r$ (Sect. 3). For example, consider Frobenius's definition of the dual system (9.11). It is based upon the existence of the $n - r$ linearly independent n-tuples $\mathbf{b}_k$ comprising B. In view of how the coordinates of $\mathbf{b}_k$ were defined above – as certain cofactors of $D(x) = \begin{pmatrix} A(x) \\ W(x) \end{pmatrix}$ – it is clear that we can choose $W(x)$ to be analytic and of full rank $n - r$ in a neighborhood of a fixed $x \in G$ but not necessarily for all $x \in \mathbf{C}^n$. In other words the desired coefficients $b_{kj}(x)$ and hence the dual system (9.11) is only defined locally, in a neighborhood of some $x \in G$. Likewise, the coordinate representation of ω_i as $\sum_{1=1}^{n} a_{ij}(x)dx_j$ is really a local one, valid in a neighborhood of $x \in G$. Thus the definition of a solution $f = C$ to a Pfaffian system is implicitly a local one as well as the attendant definition of completeness, meaning that at each $x \in G$ there is a neighborhood in which independent functions $f_1,\ldots,f_r$ exist such that $f_i = C_i$ are integrals of the Pfaffian system in the sense defined by Frobenius and given above. The integrability condition must hold for all $x \in G$, where $\omega_i(\mathbf{b})$, $\omega_i(\mathbf{c})$, and $\Omega_i(\mathbf{b}, \mathbf{c})$ are computed for vectors $\mathbf{b}$, $\mathbf{c}$ in the tangent space of G at x, using the local coordinate representations of ω_i and Ω_i in a neighborhood of x. Of course, none of this is in Frobenius's paper, but the point here is that in his Theorem 9.2 we have in tacitly local form what was to become, in modern times, the Frobenius integrability theorem. Indeed, the direct proof of Theorem 9.2 given by Frobenius [26, pp. 291–5] can be made rigorous by interpreting it locally along the lines indicated above. Although the modern Frobenius integrability theorem is essentially a local result, it is usually set within the global context of a manifold. Such a notion seems to have been foreign to Frobenius's way of thinking in 1876, as it was to most mathematicians at that time.[40]

Theorem 9.2 is the source of the appellation "Frobenius's Theorem" which is now commonplace. The first to so name it appears to have been Élie Cartan in his 1922 book on invariant integrals.[41] Singling out Frobenius's name to attach to this theorem is a bit unfair to Jacobi and Clebsch, since the above theorem is just the dual of the Jacobi–Clebsch Theorem 5.3. Furthermore, although Frobenius also gave a proof of the theorem that is independent of the Jacobi–Clebsch Theorem and the consideration of partial differential equations, that proof was, as he explained [26, p. 291], simply a more algebraic and "symmetrical" version of one by Heinrich Deahna (1815–1844) that he had discovered in Crelle's *Journal* for 1840 [20].

Deahna's name is not a familiar one to mathematicians nowadays, although it seems likely this is largely due to his untimely death at age 29. He had studied at the universities in Göttingen and Marburg and at Göttingen won a prize for work on the moments of inertia of regular solids. He was an assistant teacher (*Hülfslehrer*) at the Gymnasium in Fulda when he died. In 1840 he published two papers in Crelle's *Journal*. One gave a new proof of the fundamental theorem of algebra and the other was the paper cited by Frobenius. That paper was motivated by the the necessary and sufficient conditions that a total differential equation $\omega = \sum_{i=1}^{n} a_i(x)dx_i = 0$ be such that it has an integral solution $f(x_1, \dots, x_n) = C$. Deahna complained that although necessary conditions for the existence of an integral solution were known (as indicated near the beginning of Sect. 4 above), their sufficiency had not been established. He proposed to fill this gap by stating and proving necessary and sufficient conditions for integrability. He then went on to do the same for a system of total differential equations $\omega_i = 0, i = 1, \dots, m$.

Deahna did not write his equation in the symmetric form $\omega = 0$ given above, but rather in the form

[40] Although a notion of a manifold can be traced back to Riemann [68], a careful definition such as Frobenius would have accepted was first given by Weyl in his 1913 book on Riemann surfaces – a book, incidentally, that Frobenius admired [23, pp. 12–15]. For a discussion of that aspect of Weyl's book, see [46, §9.5].

[41] Chapter X is on completely integrable Pfaffian systems and the first section is entitled "Le théorème de Frobenius" [11, p. 99ff]. Immediately after stating Frobenius's theorem (as above) Cartan pointed out that the integrability condition could be formulated differently using his calculus of differential forms. This formulation, which he first published in 1901, is given in (11.5) below. In his monograph of 1945 on differential forms [12, p. 49], Cartan gave the necessary and sufficient condition for completeness in his own form (11.5) and no mention is made of Frobenius.

$$dx = \sum_{i=1}^{n} X_i dx_i, \quad X_i = X_i(x, x_1, \dots, x_n).$$

He articulated the completeness condition as follows: The variation of $X_1, \dots, X_n$ should vanish "when $x_1, \dots, x_n$ increase by arbitrary variations $\delta x_1, \dots, \delta x_n$ while x increases by $\delta x = \sum_{i=1}^{n} X_i \delta x_i$" [20, p. 340]. Likewise, for a system of total differential equations, Deahna articulated the completeness condition in terms of variational notions [20, pp. 333–4]. Thus there is no concept of a bilinear covariant in Deahna's work, although as noted earlier, Lipschitz had also used variational principles to show in Theorem 6.1 that $\Omega(\mathbf{b}, \mathbf{c})$ was indeed the analytical analog of a bilinear covariant. Frobenius, equipped with the concept of a bilinear covariant, could see that Deahna's condition was equivalent to his own and that, in addition, Deahna's proof was viable.

There are thus bona fide historical reasons for renaming Frobenius's Theorem 9.2 the Jacobi–Clebsch–Deahna–Frobenius Theorem.[42] That Cartan should focus upon Frobenius is nonetheless also understandable on historical grounds. As we shall see in Sect. 11, from Cartan's perspective it was Frobenius who had first revealed the important role the bilinear covariant can play in the theory of Pfaffian equations. Such a role is exemplified by the above theorem, and inspired by such applications Cartan sought to apply the bilinear covariant – the derivative of a 1-form in his calculus – to a wider range of problems. In the concluding section I will offer some tentative arguments to the effect that the phenomenon of the Frobenius Integrability Theorem is paradigmatic of one of the principal ways in which the work of Frobenius has affected the emergence of present-day mathematics.

Let us now briefly return to the reason Frobenius had established his complete integrability theorem. As we saw at the end of Sect. 8, Frobenius's proof of the Analytical Classification Theorem 8.1 required establishing the completeness of a certain system of partial differential equations $A_i(f) = 0$, $i = 1, \dots, d$. By means of a rather involved line of reasoning involving duality considerations among other things, Frobenius argued [26, pp. 313–16] that the desired completeness would follow provided the following special type of Pfaffian system is complete. Let $\omega = \sum_{i=1}^{n} a_i dx_i$ denote any Pfaffian expression in n variables $x_1, \dots, x_n$ with associated bilinear covariant $\Omega(\mathbf{u}, \mathbf{v}) = \sum_{i,j=1}^{n} a_{ij} u_i v_j$. The coefficients of Ω define a system of Pfaffian equations, namely $\omega_i = \sum_{j=1}^{n} a_{ij} dx_j = 0$, $i = 1, \dots, n$. These equations are not necessarily linearly independent but if $r = \text{rank } (a_{ij})$, then r of them are, and it is this system of r independent equations that needs to be complete. In other words, the system of all n equations needs to have r independent integrals. Clearly this will follow from Frobenius's Theorem 9.2 if $\omega_i(\mathbf{b}) = 0$ and $\omega_i(\mathbf{c}) = 0$ for $i = 1, \dots, n$, implies that $\Omega_i(\mathbf{b}, \mathbf{c}) = 0$ for $i = 1, \dots, n$, where Ω_i is the binary covariant of ω_i.

Frobenius proved this as follows [26, §22]. By virtue of the special nature of the system $\omega_i = 0$, the relations $\omega_i(\mathbf{b}) = 0$, $\omega_i(\mathbf{c}) = 0$ can be written in the form

[42] Upon reading an early draft of this essay, the late Armand Borel suggested that such a renaming would be more appropriate were it not too cumbersome. At his suggestion I have now accorded more space to Deahna's work and its motivation and wish to register my gratitude to Borel for yet another well-informed suggestion aimed at improving my work.

$$(a) \quad \sum_{i=1}^{n} a_{ij} b_i = 0 \quad \text{and} \quad (b) \quad \sum_{j=1}^{n} a_{ij} c_j = 0,$$

where the skew symmetry $a_{ji} = -a_{ij}$ has been used in (a). Then if (b) is differentiated with respect to x_k and the resulting equation multiplied by b_i and summed over i one gets using (a)

$$\sum_{i,j=1}^{n} \frac{\partial a_{ij}}{\partial x_k} b_i c_j = 0. \tag{9.15}$$

Clebsch [15, p. 162] had introduced the identity

$$\frac{\partial a_{ij}}{\partial x_k} + \frac{\partial a_{jk}}{\partial x_i} + \frac{\partial a_{ki}}{\partial xj} = 0,$$

to obtain the key relation (5.11) of his direct method. Frobenius used the identity to rewrite (9.15) as

$$\sum_{i,j=1}^{n} \left(\frac{\partial a_{ki}}{\partial x_j} - \frac{\partial a_{kj}}{\partial x_i} \right) b_i c_j = 0.$$

This of course shows that $\Omega_k(\mathbf{b}, \mathbf{c}) = 0$, where Ω_k is the bilinear covariant of ω_k, thereby fulfilling the integrability condition of Theorem 9.2.

With the proof of the Existence Theorem, Frobenius's proof of the Analytical Classification Theorem was accomplished for the case of a Pfaffian of class $p = 2m$. To handle the case in which ω is of class $p = 2m+1$ – meaning rank $A = 2m$ and rank $\widehat{A} = 2m+2$ – he sought to slightly modify ω by replacing $a_i(x)$ with $a_i^*(x) = a_i(x) + \partial f(x)/\partial x_i$ for some function f. For any such modification the bilinear covariant matrix A is easily seen to be unaltered, and the idea is to choose f so that the rank of $\widehat{A}$ would drop to $2m$, thereby making the modified Pfaffian of class $p = 2m$. In order to establish the existence of a suitable function f Frobenius once again applied his integrability theorem to a system $\omega_i = 0$ generated by the bilinear covariant of a Pfaffian ω [26, pp. 323–31].

10. Initial reactions

Compared to the work on Pfaff's problem by his predecessors, Frobenius's contribution was unique in two fundamental respects. It was the first clear and systematic attempt to deal with the problem in complete algebraic generality and by methods – dominated by rank considerations – well-suited to such generality. It was a Berlin-style solution to the problem of Pfaff. The other unique feature was the introduction of the bilinear covariant of a Pfaffian as a key theoretical tool. Of course, the skew symmetric Jacobi matrix had been central to the theory since its introduction in 1827, but by thinking of it as defining a bilinear form Ω associated to a linear form ω and by establishing the importance to the theory of its invariance under variable changes as in the Analytical Classification Theorem 8.1 and the integrability theorem (Theorem 9.2), Frobenius had added a new dimension to the theory that was eventually explored more deeply and broadly by Cartan starting in 1899 as will be seen in the next section. Here I consider briefly the reaction to Frobenius's paper in the intervening years.

The appearance of Frobenius's paper prompted two mathematicians, Sophus Lie (1842–1899) and Gaston Darboux (1842–1917), to publish papers containing some analogous results that had been discovered independently of his work.

Lie's interest in the problem of Pfaff was a natural part of his interest during the early 1870s in the theory of first order partial differential equations. Indeed, the theory of contact transformations that he developed in this connection was directly related to Pfaffian equations, since by 1873 he was characterizing contact transformations as transformations in $2n+1$ variables $z, x_1, \ldots, x_n, p_1, \ldots, p_n$ which leave the Pfaffian equation $dz - \sum_{i=1}^{n} p_i dx_i = 0$ invariant.[43] In this period Lie claimed that all his work on partial differential equations and contact transformations could be extended to the general problem of Pfaff, but the only work he published concerned an efficient method of integrating a Pfaffian equation in an even number of variables in the generic case. It was not until Klein called his attention to Frobenius's paper [26] on the general problem that Lie composed his own paper on the subject, which he published in 1877 in the Norwegian journal that he edited [60].

Invoking without proof theorems "from the theory of partial differential equations" that were actually based more specifically on his (largely unpublished) theory of contact transformations, Lie quickly arrived at Clebsch's Theorem 5.1 as his Theorem I: either (I) $2m$ independent functions F_i, f_i exist such that $\omega = \sum_{i=1}^{m} F_i df_i$ or (II) $2m+1$ independent functions Φ_i, φ_i exist such that $\omega = d\varphi_0 + \sum_{i=1}^{m} \Phi_i d\varphi_i$. Lie called these expressions normal forms of ω. He focused upon the number of functions in a normal form. This number is of course precisely the Frobenius class number p, although Lie made no reference to Frobenius's paper, even though he knew through Klein of its existence.[44] Whether Lie ever looked at Frobenius's paper is unclear, but he gave his own proof that p is the sole invariant of ω, i.e., that any two normal forms of ω have the same p and two Pfaffians ω and ω' with the same p can be transformed into one another [60, §2]. Lie also described a procedure for determining p for a given Pfaffian ω that was a development of observations made by Jacobi [51, §22] and Natani [65, §8] in case (I) [60, §4]. Like the formulations of his predecessors, Lie's was cumbersome. In the light of Frobenius's paper it is easily seen that if $A = (a_{ij})$ is the Jacobi skew symmetric matrix associated to $\omega = \sum_{i=1}^{n} a_i dx_i$, and $a^t = (a_1, \ldots, a_n)$, then one always has

[43] A discussion of the evolution of Lie's theory of contact transformations and its relation to the origins of his theory of transformation groups is given in Chapters 1–2 of my book [46].

[44] In a letter to Mayer in March 1873 he wrote, "Frobenius's work is probably very good? Since I have little time, I have not been able to bring myself to read it" [21, p. 713]. In his reply [21, p. 714] Mayer pointed out that it was rather strange that Frobenius did not mention work of either of them in his paper, and this may have prompted Lie's decision not to cite Frobenius's paper.

$p = \text{rank} \begin{pmatrix} A \\ a^t \end{pmatrix}$ (as Frobenius surely realized[45]), and this is what Lie was getting at by his procedure, albeit without seeming to fully realize it.

The title of Lie's paper, "Theory of the Pfaffian Problem I" suggested a sequel, and Lie had indeed entertained the idea of a sizable second part containing an extension to the general problem of Pfaff of his theory of contact transformations, first order partial differential equations, and transformation groups. This part never materialized. In 1883 Lie explained to Mayer that "Because of Frobenius I have lost interest in the problem of Pfaff .... I have already written too much that goes unread" [21, p. 714]. No doubt to Lie Frobenius's paper was representative of the "analytical" mathematical style of the Berlin school, which he and Klein opposed in favor of a more intuitive, geometrical or "synthetic" approach. Since in the 1870s the Berlin school was one of the most prestigious and influential centers for mathematics, Lie's remarks probably reflect the seeming hopelessness of competing with the Berlin treatment of the problem of Pfaff presented by Frobenius.

While Frobenius was working on the problem in 1876, so was Gaston Darboux, although what he wrote up was not immediately submitted for publication. Instead he gave his notes to Bertrand, who wished to incorporate them into his lectures at the Collège de France during January 1877. As Darboux explained in 1882 when he eventually published his work [19, p. 15n],

> Shortly thereafter a beautiful memoir by Mr. Frobenius appeared ... bearing a date earlier than that of January 1877 (September 1876) and there this learned geometer proceeded in a manner somewhat analogous to what I had communicated to Mr. Bertrand in the sense that it was based upon the use of invariants and of the bilinear covariant of Mr. Lipschitz. Upon returning recently to my work, it seemed to me that my exposition was more calculation-free and, in view of the importance the method of Pfaff has assumed, that it would be of interest to make it known.

This passage makes it fairly certain that it was by virtue of Frobenius's "beautiful" (but calculation-laden) treatment of the problem of Pfaff that Darboux was now publishing his own approach. By considering the problem in such great generality and detail, I suspect that Frobenius had contributed greatly to the perception of "the importance the method of Pfaff has assumed," and it was because Darboux's own approach was more calculation-free than Frobenius's that he now thought it worthwhile to publish it. It would seem from Darboux's remarks that the first part of his memoir, discussed below, represented what he had written in 1876 and given to Bertrand.

---

[45] Frobenius apparently never explicitly gave the above expression for p, but it is an immediate consequence of his remark directly following the Even Rank Theorem, namely that Theorem 7.1 implies that $\widehat{A} = \begin{pmatrix} A & a \\ -a^t & 0 \end{pmatrix}$ has the same rank $2m$ as A (whence $p = 2m$) if and only if this is true of the rank of $\begin{pmatrix} A \\ a^t \end{pmatrix}$ [26, p. 263]. Thus the latter matrix has rank $2m + 1$ precisely when $p = 2m + 1$. At this point in his paper Frobenius had not yet introduced the class number p and so was not in a position to explain this point to his readers.

Like Frobenius, Darboux began by establishing the "fundamental formula" that shows that $\sum_{i,j=1}^{n} a_{ij} dx_i \delta_j$ is a bilinear covariant, i.e., Lipschitz's Theorem 6.1.[46] However, he did not make Frobenius's bilinear covariant $\Omega(u, v)$ the conceptual basis for his theory but rather focused on the associated system of linear differential equations

$$a_{i1}(x)dx_1 + \cdots + a_{in}(x)dx_n = \lambda a_i(x)dt, \quad i = 1, \ldots, n,$$

where t denotes an auxiliary variable (so the x_i can be regarded as any functions of t) and λ is "a quantity that could be chosen arbitrarily as 0, a constant, or a function of t, according to the case" [19, p. 19]. For brevity I will denote this system in matrix notation by $Adx = \lambda adt$. Given two Pfaffians ω and ω' in variables x_i and x_i', respectively, and with respective systems $Adx = \lambda adt$ and $A'dx' = \lambda a'dt$, Lipschitz's Theorem 6.1 then implied that if $\omega = \omega'$ by virtue of $x = \varphi(x')$ and $d\varphi : dx \to dx'$, then the corresponding systems are likewise equivalent in the sense that $Adx = \lambda adt$ transforms into $A'dx' = \lambda a'dt$.

Like Lie, Darboux first quickly established Clebsch's Theorem 5.1. His proof was based upon a theorem (stated at the end of Sect. IV) that he proved carefully in the two generic cases, namely n is even and $\det A \neq 0$ and n is odd (so $\det A = 0$) and rank $A = n - 1$ [19, §III]. He considered briefly only the nongeneric case when n is even [19, §IV], and in this connection he made an assumption that can be briefly summarized as follows. Suppose $C = C(x)$ is an $n \times n$ matrix of rank $r < n$. (Here, as with Frobenius, we grant Darboux the tacit assumption that r is the maximal rank and holds in a neighborhood of some fixed x.) Then without loss of generality we can assume that in this neighborhood a minor involving the first r row indices does not vanish, so that the first r rows of $C(x)$ are linearly independent – or "distinct" as Darboux said. Darboux then took for granted that $n-r$ functions $\varphi_1(x), \ldots, \varphi_{n-r}(x)$ may be chosen so that the first r rows of $C(x)$ together with the $n-r$ gradients $\nabla\varphi_i(x)$ form a linearly independent set, i.e., that the $n \times n$ matrix formed from $\text{Row}_1(C), \ldots, \text{Row}_r(C), \nabla\varphi_1, \ldots, \nabla\varphi_{n-r}$ does not vanish (in a neighborhood of x). Certainly it is easy to pick the φ_i so the $n - r$ gradients $\nabla\varphi_i$ are linearly independent, but how to pick the φ_i so that each $\nabla\varphi_i$ is also independent of the first r rows of $C(x)$? One way is to pick them in the orthogonal complement of these rows, which means that $X_j(\varphi) = \text{Row}_j(C) \cdot \nabla\varphi = 0,\ j = 1, \ldots, r$ must have $n-r$ independent solutions. By the Jacobi–Clebsch Theorem 5.3 this requires that the system $X_j(\varphi) = 0,\ j = 1, \ldots, r$ be complete, but this is not at all apparent. Darboux thus glossed over precisely the same sort of problem that (as indicated at the end of Sect. 8) had compelled Frobenius into a lengthy detour that involved among other things duality considerations and his Integrability Theorem 9.2.

From the above-mentioned theorem at the end of Sect. IV, Clebsch's Theorem 5.1 then followed readily but of course was not rigorously established due to its dependence on that theorem. Darboux expressed the forms (I) and (II) of Clebsch's Theorem as follows:[47] Let $x \to y, z$ denote the variable change such that either

---

[46] Darboux referred to Lipschitz's Theorem in an 1882 footnote added to Part I of his paper, but there is no such reference in the paper itself so that Darboux may have rediscovered Lipschitz's Theorem in 1876.

[47] For notational consistency I have ordered the two cases as in the earlier discussion of Clebsch and Frobenius; Darboux reversed the order.

$$(I)\ \omega = z_1 dy_1 + \cdots + z_m dy_m \quad \text{or} \quad (II)\ \omega = dy - [z_1 dy_1 + \cdots + z_m dy_m].$$

He then asked: given a Pfaffian ω in the original variables x (a) how can one determine whether (I) or (II) holds, and (b) how can the number m be determined [19, p. 27]? He claimed that by putting ω in one of the two forms (I) or (II) as per Clebsch's Theorem, it is easy to answer these questions since the answers are "invariant" under coordinate change and are immediately clear in the y, z variables.

Consider, e.g., the case (II) of Clebsch's Theorem. Then in the new variables $A'dx' = \lambda a'dt$ becomes upon calculation

$$dy_i = 0, \quad dz_i = -\lambda dz_i, \ i = 1, \ldots, m, \text{ and } 0 = \lambda dt. \tag{10.1}$$

Thus a solution can only exist for $\lambda = 0$ in case (II), and the system $A'dx' = \lambda a'dt$ reduces to a completely integrable system of $2m$ equations $dy_i = 0, dz_i = 0, i = 1, \ldots m$. (If $y_i = \phi_i(x)$, and $z_i = \psi_i(x)$ give the variable change, then the solutions are $\phi_i(x) = C_i, \psi_i(x) = D_i, i = 1, \ldots m$.) Likewise in case (I) in the new variables $A'dx' = \lambda a'dt$ takes the form

$$dy_i = 0, \quad dz_i = \lambda z_i dt, \ i = 1, \ldots, m, \tag{10.2}$$

which has solutions with $\lambda = \text{const.} \neq 0$, namely $y_i = C_i, i = 1, \ldots m$ and, since $z_i = C_i' e^{\lambda t}, z_i/z_1 = D_i, i = 2, \ldots m$ upon elimination of t [19, pp. 28–9].

Although Darboux didn't mention it explicitly, note that in case (I) $Adx = \lambda adt$ in its transformed form (10.2) also has a solution for $\lambda = 0$, namely $y_i = C_i$ and $z_i = D_i$. Also note that in case (I) for both $\lambda = 0$ and $\lambda = 1$ the system $Adx = \lambda adt$ reduces to $2m$ equations, whereas in case (II) the number of equations is $2m$ for $\lambda = 0$ and $2m + 1$ in the impossible case $\lambda = 1$. Given the invariant property of these conclusions the same results hold regarding the number of distinct equations in $Adx = \lambda adt$ for $\lambda = 0$ and $\lambda = 1$ in the two cases.

Darboux's main theorem [19, pp. 28–9] may be summarized as follows.

Theorem 10.1 (Darboux) *If the system of differential equations $Adx = \lambda adt$ associated to the Pfaffian equation $\omega = 0$ has solutions only when $\lambda = 0$, then a variable change $x \to y, z$ is possible so that*

$$(II) \quad \omega = dy - [z_1 dy_1 + \cdots + z_m dy_m],$$

and hence the number of distinct equations to which $Adx = \lambda adt$ reduces when $\lambda = 0$ is $2m$ [but this number is $2m + 1$ for $\lambda = 1$]. If $Adx = \lambda adt$ has solutions for $\lambda \neq 0$, then ω may be put in the form

$$(I) \quad \omega = z_1 dy_1 + \cdots z_m dy_m,$$

and the number of equations to which $Adx = \lambda adt$ reduces is $2m$ [regardless of whether $\lambda = 0$ or $\lambda = 1$].

The parts of the theorem in square brackets were not stated explicitly by Darboux, although they are immediate consequences of his (10.1)–(10.2).

By way of an example, Darboux applied his results to the "most general case" by which he meant what I have called the generic case [19, pp. 29–30] – the sole case for which Darboux's Theorem was proved rigorously. That is, he supposed first that: (a) n is even and that det $A \neq 0$ (the case dealt with by Pfaff and Jacobi). Then "one can solve the equations ... $[Adx = \lambda adt]$ ... for the differentials dx_i." That is, $dx/dt = \lambda A^{-1}a$ and so by the theory of ordinary differential equations a solution with $\lambda \neq 0$ exists and case (I) obtains. Hence by Darboux's Theorem, $2m = \text{rank} \begin{pmatrix} A & \lambda a \end{pmatrix} = \text{rank } A = n$, and so $m = n/2$. Next he supposed that: (b) n is odd so that necessarily the skew symmetry of A forces det $A = 0$. "[B]ut its minors of the first order are not zero in general. As we have seen, it is thus necessary, save for an exceptional case, that $\lambda = 0$ and thus that the equations reduce to $n - 1$ distinct ones ... ." This statement is full of the ambiguity that attends generic reasoning (which, as noted above, Darboux tended to favor). The "exceptional case" could be interpreted as the case in which rank $A < n-1$, but I believe Darboux must have meant that, assuming rank $A = n-1$, then "in general" a will not lie in the $(n-1)$-dimensional column space of A, i.e., "in general" rank $(A \quad a) >$ rank A. The exceptional case ignored by Darboux would then be when rank $(A \quad a) =$ rank $A = n - 1$. (This exception occurs, e.g., for $\omega = 2(x_1 dx_1 - x_1 dx_2 + x_1 dx_3)$.) If, following Darboux, we ignore this exceptional case, then the case $\lambda \neq 0$ could not hold. For if it did, then by Darboux's Theorem we would have $2m = \text{rank } (A \quad \lambda a) = \text{rank } (A \quad a) = n$, which is impossible since n is odd. Thus we are in the $\lambda = 0$ case, as Darboux concluded.

The above application of Darboux's Theorem did not make use of the parts in square brackets, the parts not stated explicitly by Darboux. With these parts included Darboux's Theorem gives a simple, essentially algebraic criterion to distinguish cases (I) and (II): *Case (I) holds if and only if the system $Adx = \lambda adt$ has the same number of distinct equations for $\lambda = 0$ and $\lambda \neq 0$, i.e., if and only if* rank $(A \quad a) =$ rank A. From this criterion, the exception to case (b), namely when rank $(A \quad a) =$ rank A, is seen to fall under case (I), i.e., the $\lambda \neq 0$ case.

It should also be noted that the above criterion is equivalent to Frobenius's Analytical Classification Theorem 8.1. That is, as already explained above in a footnote to the discussion of Lie's work, it follows readily from Frobenius's results that his class number p is given by $p = \text{rank}\begin{pmatrix} A \\ -a^t \end{pmatrix} = \text{rank } (A \quad a)$. His Analytical Classification Theorem thus says that case (I) occurs if and only if $p = \text{rank } (A \quad a)$ is even. The above criterion states that case (I) occurs if and only if rank $(A \quad a) =$ rank A and Darboux's Theorem then implies that this common number is $2m$ and so is even.

Even if we grant Darboux the parts of Theorem 10.1 left unstated by him – and hence the above criterion – it must be kept in mind that Darboux's proof of his theorem in the nongeneric cases was fatally flawed. Compared to Frobenius's proof of the Analytical Classification Theorem, Darboux had achieved great brevity but he had not achieved it with a comparable degree of rigor. Darboux's predilection for thinking in terms of differential equations and their solutions may have caused him to gloss over the above essentially algebraic criterion. That same predilection prompted him to discuss the "efficiency issues" involved in integrating a Pfaffian system [19, §VII], something Frobenius, with his own predilection for algebra, ignored.

What seems to have impressed Cartan about Darboux's paper is the idea that if one can rigorously deduce the canonical forms (I) and (II) of Clebsch's Theorem 5.1, then the invariance of the binary covariant under coordinate transformations allows one to use the algebraic simplicity of these forms to avoid the extensive calculations of Frobenius, who did not prove Clebsch's Theorem as a first step but rather incorporated it into his Analytical Classification Theorem, the proof of which is extremely lengthy and laden with calculations.

Although analysts appear to have been generally impressed by the masterful command of algebra manifested in Frobenius's lengthy paper, some found his entire approach too algebraic and regretted his focus upon the equivalence theory of Pfaffian forms to the neglect of the issue of efficient methods of actually integrating Pfaffian equations. Mayer's review of Frobenius's paper in the 1880 volume of *Fortschritte der Mathematik* [63] was along these lines. Although admitting that Frobenius's paper was "ample (*reich*) in both scope and content," he proceeded to contrast, somewhat unfavorably, Frobenius's approach with the "entirely different one" of Lie. Lie "at the very outset with the help of a few simple theorems on partial differential equations" quickly established Clebsch's Theorem 5.1 and the necessary and sufficient conditions for equivalence of two Pfaffians and showed how one can algebraically determine the number p of variables in the normal form of ω. These matters thus quickly dispatched, he then "focused his attention primarily on the reduction to normal form by the smallest possible number of integrations . . . ." But Frobenius "conceives of the problem in purely algebraic terms and, as a consequence of this, while the algebraic side of the problem is more deeply grounded, leaves the question of how to best integrate each successively occurring complete system entirely untouched" [63, p. 250]

Mayer's attitude seems to have been fairly typical of most analysts of the period who were primarily interested in the integration of differential equations. Thus A. R. Forsyth in the historical remarks to the 1890 volume of his *Theory of Differential Equations*, which was devoted to Pfaffian equations, wrote that "Lie's results constitute a distinct addition to the theory. . . . . About the time of publication of the memoir by Lie just referred to, Frobenius had . . . completed his memoir dealing with Pfaff's problem. He discusses the theory of the normal form rather than the integration of the equation; and the analysis is more algebraic than differential" [22, p. 87]. Not surprisingly, only a very slim chapter – Chapter X – was devoted to "Frobenius' Method," and it began with a justification for its brevity: "The investigations of Frobenius . . . deal rather with the general theory of the reduction of the [Pfaffian] expression to a normal form than with any process for the integration of equations which occur in the reduction . . . [22, p. 272]

Forsyth's book was more a compendium of diverse methods for treating Pfaff's problem than a synthesis of those methods into a coherent theory. The first attempt at such a synthesis was made in 1900 by Eduard von Weber [71], an instructor (*Privatdozent*) at the University of Munich. In von Weber's book Frobenius's work was given a more fundamental role to play, but in reality von Weber's whole approach was destined for obsolescence because in 1899 Élie Cartan had begun to develop an entirely new approach to the theory of Pfaffian equations.

11. Cartan's calculus of differential forms

Cartan had obtained his doctorate in 1894 with a brilliant thesis that provided a rigorous foundation for Killing's ground-breaking theory of finite-dimensional complex semisimple Lie algebras and its most impressive consequence – a complete classification of simple Lie algebras.[48] During the following four years most of his attention was focused on developing applications of the ideas and results of his thesis, but he was also on the lookout for new areas of research. One such area he seems to have been considering at the time was the application of Lie's theory of groups to Poincaré's theory of invariant integrals. This comes out in a little paper of 1896 on Darboux-style differential geometry [7], which is especially relevant to the present account because in it we see that Cartan already realized that the variable change formulas in multiple integrals could be derived by submitting the differentials involved to certain rules of calculation. In a footnote he made the following observation.[49] Consider an oriented surface integral in n-dimensional space such as $\iint_\Sigma dx_i dx_j$ where Σ denotes a 2-manifold. Then if new variables $x_i = f_i(y_1, \ldots, y_n)$ are introduced the multiple integral transformation formula may be derived by formally multiplying out $dx_i dx_j = [\sum_{k=1}^n (\partial x_i/\partial y_k) dy_k][\sum_{l=1}^n (\partial x_j/\partial y_l) dy_l]$ using the rules

$$dy_k dy_k = 0 \quad \text{and} \quad dy_l dy_k = -dy_k dy_l \ (k \neq l) \tag{11.1}$$

to obtain for $dx_i dx_j$ the expression

$$\sum_{k,l=1}^{n} \frac{\partial x_i}{\partial y_k}\frac{\partial x_j}{\partial y_l} dy_k dy_l = \sum_{k<l} \left(\frac{\partial x_i}{\partial y_k}\frac{\partial x_j}{\partial y_l} - \frac{\partial x_i}{\partial y_l}\frac{\partial x_j}{\partial y_k}\right) dy_k dy_l = \sum_{k<l} \frac{\partial(x_i, x_j)}{\partial(y_k, y_l)} dy_k dy_l.$$

Since this formalism is to be applied to an integral over Σ, which is represented parametrically by $y_k = \varphi_k(s, t)$ for all k, Cartan added another rule:

$$dy_k dy_l = \frac{\partial(y_k, y_l)}{\partial(s, t)} ds dt, \tag{11.2}$$

so that

$$\iint_\Sigma dx_i dx_j = \iint_D \sum_{k<l} \frac{\partial(x_i, x_j)}{\partial(y_k, y_l)} \frac{\partial(y_k, y_l)}{\partial(s, t)} ds dt,$$

where D is the parameter domain in the st-plane.

Thus already in 1896 we see that Cartan was in possession of the fundamental rules upon which the now-familiar exterior algebra of differential forms is built. But it was not until 1899 that he found in the problem of Pfaff a sufficiently good reason to systematically develop a differential calculus based upon those rules when they are combined with the key new notion of exterior differentiation. As he explained in the introductory remarks to an 1899 paper on the problem [8, pp. 241–2],

> The present work constitutes an exposition of the problem of Pfaff based upon the consideration of certain differential expressions . . . [which] . . . are subject to the usual rules

[48] Cartan's thesis work is the subject of Chapter 6 of [46].

[49] See [7, p. 143n]. For expository reasons, I have changed Cartan's notation and expanded his brief remarks. On earlier work, e.g., by Poincaré, on the transformation of multiple integrals and the relation of this work to the generalized Stokes theorem of the exterior calculus of differential forms, see [56, 57].

of calculation as long as the order of differentials in a product is not permuted. In sum the calculus of these quantities is that of the differential expressions that are placed under the multiple integral sign. This calculus has many analogies with that of Grassmann; in addition it is identical to the geometrical calculus used by Mr. Burali–Forti in a recent book.

The similarity of his calculus with the "Grassmann algebra" underlying Grassmann's theory of extension (*Ausdehnungslehre*) was thus recognized by him by 1899, although the above quotation together with Cartan's observations of 1896 make it clear that Cartan's calculus was not suggested by Grassmann's work but by the consideration of multiple integrals. In fact, like most 19th century mathematicians, Cartan may not have attempted to penetrate the unusual notation of the *Ausdehnungslehre* (1862) so as to digest it contents. His knowledge of the latter may have been indirectly obtained through the above-mentioned Burali–Forti book, *Introduction à la géométrie différentielle suivant la méthode de H. Grassmann*, which appeared in 1897 [6].

It is possible, however, that Grassmann's work helped encourage Cartan to apply his calculus of differential forms to the problem of Pfaff because through his familiarity with Forsyth's 1890 volume on the problem he would have known that Grassmann had considered the problem in the 1862 version of his *Ausdehnungslehre* [33, §§500–527].[50] Noting the "remarkable formal conciseness" that Grassmann had achieved with his algebraic apparatus [22, p. 84], Forsyth included a brief chapter on Grassmann's work despite his suspicion that it would be unintelligible to most readers because "adequate knowledge of the analytical method of the Ausdehnungslehre" was "probably not common at present" [22, p. 120n]. The possibility that his calculus of differential forms might also lend itself to a concise treatment of the problem of Pfaff, may thus have been suggested by Grassmann's work. But Cartan's overall approach drew more definite inspiration from Frobenius, who, Cartan wrote, "in his beautiful memoir in *Crelle's Journal* employs an entirely new method. It is based on consideration of what he calls the *bilinear covariant* of a Pfaffian expression" [8, p. 241].

It was the bilinear covariant introduced by Frobenius and then emphasized from a slightly different perspective by Darboux that prompted Cartan to introduce something not found in Grassmann or Burali–Forti: the idea of what is now called the exterior derivative of a differential form. Indeed, if the rules (11.1) Cartan gave in 1896 are applied provisionally in the spirit of the above multiple integral calculations to formally calculate a derivative $d\omega$ of the Pfaffian $\omega = \sum_{i=1}^n a_i dx_i$, the result is

$$\begin{aligned} d\omega &= \textstyle\sum_{i=1}^n d(a_i dx_i) = \sum_{i=1}^n d(a_i)dx_i + \sum_{i=1}^n a_i d(dx_i) \\ &= \textstyle\sum_{i=1}^n \left\{\sum_{j=1}^n \frac{\partial a_i}{\partial x_j} dx_j\right\} dx_i + \sum_{i=1}^n a_i d(dx_i) \\ &= -\textstyle\sum_{i<j}^n \left(\frac{\partial a_i}{\partial x_j} - \frac{\partial a_j}{\partial x_i}\right) dx_i dx_j + \sum_{i=1}^n a_i d(dx_i). \end{aligned}$$

This suggests defining $d(dx_i) = 0$ so that

$$d\omega \equiv \sum_{i=1}^{n} d(a_i)dx_i = -\sum_{i<j}^{n} a_{ij} dx_i dx_j. \qquad (11.3)$$

[50] Forsyth's book [22] is cited by Cartan [8, p. 239n1] as a source of information on the problem of Pfaff.

For Cartan an expression such as that in (11.3) was regarded as symbolical. To evaluate it, the rule (11.2) is applied to the $dx_i dx_j$ to obtain

$$dx_i dx_j = \frac{\partial(x_i, x_j)}{\partial(s,t)} ds dt = \left(\frac{\partial x_i}{\partial s} \frac{\partial x_j}{\partial t} - \frac{\partial x_i}{\partial t} \frac{\partial x_j}{\partial s} \right) ds dt = u_i v_j - u_j v_i,$$

where $u_k = (\partial x_k / \partial s) ds$ and $v_k = (\partial x_k / \partial t) dt$.[51] In effect $dx_i dx_j$ was for Cartan an alternating bilinear function of u, v (as is $dx_i \wedge dx_j$ for us today); and the identification of $d\omega$ as given in (11.3) with $-\Omega(u, v)$ would thus have been evident to Cartan – as well as the all-important fact that differentiation is invariant under coordinate changes so that if $\omega = \omega'$ under $x = \varphi(x')$, then also $d\omega = d\omega'$ [8, pp. 251–2].

In view of the above considerations it is not surprising to find Cartan defining the derivative of a Pfaffian expression (1-form) as – in his notation – $\omega' = \sum_{i=1}^{n} d(a_i) dx_i$.[52] Indeed, immediately after presenting this definition he remarked that "consideration of ω' or, what amounts to the same thing, of the bilinear covariant of ω, forms the basis of the beautiful investigations on the problem of Pfaff by Frobenius and Darboux" [8, p. 252n].

By viewing Lipschitz's bilinear covariant as an alternating bilinear form Ω, Frobenius certainly made the connection between ω and the idea of its derivative $d\omega$ more readily apparent than did Darboux with his focus upon the system $Adx = \lambda a dt$ (Sect. 10). But it was Darboux who pointed out how the invariance of this system under coordinate changes to canonical forms can be used to avoid many of the complicated calculations that had attended the approach of Frobenius. Cartan developed this idea of Darboux's by defining the class number p of a Pfaffian ω as the smallest integer p for which a coordinate change $x = \varphi(y)$ exists such that $\omega = \sum_{i=1}^{p} B_i(y) dy_i$, where the B_i depend only on the first p variables $y_1, \ldots, y_p$. Cartan's rendition of the problem of Pfaff then consisted in proving that ω can be transformed into the form (I) or (II) of Frobenius's Analytical Classification Theorem 8.1 according to whether p is even or odd, respectively. This not only established the equivalence of his definition of p with that of Frobenius, it also established the end result of Frobenius's memoir, his analytical classification theorem; and the route there, which is outlined below, was far more calculation-free due to Cartan's use of invariance in the spirit Darboux's paper.

For 1-forms ω Cartan defined what he called higher derivatives as follows.[53] With $\omega' = d\omega$, let $\omega'' = \omega \wedge \omega'$, $\omega''' = (1/2)(\omega')^2$, and in general $\omega^{(2m-1)} = (1/m!)(\omega^{(m)})^2$ while $\omega^{(2m)} = \omega \wedge \omega^{(2m-1)}$. If followed from the invariance properties of his calculus that if $\omega \to \varpi$ under $x = \varphi(y)$ then $\omega^{(q)} \to \varpi^{(q)}$, and Cartan proved that ω is of class p if and only if p is the smallest integer such that $\omega^{(p)} = 0$, thereby demonstrating the succinctness possible with his notational apparatus [8, p. 255]. As with the method of Frobenius, Cartan's route to the analytical classification theorem also required the

[51] The evaluation of differential expressions, a natural consequence of their origin under multiple integral signs, became essential to Cartan's method of establishing relations among differential expressions.

[52] In his 1945 monograph [12] Cartan followed the lead of Kähler [55] and replaced the notation ω' with $d\omega$.

[53] The wedge product notation was not used by Cartan.

solutions to successive systems of partial differential equations $A_k(f) = 0$ [8, p. 258]. We saw that the completeness of the systems $A_k(f) = 0$ was not directly evident in Frobenius's method and required extensive mathematics, including his integrability theorem, to establish. Not so with Cartan. For example, the first such system arises from the equation $\omega^{(p-2)} \wedge df = 0$. Since $\omega^{(p-2)}$ is a $(p-1)$-form, $\omega^{(p-2)} \wedge df$ is a p-form, and if its coefficients are all set equal to zero the system $A_k(f) = 0$ results. To see that this system is complete, it is only necessary to make the variable change $x = \varphi(y)$ that puts ω in the form $\varpi = \sum_{i=1}^{p} B_i(y)dy_p$ in accordance with the definition of p. Then $\omega^{(p-2)} \wedge df = 0$ takes the form $\varpi^{(p-2)} \wedge dg = 0$, where $g(y) = f(\varphi(x))$, and by virtue of the simple nature of ϖ it is easily seen that the system corresponding to $A_k(f) = 0$ in the new variables satisfies Jacobi's integrability condition and so is complete by the Jacobi–Clebsch Theorem 5.3.

Although Cartan had no need of Frobenius's Integrability Theorem 9.2 in his proof of the analytical classification theorem, he incorporated Theorem 9.2 and, more generally, its characterization of complete integrability into his papers of 1901 [9, 10] on incomplete systems of Pfaffian equations

$$\omega^{(i)} = \sum_{j=1}^{n} a_j^{(i)}(x)dx_i = 0, \quad i = 1, \ldots, r. \tag{11.4}$$

Central to his study and classification of such systems was the assumption of a chain of integral manifolds through a general point x: $x \in \mathcal{M}_1 \subset \mathcal{M}_2 \subset \cdots \subset \mathcal{M}_g$, where $\dim \mathcal{M}_d = d$ and the integer g, called the *genre* of system (11.4), is such that there is no integral manifold $\mathcal{M}_{g+1}$ through x and containing $\mathcal{M}_g$ [9, pp. 254–63]. To obtain conditions guaranteeing the existence of such a chain, Cartan turned to the bilinear covariants of the $\omega^{(i)}$ [9, p. 249–50]. As Lipschitz [61, p. 77], Frobenius [26, p. 254], and Darboux [19, p. 17] had all pointed out, if d and δ denote differentials in two different directions in the tangent plane, then $\delta(\sum_j a_j^{(i)}dx_j) - d(\sum_j a_j^{(i)}\delta_j) = \sum_{jk} a_{jk}^{(i)}dx_j\delta_k$. Cartan concluded that if b and c are two vectors in the tangent plane determining these two directions, then $\omega^{(i)}(b) = 0$ and $\omega^{(i)}(c) = 0$, and so the above relation implies that $\Omega^{(i)}(b, c) = 0$, where as usual $\Omega^{(i)}$ denotes the bilinear covariant of $\omega^{(i)}$. Thus corresponding to point x is a chain of vector spaces (the tangent spaces at x to the $\mathcal{M}_d$) $\mathcal{V}_1 \subset \mathcal{V}_2 \subset \cdots \subset \mathcal{V}_g$ such that each $\mathcal{V}_d$ is integral in the sense that $\omega^{(i)}(b) = 0$ for all $b \in \mathcal{V}_d$ and in addition any two vectors $b, c \in \mathcal{V}_d$ are in involution[54] in the sense that $\Omega^{(i)}(b, c) = 0$ for all i. Also there is no $\mathcal{V}_{g+1} \supset \mathcal{V}_g$ consisting entirely of integral vectors that are pairwise in involution. Cartan then argued that the existence of such a chain of vector spaces at each x is not only necessary but sufficient for the existence of the above chain of integral manifolds.[55] In this manner the integrability condition in Frobenius's Theorem 9.2 was transformed by Cartan into the integrability condition for his own theorem.

Cartan found a use for Frobenius's integrability theorem per se in his study of incomplete systems (11.4) with "characteristic elements" passing through any point x, thereby affording a simplified integration process. The characteristic elements are determined

54 Cartan spoke of b and c as associated [9, p. 250].
55 Actually, certain regularity conditions need to be imposed [1, p. 132].

by a system of Pfaffian equations, which I will call the characteristic equations. The fundamental theorem of Cartan's theory of characteristics is that the characteristic equations form a complete system. Cartan's initial proof of this theorem [9, p. 304] was brief and did not use Frobenius's Integrability Theorem; but shortly thereafter he presented an exposition of his results within the framework of his calculus of differential forms [10, Ch. 1–2], and then he used Frobenius's Integrability Theorem to establish the completeness of the characteristic equations. To do this, however, he first translated the integrability condition of Frobenius's theorem into a form more congenial to his calculus of differential forms.

He did this as follows [10, p. 496]. If $\omega_i = 0$, $i = 1, \dots, r$, is a given system of linearly independent Pfaffians, $n - r$ additional Pfaffians ϖ_j, $j = 1, \dots, n - r$, may be added so as to obtain n linearly independent Pfaffians ω_i, ϖ_j. From the calculus of these forms it followed that any 2-form could be expressed as a linear combination of two-fold products of these n 1-forms, namely $\varpi_j \wedge \varpi_k$, $\varpi_j \wedge \omega_i$, and $\omega_i \wedge \omega_j$. This applies in particular to the 2-forms $d\omega_i$, and so

$$d\omega_i = \sum_{j<k} b_{ijk} \varpi_j \wedge \varpi_k + \sum_{j=1}^{r} \theta_{ij} \wedge \omega_j,$$

where the θ_{ij} are the 1-forms obtained by factoring out ω_j from the terms of $d\omega_i$ involving $\varpi_k \wedge \omega_j$ or $\omega_i \wedge \omega_j$. According to Frobenius's integrability condition $d\omega_i(b, c) = 0$ whenever $\omega_i(b) = 0$ and $\omega_i(c) = 0$ for all i. Since the ϖ_j are independent of the ω_i, the only way this can happen is if all the coefficients b_{ijk} are identically zero. The integrability condition in Cartan's rendition of Frobenius's theorem thus becomes the condition that

$$d\omega_i = \sum_{j=1}^{r} \theta_{ij} \wedge \omega_j, \quad i = 1, \dots, r, \tag{11.5}$$

or, as he expressed it, $d\omega_i \equiv 0 \pmod{\omega_1, \dots, \omega_r}$.

Although the problem of Pfaff had supplied the initial impetus for Cartan to develop his calculus of differential forms, his contribution to that problem pales in magnitude and significance with the subsequent applications he made of his calculus. Even his paper of 1899 contains considerably more than his new derivation of Frobenius's Analytical Classification Theorem – the theorem that had formed the goal of Frobenius's own paper. From the outset Cartan seems to have been interested in Pfaffian equations as a new means of dealing with partial differential equations of any order, and his 1899 paper contains much along these lines. Likewise Cartan's papers of 1901 on incomplete systems of Pfaffian equations (discussed briefly above) provided a new, geometrically informed approach to systems of partial differential equations, which in 1934 was extended by Kähler [55] to include systems involving k-forms with $k > 1$.[56] Cartan also developed his theory of Pfaffian systems into the basis for his theory of the structure of infinite-dimensional Lie transformation groups, which concluded in 1909 with a classification of the simple types.[57]

---

[56] In 1945 Cartan gave his own exposition of his and Kähler's results in his monograph on the exterior calculus of differential forms and its geometrical applications [12].

[57] See Sects. 8.1–8.2 of my book [46].

In all of this work by Cartan the consideration of bilinear covariants – derivatives of 1-forms – played a key role, and the idea that they should play such a role was something that Cartan evidently took away from the papers of Frobenius and Darboux. The following passage from Cartan's first paper of 1901 on incomplete systems is typical of the many such remarks in his papers which signal his debt to them. Before Cartan incomplete systems had been considered by Otto Biermann in 1885 [4] but only in the generic case. Since then, Cartan wrote [9, p. 241],

> nothing has been done except to demonstrate the same results in another form but without ever achieving perfect rigor, and almost nothing has been done on the case in which the coefficients of the system are not generic [*quelconque*].
> Precise and general results can be achieved by taking into consideration the bilinear covariants of the right-hand sides of the equations of the system, whose introduction by Frobenius and Darboux has proved to be so fertile in the theory of a single Pfaffian equation.

The second paragraph of this passage indicates that Cartan had come to see the bilinear covariant $d\omega$ as a key mathematical tool. It also suggests that in seeking to assess the influence that Frobenius's paper had on Cartan, it is impossible to fully extricate the influence of Frobenius from that of Darboux because invariably both men are mentioned together when a reference to the introduction of $d\omega$ is made. Granted that caveat, the contents of the two papers in question are sufficiently different that it seems reasonable to assume they impressed Cartan in different ways. From Darboux's paper Cartan got the idea of avoiding complicated calculations by invoking the invariance of his forms and their derivatives under variable change to canonical forms – as in his proof (sketched above) of Frobenius's analytical classification theorem. That constitutes the principal contribution made exclusively by Darboux. And what would have impressed Cartan when reading Frobenius's paper? It is there that the bilinear covariant as alternating bilinear form Ω actually occurs as a central concept (as opposed to Darboux's systems $Adx = \lambda adt$), thereby making the idea of introducing the derivative of a Pfaffian more palpable. Furthermore, through his Integrability Theorem and the applications he made of it, Frobenius did far more than Darboux to suggest the idea that the bilinear covariant is the key to the study of Pfaffian equations. Indeed as we have seen, Frobenius's manner of characterizing integrability is central to the constructs by means of which Cartan obtained "precise and general results ... by taking into consideration the bilinear covariants" in his work on incomplete systems.

Frobenius's treatment of the problem of Pfaff, although encumbered more by calculations, was also more carefully worked out and self-contained than Darboux's. Frobenius was the first and only mathematician prior to Cartan to correctly understand and systematically analyze the completely general case of the problem of Pfaff, i.e., completely general in the algebraic sense as per Sect. 3; and this precedent setting work must have impressed Cartan, who saw himself faced with the same sort of challenge with regard to the theory of incomplete Pfaffian systems, as indicated in the preceding quotation. Whereas Frobenius had systematically applied Berlin-style linear algebra, Cartan drew upon the multilinear algebra behind his calculus of differential forms – as well as his fertile geometrical imagination – to deal with the more formidable problems he faced.

12. Frobenius's mathematics & its impact

Frobenius was the founder of one major mathematical theory of lasting importance, the theory of characters and representations of finite groups, and yet the many theorems, concepts and constructs that bear his name today and are unrelated to group representation theory suggest that his influence upon present-day mathematics goes far beyond his role in creating that theory. What was it about his mathematics that caused this to be the case? No doubt there are several reasons for this phenomenon. One such reason, which I offer here on a tentative basis, is supported by my studies of various of Frobenius's publications and is subject to any modifications that arise as I continue my study of his work in preparation for a book on his role in the history of linear algebra.

Frobenius's paper on the problem of Pfaff is, I suggest, paradigmatic of one of the principal ways in which his work has left its mark on present-day mathematics. That work was initiated, rather typically, by a specific problem, a problem of the sort encouraged by his experiences at Berlin: to deal successfully with the problem of Pfaff on the nongeneric level, something that Clebsch had first attempted but without complete success. It was, again as it typically was for Frobenius, a problem formulated within the context of a body of earlier mathematical results on, in this case, the problem of Pfaff – the work of Pfaff, Jacobi, Clebsch and Natani being the most significant – and Frobenius read the literature as a scholar, carefully and thoroughly. The novel approach to the problem he posed was also very much in harmony with his Berlin schooling: By means of Lipschitz's passing observations on the bilinear covariant of a 1-form (Theorem 6.1), Frobenius was able to view his problem within the context of the simultaneous transformation of a 1-form and its associated binary covariant and so to use the construct his friend and fellow student Stickelberger had introduced in his Berlin doctoral dissertation. Then in accordance with the procedure used by Christoffel and the disciplinary ideal articulated by Kronecker, he sought to reduce the problem to an algebraic one involving the canonical forms of a pair consisting of a linear and a skew symmetric bilinear form. Having thus formulated an appropriate way to go about resolving the problem he had posed, Frobenius then applied his creative talent to develop the new approach systematically so that the end result – his paper on the problem of Pfaff – was essentially a carefully worked out and original "monograph" on the problem.

Like Weierstrass's cycle of lectures on analysis, Frobenius's publications convey the conviction that the clear and systematic presentation of results in the proper manner, i.e., developed from the proper unifying mathematical viewpoint, was just as important as the discovery of new results by whatever means.[58] This penchant on Frobenius's part is clearly set forth in a paper of 1875 (thus predating his Pfaff paper) on the application of the theory of determinants – the foundation of Berlin-style linear algebra – to metric geometry [25]. Frobenius's own account of why he published this paper is worth quoting in full:

---

[58] As Weierstrass wrote to Sonja Kowalewsky, "was ich aber von einer wissenschaftlichen Arbeit verlange, ist Einheit der Methode, consequente Verfolgung eines bestimmten Plans, gehörige Durcharbietung des Details und, dass ihr der Stempel selbständiger Forschung aufgeprägt sei." Letter dated 1 January 1875 and published in [5]. Also quoted by Mittag-Leffler [64, p. 155].

> In 1868 I was induced by a prize question set by the philosophical faculty of the University of Berlin to concern myself with the application of the theory of determinants to metric geometry, and at that time I wrote some works on this subject, which have hitherto been held back from publication by other work. In the meantime the work of Mr. Darboux ... [18] ... came to my attention, just as I was putting together a brief sketch of my investigations. In it I found a large part of the metric relations and geometrical constructions treated by me developed in a very elegant and original manner. Whereas with Mr. Darboux there is no connection between the metric relations and the geometrical constructions, a main objective of my work is to show how the solution to complicated geometrical problems can be read off simply and easily from a few metrical relations.
>
> Partly on these grounds and partly on account of the difficulty of presenting those of my results that are not found in that work [by Darboux] separated from the rest, I wish to communicate here in abbreviated form [*im Auszuge*] my developments.

Thus despite the elegant approach of Darboux and the fact that his memoir contained many of the same results, Frobenius felt compelled and justified in presenting those same results by means of his own approach, which he clearly deemed superior for the reason given, even though it required sixty-two pages to do this.

Frobenius had a genuinely scholarly approach to his mathematics in the sense that once interested in a particular problem he made a thorough search of the literature, which he then creatively viewed and developed from the unifying approach that he deemed the proper one for the subject at hand. In the case of his paper on metrical geometry it was the idea of a small set of geometrical relations from which, with the aid of the theory of determinants, a multitude of complicated geometrical constructions could be immediately obtained. In the case of the problem of Pfaff the approach was that of the transformational equivalence of Pfaffian expressions (or 1-forms) and the key unifying concept was the bilinear covariant. In developing his own approach to the material he of course drew as needed upon his broad-based knowledge of the mathematical literature. Thus he borrowed the key notion of the bilinear covariant from Lipschitz, which thanks to Jacobi's introduction of his skew-symmetric matrix into the problem of Pfaff, Frobenius could see as central to the problem of Pfaff. And of course from Clebsch's work he obtained not only the mathematical problem that motivated his work, but also the two canonical forms I and II of his equivalence theory. Likewise, as noted in Sect. 9, his complete integrability theorem drew upon the results of Jacobi, Clebsch, and Deahna, as well as the general realization in the literature of a duality between systems of linear homogeneous differential equations and systems of Pfaffian equations.

Frobenius's genius was to combine these elements in a clear, systematic manner unified by a central concept, that of the bilinear covariant. He had a considerable talent for clear mathematical exposition – more so than his mentor Weierstrass. His contemporaries, including Cartan, regarded his work as exceptionally thorough and rigorous. The resulting "monograph" that Frobenius produced did not appeal to everyone, as we saw in Sect. 10. Yet top mathematicians such as Darboux and Cartan were impressed by Frobenius's "beautiful" essay on the problem of Pfaff, and they paid him the compliment by seeking to improve and build upon it.

Thus Darboux pointed out the value of quickly establishing and then utilizing the covariance of the bilinear covariant as a means of avoiding some of Frobenius's extensive algebraic calculations; and Cartan, besides taking full advantage of Darboux's suggestion, saw that the central role played in Frobenius's work by the bilinear covariant could be incorporated into his still fragmentary, incomplete and seemingly insignificant calculus of differential forms, thereby providing it with the central notion of the exterior derivative of a 1-form so as to obtain the associated bilinear covariant and make his calculus applicable to the problem of Pfaff. Frobenius had not only formulated the results of Jacobi, Clebsch and Deahna in terms of the bilinear covariant as the Integrability Theorem 9.2 but he also made several applications of it, thereby suggesting its further utility; and Cartan, after reformulating it within the context of his now-complete calculus of differential forms, went on to demonstrate its utility within the framework of the problems of interest to him in the theory of partial differential equations.

Of course, as indicated in Sect. 11, Cartan went on to apply his calculus of differential forms to more than just the problem of Pfaff, thereby establishing that calculus as a basic tool in the repertory of present-day mathematics and advancing the theory of partial differential equations, Lie group theory, and algebraic topology in brilliant ways far removed from the work and interests of Frobenius. Yet it seems to me that Frobenius's "monograph" on the problem of Pfaff, which by 19th century standards was remarkably clear and conceptually coherent, was, for the above reasons, as well as by virtue of Frobenius's attempt (like Cartan's) to deal with problems on an algebraically nongeneric level, a major influence upon Cartan as he initiated the theory and application of his calculus of differential forms. That Frobenius's name alone has remained attached to his complete integrability theorem, although not at all historically accurate, has a kind of historical validity in the sense that it is a reflection of the way Frobenius's work influenced Cartan and, through him, present-day mathematics. Cartan's work rendered Frobenius's "monograph" obsolete, but some of its central ideas were subsumed in the work of Cartan and thereby eventually into contemporary mathematics. The association of Frobenius's name with the Integrability Theorem is a telltale of this phenomenon. Many of Frobenius's other monographic essays on various mathematical problems, I believe, also influenced the development of mathematics in similar ways, as I hope to show in the above-mentioned prospective book.

References

1. M. Akivis and B. Rosenfeld. *Élie Cartan (1869–1951)*, volume 123 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1993.
2. Anonymous. Zum Andenken an Rudolf Friedrich Alfred Clebsch. *Math. Ann.*, 6:197–202, 1873.
3. R. Baltzer. *Theorie und Anwendungen der Determinanten*. S. Hirzel, Leipzig, 3rd edition, 1870.
4. O. Biermann. Über n simultane Differentialgleichungen der Form $\sum_{\mu=1}^{n+m} X_\mu dx_\mu$. *Zeitschrift für Mathematik und Physik*, 30:234–244, 1885.
5. R. Bölling, editor. *Briefwechsel zwischen Karl Weierstrass und Sofja Kowalewskaja*. Akademie Verlag, Berlin, 1993.
6. C. Burali–Forti. *Introduction à la géométrie différentielle suivant la méthode de H. Grassmann*. Gauthier–Villars, 1897.
7. E. Cartan. Le principe de dualité et certaines intégrales multiples de l'espace tangentiel et de l'espace réglé. *Bull. Soc. Math. France*, 24:140–177, 1896. Reprinted in *Oeuvres* II.1, 265–302.

8. E. Cartan. Sur certaines expressions différentielles et le problème de Pfaff. *Annales scientifiques École Normale Sup. Paris*, 16:239–332, 1899. Reprinted in *Oeuvres* II.1, 303–396.
9. E. Cartan. Sur l'intégration des systèmes d'équations aux différentielles totales. *Annales scientifiques École Normale Sup. Paris*, 18:241–311, 1901. Reprinted in *Oeuvres* II.1, 411–481.
10. E. Cartan. Sur l'intégration de certains systèmes de Pfaff de caractère deux. *Bulletin des sciences mathématiques*, 29:233–302, 1901. Reprinted in *Oeuvres* II.1, 483–553.
11. E. Cartan. *Leçons sur les invariants intégraux*. A. Hermann & fils, Paris, 1922.
12. E. Cartan. *Les systèmes différentiels extérieurs et leur applications géométriques*. Actualités scientifiques et industrielles, no. 994. Hermann, Paris, 1945.
13. E. B. Christoffel. Transformation der homogenen Differentialausdrücke zweiten Grades. *Jl. für die reine u. angew. Math.*, 70:46–70, 1869. Reprinted in *Abhandlungen* 1, 352–377
14. A. Clebsch. Über das Pfaffsche Problem. *Jl. für die reine u. angew. Math.*, 60:193–251, 1861.
15. A. Clebsch. Über das Pfaffsche Problem. Zweite Abhandlung. *Jl. für die reine u. angew. Math.*, 61:146–179, 1863.
16. A. Clebsch. Über die simultane Integration linearer partieller Differentialgleichungen. *Jl. für die reine u. angew. Math.*, 65:257–268, 1866.
17. C. W. Curtis. *Pioneers of Representation Theory: Frobenius, Burnside, Schur and Brauer*. American Mathematical Society, 1999.
18. G. Darboux. Sur les relations entre groupes de points, de cercles et de sphères dans le plan et dans l'espace. *Annales scientifiques École Normale Sup. Paris*, (2) 1:323ff., 1872.
19. G. Darboux. Sur le probème de Pfaff. *Bulletin des sciences mathématiques*, (2)6:14–68, 1882.
20. H. Deahna. Über die Bedingungen der Integrabilität linearer Differentialgleichungen erster Ordnung zwischen einer beliebigen Anzahl veränderlicher Grössen. *Jl. für die reine u. angew. Math.*, 20:340–349, 1840.
21. F. Engel. Anmerkungen. In *Sophus Lie Gesammelte Abhundlungen*, volume 3, pages 585–789. Teubner, Leipzig, 1922.
22. A. R. Forsyth. *Theory of Differential Equations. Part I. Exact Equations and Pfaff's Problem*. Cambridge University Press, 1890.
23. G. Frei and U. Stammbach. *Hermann Weyl und die Mathematik an der ETH Zürich, 1913–1930*. Birkhäuser, Basel, 1992.
24. G. Frobenius. Über die Entwicklung analytischer Functionen in Reihen, die nach gegebenen Functionen fortschreiten. *Jl. für die reine u. angew. Math.*, 73:1–30, 1871. Reprinted in *Abhandlungen* 1, 35–64. Essentially a German reworking of his Berlin dissertation, which is reprinted in *Abhandlungen* 1, 1–34.
25. G. Frobenius. Anwendungen der Determinantentheorie auf die Geometrie des Maasses. *Jl. für die reine u. angew. Math.*, 79:184–247, 1875. Reprinted in *Abhandlungen* 1, 158–231.
26. G. Frobenius. Über das Pfaffsche Problem. *Jl. für die reine u. angew. Math.*, 82:230–315, 1877. Reprinted in *Abhandlungen* 1, 249–334.
27. G. Frobenius. Über lineare Substitutionen und bilineare Formen. *Jl. für die reine u. angew. Math.*, 84:1–63, 1878. Reprinted in *Abhandlungen* 1, 343-405.
28. G. Frobenius. Theorie der linearen Formen mit ganzen Coefficienten. *Jl. für die reine u. angew. Math.*, 86:146–208, 1879. Reprinted in *Abhandlungen* 1, 482–544.
29. G. Frobenius. Über homogene totale Differentialgleichungen. *Jl. für die reine u. angew. Math.*, 86:1–19, 1879. Reprinted in *Abhandlungen* 1, 435–453.
30. G. Frobenius. Theorie der linearen Formen mit ganzen Coefficienten (Forts.). *Jl. für die reine u. angew. Math.*, 88:96–116, 1880. Reprinted in *Abhandlungen* 1, 591–611.
31. G. Frobenius. Zur Theorie der linearen Gleichungen. *Jl. für die reine u. angew. Math.*, 129:175–180, 1905.

32. C. Gauss. Review of [66]. *Göttingische gelehrte Anzeigen*, pages 1025–1038, 1831.
33. H. Grassmann. *Die lineale Ausdehnungslehre*, volume 1, Pt. 2. Enslin, Berlin, 1862. Reprinted with extensive notes by F. Engel in [34].
34. H. Grassmann. *Hermann Grassmanns gesammelte mathematische und physikalische Werke ... herausgegeben von Friedrich Engel*, volume I_2. Teubner, 1896.
35. J. Gray. *Linear Differential Equations and Group Theory from Riemann to Poincaré*. Birkhäuser, Boston, 2nd. edition, 2000.
36. M. Hamburger. Zur Theorie der Integration eines Systems von n linearen partiellen Differentialgleichungen erster Ordnung mit zwei unabhängigen und n abhängigen Veränderlichen. *Jl. für die reine u. angew. Math.*, 81:243–280, 1876.
37. M. Hamburger. Über das Pfaff'sche Problem. *Archiv der Mathematik und Physik*, 60:185–214, 1877.
38. T. Hawkins. The origins of the theory of group characters. *Arch. Hist. Exact. Sci.*, 7:142–170, 1971.
39. T. Hawkins. New light on Frobenius' creation of the theory of group characters. *Arch. Hist. Exact Sci.*, 12:217–243, 1974.
40. T. Hawkins. Another look at Cayley and the theory of matrices. *Archives internationales d'histoire des sciences*, 26:82–112, 1977.
41. T. Hawkins. Weierstrass and the theory of matrices. *Archive for History of Exact Sciences*, 17:119–163, 1977.
42. T. Hawkins. Non-Euclidean geometry and Weierstrassian mathematics: The background to Killing's work on Lie algebras. *Historia Mathematica*, 7:289–342, 1980.
43. T. Hawkins. The Berlin School. In H. Mertens, et al., editor, *Social History of Nineteenth Century Mathematics*, pages 233–245. Birkhäuser, Basel and Boston, 1981.
44. T. Hawkins. Jacobi and the birth of Lie's theory of groups. *Archive for History of Exact Sciences*, 42:187–278, 1991.
45. T. Hawkins. Weyl and the topology of continuous groups. In I. M. James, editor, *History of Topology*, pages 169–198. North-Holland, 1999.
46. T. Hawkins. *Emergence of the Theory of Lie Groups. An Essay on the History of Mathematics 1869-1926*. Springer, New York, 2000.
47. K. Itô, editor. *Encyclopedic Dictionary of Mathematics*, volume 3. MIT Press, 1987.
48. C. G. J. Jacobi. Über die Integration der partiellen Differentialgleichungen erster Ordnung. *Jl. für die reine u. angew. Math.*, 2:317–329, 1827. Reprinted in *Werke* 4, 1–15.
49. C. G. J. Jacobi. Über die Pfaffsche Methode, eine gewöhnliche lineare Differentialglei-chung zwischen $2n$ Variabeln durch ein System von n Gleichungen zu integriren. *Jl. für die reine u. angew. Math.*, 2:347–357, 1827. Reprinted in *Werke* 4, 19–29.
50. C. G. J. Jacobi. De binis quibuslibet functionibus homogeneis secundi ordinis per substitutiones lineares in alias binas transformandis, quae solis quadratis variabilium constant; .... *Jl. für die reine u. angew. Math.*, 12:1–69, 1834. Reprinted in *Werke* 3, 191–268.
51. C. G. J. Jacobi. Theoria novi multiplicatoris systemati aequationum differentialum vulgarium applicandi. *Jl. für die reine u. angew. Math.*, 27 & 29:199–268 & 213–279; 333–376, 1845. Reprinted in *Werke* 4, 317–509.
52. C. G. J. Jacobi. Über eine elementare Transformation eines in Bezug auf jedes von zwei Variablen-systemen linearen und homogenen Ausdrucks. *Jl. für die reine u. angew. Math.*, 53:265–270, 1857. Reprinted in *Werke* 3, 583–590.
53. C. G. J. Jacobi. Nova methodus, aequationes differentiales partiales primi ordinis inter numerum variabilium quemcumque propositas integrandi. *Jl. für die reine u. angew. Math.*, 60:1–181, 1862. Reprinted in *Werke* 4, 3–189. Translated into German and annoted by G. Kowalewski as *Ostwald's Klassiker der exakten Wissenschaften*, Nr. 156 (Leipzig, 1906).
54. C. Jordan. Sur les polynômes bilinéaires. *Comptes Rendus, Acad. Sci. Paris*, 77:1487–1491, 1873. Reprinted in *Oeuvres* 3, 7–11.

55. E. Kähler. *Einführung in die Theorie der Systeme von Differential-gleichungen*. Teubner, Leipzig, 1934.
56. V. Katz. The history of differential forms from Clairaut to Poincaré. *Historia Mathematica*, 8:161–188, 1981.
57. V. Katz. Differential forms – Cartan to De Rham. *Archive for History of Exact Sciences*, 33:321–336, 1985.
58. L. Kronecker. Über Schaaren von quadratischen und bilinearen Formen. *Monatsberichte der Akademie der Wiss. zu Berlin*, pages 59–76, 1874. Reprinted in *Werke* 1, 349–372.
59. J. L. Lagrange. Sur différentes questions d'analyse relatives à la théorie des integrales particulières. *Nouv. Mém. Acad. Sci. Berlin*, 1781. Reprinted in *Oeuvres* 4, 585–634.
60. S. Lie. Theorie des Pfaffschen Problems I. *Archiv for Mathematik*, 2:338–379, 1877. Reprinted in *Abhandlungen* 3, 320–351.
61. R. Lipschitz. Untersuchungen in Betreff der ganzen homogenen Functionen von n Differentialen. *Jl. für die reine u. angew. Math.*, 70:71–102, 1869.
62. A. Mayer. Über unbeschränkt integrable Systeme von linearen totalen Differentialgleichungen und die simultane Integration linearer partieller Differentialgleichungen. *Math. Ann.*, 5:448–470, 1872.
63. A. Mayer. Review of [26]. *Jahrbuch über die Fortschritte der Mathematik*, 9:249–254, 1880.
64. G. Mittag-Leffler. Weierstrass et Sonja Kowalewsky. *Acta mathematica*, 39:133–198, 1923.
65. L. Natani. Über totale und partielle Differentialgleichungen. *Jl. für die reine u. angew. Math.*, 58:301–328, 1861.
66. J. F. Pfaff. Methodus generalis aequationes differentiarum partialium, nec non aequationes differentiales vulgares, utrasque primi ordinis, inter quotcunque variabiles, complete integrandi. *Abhandlungen d. Akad. der Wiss. zu Berlin 1814–15*, pages 76–136, 1818. All references are to the annoted German translation by G. Kowalewski published as *Ostwald's Klassiker der exakten Wissenschaften, Nr. 129*, W. Engelmann, Leipzig, 1902.
67. B. Riemann. Über die Hypothesen, welche der Geometrie zu Grunde liegen. *Abhandlungen der K. Gesellschaft der Wissenschaften zu Göttingen*, 1868. Reprinted in *Werke*, 272–287.
68. E. Scholz. *Geschichte des Mannigfaltigkeitsbegriffes von Riemann bis Poincaré*. Birkhäuser, Basel and Boston, 1980.
69. L. Stickelberger. *De problemate quodam ad duarum bilinearium vel quadraticarium transformationum pertinente*. G. Schade, Berlin, 1874.
70. R. Tobies and D. Rowe, editor. *Korrespondenz Felix Klein-Adolph Mayer. Auswahl aus den Jahren 1871–1907*. Teubner, Leipzig, 1990.
71. E. von Weber. *Vorlesungen über das Pfaffsche Problem und die Theorie der partiellen Differentialglei-chungen erster Ordnung*. Teubner, Leipzig, 1900.
72. K. Weierstrass. Über ein die homogenen Functionen zweiten Grades betreffendes Theorem. *Monatsberichte der Akademie der Wiss. zu Berlin*, 1858. Reprinted in *Werke* 1, 233–246.
73. K. Weierstrass. Zur Theorie des quadratischen und bilinearen Formen. *Monatsberichte der Akademie der Wiss. zu Berlin*, pages 311–338, 1868. Reprinted with modifications in *Werke* 2, 19–44.

Department of Mathematics and Statistics
Boston University
111 Cummington Street
Boston, MA 02215
USA
twh@math.bu.edu

(Received May 15, 2004)
Published online March 5, 2005 –