Archive for History of Exact Sciences (2023) 77:241–343
https://doi.org/10.1007/s00407-022-00303-9

# Canonical transformations from Jacobi to Whittaker

Craig Fraser[1] · Michiyo Nakane[2]

Received: 21 June 2022 / Accepted: 10 December 2022 / Published online: 31 January 2023


## Abstract

The idea of a canonical transformation emerged in 1837 in the course of Carl Jacobi's researches in analytical dynamics. To understand Jacobi's moment of discovery it is necessary to examine some background, especially the work of Joseph Lagrange and Siméon Poisson on the variation of arbitrary constants as well as some of the dynamical discoveries of William Rowan Hamilton. Significant figures following Jacobi in the middle of the century were Adolphe Desboves and William Donkin, while the delayed posthumous publication in 1866 of Jacobi's full dynamical corpus was a critical event. François Tisserand's doctoral dissertation of 1868 was devoted primarily to lunar and planetary theory but placed Hamilton–Jacobi mathematical methods at the forefront of the investigation. Henri Poincaré's writings on celestial mechanics in the period 1890–1910 succeeded in making canonical transformations a fundamental part of the dynamical theory. Poincaré offered a mathematical vision of the subject that differed from Jacobi's and would become influential in subsequent research. Two prominent researchers around 1900 were Carl Charlier and Edmund Whittaker, and their books included chapters devoted explicitly to transformation theory. In the first three decades of the twentieth century Hamilton–Jacobi theory in general and canonical transformations in particular would be embraced by a range of researchers in astronomy, physics and mathematics.

Communicated by Jesper Lützen.

✉ Craig Fraser
craig.fraser@utoronto.ca

Michiyo Nakane
michiyo.nakane@nifty.com

1 University of Toronto, Toronto, Canada

2 Seijo University, Tokyo, Japan

## 1 Introduction

The idea of a canonical transformation emerged in 1837 almost from nowhere in Carl Jacobi's researches and became the subject of episodic investigation for the next fifty years. Significant figures in the middle of the century were Adolphe Desboves and William Donkin, while the delayed posthumous publication in 1866 of Jacobi's full dynamical corpus was a critical event. François Tisserand's doctoral dissertation of 1868 was devoted primarily to lunar and planetary theory but placed Hamilton–Jacobi mathematical methods at the forefront of the investigation. Henri Poincaré's writings on celestial mechanics in the period 1890–1910 succeeded in making canonical transformations a fundamental part of a dynamical theory. Poincaré offered a mathematical vision of the theory that differed fundamentally from Jacobi's and would become influential in subsequent research. Two prominent researchers around 1900 were Carl Charlier and Edmund Whittaker, and their books included chapters devoted explicitly to transformation theory.

To understand Jacobi's 1837 moment of discovery it is necessary to examine some background, especially the work of the French school of Joseph Lagrange and Siméon Poisson on the variation of arbitrary constants. Although this subject is relegated to the special field of perturbations in modern textbooks, it occupied a much more prominent place in the nineteenth century. It was certainly central to Jacobi's investigations throughout his dynamical researches.

The development of Hamilton–Jacobi theory from its inception up to 1910 was closely tied to celestial mechanics or what in the nineteenth century was sometimes called physical astronomy. The two most important figures in this history—Jacobi and Poincaré—were mathematicians first and foremost who also made fundamental contributions to planetary dynamics and the theory of perturbations. The older development of special techniques and problems in celestial mechanics is generally of somewhat limited interest to a modern reader. By contrast, the mathematical side of the subject endures as a source of interest and remains very much alive intellectually to a reader today. The emphasis of the present study is on the mathematical subject, although a certain amount of detail about astronomy is necessary to appreciate the concrete context within which the theoretical advances took place.

In 1927 Lothar Nordheim and Erwin Fues published an essay in the *Handbuch der Physik* titled "Die Prinzipe der Dynamik. Die Hamilton-Jacobische Theorie der Dynamik."[1] They divided the theory into three parts: Hamilton's equations; the study

[1] The idea of Hamilton-Jacobi theory as a distinct subject area seems to have taken hold in the last few decades of the nineteenth century. Schering (1873) published "Hamilton-Jacobische Theorie für Kräfte" and von Weber (1900) in a substantial article on partial differential equations for the *Encyklopädie* included a section on "Die Hamilton-Jacobi'sche Theorie". In the calculus of variations one also finds "Hamilton-Jacobi theory" as a subject area: see Bolza (1909, 595–601), Bliss (1946, Chapter III) and Courant (1953, II 1–35). (Concerning the term "theory" Demidov (1982, 325–326) writes: "Mathematicians use the word *theory* in two essentially different meanings. In a narrow sense, it denotes a complex structure based on definite ideas and methods and covering a certain range of studies (thus, the theory of Galois, or Lagrange's theory of first-order partial differential equations). In a broad sense, the word theory designates a province of thought (e.g., theory of numbers; of differential equations." In the case of Hamilton-Jacobi theory we are evidently using the first meaning of the word "theory.").

of canonical transformations and invariance; and the integration of Hamilton's equations by means of a canonical transformation using a solution of the Hamilton–Jacobi equation as a generating function. Historically, the third part of the theory developed fairly late, in the 1890s and the early years of the twentieth century. Nevertheless, in his textbook *Classical Mechanics* (1950) the young Harvard physicist Herbert Goldstein (born 1922) would define the subject exclusively in terms of this third part. He did so with no apparent awareness of the historical incongruity of such a conception.[2] In focusing on canonical transformations the present paper identifies only one thread in history, although a critical one in the evolution of Hamilton–Jacobi theory.

Our history will overlook subjects such as the theory of the last multiplier, so cherished by Jacobi but not well known today, and pay only limited attention to techniques such as separability invented to integrate the Hamilton–Jacobi partial differential equation. We have also excluded the stream of research that originated with Sophus Lie and which became influential with the more abstract development of the theory after 1940.

For the primary figures of interest in this study (Jacobi, Desboves, Donkin, Tisserand, Poincaré, Charlier, Whittaker) we have adhered in our exposition to the original notation used by the authors. In the prolegomenon we have departed from this practice, presenting the relevant results of Lagrange and Poisson in terms of Jacobi's notation. This decision was based on two considerations: first, it would be cumbersome to follow the authors' notation, which involved looking at general cases by considering three variables or terms and writing out formulas and equations in laborious and extended detail; and second, the primary interest for the present study of the results of Lagrange and Poisson is as background for understanding Jacobi's development of the theory and the work of his successors.

## 2 Prolegomenon: Lagrange (1809, 1811), Poisson (1809a, b) and the theory of variation of constants

The method of variation of arbitrary constants does not figure prominently in modern expositions of Hamilton–Jacobi theory. Nevertheless, it was of central concern in the nineteenth century and much of the work on transformations unfolded in investigations involving this method. To follow this history some acquaintance is necessary with the work of the original architects of the subject, Joseph Lagrange and Siméon Poisson in the years from 1808 to 1811. Our purpose is not to provide a detailed historical account of this episode, but to identify the key results that are necessary to understand the work of Jacobi and subsequent researchers leading up to the twentieth century.

In papers published in 1783 and 1784 Lagrange had introduced a general mathematical method to investigate constants that appeared in the integration of differential equations for perturbed planetary motion.[3] This method would provide the basis for

[2] Nordheim and Fues' tripartite division of Hamilton–Jacobi theory was also adopted in such standard mid-century textbooks as Lanczos (1949) and Corben and Stehle (1950). What distinguished Goldstein's perspective was the notion that the term Hamilton–Jacobi theory properly only applied to the third part of this division.

[3] The account which follows draws on Lützen (1990, 638–642) and Cayley (1858, 3–9).

Simon Laplace's subsequent successful analysis of the inequalities in the Jupiter–Saturn–Sun system. Consider an *n*th-order ordinary differential equation

$$x^{(n)} = P\Big(t, x, x', x'' + \cdots + x^{(n-1)}\Big), \tag{1}$$

where $x$, $x''$, …, $x^{(n)}$ are the successive derivatives of $x$ with respect to $t$. Assume a solution $x = F(t,a_1,\ldots,a_n)$ containing $n$ arbitrary constants $a_i$ is known. Let us modify Eq. (1) by adding a perturbing term $Q(t, x, x', x'' + \cdots + x^{(n-1)})$ to $P$:

$$x^{(n)} = P\Big(t, x, x', x'' + \cdots + x^{(n-1)}\Big) + Q\Big(t, x, x', x'' + \cdots + x^{(n-1)}\Big). \tag{2}$$

We consider solutions of (2) that have the form $x = F(t,a_1(t),\ldots,a_n(t))$, where the $a_i$ which were constant before are now allowed to vary with time. The problem reduces to determining the $a_i$ as functions of time. One condition for the solution is obtained by assuming that the derivatives of $x$ when the $a_i$ are allowed to vary remain equal to the original expressions for these derivatives when the $a_i$ are assumed constant. The resulting method of integration is known as the method of variation of arbitrary constants.

On the 20th of June 1808 Poisson presented "Sur les inégalités des moyens mouvements des planètes" to the Paris Academy of Sciences. This paper was published by Poisson (1809a). Here Poisson explored the variation of the elements that occur in the solutions of the differential equations of perturbed planetary motion. Lagrange was one of the reviewers of the memoir and was stimulated to do some investigations of his own which he presented in three memoirs, the most important one being "Mémoire sur la théorie générale de la variation des constantes arbitraires dans tous les problèmes de la mécanique," presented to the Academy on March 13 1809 and published as Lagrange (1809b).[4] Continuing the work done in Lagrange (1809a) he here introduced new ideas and methods to the subject. Poisson in turn responded to Lagrange's theory with some additions of his own, that appeared in Poisson (1809b) with a title that is very similar to Lagrange's (1809b). Lagrange refined and systematically presented his findings in volume one of the second edition of the *Mécanique analytique*, published in 1811. Lagrange at this time was in his seventies, while Poisson was not yet thirty. Lagrange's research on the variation of constants was his last contribution to mathematical science; he died in 1813, aged 77.

Lagrange (1809b) derived the standard formula for the perturbation function using the Lagrange bracket formalism and introduced symbols for the conjugate momenta. Poisson (1809b) adopted the conjugate momenta but treated them not simply as notational symbols but more conceptually as variables on a par with the actual dynamical variables. Poisson also derived the symmetric or canonical form of the differential equations for the variable constants of integration. Finally, Poisson introduced the Poisson brackets and derived some of their basic properties. Lagrange and Poisson developed new methods within the formalism of the *Méchanique analitique* (1788).

[4] For the related researches of Lagrange and Poisson on the variation of constants at this time see Costabel (1981, 484–485).

The resulting theory was the achievement of both men, although Poisson (1809b) provided the most lucid account of the new results; this memoir was a seminal contribution to the history of mathematical mechanics. (Costabel (1981, 405) writes "His [Poisson's] sense of formalization led him to discover analogies, to unify problems and topics previously considered distinct, and to extend definitively "the domain of the calculus" (*l'emprise du calcul*).") The notation used by Poisson and Lagrange is different from the modern one, which essentially originated with Hamilton and Jacobi. Poisson wrote $\varphi$, $\psi$ and $\theta$ for the generalized coordinates $q_1$,$q_2$, and $q_3$, and $s$,$u$ and $v$ for the generalized momenta $p_1$,$p_2$, and $p_3$. The three symbols in each of these cases were understood to stand in for a full set of $n$ variables. The Lagrangian $L$ was denoted by $R$. The symbol "$d$" denoted both ordinary and partial differentiation. The equational forms presented by Poisson and Lagrange were unwieldy and long. Because our story begins primarily with Jacobi, we shall adopt Jacobi-style notation (see Sect. 3.3) in our account of their work. (Also, we do not explore the interesting results Poisson obtained involving his eponymous brackets.)

We begin with the standard Lagrangian equations of motion

$$\frac{\mathrm{d}\frac{\partial T}{\partial q_i'}}{\mathrm{d}t} - \frac{\partial T}{\partial q_i} + \frac{\partial V}{\partial q_i} = 0, \quad (i = 1, \ldots, n), \tag{3}$$

where the $q_i$ are generalized coordinates, $T$ is one-half the live force or kinetic energy and $-V$ is what Hamilton would later call the force function; $V$ is today called the potential. Lagrange assumed that a perturbing term or potential $-\Omega$ is added to $V$ in (3), where $\Omega = \Omega(q_1, q_2, \ldots, q_n)$. Equation (3) then becomes

$$\frac{\mathrm{d}\frac{\partial T}{\partial q_i'}}{\mathrm{d}t} - \frac{\partial T}{\partial q_i} + \frac{\partial V}{\partial q_i} = \frac{\partial \Omega}{\partial q_i}. \quad (i = 1, \ldots, n) \tag{4}$$

Let $q_i = q_i(t, a_1, \ldots, a_{2n})$ be a complete solution of (3) without the disturbing function, containing the $2n$ arbitrary constants $a_k$. We now let $a_k$ be variable functions of time, $a_k = a_k(t)$, and suppose that the resulting $q_i = q_i(t, a_1(t), \ldots, a_{2n}(t))$ are a solution of (4). We introduce the new variable $p_i$ defined as $p_i = \frac{\partial T}{\partial q_i'}$. Lagrange derived the general identity

$$\frac{\partial \Omega}{\partial a_k} = \sum_{j=1}^{2n} [a_k, a_j] \frac{\mathrm{d}a_j}{\mathrm{d}t} \quad (k = 1, \ldots, 2n). \tag{5}$$

Here what would later be called the "Lagrange bracket" $[a_k, a_j]$ is defined as

$$[a_k, a_j] = \sum_{s=1}^{n} \left( \frac{\partial q_s}{\partial a_k} \frac{\partial p_s}{\partial a_j} - \frac{\partial q_s}{\partial a_j} \frac{\partial p_s}{\partial a_k} \right). \tag{6}$$

Lagrange (1809b, Sect. 26) seemed to have introduced the $p_i$ as a matter of notation to write down the general identity (5). The idea of setting $p_i = \frac{\partial T}{\partial q_i'}$ and using $p_i$ as a new dynamical variable was a further step taken by Poisson (1809b, Sect. 1). This was the first appearance of the "conjugate momentum," which later would also be adopted by Hamilton (who referred to Poisson) and become a standard part of Hamilton–Jacobi theory.

Equation (5) was the foundational identity for the Lagrange-Poisson theory of variation of constants. We indicate here the main idea involved in Lagrange's derivation of (5) for the case of a one-dimensional system with coordinate $q$ and generalized momentum $p = \frac{\partial T}{\partial q'}$. There would appear to be only one way of proving the result. The perturbed Lagrangian equation is

$$\frac{\mathrm{d}p}{\mathrm{d}t} - \frac{\partial T}{\partial q} + \frac{\partial V}{\partial q} - \frac{\partial \Omega}{\partial q} = 0, \tag{7}$$

where the perturbation potential $\Omega$ is a function of $t$ and $q$ (but not $p$). We have

$$\frac{\mathrm{d}p}{\mathrm{d}t} = \frac{\partial p}{\partial t} + \frac{\partial p}{\partial a_1}{a_1}' + \frac{\partial p}{\partial a_2}{a_2}'. \tag{8}$$

By the method of variation of constants we suppose that $\frac{\partial p}{\partial t} - \frac{\partial T}{\partial q} + \frac{\partial V}{\partial q} = 0$. Hence we obtain

$$\frac{\partial p}{\partial a_1}{a_1}' + \frac{\partial p}{\partial a_2}{a_2}' = \frac{\partial \Omega}{\partial q}. \tag{9}$$

Again by the method of variation of constants we assume the velocity is not altered in the perturbed motion and so $\frac{\mathrm{d}q}{\mathrm{d}t} = \frac{\partial q}{\partial t}$. Because $\frac{\mathrm{d}q}{\mathrm{d}t} = \frac{\partial q}{\partial t} + \frac{\partial q}{\partial a_1}a_1' + \frac{\partial q}{\partial a_2}a_2'$ there follows

$$\frac{\partial q}{\partial a_1}{a_1}' + \frac{\partial q}{\partial a_2}{a_2}' = 0. \tag{10}$$

$\Omega$ is a function of $q$ alone and so from (9) we have

$$\frac{\partial \Omega}{\partial a_1} = \frac{\partial \Omega}{\partial q}\frac{\partial q}{\partial a_1} = \left(\frac{\partial p}{\partial a_1}{a_1}' + \frac{\partial p}{\partial a_2}{a_2}'\right)\frac{\partial q}{\partial a_1}. \tag{11}$$

with a similar expression for $\frac{\partial \Omega}{\partial a_2}$. Now from (10) we have ${a_1}' = -\left(\frac{\partial q}{\partial a_2}/\frac{\partial q}{\partial a_1}\right){a_2}'$. Substitution into (11) then gives

$$\frac{\partial \Omega}{\partial a_1} = \frac{\partial \Omega}{\partial q}\frac{\partial q}{\partial a_1} = \left(\frac{\partial q}{\partial a_1}\frac{\partial p}{\partial a_2} - \frac{\partial q}{\partial a_2}\frac{\partial p}{\partial a_1}\right){a_2}' = [a_1, a_2]{a_2}'. \tag{12}$$

Similarly, $\frac{\partial \Omega}{\partial a_2} = [a_2, a_1]{a_1}'$. The extension of the derivation in the case of $n$ coordinates is a matter of arranging the notation.

Lagrange also showed that $\left[a_k, a_j\right]$ is independent of time, an important theorem that he proved in the second edition of the *Mécanique analytique* (1811, part 2, Sect. 5, article 7). Writing in the late 1830s Jacobi (1866b, 419) would refer to this result as the "berühmte *Lagrangesche* Satz."

Poisson (1809b, 313) and Lagrange (1810, Sect. 2) following him presented first-order differential equations in symmetric form for the variable constants $a_i = a_i(t)$ and $b_i = b_i(t)$ involving the partial derivatives of the perturbing function with respect to these quantities. Poisson provided no proof (possibly because he viewed it as being fairly evident) and regarded the equations themselves as lacking any "great utility." Lagrange by contrast admired the symmetric form of the equations and provided an explicit derivation in his "Second Mémoire" of 1810 and in the *Mécanique analytique* (1811, part 2, Sect. 5, article 14). We let the $2n$ arbitrary constants be the initial values of the coordinates $q_i$ and the momenta $p_i$: $a_1, \ldots, a_n$ are the values of $q_1, \ldots, q_n$ at $t = 0$ and $b_1, \ldots, b_n$ are the values of $p_1, \ldots, p_n$ at $t = 0$. The differential equations for the perturbed motion are obtained in the form

$$\begin{array}{l} \frac{\mathrm{d}a_i}{\mathrm{d}t} = -\frac{\partial\Omega}{\partial b_i} \\ \frac{\mathrm{d}b_i}{\mathrm{d}t} = \frac{\partial\Omega}{\partial a_i} \end{array} (i = 1, \ldots, n). \tag{13}$$

This marked the first appearance of the "canonical" equations, although (13) was expressed in terms of the variable constants of integration and not as Hamilton would later do in terms of dynamical coordinate variables.

Equation (13) follow directly from the main identity (5) and are basic to the Lagrange-Poisson theory of perturbations. As before, we give the derivation for the one-dimensional case with variables $p$ and $q$. These variables are expanded as power series about zero:

$$\begin{array}{l} q(t) = q(0) + tq'(0) + \ldots \\ p(t) = p(0) + tq'(0) + \ldots \end{array} \tag{14}$$

The values of $a$ and $b$ are $a = q(0)$ and $b = p(0)$. In terms of the notation in (5) we have $a = a_1$ and $b = a_2$. Equation (5) then becomes $\frac{\partial\Omega}{\partial b} = [b, a]\frac{\mathrm{d}a}{\mathrm{d}t}$ and $\frac{\partial\Omega}{\partial a} = [a, b]\frac{\mathrm{d}b}{\mathrm{d}t}$. By definition $[a, b] = \frac{\partial q}{\partial a}\frac{\partial p}{\partial b} - \frac{\partial q}{\partial b}\frac{\partial p}{\partial a}$. Letting $a = q(0)$ and $b = p(0)$ it then follows from (2.14) that $\frac{\partial q}{\partial a} = \frac{\partial p}{\partial b} = 1$ and $\frac{\partial q}{\partial b} = \frac{\partial p}{\partial a} = 0$. Hence [a,b] = 1 and we have $\frac{\partial\Omega}{\partial a} = \frac{\mathrm{d}b}{\mathrm{d}t}$. Similarly, there follows $\frac{\partial\Omega}{\partial b} = -\frac{\mathrm{d}a}{\mathrm{d}t}$.

## 3 Jacobi's reshaping (1837, 1866a, b) of Hamilton's theory: the three Jacobian theorems

### 3.1 Introduction

Jacobi tended to emphasize the more mathematical dimension of Hamilton's ground-breaking dynamical researches. Even when the primary focus was on mechanics, he was appreciative of the analytical possibilities of Hamilton's discoveries related to partial differential equations. Concerning the theory of partial differential equations Jacobi wrote,

> Hamilton's theorems themselves contribute to the perfection of this theory in a significant and unexpected way, although the author has not emphasized this purely analytical interest (Jacobi 1866b, 304)

Working from established researches by Lagrange, Poisson, Hamilton, Johann Pfaff and Augustin Cauchy, Jacobi extended and deepened their insights. He blended dynamics, the calculus of variations and the theory of differential equations into a coherent program of research.

In focusing on Jacobi's development of Hamilton–Jacobi theory it should be emphasized that we are being selective and that his own mathematical tastes encompassed other subjects that may not be of compelling interest to us but were an integral part of his conception of dynamical science. Eight lectures of the *Vorlesungen* are devoted to an investigation of the principle of the last multiplier. The topic receives more attention than any other in the work. Jacobi also explored this subject in two lengthy articles published in Crelle's journal in 1844 and 1845. (For references to these articles and related publications see Jacobi (1996, 319–320).) He attached great importance to the theory of the last multiplier and believed that it was of fundamental dynamical significance. It is discussed in books on differential equations at the end of the nineteenth where it is presented as a noteworthy subject. Insofar as later mechanics is concerned, the principle is not emphasized and occupies a peripheral place in modern mathematical physics.

### 3.2 Jacobi's writings

Jacobi's development of what is known as Hamilton–Jacobi theory took place in the years from 1837 to 1843, with some further refinements in his Berlin lectures of 1847–1848, which, however, remained unpublished until 1996. The works in question are

1. The article "Ueber die Reduction der Integration der partielle Differentialgleichungen erster Ordnung zwischen irgend einer Zahl Variabeln auf die Integration eines einzigen Systemes gewöhnlicher Differentialgleichungen." In Crelle's journal in 1837, followed by a French translation in Liouville's journal the next year. (1837a).
2. The short piece "Note sur l'intégration des équations différentielles de la dynamique" in 1837 in the *Comptes rendus* of the Paris Academy of Sciences. (1837b)
3. *Ueber diejenigen Probleme der Mechanik, in welchem eine Kräftefunction existiert und über die Theorie der Störungen*.[5] This treatise was the first of five works that were drawn from Jacobi's *Nachlass* and published as supplements to the 1866 edition of his *Vorlesungen*. A sprawling treatise, it occupies 167 printed pages and includes subjects that were covered in the *Vorlesungen*, as well as other subjects

[5] The first word of the title is spelled "Ueber" as in German spelling at the time. In the 1890 edition in Jacobi's *Gesammelte Werke 5* the word is spelled "Über". We have adopted the original spelling in our study.

that were not explored in any detail in his other writings or lectures. It was reprinted with some minor emendations under the direction of Karl Weierstrass in 1890 in volume 5 of Jacobi's *Werke*, and so would have enjoyed some circulation among researchers at the end of the century (Jacobi 1866b).

In the note in the *Comptes rendus* Jacobi (1837b) referred to this treatise as a "memoir" that he had been working on, but had been drawn away from by researches in the theory of numbers. The purpose of the Paris note was simply to state some key results that would be presented and derived in detail in the completed memoir. In some editorial remarks, Weierstrass (or his assistant Fritz Kötter) called attention to the preliminary character of the treatise, and suggested it was kind of a draft work that Jacobi himself may not have been completely satisfied with.[6] It is noteworthy that (1837b) presents the equations of motion in Newtonian form, whereas in (Jacobi 1866b) these equations are given in both Newtonian and canonical form.

The editors of the 1890 edition state that the work probably dates from 1836 or 1837 at the latest (an estimate no doubt based on the remarks in Jacobi (1837b)).[7] There is in fact some textual evidence for this dating. The opening part of *Ueber diejenigen Probleme* is similar in content to that of his 1837 article "Ueber die Reduction …" In the Paris Academy "Note sur L'intégration …" of 1837 it was referred to as a work in progress. In addition, the proof of Jacobi's integration theorem in *Ueber diejenigen Probleme* uses the $d$ and $\delta$ operations in a way that is similar to his work from the same period on the theory of the second variation in the calculus of variations. It is different from the proof in either the 1842–43 Königsberg lectures or the 1847–48 Berlin lectures. Furthermore, Jacobi's introduction in the Paris note of transformations for the constants of integration is dealt with in some detail in *Ueber diejenigen Probleme* in the later part of this work on perturbations. Here Jacobi provides proofs for the results announced in the Paris note. Since he did not concern himself with this subject in his later published writings or in the 1842–43 Königsberg lectures and the 1847–48 Berlin lectures, it would seem that the researches in the *Comptes rendus* note and in *Ueber diejenigen Probleme* were carried out at about the same time, which would have been in the late 1830s.

4. *Vorlesungen über Dynamik* (1866a). Lectures given at the University of Königsberg in the winter semester of 1842–1843. Alfred Clebsch was the nominal editor of the work; the notes were taken and redacted by Carl Borchardt, who by the 1860s was a prominent figure in German mathematics and editor of *Crelle*. Included also are five supplementary essays, the first of which is the aforementioned *Ueber diejenigen Probleme der Mechanik* (1866b).

[6] Weierstrass/Kotter writes "it would be highly unusual for Jacobi to have left such an important essay unpublished, unless he believed that it needed to be revised in some places." (Jacobi 1890, *Werke 5*, 514).

[7] In the foreword to the *Vorlesungen* (1866a) it is stated that there is some uncertainty concerning the dating of the five supplementary works. The first four are said to be probably ("wahrscheinlich") composed sometime after the Königsberg lectures of 1842-1843. However, no evidence is given for this judgement. Helmut Pulte (Jacobi (1996, xxxviii n. 110)) dates *Ueber diejenigen Probleme* to 1836 or into 1837, although it is possible the manuscript may have been completed somewhat later than this, although definitely before 1842.

5. *Vorlesungen über analytische Mechanik* (1996). Lectures delivered at the University of Berlin in 1847–1848 and recorded by Wilhelm Scheibner; published in 1996 with the introduction, editing and commentary by Helmut Pulte.

Our primary concern will be with three results that are presented with proof in the 1866 volume. The first is what is sometimes known as Jacobi's theorem in modern textbooks, and was called "Jacobi's first theorem" by Poincaré. We shall refer to it as **Jacobi's integration theorem**. The result first appeared research stimulated by Jacobi's discovery of Hamilton's papers in the *Philosophical Transactions*, and was published in "Ueber die Reduction der Integration." It is also the subject of lectures 20 and 21 of the Königsberg *Vorlesungen* (1866a), appears as Theorem VI of *Ueber diejenigen Probleme* and is covered again in the Berlin *Vorlesungen* (1996). There are some important differences in the formulation and proof of the theorem in these works. In terms of modern theory, it was the treatment in the *Vorlesungen* (1866a) that is most relevant.

The second result is **Jacobi's theorem on canonical elements**, formulated for perturbed systems within the theory of variation of constants. This result may be regarded as an extension of the Lagrange-Poisson theorem on canonical constants using a solution of the Hamilton–Jacobi partial differential equation. A statement of the result was given in his 1837 note in the *Comptes rendus*. The theorem was stated and proved as Theorem IX of *Ueber diejenigen Probleme* (1866b). The theorem was also dealt with in a somewhat different way in lecture 36 of the Königsberg *Vorlesungen* (1866a) and again in lecture 43 of the Berlin *Vorlesungen* (1996).

The third result is **Jacobi's theorem on canonical transformations**, formulated for perturbed systems within the theory of variation of constants; it is what is called "Jacobi's second theorem" by Poincaré. It was stated by Jacobi in 1837 in the *Comptes rendus* and proved as Theorem X of *Ueber diejenigen Probleme* (1866b). Although it became one of Jacobi's signature contributions to dynamical analysis, it was not taken up in either his Königsberg or Berlin lectures. Jacobi formulated it as a result in the theory of variation of constants, but he asserted that it was logically more general and not restricted to the theory of perturbations. In the final articles of *Ueber diejenigen Probleme* he made some concrete first steps toward developing a general theory of transformations based on Theorem X.

### 3.3 Jacobi's notation

Jacobi initiated the practice of using distinct symbols for ordinary and partial differentiation, $d$ for ordinary differentiation and $\partial$ for partial differentiation, just as we do today. He used $q$ to denote generalized coordinates and $p$ to denote conjugate momenta, as is common in modern textbooks. In comparison with the writings of other mathematicians of the nineteenth century, his work, at least at the level of presentation and notation, seem more familiar to a modern reader.

In his Crelle article "Ueber die Reduction" Jacobi (1837a, 116) presented the Newtonian equations of motion in the form:

$$m_i \frac{\mathrm{d}^2 x_i}{\mathrm{d}t^2} = \frac{\partial U}{\mathrm{d}x_i}, \quad m_i \frac{\mathrm{d}^2 y_i}{\mathrm{d}t^2} = \frac{\partial U}{\mathrm{d}y_i}, \quad m_i \frac{\mathrm{d}^2 z_i}{\mathrm{d}t^2} = \frac{\partial U}{\mathrm{d}z_i}.$$

The Hamilton–Jacobi equation for conservative systems is given on p. 113 in the form:

$$\frac{1}{2} \sum \frac{1}{m_i} \left[ \left( \frac{\partial S}{\partial x_i} \right)^2 + \left( \frac{\partial S}{\partial y_i} \right)^2 + \left( \frac{\partial S}{\partial z_i} \right)^2 \right] - U = H.$$

In Jacobi's (1837b) note for the *Comptes rendus* of the same year these equations were given (p. 65) in the form.[8]

$$m \frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = \frac{\mathrm{d}U}{\mathrm{d}x}, \quad m \frac{\mathrm{d}^2 y}{\mathrm{d}t^2} = \frac{\mathrm{d}U}{\mathrm{d}y}, \quad m \frac{\mathrm{d}^2 z}{\mathrm{d}t^2} = \frac{\mathrm{d}U}{\mathrm{d}z_i}.$$

$$\frac{1}{2} \sum \frac{1}{m} \left[ \left( \frac{\mathrm{d}V}{\mathrm{d}x} \right)^2 + \left( \frac{\mathrm{d}V}{\mathrm{d}y} \right)^2 + \left( \frac{\mathrm{d}V}{\mathrm{d}z} \right)^2 \right] = U + h.$$

In French and English journals of the time and until much later it was customary to use $d$ to denote both ordinary and partial differentiation. The editors of the *Comptes rendus* rendered Jacobi's equations in the common way.

In a paper on determinants published in 1841 Jacobi wrote:

> To distinguish the partial derivatives from the ordinary, where all variable quantities are regarded as functions of a single one … I have preferred to designate ordinary differentials by the character $d$ and the partial differential by $\partial$. (Jacobi 1841, 320)[9]

Throughout the *Vorlesungen* both symbols $d$ and $\partial$ are used.[10] However, Cajori (1929, 238) writes: "The notation $\frac{\partial u}{\partial x}$ did not meet with immediate adoption. It took over half a century for it to secure a generally accepted place in mathematical writing."[11] Indeed, as we shall see, traditional notation was employed by Desboves (1848), Donkin (1855), Cayley (1858), Tisserand (1868) and Poincaré (1892–1905). In Charlier (1902) and Whittaker (1904) the modern dual notation is used throughout.

---

[8] There is a misprint in the third equation of the first line, where $\frac{\mathrm{d}U}{\mathrm{d}t}$ should be $\frac{\mathrm{d}U}{\mathrm{d}z}$.

[9] English translation by Cajori (1929, 236).

[10] There is the possibility that Borchardt or Clebsch added the two symbols ($d$ and $\partial$) for regular and partial differentiation in the 1866 transcription of Jacobi's Königsberg lectures, where only one symbol may have been used. There is no evidence that this was the case. Indeed, the available evidence points to the opposite conclusion. The two symbols are used in Jacobi's (1837a) as we saw above. Both symbols appear as well as in Jacobi (1846), written in 1838. Jacobi (1841) publicly advocated for using both symbols. Jacobi's (1837c) famous article on the calculus of variations, "Zur Theorie der Variationsrechnung und der Differentialgleichungen," uses the cursive $\partial$ for both ordinary and partial differentiation. The two symbols appear in all of Jacobi's German-language publications on mathematical dynamics. Furthermore, in Pulte's 1996 edition of the 1847-1848 Berlin lectures he includes some copies from the original handwritten notes taken by Scheibner. Consider in particular p. LX where the two symbols for ordinary and partial differentiation are employed; the corresponding typeset text is on p. 188.

[11] Cajori (1929, 235) writes of the "notation of partial derivatives which has become so popular in recent years." Even at this late date the dual notation *d/∂* was regarded as fairly novel.

### 3.4 Jacobi's two papers of 1837

Jacobi's "Ueber die Reduction" of 1837 occupies sixty-six journal pages and deals with both dynamics and more purely mathematical matters involving the solutions of ordinary and partial differential equations.[12] These researches were stimulated by results that Hamilton (1834, 307) and (1835, 99–101) had obtained, although Jacobi's formulation was mathematically original and more far-reaching. (For Hamilton's investigation see Nakane and Fraser (2002, 180–183).) The main result of the memoir is Jacobi's integration theorem, which possesses a more particular form than the corresponding result in the *Vorlesungen*.[13] At this stage, Jacobi had embraced Hamilton's partial differential equation but had not yet adopted the "canonical momenta" or Hamilton's first-order dynamical equations. The Hamiltonian was given in the usual way as the sum (in modern terminology) of the kinetic and potential energies. The analysis was formulated within a Cartesian coordinate system involving $3n$ variables $x_i$, $y_i$ and $z_i$, where these variables are functions of the time $t$. An account of Jacobi's statement and proof of the integration theorem is given in Nakane and Fraser (2002, 206–208).

In the Paris note of 1837 Jacobi set out two new theorems. The first, the theorem on canonical elements, was a result about perturbed motion. One begins with an unperturbed system, takes a complete integral of the Hamilton–Jacobi equation, and applies the integration theorem to get a solution to the equations of motion in terms of the complete integral and $2n$ constants $\alpha$ and $\beta$. One then assumes the system is perturbed, giving rise to the addition to the potential or force function $U$ of a second term $\Omega$, so that the total force function becomes $U + \Omega$. It is also assumed that a solution to the perturbed system may be obtained from the integral of the unperturbed system by making the constants $\alpha$ and $\beta$ variable functions of time. Jacobi asserted that these variables will satisfy equations with respect to $\Omega$ of the same form as the Lagrange-Poisson equations in the theory of variation of constants. Jacobi referred to $\alpha(t)$ and $\beta(t)$ as "canonical elements"; this is the first appearance of the word "canonical" in his dynamical writings. It later became standard to refer to such equations as canonical, whether they involved variable constants or dynamical variables. No proof was given of this result.

Suppose now that we have a set of variable constants satisfying equations in canonical form. A new theorem that Jacobi presented in the Paris note involved transformations from this set of variable constants to another set of variable constants that preserves the canonical form of the differential equations. Again, no proof was given of the result.

The brevity of Jacobi's Paris memoir belied its seminal character and the significance of the results he presented there. In the decades which followed, it stimulated

[12] The title of the paper, "On the reduction of the integration of partial differential equations of the first order between any number of variables to the integration of a single system of ordinary differential equations," is puzzling, since the integration theorem at the centre of the paper is the converse of what is stated in the title. The title better describes Cauchy's (1819) paper on first-order partial differential equations.

[13] In *Ueber diejenigen Probleme* Jacobi (1866b, 355) refers to this theorem as a generalization of Hamilton's theorem. Courant and Hilbert (1962, 127) call Jacobi's integration theorem the Hamilton–Jacobi theorem, although it is generally called Jacobi's theorem in modern textbooks.

researches in France and England that we examine in Sect. 4. In the late 1830s (not long after he wrote the Paris memoir) Jacobi (1866b) gave a derivation of these two results in Sects. 27–38 of *Ueber diejenigen Probleme*, where they appear as Theorems IX and X.

### 3.5 Jacobi's derivation of the canonical equations and the Hamilton–Jacobi equation

We turn now to a detailed consideration of Hamilton's dynamical equations and Jacobi's derivation of the Hamilton–Jacobi partial differential equation as they were set out in Jacobi's (1866a) Königsberg lectures.

In the eighth lecture Jacobi (1866a, 58) laid down a variational principle that was related to but distinct from the principle of least action. He attributed it to Hamilton; it is indeed known today as Hamilton's principle, although neither Hamilton nor Jacobi referred to in this way.[14] Consider a system with $\mu$ quantities $q_i$ ($i = 1, \ldots, \mu$). The $q_i$ are functions of time and are called generalized coordinates in the modern subject. Jacobi denoted the time derivative of $q_i$ by $q_i{}'$. We are given the quantity $T$ which is one-half the living force, or the kinetic energy in modern parlance. $T$ is a function of $q_i$, and $q_i{}'$. We also have the quantity $U$ called the force function; $-U$ is what is known today as the potential energy. Whereas Hamilton had only considered functions $U$ of $q_i$ that do not contain the time, Jacobi supposed that $U$ was a function of both $q_i$ and $t$. The principle of stationary action asserts that the variation of $T + U$ is equal to zero:

$$\delta \int (T + U)dt = 0, \tag{15}$$

where the integration is taken from the initial and final values of $t$. Jacobi (1866a, 59) observed that with this principle "in general one does not obtain a minimum with the vanishing of the variation." The principle provides an effective tool that is not necessarily based on the optimization of a physical quantity. (The principle in this sense would later be exploited in Poincaré's incisive development of transformation theory, as we shall see in Sects. 6.3 and 6.4.)[15]

From (15) Jacobi derived the dynamical equations

$$\frac{\mathrm{d}\frac{\partial T}{\partial q_i'}}{\mathrm{d}t} - \frac{\partial T}{\partial q_i} = \frac{\partial U}{\partial q_i}. \tag{16}$$

Jacobi called (16) the Lagrangian form of the differential equations of motion and referenced Lagrange's *Mécanique analytique*. He credited Hamilton for the derivation although it actually went back to Lagrange's earliest dynamical work in 1762. (See Fraser (1983) for details.)

14 See Nakane and Fraser (2002, 184–185).

15 In discussing fundamental laws of mechanics, Truesdell (1968, 242 n. 4) calls attention to the utility of variational principles: "formal rearrangements are possible. E.g., a sufficiently general "variational" principle (i.e., a formal expression in variations, not a true minimum principle) can be made equivalent [to these laws]".

Hamilton had shown that the $\mu$ second-order differential equations of motion (16) can be expressed in a new way as a system of $2\mu$ first-order equations. In the ninth lecture, Jacobi stated and derived these equations, referring to them as Hamilton's equations. We introduce the new quantity $p_i$ defined as

$$p_i = \frac{\partial T}{\partial q_i'}. \tag{17}$$

Hamilton's equations are then given in the symmetric form:

$$\begin{aligned} \frac{\mathrm{d}q_i}{\mathrm{d}t} &= \frac{\partial H}{\partial p_i}, \\ \frac{\mathrm{d}p_i}{\mathrm{d}t} &= -\frac{\partial H}{\partial q_i}, \end{aligned} \quad (i = 1, \ldots, \mu) \tag{18}$$

where $H = T - U$ is a function of $t$ and $q_i$and$p_i (i = 1, \ldots, \mu)$. $H$ is called the Hamiltonian in modern dynamics. The derivation of (18) is done directly by Jacobi from (16) and (17). It is the same derivation as Hamilton's and is found in modern textbooks.[16] Note that it is possible to obtain (18) from an integral variational principle, as Poincaré would later show (see Sect. 6.4.1).

In the nineteenth lesson, Jacobi returned to a consideration of the integral $\int (T + U)dt$. A full account of this lecture would require looking in some detail at his understanding of the foundations of the calculus of variations. For example, he had a somewhat different concept of variation than is standard in modern textbooks. He took a general solution to the Euler–Lagrange equation and then varied the arbitrary constants in this solution to obtain the variations. We will here only explore the main results insofar as Hamilton–Jacobi theory itself is concerned.

Jacobi designated $T + U$ by $\varphi$ (what is known in modern mechanics as the Lagrangian). Using the same procedure involved in the derivation of (16) he arrived at the equation

$$\delta \int \varphi dt = \sum \frac{\partial \varphi}{\partial q_i}\delta q_i - \sum \frac{\partial \varphi}{\partial q_i^0}\delta q_i^0 + \int \sum \left( \frac{\partial \varphi}{\partial q_i} - \frac{\mathrm{d}\frac{\partial \varphi}{\partial q_i'}}{\mathrm{d}t} \right) \delta q_i dt. \tag{19}$$

The integration is presumed to take place along the arcs $q_i = q_i(t)$ for which Lagrange's equation (3) hold.[17] Hence the term involving the integral on the right side of (19) is zero. As Jacobi (1866a, 145) stated, "the part under the integral sign of the variations sought for vanishes only by virtue of the differential equations of motion which are assumed to be satisfied." We now define $p_i$ as

[16] Hamilton's derivation is given by Nakane and Fraser (2002, 181). For this derivation in modern textbooks see Goldstein (1950, 217), Fox (1954, 144-145), Gelfand and Fomin (1963, 67-70), Landau and Lifshitz (1969, 131-132) or Arthurs (1975, 19-20).

[17] Jacobi implicitly assumed that for each point $t$, $q_i$, one and only one extremal connects the initial time and position to $t$, $q_i$. In modern calculus of variations $q = q_i(t)$ is said to be embedded in a field of extremals containing the initial and final values of $t$, $q_i$. The function $\int \varphi dt$ in Eq. (19) is known as a field integral.

$$\frac{\partial \varphi}{\partial q_i'} = p_i.(i = 1, \ldots, \mu) \tag{20}$$

Equation (19) then becomes

$$\delta \int \varphi dt = \sum p_i \delta q_i - \sum p_i^0 \delta q_i^0. \tag{21}$$

We infer from (21) that the integral $\int \varphi dt$ is a well-defined function $V$ of $t$ and the $q_i$

$$V = \int \varphi dt. \tag{22}$$

Equation (21) may now be written

$$\delta V = \sum \frac{\partial V}{\partial q_i} \delta q_i - \sum \frac{\partial V}{\partial q_i^0} \delta q_i^0. \tag{23}$$

Comparing (21) and (23) then gives

$$\frac{\partial V}{\partial q_i} = p_i (i = 1, \ldots, \mu). \tag{24}$$

Further it is evident from (22) that

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \varphi.$$

Hence, we obtain

$$\varphi = \frac{\mathrm{d}V}{\mathrm{d}t} = \frac{\partial V}{\partial t} + \sum \frac{\partial V}{\partial q_i} {q_i}' = \frac{\partial V}{\partial t} + \sum p_i {q_i}',$$

or

$$0 = \frac{\partial V}{\partial t} + \sum p_i {q_i}' - \varphi. \tag{25}$$

Letting

$$\psi = \sum p_i {q_i}' - \varphi, \tag{26}$$

which is the origin of what in modern dynamics is known as the Hamiltonian. Equation (25) is written

$$\frac{\partial V}{\partial t} + \psi = 0. \tag{27}$$

Using (20) we may express $q_1', q_2', \ldots, q_\mu'$ as functions of $t$, $q_1, q_2, \ldots, q_\mu, p_1, p_2, \ldots, p_\mu$. Hence $\psi$ is a function of $t$, $q_1, q_2, \ldots, q_\mu, p_1, p_2, \ldots, p_\mu$. Also, from (24) we have $\frac{\partial V}{\partial q_i} = p_i$. Hence $\psi$ becomes a function of $t$ and $q_1, q_2, \ldots, q_\mu : \psi = \psi\left(t, q_1, q_2, \ldots, q_\mu, \frac{\partial V}{\partial q_1}, \frac{\partial V}{\partial q_2}, \ldots, \frac{\partial V}{\partial q_\mu}\right)$. Thus (27) may be written

$$\frac{\partial V}{\partial t} + \psi\left(t, q_1, q_2, \ldots, q_\mu, \frac{\partial V}{\partial q_1}, \frac{\partial V}{\partial q_2}, \ldots, \frac{\partial V}{\partial q_\mu}\right) = 0, \tag{28}$$

a first-order non-linear partial differential equation for $V$ in which $V$ itself does not appear. Jacobi referred to (28) as the Hamiltonian partial differential equation. It is known in the modern subject as the Hamilton–Jacobi equation. He noted that an integration of the differential equations of motion yields a general solution to (28) containing as arbitrary constants the initial values of $q_i$.

### 3.6 Jacobi's integration theorem in the *Vorlesungen* (1866a)

In lecture 20 Jacobi stated and proved the integration theorem that he had originally presented in his 1837 paper in Crelle. While the core of Jacobi's 1837 proof and the proof in the *Vorlesungen* are the same, there are two respects in which the two versions differ. First, in the *Vorlesungen* the equations of motion are given in canonical rather than Newtonian form. Second, the theorem in the *Vorlesungen* is valid in a wider setting than the mechanical problem. Jacobi asserted that the theorem was of general mathematical significance, although dynamics was the ostensible subject of the investigation. He wrote (1866a, 147).

Everything that has been said up to now holds not only for problems of mechanics but also when $\varphi$, instead of being equal to $T + U$, is an arbitrary function of $t, q_1, q_2, \ldots, q_\mu, q_1', q_2', \ldots, q_\mu'$. In problems of mechanics, however, $\psi$ acquires a simple significance …

Unlike earlier in the *Vorlesungen*, where $\psi$ was given as $\psi = \sum p_i q_i' - \varphi$, in lecture 20 $\psi$ is any function of $t, q_1, q_2, \ldots, q_\mu, p_1, p_2, \ldots, p_\mu$. It is stipulated that the following relations hold:

$$\frac{\partial V}{\partial q_i} = p_i (i = 1, \ldots, \mu) \tag{29}$$

One replaces $p_i$ by $\frac{\partial V}{\partial q_i}$ in $\psi$ and forms the partial differential equation

$$\frac{\partial V}{\partial t} + \psi\left(t, q_1, q_2, \ldots, q_\mu, \frac{\partial V}{\partial q_1}, \frac{\partial V}{\partial q_2}, \ldots, \frac{\partial V}{\partial q_\mu}\right) = 0, \tag{30}$$

A complete solution of (30) will be a solution (up to an additive constant) of the form:

$$V = V\big(t, q_1, \ldots, q_\mu, \alpha_1, \ldots, \alpha_\mu\big) + C, \tag{31}$$

where $\alpha_1, \ldots, \alpha_\mu, C$ are constants of integration. (There are $\mu + 1$ constants corresponding to the $\mu + 1$ variables $V, q_i$.) We now introduce $\mu$ new constants $\beta_1, \ldots, \beta_\mu$ and stipulate that the following equations hold:

$$\frac{\partial V}{\partial \alpha_i} = \beta_i. \tag{32}$$

Using (32) we obtain the $q_i$ as a function of $t, \alpha_1, \ldots, \alpha_\mu$ and $\beta_1, \ldots, \beta_\mu$:

$$q_i = q_i\big(t, \alpha_1, \ldots, \alpha_\mu, \beta_1, \ldots, \beta_\mu\big). \tag{33}$$

By (29) the $p_i$ are in turn given as functions of $t, \alpha_1, \ldots, \alpha_\mu$ and $\beta_1, \ldots, \beta_\mu$:

$$p_i = p_i\big(t, \alpha_1, \ldots, \alpha_\mu, \beta_1, \ldots, \beta_\mu\big). \tag{34}$$

Jacobi's integration theorem states that $q_i$ given by (33) and $p_i$ given by (34) yield solutions to Hamilton's equations, which take the form

$$\begin{aligned} \tfrac{\mathrm{d}q_i}{\mathrm{d}t} &= \tfrac{\partial \psi}{\partial p_i}, \\ \tfrac{\mathrm{d}p_i}{\mathrm{d}t} &= -\tfrac{\partial \psi}{\partial q_i}. \end{aligned} \tag{35}$$

By integrating the partial differential Eq. (30) and stipulating that (29) and (33) hold we are led to a solution of (35).

In the proof Jacobi wrote equations without using summation indexes; for purposes of clarity, we supply this notation. We begin by differentiating (32) with respect to $t$:

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial V}{\partial \alpha_i}\right) = \frac{\partial^2 V}{\partial t \partial \alpha_i} + \sum_j \frac{\partial V}{\partial q_j \partial \alpha_i} q_j{}', \tag{36}$$

where $i = 1, \ldots, \mu$. We now differentiate the Hamilton–Jacobi Eq. (30) with respect to $\alpha_i$:

$$\frac{\partial V}{\partial \alpha_i \partial t} + \sum_j \frac{\partial \psi}{\partial p_j} \frac{\partial^2 V}{\partial \alpha_i \partial q_j} = 0. \tag{37}$$

We combine (36) and (37) to obtain

$$0 = \sum_j \frac{\partial^2 V}{\partial q_j \partial \alpha_i}\left(q_j' - \frac{\partial \psi}{\partial p_j}\right), \tag{38}$$

for $i = 1, \ldots, \mu$. It is assumed that the determinant of the coefficient matrix in (37) is non-zero:

$$\mathrm{d}et\left(\frac{\partial^2 V}{\partial q_j \partial \alpha_i}\right) \neq 0. \tag{39}$$

We then conclude from (38) that

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = \frac{\partial \psi}{\partial p_i}(i = 1, \ldots, \mu), \tag{40}$$

and the first of Hamilton's equations holds.

We proceed to differentiate (34) with respect to $t$

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = \frac{\partial^2 V}{\partial t \partial q_i} + \sum_j \frac{\partial^2 V}{\partial q_j \partial q_i} q_j', \tag{41}$$

which becomes from (40),

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = \frac{\partial^2 V}{\partial t \partial q_i} + \sum_j \frac{\partial^2 V}{\partial q_j \partial q_i} \frac{\partial \psi}{\partial p_j}. \tag{42}$$

Differentiating the Hamilton–Jacobi Eq. (30) with respect to $q_i$ gives

$$\frac{\partial^2 V}{\partial q_i \partial t} + \frac{\partial \psi}{\partial q_i} + \sum_j \frac{\partial \psi}{\partial p_j} \frac{\partial^2 V}{\partial q_j \partial q_i} = 0. \tag{43}$$

Combining (42) and (43) we obtain finally

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial \psi}{\partial q_i}, \tag{44}$$

which is the second of Hamilton's equations.

Jacobi supplemented this proof with a discussion of condition (39), required to ensure that one can pass from Eqs. (38) to (40). He developed an argument that if the determinant in (39) were zero, it would be possible to derive a second partial differential equation, a fact that is not consistent with the existing partial differential equation for the problem. Apart from the question of the validity of this argument, it is noteworthy that Jacobi was attentive to such points of rigor in his development of the theory.

In the dynamical problem where $\psi = \sum p_i q_i{}' - \varphi$, Eq. (29) were obtained in the course of the derivation of the Hamilton–Jacobi equation. In the integration theorem, the Hamilton–Jacobi equation is given from the outset. If $\psi$ has the form $\psi = \sum p_i q_i{}' - \varphi$ then one is able to derive (29) along with (32) in the proof of the theorem. This is what Jacobi (1837a) did in his Crelle paper.[18] By contrast, in

[18] In the Crelle article Jacobi (1837a) did not use canonical coordinates and Eqs. (29) are derived in the form.

$$\frac{\partial S}{\partial x_i} = m_i \dot{x}_i, \frac{\partial S}{\partial y_i} = m_i \dot{y}_i, \frac{\partial S}{\partial z_i} = m_i \dot{z}_i (i = 1, \ldots, \mu),$$

where $S$ is the principal function that is denoted by $V$ in the *Vorlesungen*.

the more general mathematical setting of Jacobi's *Vorlesungen* $\psi$ is any function of $t, q_1, q_2, \ldots, q_\mu, p_1, p_2, \ldots, p_\mu$. It is here necessary in the statement of the theorem to posit (29) as a condition that must be satisfied along with (32).

Jacobi turned in the twenty-first lesson to the dynamical case of the integration theorem where the function $\psi$ does not explicitly contain the time. In dynamics, this would involve a conservative system in which the total energy $\psi = H$ is conserved, a problem that occurs commonly. Hamilton's investigation was restricted to this case. In the 1837 memoir in the *Comptes rendus* Jacobi had stated this result for the Newtonian case, but omitted the proof.

$\psi$ is now a function of the $q_i$ $(i = 1, \ldots, \mu)$ alone and the Hamilton–Jacobi Eq. (29) with $p_i = \frac{\partial V}{\partial q_i}$ becomes

$$\frac{\partial V}{\partial t} + \psi\left(q_1, q_2, \ldots, q_\mu, \frac{\partial V}{\partial q_1}, \frac{\partial V}{\partial q_2}, \ldots, \frac{\partial V}{\partial q_\mu}\right) = 0, \tag{45}$$

Let $W$ be a function of $q_1, q_2, \ldots, q_\mu$ and consider a solution of (45) of the form

$$V\big(t, q_1, q_2, \ldots, q_\mu\big) = \alpha t + W\big(q_1, q_2, \ldots, q_\mu\big), \tag{46}$$

where $\alpha$ is a constant. We have $\frac{\partial V}{\partial q_i} = \frac{\partial W}{\partial q_i}$ and Eq. (45) becomes

$$\alpha + \psi\left(q_1, q_2, \ldots, q_\mu, \frac{\partial W}{\partial q_1}, \frac{\partial W}{\partial q_2}, \ldots, \frac{\partial W}{\partial q_\mu}\right) = 0, \tag{47}$$

with solution $W = W(q_1, q_2, \ldots, q_\mu)$. (In dynamics this would occur for a system in which the total energy $\psi = h$ is a constant equal to$-\alpha$.) Consider a complete solution of (45) of the form$V = V\big(t, q_1, \ldots, q_\mu, \alpha_1, \ldots, \alpha_{\mu-1}, \alpha\big)$, where the last arbitrary constant $\alpha_\mu$ is taken to be $\alpha$. Applying Jacobi's integration theorem to $V$, we set$\frac{\partial V}{\partial \alpha_i} = \frac{\partial W}{\partial \alpha_i} = \beta_i$, for $i = 1,\ldots, \mu - 1$. For $\alpha_\mu = \alpha$ we set

$$\frac{\partial V}{\partial \alpha} = \frac{\partial}{\partial \alpha}\big(\alpha t + W\big(q_1, q_2, \ldots, q_\mu\big)\big) = t + \frac{\partial W\big(q_1, q_2, \ldots, q_\mu\big)}{\partial \alpha} = \tau, \tag{48}$$

where $\tau$ is the arbitrary constant $\beta_\mu$ in Jacobi's integration theorem. We have the equations

$$\begin{aligned} \frac{\partial W}{\partial q_i} &= p_i, \, (i = 1, \ldots, \mu) \\ \frac{\partial W}{\partial \alpha_i} &= \beta_i, \, (i = 1, \ldots, \mu - 1) \\ \frac{\partial W}{\partial \alpha} &= \tau - t, \end{aligned} \tag{49}$$

where $W = W\big(q_1, q_2, \ldots, q_\mu, \alpha_1, \ldots, \alpha_{\mu-1}, \alpha\big)$ is a complete solution of (47). Equation (49) allow us to express $q_i$ and $p_i$ as functions of $t$ and the $2n$ arbitrary constants

$\alpha_1, \ldots, \alpha_{\mu-1}, \alpha, \beta_1, \ldots, \beta_{\mu-1}, \tau$ and so obtain a solution of the original canonical Eq. (35).

Augustin Cauchy in 1819 had developed a method of solving a given partial differential equation by reducing this problem to solving a single system of ordinary differential equations. (See Demidov (1982, 334–336) for details.) In the context of dynamical analysis, Jacobi arrived with his integration theorem at the converse of this result. The utility of this theorem is highlighted in the following remarks of Courant and Hilbert (1962, 107), who observe that ordinary differential equations "may be difficult to integrate by elementary methods, while the corresponding partial differential equation is manageable; in particular, it may happen that a complete integral is easily obtained, e.g., with the help of separation of variables … Knowing the complete integral, one can then solve the corresponding system of characteristic ordinary differential equations by processes of differentiation and elimination."

The point here is clearly evident in lecture 24 of the *Vorlesungen*, where the integration theorem is used to analyze the motion of a planet around the sun. The planet's heliocentric coordinates are $x$, $y$, $z$ and its mass is taken to be equal to 1. The force function (potential) is $U = \frac{k^2}{r}$, where $k$ is a constant and $r$ is the distance of the planet from the sun. In lecture 21 Jacobi had derived the Hamilton–Jacobi equation for the problem in the form

$$T = \frac{1}{2}\left\{\left(\frac{\partial W}{\partial x}\right)^2 + \left(\frac{\partial W}{\partial y}\right)^2 + \left(\frac{\partial W}{\partial z}\right)^2\right\} = \frac{k^2}{r} - \alpha, \tag{50}$$

where $T$ is one-half the live force (kinetic energy), $W = W(x,y,z)$ and $\alpha$ is a constant. Jacobi replaced the rectangular coordinates $x$, $y$, $z$ by spherical coordinates $r$, $\phi$, $\psi$:

$$x = r\cos\varphi,\; y = r\sin\varphi\cos\psi,\; z = r\sin\varphi\sin\psi. \tag{51}$$

The expression for $T$ becomes[19]

$$T = \frac{1}{2}\left(x'^2 + y'^2 + z'^2\right) = \frac{1}{2}\left(r'^2 + r^2\varphi'^2 + r^2\sin\varphi^2\psi'^2\right). \tag{52}$$

The conjugate momenta are then

$$\frac{\partial T}{\partial r'} = r',\; \frac{\partial T}{\partial \varphi'} = r^2\varphi',\; \frac{\partial T}{\partial \psi'} = r^2\sin\varphi\psi'. \tag{53}$$

The conjugate momenta are equal to the partial derivatives of $W$ with respect to the variables:

$$\frac{\partial T}{\partial r'} = \frac{\partial W}{\partial r},\; \frac{\partial T}{\partial \varphi'} = \frac{\partial W}{\partial \varphi},\; \frac{\partial T}{\partial \psi'} = \frac{\partial W}{\partial \psi}. \tag{54}$$

[19] Jacobi writes $\sin\phi^2$ for $\sin^2\phi$.

Combining (52), (53) and (54) in $T = \frac{k^2}{r} - \alpha$ we obtain the Hamilton–Jacobi equation in the form

$$\frac{1}{2}\left\{\left(\frac{\partial W}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial W}{\partial \varphi}\right)^2 + \frac{1}{r^2 \sin \varphi^2}\left(\frac{\partial W}{\partial \psi}\right)^2\right\} = \frac{k^2}{r} - \alpha \tag{55}$$

Jacobi integrated (55) by the method of separation of variables. He set down the equations

$$\begin{gathered} \tfrac{1}{2}\left(\tfrac{\partial W}{\partial r}\right)^2 = \tfrac{k^2}{r} - \alpha, \\ \left(\tfrac{\partial W}{\partial \varphi}\right)^2 + \tfrac{1}{\sin \varphi^2}\left(\tfrac{\partial W}{\partial \psi}\right)^2 = 0. \end{gathered} \tag{56}$$

Integrating these equations he arrived at the following solution of (55):

$$W = \int \sqrt{\frac{2k^2}{r} - 2\alpha - \frac{2\beta}{r^2}}\mathrm{d}r + \int \sqrt{2\beta - \frac{2\gamma}{\sin \varphi^2}}\mathrm{d}\varphi + \sqrt{2\gamma}, \tag{57}$$

where $\alpha, \beta, \gamma$ are arbitrary constants. By Jacobi's integration theorem one obtains an integral of the equations of motion from

$$\frac{\partial W}{\partial \alpha} = \alpha' - t,\ \frac{\partial W}{\partial \beta} = \beta',\ \frac{\partial W}{\partial \gamma} = \gamma', \tag{58}$$

where $\alpha', \beta', \gamma'$ are new arbitrary constants. From (57) and (58) we obtain finally

$$\begin{aligned} t - \alpha' &= \int \frac{\mathrm{d}r}{\sqrt{\frac{2k^2}{r} - 2\alpha - \frac{2\beta}{r^2}}}, \\ \beta' &= -\int \frac{\mathrm{d}r}{r^2\sqrt{\frac{2k^2}{r} - 2\alpha - \frac{2\beta}{r^2}}} + \int \frac{\mathrm{d}\varphi}{\sqrt{2\beta - \frac{2\gamma}{\sin \varphi^2}}}, \\ \gamma' &= -\int \frac{\mathrm{d}\varphi}{\sin \varphi^2 \sqrt{2\beta - \frac{2\gamma}{\sin \varphi^2}}} + \frac{1}{\sqrt{2\gamma}}\psi. \end{aligned} \tag{59}$$

The remainder of the lecture is devoted to a discussion of the geometric significance of the constants that appear in (59).

The problem and solution presented in lecture 24 would appear in Tisserand's (1868) doctoral dissertation and in his 1889 treatise on celestial mechanics. Poincaré (1899, Chapter 1 Sects. 8 and 10) would make use of this solution in his investigation of perturbed Keplerian motion, where $W$ was employed as a generating function for a transformation that led from the canonical equations of motion to canonical equations for the variable constants arising in the perturbed system. (See our discussion above in Sect. 5.3 (Tisserand) and Sect. 6.2.2 (Poincaré).)

### 3.7 Canonical elements and transformations

We turn now to the proofs in Jacobi's (1866b) *Ueber diejenigen Probleme* of two theorems stated without proof in his Paris note of 1837. In 1837 Jacobi introduced the equations of motion in traditional Newtonian form; in the 1866 treatise (written sometime around 1840) both Newtonian and canonical formalism are used.[20]

#### 3.7.1 Jacobi's theorem on canonical elements

In Sect. 27 of *Ueber diejenigen Probleme* Jacobi (1866b, 416–419) proved a more general form of Lagrange's fundamental identity (5) in the wider setting of dynamical theory involving $m$ generalized coordinates $q_i$ and $m$ conjugate momenta $p_i$. The undisturbed motion of the system is described by Hamilton's equations:

$$\begin{aligned}\frac{dq_i}{dt} &= \frac{\partial H}{\partial p_i},\\ \frac{dp_i}{dt} &= -\frac{\partial H}{\partial q_i}.\end{aligned} \tag{60}$$

Let $q_i = q_i(t, \alpha_1, \ldots, \alpha_{2m})$ be a complete solution of (60) containing the $2m$ arbitrary constants $\alpha_j$. We let $\alpha_j$ be variable functions of time, $\alpha_j = \alpha_j(t)$, and suppose that the resulting $q_i = q_i(t, \alpha_1(t), \ldots, \alpha_{2m}(t))$ and $p_i = p_i(t, \alpha_1(t), \ldots, \alpha_{2m}(t))$ are a solution of the perturbed system

$$\begin{aligned}\frac{dq_i}{dt} &= \frac{\partial H}{\partial p_i} + \frac{\partial H_1}{\partial p_i},\\ \frac{dp_i}{dt} &= -\frac{\partial H}{\partial q_i} - \frac{\partial H_1}{\partial q_i}.\end{aligned} \tag{61}$$

Here $H_1$ is a term that is added to $H$ to represent the perturbation and is a function of $t$, $q_i$ and $p_i$.. (Note that in Lagrange's analysis the perturbation function $\Omega$ was a function of the $q_i$ alone.) Differentiating the coordinates $q_i$ with respect to time we have

$$\begin{aligned}\frac{dq_i}{dt} &= \frac{\partial q_i}{\partial t} + \sum_{k=1}^{2m} \frac{\partial q_i}{\partial \alpha_k}\frac{d\alpha_k}{dt},\\ \frac{dp_i}{dt} &= \frac{\partial p_i}{\partial t} + \sum_{k=1}^{2m} \frac{\partial p_i}{\partial \alpha_k}\frac{d\alpha_k}{dt}.\end{aligned} \tag{62}$$

Following the method of variation of constants, we assume that Hamilton's original Eq. (60) continue to hold for the parts of the time derivatives in (62) containing the partial time derivatives:

[20] The term "Newtonian" here refers in the usual way to force equations given in terms of second derivatives of the coordinate variables.

$$\begin{aligned}\frac{\partial q_i}{\partial t} &= \frac{\partial H}{\partial p_i},\\ \frac{\partial p_i}{\partial t} &= -\frac{\partial H}{\partial q_i}.\end{aligned} \tag{63}$$

Combining (61), (62) and (63) we find that

$$\begin{aligned}\sum_{k=1}^{2m} \frac{\partial q_i}{\partial \alpha_k}\frac{\mathrm{d}\alpha_k}{\mathrm{d}t} &= \frac{\partial H_1}{\partial p_i},\\ \sum_{k=1}^{2m} \frac{\partial p_i}{\partial \alpha_k}\frac{\mathrm{d}\alpha_k}{\mathrm{d}t} &= -\frac{\partial H_1}{\partial q_i}.\end{aligned} \tag{64}$$

We have as well the equations

$$\frac{\partial H_1}{\partial \alpha_i} = \sum_{j=1}^{m}\left(\frac{\partial H_1}{\partial q_j}\frac{\partial q_j}{\partial \alpha_i} + \frac{\partial H_1}{\partial p_j}\frac{\partial p_j}{\partial \alpha_i}\right). \tag{65}$$

Substituting (64) into (65) leads to

$$\frac{\partial H_1}{\partial \alpha_r} = \sum_{j=1}^{m}\left(-\left(\sum_{k=1}^{2m}\frac{\partial p_j}{\partial \alpha_k}\frac{\mathrm{d}\alpha_k}{\mathrm{d}t}\right)\frac{\partial q_j}{\partial \alpha_r} + \sum_{k=1}^{2m}\left(\frac{\partial q_j}{\partial \alpha_k}\frac{\mathrm{d}\alpha_k}{\mathrm{d}t}\right)\frac{\partial p_j}{\partial \alpha_r}\right),$$

which simplifies to

$$\frac{\partial H_1}{\partial a_r} = \sum_{k=1}^{2m}\sum_{j=1}^{m}\left(\frac{\partial q_j}{\partial \alpha_k}\frac{\partial p_j}{\partial \alpha_r} - \frac{\partial q_j}{\partial \alpha_r}\frac{\partial p_j}{\partial \alpha_k}\right)\frac{\mathrm{d}\alpha_k}{\mathrm{d}t}. \tag{66}$$

We rewrite (66) as

$$\frac{\partial H_1}{\partial a_r} = \sum_{k=1}^{2m}(\alpha_k, \alpha_r)\frac{\mathrm{d}\alpha_k}{\mathrm{d}t},\ (r = 1, \ldots, 2m) \tag{67}$$

where the Lagrange bracket (which Jacobi writes with curved brackets[21]) is

$$(\alpha_k, \alpha_r) = \sum_{j=1}^{m}\left(\frac{\partial q_j}{\partial \alpha_k}\frac{\partial p_j}{\partial \alpha_r} - \frac{\partial q_j}{\partial \alpha_r}\frac{\partial p_j}{\partial \alpha_k}\right)(k, r = 1, \ldots, 2m) \tag{68}$$

[21] There is a fair degree of variation in the notation adopted by researchers for Lagrange and Poisson brackets. The appendix at the end of this article is a table giving the notation used by each researcher.

Equation (67) is Jacobi's statement of Lagrange's fundamental identity (5). We have provided more detail in the derivation than Jacobi did, who went without comment directly from (63) to (66).

In Theorem IX Jacobi (1866b, 432–436) proceeded to prove the theorem on canonical elements. We have as before Hamilton's equation (60). The Hamilton–Jacobi equation is set down in the form

$$\frac{\partial W}{\partial t} + f\left(t, q_1, q_2, \ldots, q_m, \frac{\partial W}{\partial q_1}, \frac{\partial W}{\partial q_2}, \ldots, \frac{\partial W}{\partial q_m}\right) = 0, \tag{69}$$

where $W = W(t, q_1, q_2, \ldots, q_m)$. Let $W = W(t, q_1, q_2, \ldots, q_m, \alpha_1, \alpha_2, \ldots, \alpha_m)$ be a complete solution of (69) containing the $m$ arbitrary constants $\alpha_1, \ldots, \alpha_m$. By Jacobi's integration theorem a solution to Hamilton's equation (60) is given by the $2m$ equations

$$\begin{aligned} \frac{\partial W}{\partial \alpha_i} &= \beta_i, \\ \frac{\partial W}{\partial q_i} &= p_i, \end{aligned} \tag{70}$$

where $\beta_1, \ldots, \beta_m$ are a second set of arbitrary constants. From (70) we obtain $q_i$ and $p_i$ in the form:

$$\begin{aligned} q_i &= q_i(t, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_m), \\ p_i &= p_i(t, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_m). \end{aligned} \tag{71}$$

Assume now that there is a perturbing force that gives rise to an increment $H_1$ added to the Hamiltonian $H$. In the new problem, we consider solutions to Hamilton's equations of the form (71) but suppose the arbitrary constants $\alpha_i$ and $\beta_i$ are functions of $t$, so that the coordinates in the perturbed motion are

$$\begin{aligned} q_i &= q_i(t, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_m), \\ p_i &= p_i(t, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_m). \end{aligned} \tag{72}$$

Theorem IX asserts that the variable elements $\alpha_i$ and $\beta_i$ satisfy the canonical equations:

$$\begin{aligned} \frac{d\alpha_i}{dt} &= -\frac{\partial H_1}{\partial \beta_i}, \\ \frac{d\beta_i}{dt} &= \frac{\partial H_1}{\partial \alpha_i}. \end{aligned} \tag{73}$$

This is the content of Jacobi's theorem on canonical elements.

The proof begins by expressing the function $W$ in the form

$$\begin{aligned} W = W\big(t, q_1(t, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_m), \ldots, \\ q_m(t, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_m), \alpha_1, \ldots, \alpha_m\big) \end{aligned} \tag{74}$$

Jacobi used parentheses to designate the total partial derivative of a function. For example, if $f = f(q, \alpha)$ where $q$ is a function of $t$ and $\alpha$ Jacobi would write

$$\left(\frac{\partial f}{\partial \alpha}\right) = \frac{\partial f}{\partial q}\frac{\partial q}{\partial \alpha} + \frac{\partial f}{\partial \alpha}.$$

Taking the total partial derivatives of (74) with respect to $\alpha_k$ and $\beta_k$ $(k = 1, \ldots, m)$ we obtain:

$$\left(\frac{\partial W}{\partial \alpha_k}\right) = \sum_{i=1}^{m} \frac{\partial W}{\partial q_i}\frac{\partial q_i}{\partial \alpha_k} + \frac{\partial W}{\partial \alpha_k}, \left(\frac{\partial W}{\partial \beta_k}\right) = \sum_{i=1}^{m} \frac{\partial W}{\partial q_i}\frac{\partial q_i}{\partial \beta_k}. \tag{75}$$

By (70) these equations become

$$\begin{aligned} \left(\tfrac{\partial W}{\partial \alpha_k}\right) &= \sum_{i=1}^{m} p_i \tfrac{\partial q_i}{\partial \alpha_k} + \beta_k (k = 1, \ldots, m), \\ \left(\tfrac{\partial W}{\partial \beta_k}\right) &= \sum_{i=1}^{m} p_i \tfrac{\partial q_i}{\partial \beta_k} (k = 1, \ldots, m). \end{aligned} \tag{76}$$

For the symbols $\alpha$ and $\beta$ the Lagrange bracket $(\alpha, \beta)$ is

$$(\alpha, \beta) = \sum_{i=1}^{m} \left( \frac{\partial q_i}{\partial \alpha}\frac{\partial p_i}{\partial \beta} - \frac{\partial q_i}{\partial \beta}\frac{\partial p_i}{\partial \alpha} \right). \tag{77}$$

In what follows we will make use of the following identities:

$$\begin{aligned} \tfrac{\partial^2 q_i}{\partial \alpha \partial \beta} &= \tfrac{\partial^2 q_i}{\partial \beta \partial \alpha} (i = 1, \ldots, m), \\ \tfrac{\partial^2 (W)}{\partial \alpha \partial \beta} &= \tfrac{\partial^2 (W)}{\partial \beta \partial \alpha}. \end{aligned} \tag{78}$$

By means of the first identity, Eq. (77) may be written

$$(\alpha, \beta) = \frac{\partial \left( \sum_{i=1}^{m} p_i \frac{\partial q_i}{\partial \alpha} \right)}{\partial \beta} - \frac{\partial \left( \sum_{i=1}^{m} p_i \frac{\partial q_i}{\partial \beta} \right)}{\partial \alpha}. \tag{79}$$

We now consider the value of $(\alpha, \beta)$ for selected choices of $\alpha$ and $\beta$. If $(\alpha, \beta) = (\alpha_k, \alpha_s)(s = 1, \ldots, m)$ then from (76), (79) and the second identity in (78) we obtain

$$(\alpha_k, \alpha_s) = \frac{\partial \left( \left( \frac{\partial W}{\partial \alpha_k} \right) - \beta_k \right)}{\partial \alpha_s} - \frac{\partial \left( \left( \frac{\partial W}{\partial \alpha_s} \right) - \beta_s \right)}{\partial \alpha_k} = 0 \tag{80}$$

for all $k$ and $s$. Similarly for all $k$ and $s$ we have

$$(\beta_k, \beta_s) = \frac{\partial\left(\left(\frac{\partial W}{\partial \beta_k}\right)\right)}{\partial \beta_s} - \frac{\partial\left(\left(\frac{\partial W}{\partial \beta_s}\right)\right)}{\partial \beta_k} = 0. \tag{81}$$

Finally, consider the case $(, \beta) = (\alpha_k, \beta_s)$. We have

$$(\alpha_k, \beta_s) = \frac{\partial\left(\left(\frac{\partial W}{\partial \alpha_k}\right) - \beta_k\right)}{\partial \beta_s} - \frac{\partial\left(\left(\frac{\partial W}{\partial \beta_s}\right)\right)}{\partial \alpha_k}. \tag{82}$$

Hence it follows that $(\alpha_k, \beta_s) = 0$ if $k \neq s$ and $(\alpha_k, \beta_k) = -1$ if $k = s$.

We now take the generalized Lagrange identity (67) which in the present derivation is presented by Jacobi in the form

$$\frac{\partial H_1}{\partial \beta} = \sum (\alpha, \beta) \frac{\mathrm{d}\alpha}{\mathrm{d}t}, \tag{83}$$

where $\beta$ denotes any of the variables $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_m$ and the summation variable $\alpha$ ranges over these $2m$ variables. We find that for $\beta = \beta_i$ (83) becomes $\frac{\partial H_1}{\partial \beta_i} = -\frac{\mathrm{d}\alpha_i}{\mathrm{d}t}$. For $\beta = \alpha_i$ (83) becomes $\frac{\partial H_1}{\partial \alpha_i} = \frac{\mathrm{d}\beta_i}{\mathrm{d}t}$. We have therefore arrived at the desired canonical Eq. (73) for the variable elements $\alpha_i$ and $\beta_i$ and the proof is complete.

It should be noted that Jacobi's identity (83) is more general than Lagrange's (5). Equation (83) holds for perturbations functions involving $t$ and the $q_i$ and $p_i$, whereas Lagrange's account was restricted to perturbation functions involving $q_i$ alone. Theorem IX is also more general than the corresponding Eq. (12) obtained by Poisson and Lagrange. The arbitrary constants in Jacobi's account can be any constants and are not restricted to the initial values of the $q_i$ and $p_i$.

Jacobi's theorem on canonical elements was also proved by Desboves (1848) and Donkin (1854), as we shall see in the next section. It would be successfully deployed in research in celestial mechanics in the second half of the century. Tisserand (1868) and Charlier (1907) used the result as the basis for their investigations of the canonical differential equations satisfied by the elements in various three-body systems (see below Sects. 5.3 and 7.2.2). The theorem also occupies a prominent place in modern literature on celestial mechanics (see for example, Vinti (1998, Chapter 7)).

### 3.7.2 Transformations and canonical equations

Jacobi (1866b, 446–470) followed Theorem IX with an investigation of changes in variables from $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_m$ to a new set of variables $\alpha_1{}', \ldots, \alpha_m{}', \beta_1{}', \ldots, \beta_m{}'$. He defined a change of variables in terms of an arbitrary function $\psi$ of $\alpha_1, \ldots, \alpha_m$ and $\alpha_1{}', \ldots, \alpha_m{}'$:

$$\psi = \psi\left(\alpha_1, \ldots, \alpha_m, \alpha_1{}', \ldots, \alpha_m{}'\right). \tag{84}$$

$\psi$ is known as a generating function in the modern subject.[22] The old and new variables are related by the equations:

$$\frac{\partial \psi}{\partial \alpha_i} = -\beta_i, \qquad \frac{\partial \psi}{\partial \alpha_i{}'} = \beta_i'. \tag{85}$$

These relations enable one to express the old variables $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_m$ in terms of the new variables $\alpha_1{}', \ldots, \alpha_m{}', \beta_1{}', \ldots, \beta_m{}'$: $\alpha_i = \alpha_i\left(\alpha_1{}', \ldots, \alpha_m{}', \beta_1{}', \ldots, \beta_m{}'\right)$ and $\beta_i = \beta_i(\alpha_1{}', \ldots, \alpha_m{}', \beta_1{}', \ldots, \beta_m{}'$. $H_1$ in turn then becomes a function of $t, \alpha_1{}', \ldots, \alpha_m{}', \beta_1{}', \ldots, \beta_m{}'$. In Theorem X Jacobi (1866b, 371–373) showed that the original canonical Eq. (73), expressed in terms of the new variables $\alpha_i{}'$ and $\beta_i{}'$ are also canonical in these variables:

$$\frac{\mathrm{d}\alpha_i{}'}{\mathrm{d}t} = -\frac{\partial H_1}{\partial \beta_i{}'} \quad (i = 1, \ldots, m) \qquad \frac{\mathrm{d}\beta_i{}'}{\mathrm{d}t} = \frac{\partial H_1}{\partial \alpha_i{}'}. \tag{86}$$

Jacobi had originally stated this theorem in his 1837 Paris note where he referred to it as "entièrement nouvelle" (1837b, 66).

A change of variables with the above property is said in the modern subject to be a canonical transformation because it preserves the canonical form of the equations. The canonical form is a fundamental invariant of the change of variables. Jacobi's Theorem X is an important result. Reformulated in a general dynamical setting for coordinate variables it is the cornerstone of the whole theory of transformations in modern dynamical theory.

Jacobi (1866b, 448) noted that Theorem X is logically independent of Theorem IX, and would be true for any function $H_1$ of $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_m$ whatever might be its origin:

> Although Theorem IX, in conjunction with the known method of deriving infinitely many other solutions from a complete solution, leads to the above theorem, the latter is by its nature independent of all previous conditions and can itself be proved directly …

This passage is of interest in indicating that although Theorem X may be proved independently of Theorem IX, the idea of a canonical transformation and the associated method of generating functions may in point of fact have been originally suggested to Jacobi by Theorem IX. We begin with the variables $q_1, \ldots, q_m, p_1, \ldots, p_m$ that satisfy canonical equations with respect to the Hamiltonian $H$. Taking a complete

[22] The term "generating function" had become standard by the 1960s. The following are the designations for generating function adopted by the researches considered in this study: Jacobi (1866a) *wilkürliche Function*; Desboves (1848) *function quelconque*; Donkin (1855) *modulus of a normal transformation*; Poincaré (1892) *function quelconque*; Charlier (1907) *Transformationsfunction*; Klein (1926) *Leitfunktion*; Born (1925); Nordheim and Fues (1927) and Carathéodory (1935), *Erzeugende Funktion der Transformation*; Goldstein (1950) *generating function*; Corben and Stehle (1950) *generator*; Gelfand and Fomin (1963) *generating function*.

solution $W = W(t, q_1, \ldots, q_m, \alpha_1, \ldots, \alpha_m)$ of the Hamilton–Jacobi partial differential Eq. (69) we consider the quantities $\alpha_i$ and $\beta_i$ that appear in (70). According to Jacobi's integration theorem (70) give rise to a solution to the canonical equations with respect to $H$. Theorem IX asserts that $\alpha_i$ and $\beta_i$ themselves will in turn satisfy canonical equations with respect to the perturbed Hamiltonian $H_1$. Here $W$ may be regarded as a generating function. Of course, in Theorem IX there are two different Hamiltonians $H$ and $H_1$ whereas in Theorem X there is one and the same Hamiltonian $H_1$. Furthermore, Jacobi did not pursue this line of reasoning, and the derivation that he did give for Theorem X is as he noted independent of the one for Theorem IX.[23] The statement and proof of Theorem X do not refer at all to partial differential equations and their solutions and are logically independent of both Jacobi's integration theorem and his theorem on canonical elements.[24]

In his original statement of this result in the 1837 Paris note, Jacobi (1837b, 67) gave no proof, asserting that "The demonstrations of these theorems offer no difficulties." Nevertheless, in *Ueber diejenigen Probleme* he developed a full derivation of Theorem X that certainly was not simple. His proof is different from the one found in many modern textbooks, the latter having originated in writings of Poincaré (see Sect. 6.4.1 above). Jacobi's proof of Theorem X unfolds as a result of differential equations and without any reference to variational concepts or principles. We now give this proof.

Jacobi began by taking the partial derivative of $H_1\big(t, \alpha_1\big(\alpha_1{}', \ldots, \alpha_m{}', \beta_1{}', \ldots, \beta_m{}'\big), \ldots, \alpha_m\big(\alpha_1{}', \ldots, \alpha_m{}', \beta_1{}', \ldots, \beta_m{}'\big), \beta_1\big(\alpha_1{}', \ldots, \alpha_m{}', \beta_1{}', \ldots, \beta_m{}'\big), \ldots, \beta_m\big(\alpha_1{}', \ldots, \alpha_m{}', \beta_1{}', \ldots, \beta_m{}'\big)\big)$ with respect to $\beta_k{}'$:

$$\frac{\partial H_1}{\partial \beta_k'} = \sum_i \frac{\partial H_1}{\partial \alpha_i}\frac{\partial \alpha_i}{\partial \beta_k'} + \sum_i \frac{\partial H_1}{\partial \beta_i}\frac{\partial \beta_i}{\partial \beta_k'}. \tag{87}$$

From the original canonical Eq. (73), (87) becomes

$$\frac{\partial H_1}{\partial \beta_k'} = \sum_i \frac{\partial \alpha_i}{\partial \beta_k'}\frac{\mathrm{d}\beta_i}{\mathrm{d}t} - \sum_i \frac{\partial \beta_i}{\partial \beta_k'}\frac{\mathrm{d}\alpha_i}{\mathrm{d}t}. \tag{88}$$

Jacobi did not indicate in the summation symbol the index with respect to which the summation was taken. For clarity of exposition, we have added these indices. He also added a new index symbol that he denoted by "$i'$". Because of possible confusion between a variable with a primed index and a primed variable itself, we will for the

[23] The quoted passage is open to other objections. Jacobi developed a proof for Theorem X that was independent of Theorem IX because in point of fact Theorem IX does not imply Theorem X. In Sect. 42 he showed that for conservative systems Theorem X implies Theorem IX but the converse does not seem to be true.

[24] Felix Klein (1926, 203 and 1969, 191-192) seems to suggest that the transformation theorem (Theorem X of *Ueber diejenigen Probleme*) was deduced by Jacobi from his integration theorem (Theorem VI of *Ueber diejenigen Probleme*). This was not the case, and Jacobi himself regarded the transformation theorem as a new result. Furthermore, whereas the integration theorem was presented by Jacobi as a result in general dynamical theory, the transformation theorem was ostensibly introduced within the theory of variation of arbitrary constants. (Klein is intent on asserting Hamilton's importance (long overlooked in his view) and seems even to attribute the idea of a canonical transformation to the Irishman. On Klein's "discovery" of Hamilton see Hankins (1980, 203–204).).

sake of clarity transcribe "$i'$"as "$s$". Taking the partial derivative of the transformation equation $\frac{\partial\psi}{\partial\alpha_i} = -\beta_i$ (85) with respect to $\beta_k{}'$ we obtain:

$$-\frac{\partial\beta_i}{\partial\beta_k'} = \sum_s \frac{\partial^2\psi}{\partial\alpha_i\partial\alpha_s}\frac{\partial\alpha_s}{\partial\beta_k'}. \tag{89}$$

Combining (88) and (89) gives

$$\frac{\partial H_1}{\partial\beta_k'} = \sum_s\left(\frac{\partial\alpha_s}{\partial\beta_k'}\left(\sum_i \frac{\partial^2\psi}{\partial\alpha_i\partial\alpha_s}\frac{\mathrm{d}\alpha_i}{\mathrm{d}t} + \frac{\mathrm{d}\beta_s}{\mathrm{d}t}\right)\right). \tag{90}$$

We now differentiate (85) with respect to time:

$$\sum_i\left(\frac{\partial^2\psi}{\partial\alpha_i\partial\alpha_s}\frac{\mathrm{d}\alpha_i}{\mathrm{d}t} + \frac{\mathrm{d}\beta_s}{\mathrm{d}t}\right) = -\sum_i \frac{\partial^2\psi}{\partial\alpha_i'\partial\alpha_s}\frac{\mathrm{d}\alpha_i'}{\mathrm{d}t}. \tag{91}$$

Combining (90) and (91) there follows

$$\frac{\partial H_1}{\partial\beta_k'} = -\sum_i\sum_s \frac{\partial^2\psi}{\partial\alpha_i'\partial\alpha_s}\frac{\partial\alpha_s}{\partial\beta_k'}\frac{\mathrm{d}\alpha_i'}{\mathrm{d}t}. \tag{92}$$

The final step is to take $\frac{\partial\psi}{\partial\alpha_i{}'} = \beta_i{}'$ (3.68b) and differentiate the left and right sides partially with respect to $\beta_k{}'$

$$\sum_s \frac{\partial^2\psi}{\partial\alpha_s\partial\alpha_i'}\frac{\partial\alpha_s}{\partial\beta_k'} = \begin{cases} 0 \text{ if } i \neq k \\ 1 \text{ if } i = k \end{cases}. \tag{93}$$

With (92), (93) simplifies to

$$\frac{\mathrm{d}\alpha_k'}{\mathrm{d}t} = -\frac{\partial H_1}{\partial\beta_k{}'}, \tag{94}$$

which is the first canonical equation for the variables $\alpha_k{}'$ and $\beta_k{}'$.

We turn now to the derivation of Eq. (86). Because of the particular form of the generating function, the derivation differs somewhat from the one for (3.69a). Jacobi gave only an abridged derivation; here we supply the missing steps. The first step is to take the partial derivative of $H_1$ with respect to $\alpha_k{}'$:

$$\frac{\partial H_1}{\partial\alpha_k'} = \sum_s \frac{\partial H_1}{\partial\alpha_s}\frac{\partial\alpha_s}{\partial\alpha_k'} + \sum_s \frac{\partial H_1}{\partial\beta_s}\frac{\partial\beta_s}{\partial\alpha_k'}. \tag{95}$$

From the canonical Eqs. (73), (95) becomes

$$\frac{\partial H_1}{\partial \alpha'_k} = \sum_s \frac{\partial \alpha_s}{\partial \alpha'_k} \frac{\mathrm{d}\beta_s}{\mathrm{d}t} - \sum_s \frac{\partial \beta_s}{\partial \alpha'_k} \frac{\mathrm{d}\alpha_s}{\mathrm{d}t}. \tag{96}$$

It is necessary to find expressions for the coefficients in (96). Equation (91) gives

$$\frac{\mathrm{d}\beta_s}{\mathrm{d}t} = -\sum_i \frac{\partial^2 \psi}{\partial \alpha_i \partial \alpha_s} \frac{\mathrm{d}\alpha_i}{\mathrm{d}t} - \sum_i \frac{\partial^2 \psi}{\partial \alpha'_i \partial \alpha_s} \frac{\mathrm{d}\alpha'_i}{\mathrm{d}t}. \tag{97}$$

Taking the partial derivative of $\beta_s = \frac{\partial \psi}{\partial \alpha_s}$ (85) with respect to $\alpha_k{}'$ we have

$$\frac{\partial \beta_s}{\partial \alpha'_k} = \sum_i \left( \frac{\partial^2 \psi}{\partial \alpha_i \partial \alpha_s} \frac{\partial \alpha_i}{\partial \alpha'_k} + \frac{\partial^2 \psi}{\partial \alpha'_k \partial \alpha_s} \right). \tag{98}$$

Using (97) and (96), (98) becomes

$$\begin{aligned} \frac{\partial H_1}{\partial \alpha'_k} = &-\sum_s \frac{\partial \alpha_s}{\partial \alpha'_k} \left( \sum_i \frac{\partial^2 \psi}{\partial \alpha_i \partial \alpha_s} \frac{\mathrm{d}\alpha_i}{\mathrm{d}t} + \sum_i \frac{\partial^2 \psi}{\partial \alpha'_i \partial \alpha_s} \frac{\mathrm{d}\alpha'_i}{\mathrm{d}t} \right) \\ &+ \sum_s \sum_i \left( \frac{\partial^2 \psi}{\partial \alpha_i \partial \alpha_s} \frac{\partial \alpha_i}{\partial \alpha'_k} + \frac{\partial^2 \psi}{\partial \alpha'_k \partial \alpha_s} \right) \frac{\mathrm{d}\alpha_s}{\mathrm{d}t}. \end{aligned} \tag{99}$$

We have the identity

$$-\sum_s \sum_i \frac{\partial \alpha_s}{\partial \alpha'_k} \frac{\partial^2 \psi}{\partial \alpha_i \partial \alpha_s} \frac{\mathrm{d}\alpha_i}{\mathrm{d}t} + \sum_s \sum_i \frac{\partial^2 \psi}{\partial \alpha_i \partial \alpha_s} \frac{\partial \alpha_i}{\partial \alpha'_k} \frac{\mathrm{d}\alpha_s}{\mathrm{d}t} = 0. \tag{100}$$

Hence (99) reduces to

$$\frac{\partial H_1}{\partial \alpha'_k} = -\sum_s \sum_i \frac{\partial^2 \psi}{\partial \alpha'_i \partial \alpha_s} \frac{\partial \alpha_s}{\partial \alpha'_k} \frac{\mathrm{d}\alpha'_i}{\mathrm{d}t} + \sum_s \frac{\partial^2 \psi}{\partial \alpha'_k \partial \alpha_s} \frac{\mathrm{d}\alpha_s}{\mathrm{d}t}. \tag{101}$$

It should be noted that Jacobi began his derivation of the second canonical equation with (101); it was left to the reader to reconstruct the steps from (95) to (101).

We proceed to take the partial derivative of $\frac{\partial \psi}{\partial \alpha_i{}'} = \beta_i{}'$ (3.68b) with respect to $\alpha_k{}'$:

$$\frac{\partial \beta'_i}{\partial \alpha'_k} = \sum_s \left( \frac{\partial^2 \psi}{\partial \alpha_s \partial \alpha'_i} \frac{\partial \alpha_s}{\partial \alpha'_k} + \frac{\partial^2 \psi}{\partial \alpha'_i \partial \alpha'_k} \right) = 0. \tag{102}$$

With (102), (101) simplifies to

$$\frac{\partial H_1}{\partial \alpha'_k} = \sum_i \left( \frac{\partial^2 \psi}{\partial \alpha'_i \partial \alpha'_k} \frac{\mathrm{d}\alpha'_i}{\mathrm{d}t} + \frac{\partial^2 \psi}{\partial \alpha_i \partial \alpha'_k} \frac{\mathrm{d}\alpha_i}{\mathrm{d}t} \right). \tag{103}$$

Taking the time derivative of $\frac{\partial \psi}{\partial \alpha_k{}'} = \beta_k{}'$ (3.68b) we see that the right side of (103) is the derivative of $\beta_k{}'$ with respect to $t$. Hence, we have finally

$$\frac{\mathrm{d}\beta_k'}{\mathrm{d}t} = \frac{\partial H_1}{\partial \alpha_k{}'}, \tag{104}$$

which is the second canonical equation for the variables $\alpha_k{}'$ and $\beta_k{}'$.

In the next two theorems, Jacobi extended this result to cases where side conditions hold among the variable elements. For example, suppose that $\psi_1$ is a function of $\alpha_1, \alpha_2, \ldots, \alpha_m, \alpha_1{}', \alpha_2{}', \ldots, \alpha_m{}'$ and the condition

$$\psi_1 = 0 \tag{105}$$

is assumed to hold. Then we introduce the multiplier function λ and stipulate that the following relations analogous to (85) hold between $\alpha_i$,$\beta_i$, $\alpha_i{}'$,$\beta_i{}'$:

$$\begin{aligned} \beta_i + \frac{\partial \psi}{\partial \alpha_i} + \lambda \frac{\partial \psi_1}{\partial \alpha_i} &= 0, \\ -\beta_i' + \frac{\partial \psi}{\partial \alpha_i'} + \lambda \frac{\partial \psi_1}{\partial \alpha_i'} &= 0. \end{aligned} \tag{106}$$

Using (105) and (106) we can eliminate λ and obtain a transformation from$\alpha_i$, $\beta_i$ to $\alpha_i{}'$,$\beta_i{}'$. Theorem XI asserts that the variables $\alpha_i{}'$, $\beta_i{}'$ satisfy Eq. (3.69) and the transformation, therefore, preserves the canonical form of the equations.

Jacobi's proof of the transformation theorem is evidently a rather involved construction, one that may be daunting for a modern reader to follow. Nevertheless, the basic ideas in the proof are clear, and much of the complexity in the proof arises from notational issues associated with handling the summation processes that appear in a general formulation of the result. If the proof is carried out for the case of a single coordinate variable and a single conjugate variable the derivation is simplified and the basic character of the proof is clear. It should also be noted that the transformation theorem is a result of the theory of differential equations and Jacobi's proof is a natural one. The popularity of Poincare's proof using variational theory arises from its simplicity in the general case and the fact that it connects—in a characteristically modern way—ideas from different parts of analysis.

### 3.7.3 A theory based on transformations

In a discussion in Sect. 41 of Theorems X–XII, Jacobi (1866b, 464) commented on the character and scope of these results. He wrote:

Theorems X–XII are completely independent of the meaning of the quantities $\alpha_1, \alpha_2 \ldots \alpha_m, \beta_1, \beta_2 \ldots, \beta_m$ as constant elements in a mechanical problem, for the theorems are removed from everything that relates to such a problem. It follows that Theorems X–XII provide the most general way of transforming a system of differential equations in canonical form by the introduction of other variables into another system with the same form.

In the discussion that followed, he gave an example of such a transformation using a generating function from the variables $q_1, q_2, \ldots, q_m, p_1, p_2, \ldots p_m$ to the variables $\alpha_1, \alpha_2, \ldots \alpha_m, \beta_1, \beta_2, \ldots \beta_m$. In these final sections of the treatise, Jacobi highlighted the fundamental character of Theorem X and even began to develop the outline of a theory of transformations. His language shifted and he began to use the verb "to transform" and the noun "transformation" in his account. Whereas in the statement of Theorem X he wrote of one set of variables determining the other, he now wrote of a transformation from one set to the other.

Jacobi's theoretical predilections involving transformations are apparent in the second part of Sect. 41, where Theorem X is used to prove the integration theorem and the theorem on canonical elements. We will focus our account on his treatment of the first of these results. The idea is to take a complete solution of the Hamilton–Jacobi equation and use it a generating function for a canonical transformation. Because Theorem X had been derived for time-independent generating functions, Jacobi restricted his analysis to conservative systems in which the partial differential equation may be reduced to one for which a complete solution does not involve time. We begin with the standard canonical equations

$$\begin{aligned} \frac{\mathrm{d}q_i}{\mathrm{d}t} &= \frac{\partial H}{\partial p_i}, \\ \frac{\mathrm{d}p_i}{\mathrm{d}t} &= -\frac{\partial H}{\partial q_i}, \end{aligned} \quad (i = 1, 2, \ldots, m) \tag{107}$$

where the Hamiltonian $H$ is a function of the $q_1, q_2, \ldots, q_m, p_1, p_2, \ldots p_m$. Let $\psi$ be a function of $q_1, q_2, \ldots, q_m, p_1, p_2, \ldots p_m$. Replace $p_i$ by $-\frac{\partial \psi}{\partial q_i}$ $(i = 1, 2, \ldots, m)$ in $H$. The Hamilton–Jacobi partial differential equation is

$$H\left(t, q_1, q_2, \ldots, q_m, -\frac{\partial \psi}{\partial q_1}, -\frac{\partial \psi}{\partial q_2}, \ldots, \frac{\partial \psi}{\partial q_m}\right) = h, \tag{108}$$

where $h$ is a constant (the total energy). Let $\psi$ be a complete solution of (108):

$$\psi = \psi(q_1, q_2, \ldots, q_m, \alpha_1, \alpha_2, \ldots, \alpha_{m-1}, h), \tag{109}$$

containing the arbitrary constants $\alpha_1, \alpha_2, \ldots, \alpha_{m-1}, h$. We now use $\psi$ as a generating function for a transformation from the variables

$$q_1, q_2, \ldots, q_m, p_1, p_2, \ldots, p_m$$

to the variables

$$\alpha_1, \alpha_2, \ldots, \alpha_{m-1}, h, \beta_1, \beta_2, \ldots, \beta_{m-1}, -(t + \tau).$$

The transformation is given by the relations

$$\begin{aligned} &\frac{\partial \psi}{\partial q_1} = -p_1, \frac{\partial \psi}{\partial q_2} = -p_2, \ldots, \frac{\partial \psi}{\partial q_m} = -p_m, \\ &\frac{\partial \psi}{\partial \alpha_1} = \beta_1, \frac{\partial \psi}{\partial \alpha_2} = \beta_2, \ldots, \frac{\partial \psi}{\partial \alpha_{m-1}} = \beta_{m-1}, \frac{\partial \psi}{\partial h} = -(t + \tau), \end{aligned} \tag{110}$$

By Theorem X Eq. (107) expressed in terms of $\alpha_i$, $\beta_i$, $h$, $t+\tau$ $(i = 1, 2, \ldots, m-1)$ remain canical. Because the transformed Hamiltonian $H = h$ is a constant we have

$$\begin{aligned} &\frac{\partial H}{\partial \alpha_i} = 0,\ \frac{\partial H}{\partial \beta_i} = 0, \\ &\frac{\partial H}{\partial h} = 1,\ \frac{\partial H}{\partial(-(t+\tau))} = 0. \end{aligned} \tag{111}$$

From (111) and the fact that (107) remain canonical in $\alpha_i$, $\beta_i$, $h$, $t+\tau$ there follows

$$\begin{aligned} &\frac{\mathrm{d}\alpha_i}{\mathrm{d}t} = 0,\ \frac{\mathrm{d}\beta_i}{\mathrm{d}t} = 0, \\ &\frac{\mathrm{d}h}{\mathrm{d}t} = 0,\ \frac{\mathrm{d}(-(t+\tau))}{\mathrm{d}t} = -1. \end{aligned} \tag{112}$$

From $\frac{\mathrm{d}(-(t+\tau))}{\mathrm{d}t} = -1$ in (112) we evidently have $-1 - \frac{\mathrm{d}\tau}{\mathrm{d}t} = -1$ or $\frac{\mathrm{d}\tau}{\mathrm{d}t} = 0$. Hence $\alpha_i$, $\beta_i h$, $\tau$ are all constant. We then use Eq. (110) to express $q_i$, $p_i$ as functions of $t, \alpha_1, \alpha_2, \ldots, \alpha_{m-1}, \beta_1, \beta_2, \ldots, \beta_{m-1}, h, \tau$, and so obtain solutions of the canonical Eq. (107) containing $2m$ arbitrary constants. Jacobi's integration theorem for the conservative case (presented in lecture 21 of the *Vorlesungen* and described above in Sect. 3.6) is proved.[25]

Jacobi also gave a fully detailed derivation of Theorem IX from Theorem X, for conservative systems. He concluded (1866b, 468),

> The foregoing considerations show that Theorem X, of which I have given a direct proof above, includes both the reduction of a mechanical problem in which the theorem of live forces holds to the complete integral of a partial differential equation of the first order, and to the most general and simplest formulas for the variation of the constants.

In the last few pages of the treatise Jacobi outlined some future directions that involved using canonical transformations to investigate partial differential equations. Demidov (1982, 339) views this account as introducing some of the themes that would be explored in the 1870s by Sophus Lie in systematic detail from a more geometric perspective (see also Hawkins (1991)).

### 3.8 The term "canonical" in Jacobi's dynamics

The word "canonical" ("canonische") and its grammatical variants appear 26 times in Jacobi (1866a, b). In introducing lecture 8 of the *Vorlesungen* (1866a) the differential Eq. (18) that are today called "canonical," Jacobi did not, in fact, use this word. He referred to them as "Hamilton's equations." The word "canonical" in the *Vorlesungen* first appears in lecture 15. Assume one has several variables $x$, $y$, $z$, etc., each of which is a function of $t$. We have a set of differential equations involving these variables. The

25 The notation in Sect. 41 of *Ueber diejenigen Probleme* and lecture 21 of the *Vorlesungen* differ. The symbols $\psi$, $W$, $\alpha$ in the *Vorlesungen* become respectively $H$,—$\psi$, -$\alpha$ in *Ueber diejenigen Probleme*.

goal is to express the derivatives of these variables in the form

$$\frac{\mathrm{d}^m x}{\mathrm{d}t^m} = A, \frac{\mathrm{d}^n y}{\mathrm{d}t^n} = B, \frac{\mathrm{d}^p z}{\mathrm{d}t^p} = C, \ldots \tag{113}$$

where $A$, $B$, $C$, ... contain derivatives of $x$ of order at most $m - 1$, of $y$ of order at most $n - 1$, of $z$ of order at most $p - 1$, and so on. Equations given in this way are said to be in canonical form. Evidently, Hamilton's equations are canonical in this sense because the first derivatives of the variables are given as functions of time and the variables themselves.

Jacobi applied the term "canonical form" to the symmetric perturbation Eq. (73) and called the variable constants in these equations "canonical constants." This usage went back to his Paris note of 1837. These equations historically preceded Hamilton's equations (having originated with Lagrange and Poisson) and involved variable constants rather than general dynamical variables, so it was presumably necessary to designate them in a different way, and so Jacobi used the term "canonical form."

The defining characteristic of the transformations Jacobi introduced in *Ueber diejenigen Probleme* was that they preserved the canonical form of the equations. As we observed in the preceding section, he envisaged applying such transformations not just to the variable elements in perturbation problems, but to any set of dynamical variables. Jacobi (1866b, 455) concluded:

> We had then several examples, in which one and the same system of differential equations in canonical form retained this form in various ways with the introduction of new variables \ldots . The formulas for the variation of the constants in the problems of mechanics, in their simplest form, are hereafter only one case of the transformation of one canonical form into another.

Jacobi recognized that the notion of canonical form was a general one that characterized a whole class of transformations in dynamical analysis.

## 4 Desboves and Donkin: canonical transformations

### 4.1 Cayley's *Report* (1858)

Arthur Cayley's 1858 "Report on the recent progress of theoretical dynamics" was presented at the meeting the previous year in Dublin of the British Association for the Advancement of Science. It was an overview of the development of mathematical dynamics since Lagrange's *Méchanique analitique* of 1788. The account is a factual report rather than a critical examination of research. Theorems are stated but the little indication is provided on how the results were proved, nor is there discussion of the particular import and character of the methods employed.

Cayley (1858, 21) drew attention to Jacobi's (1837b) note in the *Comptes rendus* and provided what amounted to an English translation of the passage where Jacobi stated the transformation theorem. Elsewhere in the report, he discussed researches of Adolphe Desboves (p. 26) and William Donkin (pp. 31–36) who published proofs of

Jacobi's results. The account of Donkin was a fairly detailed one, although only the statement of theorems is given.

There is the question of whether there was any dissemination of Jacobi's results in the 1842/43 Königsberg lectures by the students in the seminar, such as Carl Borchardt, in the period leading up to 1866 when the lectures were actually published. However, there does not seem to be evidence in the journal record of any such dissemination having taken place.[26] There is nothing in Cayley's 1858 report, which was fairly detailed, on published research that might have been stimulated by these lectures.

## 4.2 Adolphe Desboves

Jacobi's "Ueber die Reduction" was published in French translation in 1838 in Liouville's journal, and the results presented in it were discussed by Jacques Binet in an article of 1841 in the *Journal de l'École Polytechnique*. There was of course the 1837 note in the Paris *Comptes rendus* in which Jacobi stated the integration theorem named after him, stated Jacobi's theorem on canonical elements and stated the fundamental transformation theorem itself. By the early 1840s knowledge of the theory developed by Hamilton and Jacobi had been disseminated in French journals.

In 1848 the young French researcher Adolphe Desboves published his doctoral dissertation at the University of Paris giving proof of Jacobi's theorem on canonical elements. In the second part of the paper, he investigated forces acting on a planetary body. Desboves began with the Lagrange-Poisson theorem (Eq. 5) in which the constants are the initial values of the coordinates and the momenta. These quantities considered as variables satisfy canonical equations (Eq. 13), expressed in terms of the perturbation potential function. Desboves then proved Jacobi's transformation theorem. He proceeded to take a solution to the Hamilton–Jacobi partial differential equation and to use it as a generating function for a canonical transformation from the Lagrangian constants (initial values of coordinates and momenta) to the arbitrary constants that appear in a complete solution of the Hamilton–Jacobi equation.

In Desboves formulation the equations for the dynamical variables are given in the Newtonian form, while the equations satisfied by the variable constants are given in Lagrange-Poisson canonical form. We begin with the Newtonian equations of motion

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2}=m\frac{\mathrm{d}U}{\mathrm{d}x},\ \frac{\mathrm{d}^2y}{\mathrm{d}t^2}=m\frac{\mathrm{d}U}{\mathrm{d}y},\ \frac{\mathrm{d}^2z}{\mathrm{d}t^2}=m\frac{\mathrm{d}U}{\mathrm{d}z},\ \text{etc.} \tag{114}$$

Here the derivatives on the right sides of the equations are partial derivatives. The solution of (114) will contain the arbitrary constants or elements $a_1, b_1, a_2, b_2, \ldots, a_n, b_n$.[27] The disturbed system is now described in terms of the perturbation function $\Omega$ by the equations

[26] Demidov (1982, 340) comments on the delay in publication of Jacobi's work and observes "Some of his achievements became known by word of mouth through his former students at Königsberg (K. W. Borchardt and others) …" However, no specific documentation is provided by Demidov for this assertion.

[27] Desboves evidently took for granted that the reader would understand that $n$ is equal to *3 m*, where $m$ is the number of particles.

$$\frac{d^2x}{dt^2} = m\frac{dU}{dx} + \frac{d\Omega}{dx},$$
$$\frac{d^2y}{dt^2} = m\frac{dU}{dy} + \frac{d\Omega}{dy}, \text{ etc.} \tag{115}$$

A solution of (115) is obtained from the solution of (114) by supposing that the elements $a_1, b_1, a_2, b_2, \ldots, a_n, b_n$ become variable functions of time. Let us assume that the constants $a_i, b_i$ in the solution to (114) are the initial values of the coordinates and velocities. In this case the variables $a_i(t), b_i(t)$ satisfy the Lagrange-Poisson equations

$$\frac{da_1}{dt} = -\frac{d\Omega}{\partial b_1}, \frac{da_2}{dt} = -\frac{d\Omega}{\partial b_2}, \ldots, \frac{da_n}{dt} = -\frac{d\Omega}{\partial b_n}.$$
$$\frac{db_1}{dt} = \frac{d\Omega}{\partial a_1}, \frac{db_2}{dt} = \frac{d\Omega}{\partial a_2}, \ldots, \frac{db_n}{dt} = \frac{d\Omega}{\partial a_n}. \tag{116}$$

Desboves stated Jacobi's integration theorem for conservative systems, which he took directly with only a slight difference in notation from Jacobi (1837b).[28] Consider the partial differential equation

$$\sum \frac{1}{m}\left[\left(\frac{d\theta}{dx}\right)^2 + \left(\frac{d\theta}{dy}\right)^2 + \left(\frac{d\theta}{dz}\right)^2\right] = 2(U + C), \tag{117}$$

where $C$ is the constant in the equation of live forces. Let $\theta = \theta(x, y, z, \alpha_1, \alpha_2, \ldots)$ be a complete solution of (117). Desboves stated that a solution to (114) is given in the form

$$\frac{d\theta}{d\alpha_1} = \beta_1, \frac{d\theta}{d\alpha_2} = \beta_2, \ldots \frac{d\theta}{dC} = t + \tau, \tag{118}$$

where the $2n$ arbitrary constants or elements are $\alpha_1, \alpha_2, \ldots \alpha_{n-1}, C, \beta_1, \beta_2, \ldots \beta_{n-1}, \tau$. (This result is discussed above Sect. 3.6) Desboves asserted that in the perturbed system these constants regarded as variable will satisfy canonical equations of the form

$$\frac{d\alpha_1}{dt} = +\frac{d\Omega}{d\beta_1}, \frac{d\alpha_2}{dt} = +\frac{d\Omega}{d\beta_2}, \ldots, \frac{dC}{dt} = +\frac{d\Omega}{d\tau}.$$
$$\frac{d\beta_1}{dt} = -\frac{d\Omega}{d\alpha_1}, \frac{d\beta_2}{dt} = -\frac{d\Omega}{d\alpha_2}, \ldots, \frac{d\tau}{dt} = -\frac{d\Omega}{dC}. \tag{119}$$

Jacobi had given no proof of this result in his Paris note. As we saw in Sect. 3.7.1, he provided a full derivation in *Ueber diejenige Probleme*, but the latter work was only published in 1866. The purpose of the first part of Desboves' paper is to derive Eq. (119).

[28] Desboves denotes the solution to the *H-J* Eq. (117) as $\theta$ where Jacobi calls it $V$, and Desboves' arbitrary constant $C$ is called $h$ by Jacobi. Also there is a minor variation in how they indicate the range of the subscripts.

Desboves stated that the proof is based on the third result stated by Jacobi (1837b) (later proved in 1866b) concerning the preservation of the canonical form of Hamilton's equations for transformations defined in terms of a generating function. Desboves adopted Jacobi's formulation and notation, although he never used the word "canonical." We are given the system of equations

$$\begin{aligned}\frac{da_1}{dt} &= -\frac{dH}{db_1}, \frac{da_2}{dt} = -\frac{dH}{db_2}, \ldots, \frac{da_n}{dt} = -\frac{dH}{db_n},\\ \frac{db_1}{dt} &= \frac{dH}{da_1}, \frac{db_2}{dt} = \frac{dH}{da_2}, \ldots, \frac{db_n}{dt} = \frac{dH}{da_n},\end{aligned} \tag{120}$$

where $H$ is a function of $t$ and the variables $a_1, b_1, a_2, b_2, \ldots, a_n, b_n$. We introduce new variables $\alpha_1, \beta_1, \alpha_2, \beta_2, \ldots, \alpha_n, \beta_n$. Consider $\psi$ which is an arbitrary function ("une fonction quelconque") of $\alpha_1, \alpha_2, \ldots, \alpha_n$ and $a_1, a_2, \ldots a_n$. The two sets of $n$ variables $a_i, b_i, a_i, \beta_i$ are assumed to be related by the equations

$$\begin{aligned}\frac{d\psi}{d\alpha_1} &= \beta_1, \frac{d\psi}{d\alpha_2} = \beta_2, \ldots, \frac{d\psi}{d\alpha_n} = \beta_n,\\ \frac{d\psi}{da_1} &= -b_1, \frac{d\psi}{da_2} = -b_2, \ldots, \frac{d\psi}{da_n} = -b_n.\end{aligned} \tag{121}$$

For a change of variables defined in this way the transformation theorem asserts that Eq. (120) become

$$\begin{aligned}\frac{d\alpha_1}{dt} &= -\frac{dH}{d\beta_1}, \frac{d\alpha_2}{dt} = -\frac{dH}{d\beta_2}, \ldots, \frac{d\alpha_n}{dt} = -\frac{dH}{d\beta_n},\\ \frac{d\beta_1}{dt} &= \frac{dH}{d\alpha_1}, \frac{d\beta_2}{dt} = \frac{dH}{d\alpha_2}, \ldots, \frac{d\beta_n}{dt} = \frac{dH}{d\alpha_n}.\end{aligned} \tag{122}$$

Hence the canonical form of the equations is preserved in the transformation.

Desboves' proof of the transformation theorem is taken up below. For now, we describe how he used it to derive the canonical equations for the variable constants in the Jacobian form (119). The key is to use the solution $\theta$ of the Hamilton–Jacobi Eq. (117) as a generating function for a canonical transformation. We have $\theta = \theta(x, y, z, \alpha_1, \alpha_2, \ldots, C)$ and Eq. (118). There are also the equations

$$\frac{d\theta}{dx} = \frac{mdx}{dt}, \frac{d\theta}{dy} = \frac{mdy}{dt}, \text{ etc.,} \tag{123}$$

that were derived in the proof of the integration theorem. If we set $t = 0$ in (118) and (123) then $x$, $y$, $z$, ... become $a_1, a_2, \ldots, a_n$ and $\frac{mdx}{dt}, \frac{mdy}{dt}, \ldots$ become$-b_1, -b_2, \ldots, -b_n$. (Here the $b_i$ now denote minus the values of $\frac{mdx}{dt}, \frac{mdy}{dt}, \ldots$ at $t = 0$.) $\theta$ is a function of$a_1, a_2, \ldots, a_n, \alpha_1, \alpha_2, \ldots \alpha_{n-1}, C$. Hence the following relations involving the generating function $\theta$ hold:

$$\begin{aligned}\frac{d\theta}{d\alpha_1} &= \beta_1, \frac{d\theta}{d\alpha_2} = \beta_2, \ldots, \frac{d\theta}{dC} = \tau,\\ \frac{d\theta}{da_1} &= -b_1, \frac{d\theta}{da_2} = -b_2, \ldots, \frac{d\theta}{da_n} = -b_n.\end{aligned} \tag{124}$$

We have the original canonical Eqs. (116), and (124) gives rise to a canonical transformation from $a_1, a_2, \ldots a_n, b_1, b_2, \ldots b_n$ to $\alpha_1, \alpha_2, \ldots \alpha_{n-1}, C, \beta_1, \beta_2, \ldots \beta_{n-1}, \tau$. It follows that (122) holds (with $\alpha_n = C$ and$\beta_n = \tau$)), with $H = -\Omega$. But this system of equations is simply (119) and the main theorem is established.[29]

We turn now to Desboves' proof of the transformation theorem itself. Using relations (121) we can express $a_i, b_i$ as functions of $\alpha_i, \beta_i$. Desboves asserted that the function H will satisfy Lagrange's fundamental identity (5), which here takes the form:

$$\begin{aligned}
\frac{\mathrm{dH}}{\mathrm{d}\alpha_1} dt &= [\alpha_1, \beta_1] d\beta_1 + [\alpha_1, \alpha_2] d\alpha_2 + [\alpha_1, \beta_2] d\beta_2 + \ldots. \\
\frac{\mathrm{dH}}{\mathrm{d}\beta_1} dt &= [\beta_1, \alpha_1] d\alpha_1 + [\beta_1, \alpha_2] d\alpha_2 + [\beta_1, \beta_2] d\beta_2 + \ldots. \\
&\ldots\ldots\ldots\ldots\ldots
\end{aligned} \tag{125}$$

Desboves wrote simply that Eq. (125) were "known." The Lagrange bracket $[\alpha_1, \beta_1]$ is given as

$$[\alpha_1, \beta_1] = \frac{\mathrm{d}a_1}{\mathrm{d}\alpha_1}\frac{\mathrm{d}b_1}{\partial\beta_1} - \frac{\mathrm{d}b_1}{\mathrm{d}\alpha_1}\frac{\mathrm{d}a_1}{\mathrm{d}\beta_1} + \frac{\mathrm{d}a_2}{\mathrm{d}\alpha_1}\frac{\mathrm{d}b_2}{\partial\beta_1} - \frac{\mathrm{d}b_2}{\mathrm{d}\alpha_1}\frac{\mathrm{d}a_2}{\mathrm{d}\beta_1} + \ldots. \tag{126}$$

with corresponding expressions for$\left[\alpha_i, \alpha_j\right]$,$\left[\beta_i, \beta_j\right]$and$\left[\alpha_i, \beta_j\right]$. It is necessary to show

$$\begin{gathered}
\left[\alpha_i, \alpha_j\right] = \left[\beta_i, \beta_j\right] = 0 \text{ for all } i \text{ and } j \\
\left[\alpha_i, \beta_j\right] = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}
\end{gathered} \tag{127}$$

Equations (122) in canonical form follow from (125) and (127) and the theorem is established.

Desboves gave a detailed proof that$[\alpha_1, \beta_1] = 1$. He adopted a convention for writing differentiation which in modern notation goes as follows. If $z = \varphi(x, y(x, w))$ then we write

$$\left(\frac{\mathrm{d}\varphi}{\mathrm{d}x}\right) = \frac{\partial\varphi}{\partial x}, \frac{\mathrm{d}\varphi}{\mathrm{d}x} = \left(\frac{\mathrm{d}\varphi}{\mathrm{d}x}\right) + \frac{\mathrm{d}\varphi}{\mathrm{d}y}\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\partial\varphi}{\partial x} + \frac{\partial\varphi}{\partial y}\frac{\partial y}{\partial x}.$$

The generating function $\psi$ is a function of $\alpha_1, \alpha_2, \ldots, \alpha_n$ and $a_1, a_2, \ldots a_n$, where $a_1, a_2, \ldots a_n$ may from (121) be regarded as functions of $\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_n$. Hence we have

[29] In (116) we set $b_i^* = -b_i$, so that (116) becomes $\frac{\mathrm{d}b_i^*}{\mathrm{d}t} = -\frac{\mathrm{d}(-\Omega)}{\partial a_1}$, $\frac{\mathrm{d}b_i^*}{\mathrm{d}t} = \frac{\mathrm{d}(-\Omega)}{\partial a_2}$. The relations (124) given in terms of the generating function $\theta$ become $\frac{\mathrm{d}\theta}{\mathrm{d}\alpha_i} = \beta_i$, $\frac{\mathrm{d}\theta}{\mathrm{d}a_1} = -b_i^*$. By the transformation theorem it then follows that (119) hold.

$$\begin{aligned} \frac{\mathrm{d}\psi}{\mathrm{d}\alpha_1} &= \left(\frac{\mathrm{d}\psi}{\mathrm{d}\alpha_1}\right) + \frac{\mathrm{d}\psi}{\mathrm{d}a_1}\frac{\mathrm{d}a_1}{\mathrm{d}\alpha_1} + \frac{\mathrm{d}\psi}{\mathrm{d}a_2}\frac{\mathrm{d}a_2}{\mathrm{d}\alpha_1} + \ldots + \frac{\mathrm{d}\psi}{\mathrm{d}a_n}\frac{\mathrm{d}a_n}{\mathrm{d}\alpha_1}. \\ \frac{\mathrm{d}\psi}{\mathrm{d}\beta_1} &= \frac{\mathrm{d}\psi}{\mathrm{d}a_1}\frac{\mathrm{d}a_1}{\mathrm{d}\beta_1} + \frac{\mathrm{d}\psi}{\mathrm{d}a_2}\frac{\mathrm{d}a_2}{\mathrm{d}\beta_1} + \ldots + \frac{\mathrm{d}\psi}{\mathrm{d}a_n}\frac{\mathrm{d}a_n}{\mathrm{d}\beta_1}. \end{aligned} \tag{128}$$

From (121), (128) becomes

$$\begin{aligned} \frac{\mathrm{d}\psi}{\mathrm{d}\alpha_1} &= \beta_1 - b_1\frac{\mathrm{d}a_1}{\mathrm{d}\alpha_1} - b_2\frac{\mathrm{d}a_2}{\mathrm{d}\alpha_1} - \ldots - b_n\frac{\mathrm{d}a_n}{\mathrm{d}\alpha_1}. \\ \frac{\mathrm{d}\psi}{\mathrm{d}\beta_1} &= -b_1\frac{\mathrm{d}a_1}{\mathrm{d}\beta_1} - b_2\frac{\mathrm{d}a_2}{\mathrm{d}\beta_1} - \ldots - b_n\frac{\mathrm{d}a_n}{\mathrm{d}\beta_1}. \end{aligned} \tag{129}$$

There is a known identity expressing the equality of mixed partial derivatives, $\frac{\mathrm{d}}{\mathrm{d}\beta_1}\frac{\mathrm{d}\psi}{\mathrm{d}\alpha_1} = \frac{\mathrm{d}}{\mathrm{d}\alpha_1}\frac{\mathrm{d}\psi}{\mathrm{d}\beta_1}$. Hence, we are able to equate the (partial derivative) of the right side of (121) with respect to $\beta_1$ to the (partial) derivative of the right side of (121) with respect to $\alpha_1$. Doing so we obtain$[\alpha_1, \beta_1] = 1$, where we use the fact that terms analogous to $\frac{b_1\mathrm{d}^2a_1}{\mathrm{d}\alpha_1 d\beta_1}$ and $\frac{b_1\mathrm{d}^2a_1}{\mathrm{d}\beta_1 d\alpha_1}$ cancel each other.

Using similar reasoning we find that $[\alpha_1, \alpha_2] = 0$. In this case, we have

$$\begin{aligned} \frac{\mathrm{d}\psi}{\mathrm{d}\alpha_1} &= \beta_1 - b_1\frac{\mathrm{d}a_1}{\mathrm{d}\alpha_1} - b_2\frac{\mathrm{d}a_2}{\mathrm{d}\alpha_1} - \ldots - b_n\frac{\mathrm{d}a_n}{\mathrm{d}\alpha_1} \\ \frac{\mathrm{d}\psi}{\mathrm{d}\alpha_2} &= \beta_2 - b_1\frac{\mathrm{d}a_1}{\mathrm{d}\alpha_2} - b_2\frac{\mathrm{d}a_2}{\mathrm{d}\alpha_2} - \ldots - b_n\frac{\mathrm{d}a_n}{\mathrm{d}\alpha_2} \end{aligned} \tag{130}$$

It is apparent that$\frac{\mathrm{d}\beta_1}{\mathrm{d}\alpha_2} = \frac{\mathrm{d}\beta_2}{\mathrm{d}\alpha_1} = 0$. From $\frac{\mathrm{d}}{\mathrm{d}\alpha_2}\frac{\mathrm{d}\psi}{\mathrm{d}\alpha_1} = \frac{\mathrm{d}}{\mathrm{d}\alpha_1}\frac{\mathrm{d}\psi}{\mathrm{d}\alpha_2}$ it then follows that$[\alpha_1, \alpha_2] = 0$. In a similar way, we verify the other cases in (127).

It is important to note that Desboves proved the transformation theorem on the assumption that the function $H$ satisfies the Lagrangian fundamental identity (125). Thus he only proved a special case of the theorem, although what he did prove was sufficient to derive the main result (119) of his investigation. Desboves' whole analysis is closely bound up with the theory of variation of constants, in a way that was not true for Jacobi. While Jacobi (1866b) obtained his transformation result within the general setting of this theory, the proof was logically independent of it, as Jacobi himself recognized. The special character of Desboves' proof may explain historically why later researchers such as Donkin did not follow in his footsteps.

It is nonetheless interesting that in 1848 Desboves took a solution of the Hamilton—Jacobi equation as a generating function and then used Jacobi's integration theorem to obtain a canonical transformation. Of course, what Desboves did is different from the modern treatment, which is developed in a general dynamical setting. By contrast, Desboves was considering perturbations and variable constants. Furthermore, whereas a common proof of Jacobi's integration theorem in modern texts uses transformations, Desboves assumed this theorem from the outset, the statement and proof having appeared in Jacobi (1837a).

The first part of Desboves memoir was notable for its succinctness and originality. The idea of using a solution of the Hamilton–Jacobi equation as the generating function for a canonical transformation had been introduced by Jacobi in a work that would not be published until 1866. It is fundamental to the modern subject. Desboves was able to put together elements from the work of Poisson and Lagrange on the one hand, and

Jacobi on the other, and develop the resulting theory conceptually in an innovative way.

Desboves went on his memoir to consider a problem in particle dynamics that had been investigated in some work of Joseph Liouville (1847) in which Jacobi's (1837a) integration theorem was used as a basis for integrating the differential equations of the problem. Desboves carried out a detailed analysis of the problem, and arrived at a set of elements he believed was particularly effective in determining the motion. To do so he used Jacobi's integration theorem and Eq. (118). The paper concluded by asserting that the theorem on canonical elements formulated and proved at the beginning of the paper could be applied to the given elements, although a such further investigation was not pursued in the memoir. Desbove's study of particle motion involved only Jacobi's integration theorem (expressed in (118)). The stage was set for the future application of the theorem on canonical elements. Desboves was presenting new results that he had obtained, one involving some proofs of Jacobian theorems, and the other on a problem in particle dynamics based on Jacobi's (1837a) integration theorem.[30]

Desboves became a teacher of mathematics at a lycée in Fontanes, and went on to publish books on trigonometry, analytic geometry and the history of seventeenth-century geometry. He also wrote a monograph on early childhood development based on observations of his granddaughter in the first two and a half years of her life. He apparently carried out no further researches in dynamical theory or mathematical astronomy.

## 4.3 William F. Donkin

### 4.3.1 Introduction

William Donkin was trained in mathematics at Oxford University and from 1842 was the Savilian Professor of Astronomy there. In 1854 and 1855 the forty-year-old professor published a two-part article in the *Philosophical Transactions* titled "On a Class of Differential Equations, including those which occur in Dynamical Problems." This substantial study occupied over one hundred pages of the journal and was devoted to an investigation of the mathematical methods of Jacobi, Poisson and Hamilton and their application to dynamics. Although Donkin did not claim to be at the same level as these illustrious researchers, he stated that his work "might be found to possess some degree of novelty and interest."

Referring to the application of the dynamical theory to planetary motion Donkin adopted the perspective of the mathematician:

> This investigation, if interesting at all, will probably be so to the mathematician rather than to the astronomer. I think, however, that if the theories of physical astronomy were more frequently treated rigorously and symmetrically, apart

[30] Cayley's (1858, 26) remarks on Desboves' memoir leave the impression that the two theorems proved by Desboves in the first part of his memoir are applied in the second part, which is not true. The theorem on canonical elements is only mentioned by Desboves at the end as a basis for further study of these elements. The transformation theorem itself had only been used to prove the theorem on canonical elements and did not come up again.

> from any approximate integrations; and if, when the latter are introduced, more care were taken to give a clear and exact view of the nature of the reasoning employed, it might be possible to draw the attention and secure the cooperation of a class of mathematicians who now may well be excused, if, after a slight trial, they turn from the subject in disgust, and prefer to expatiate in those beautiful fields of speculation which are offered to them by other branches of modern geometry and analysis. (Donkin 1855, 300–301)

Donkin's approach was formal and, in some respects, resembled more the style of a writer on determinants than it did an investigation in dynamical analysis.[31] He seemed in a self-conscious way to have valued generality in mathematics, and his formal inclinations and originality were no doubt connected to this. Nevertheless, he produced a systematic and substantial work that included proofs all three of the Jacobian theorems discussed above in Sect. 3.[32] It should be noted that in Jacobi's original publications (1837ab), the equations of motion were given in Newtonian rather than canonical form. This was the formulation adopted by researchers such as Liouville (1847) and Desboves (1848). By contrast, Donkin (1854, 87–89) stated and proved the result for dynamical equations of motion given in the Hamiltonian canonical form. He even developed an argument that showed that in order for the proof to work it had to be the case that the given solution of the partial differential equation was a complete solution. Donkin did all of this some twelve years before the publication of Jacobi's posthumous *magnum opus*.

Although Donkin largely followed the tradition of French and British notation of the period, at times he adopted unconventional terminology that was not followed in the later subject. Donkin (1854, 103) himself stated, "I regret to use symbols with a meaning different from that which custom has to some extent sanctioned but there seemed to be only a choice of difficulties."

In Donkin's writing the symbol $d$ denotes partial differentiation, and the prime symbol $'$ denotes ordinary differentiation. Suppose $p$ is a function of $t$, $x$, $y$, where $x$, $y$ are themselves functions of $t$. Whereas today we would write

$$\frac{\mathrm{d}p}{\mathrm{d}t} = \frac{\partial p}{\partial t} + \frac{\partial p}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial p}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t},$$

[31] Referring to English mathematicians of note, Hamilton in 1851 mentioned Herschel, Cayley, Donkin, Peacock and De Morgan (Crilly 2006, 178). Peacock and De Morgan were formalists and Cayley's mathematical orientation was algebraic and formalistic. Donkin himself published a paper in 1850 titled "On Certain Theorems in the Calculus of Operations," research which was informed by Boole's formal operational approach to analysis.

[32] Jacobi's integration theorem appears in Donkin (1854, 87–99); Jacobi's theorem on canonical elements in Donkin (1854, 105–106); and Jacobi's transformation theorem in Donkin (1855, 315–316). Concerning Jacobi's theorem on canonical elements, Donkin (1854, 106 note) stated that the result was given by Jacobi (no specific citation is provided) and proved by Desboves (1848) but that his own proof was different from Desboves'. In his proof Donkin did not (as Desboves had) use a canonical transformation with a generating function that is a solution to a partial differential equation. Instead, Donkin's proof uses Poisson bracket methods.

Donkin wrote

$$p' = \frac{\mathrm{d}p}{\mathrm{d}t} + \frac{\mathrm{d}p}{\mathrm{d}x}x' + \frac{\mathrm{d}p}{\mathrm{d}y}y'.$$

Donkin introduced new notation for the terms that appear in Lagrange brackets. The expression $\frac{\mathrm{d}(y_i, x_i)}{\mathrm{d}(h,k)}$ is defined as

$$\frac{\mathrm{d}(y_i, x_i)}{\mathrm{d}(h, k)} = \frac{\mathrm{d}y_i}{\mathrm{d}h}\frac{\mathrm{d}x_i}{\mathrm{d}k} - \frac{\mathrm{d}y_i}{\mathrm{d}k}\frac{\mathrm{d}x_i}{\mathrm{d}h}.$$

The Lagrange bracket $[k, h]$ is then given as

$$[k, h] = \sum_{i=1}^{n} \frac{\mathrm{d}(y_i, x_i)}{\mathrm{d}(h, k)}.$$

Donkin used this notation extensively in his investigation, although it did not entirely catch on in subsequent mathematics.

#### 4.3.2 Donkin's analysis

The purpose of the present article is not to give a complete account of Donkin's papers, something that would entail an extended study. The scope of Donkin's investigation much surpasses that of Desboves. We will consider only his treatment of transformations of variables and coordinates, a subject that occupies the better part of the 1855s paper. Our concern is with Donkin's development of the theory rather than the applications he made to celestial mechanics. In particular, we focus on the proof of Jacobi's theorem on canonical transformations that formed the centrepiece of his investigation. This result appears as Theorem VIII of Sect. 6 of Part Two. Donkin in fact did not cite Jacobi (1837b), where the result had originally been stated, but did refer to Desboves' 1848 "Deux théorèmes de M. Jacobi" as having provided proof of the result, in a less general form and by a different method of demonstration than his own.

Donkin's proof of the transformation theorem relies on a result that had appeared with a different notation in part one. We begin by examining the latter. In Sect. 5 of Part One Donkin (1854, 79) considered a set of variables $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n$ and a corresponding set of constants $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n$. We are given a function $X$ that is a function of $t, x_1, x_2, \ldots, x_n, a_1, a_2, \ldots, a_n$. (In what follows recall that for Donkin $d$ denotes partial differentiation). The variables and constants are related by the equations

$$\frac{\mathrm{d}X}{\mathrm{d}x_i} = y_i, \; \frac{\mathrm{d}X}{\mathrm{d}a_i} = b_i. \tag{131}$$

Equation (131) defines a transformation from the variables $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n$ to $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n.$ In part two Donkin (1855, 313) would refer to such a change of variables as a "normal" transformation and call $X$ the "modulus" of the transformation. In the modern subject, one speaks of a canonical transformation and $X$ is called a generating function. Unlike the generating functions considered by Jacobi and Desboves, $X$ is a function of time $t$ as well of the quantities $x_1, x_2, \ldots, x_n$ and $a_1, a_2, \ldots, a_n$. Donkin proceeded to derive from (131) the identity

$$\frac{\mathrm{d}b_j}{\mathrm{d}y_k} = -\frac{\mathrm{d}x_k}{\mathrm{d}a_j}(j = 1, \ldots, n \text{and} k = 1, \ldots n) \tag{132}$$

In (132) the $b_j$ on the left side are functions of $t$, $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n$ while the $x_k$ on the right side are functions of $t$,$a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n$. To obtain (132) we begin by differentiating (4.18b) partially with respect to $a_j$:

$$\frac{\mathrm{d}^2X}{\mathrm{d}a_i da_j} + \frac{\mathrm{d}^2X}{\mathrm{d}a_i \mathrm{d}x_1}\frac{\mathrm{d}x_1}{\mathrm{d}a_j} + \frac{\mathrm{d}^2X}{\mathrm{d}a_i \mathrm{d}x_2}\frac{\mathrm{d}x_2}{\mathrm{d}a_j} + \ldots = 0. \tag{133}$$

Using (131) and substituting into (133) we obtain

$$\frac{\mathrm{d}b_j}{\mathrm{d}a_i} + \frac{\mathrm{d}y_1}{\mathrm{d}a_i}\frac{\mathrm{d}x_1}{\mathrm{d}a_j} + \frac{\mathrm{d}y_2}{\mathrm{d}a_i}\frac{\mathrm{d}x_2}{\mathrm{d}a_j} + \cdots + \frac{\mathrm{d}y_n}{\mathrm{d}a_i}\frac{\mathrm{d}x_n}{\mathrm{d}a_j} = 0. \tag{134}$$

The next step is to multiply (134) by $\frac{\mathrm{d}a_i}{\mathrm{d}y_k}$ and sum the resulting $n$ equations over $i$. The first term of this sum is

$$\sum_{i=1}^{n}\frac{\mathrm{d}b_j}{\mathrm{d}a_i}\frac{\mathrm{d}a_i}{\mathrm{d}y_k} = \frac{\mathrm{d}b_j}{\mathrm{d}y_k}. \tag{135}$$

In the remaining terms of the sum, the coefficient of each $\frac{\mathrm{d}x_s}{\mathrm{d}a_j}$ $(s = 1, \ldots, n)$ is

$$\left.\sum_{i=1}^{n}\frac{\mathrm{d}a_i}{\mathrm{d}y_k}\frac{\mathrm{d}y_s}{\mathrm{d}a_i}\right\} = \begin{matrix} 1 if s = k \\ 0 if s \neq k \end{matrix}. \tag{136}$$

It follows that the total sum equated to zero is

$$\frac{\mathrm{d}b_j}{\mathrm{d}y_k} + \frac{\mathrm{d}x_k}{\mathrm{d}a_j} = 0,$$

which is simply (132).

It should be noted that in Donkin's rather formal derivation the same symbol may be used in slightly different ways. For example, in deriving (133) from (131) $a_i$ and $b_i$ are taken to be independent variables and $\frac{\mathrm{d}b_i}{\mathrm{d}a_i} = 0$. On the other hand, in the quotient $\frac{\mathrm{d}b_j}{\mathrm{d}a_i}$ in (134) $b_j$ is a function of $t$, $x_1, \ldots, x_n$ and $a_1, \ldots, a_n$. Also in (135) the $\frac{\mathrm{d}b_j}{\mathrm{d}a_i}$ on the left is a function of $t$, $x_1, \ldots, x_n$ and $a_1, \ldots, a_n$ while the $\frac{\mathrm{d}b_j}{\mathrm{d}y_k}$ on the right is a function of $t, x_1, \ldots, x_n$ and $y_1, \ldots, y_n$. The reasoning is apparently valid but more explanation and greater clarity and detail in notation would be desirable.[33]

Donkin proceeded to derive other identities similar to (132). In all, he showed that relations (131) give rise to the following identities

$$\begin{aligned}\frac{dx_i}{da_j} &= -\frac{db_j}{dy_i}, \frac{dx_i}{db_j} = \frac{da_j}{dy_i} \\ \frac{dy_i}{da_j} &= \frac{db_j}{dx_i}, \frac{dy_i}{db_j} = -\frac{da_j}{dx_i}\end{aligned} \quad (i = 1, \ldots, n \text{ and } j = 1, \ldots n). \tag{137}$$

In these identities the $x_i$ and $y_i$ on the left sides are functions of $t$, $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n$ while the $b_j$ and $a_j$ on the right sides are functions of $t, x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n$.

In Theorem VIII of Part Two Donkin stated and proved Jacobi's theorem on canonical transformations.[34] He began his account by introducing the Hamiltonian function $Z$ of $t, x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n$. (Donkin did not use the modern term Hamiltonian.) These variables satisfy the equations

$$x_i{}' = \frac{dZ}{dy_i}, y_i{}' = -\frac{dZ}{dx_i}. \tag{138}$$

Such equations were said by Donkin to be in *canonical form* (italics in the original).[35]

Consider now the second set of variables $\xi_1, \ldots, \xi_n, \eta_1, \ldots, \eta_n$. We are given the function $P$ that is a function of $t, \xi_1, \ldots, \xi_n \text{and} y_1, \ldots, y_n$. The two sets of variables

[33] The reasoning is verified in simple examples. Suppose for instance that $n = 1$ and we have the two pairs of variables *x,y* and *a,b*. Let $X = a^2x^2$. Equations (131) give $y(x,a) = 2xa^2$ and $b(x,a) = 2\,ax^2$. Then we find that $a(x,y) = \sqrt{\frac{y}{2x}}$ and $b(x,y) = \sqrt{2x^3y}$. The identity

$$\frac{\partial b(x,a)}{\partial a} \bullet \frac{\partial a(x,y)}{\partial y} = \frac{\partial b(x,y)}{\partial y}$$

then becomes

$$2x^2 \bullet \frac{1}{\sqrt{8xy}} = \sqrt{\frac{x^3}{2y}},$$

and the left and right sides of this equation are indeed equal. (We have here used the modern symbol for the partial derivative.).

[34] One might fault Donkin for not having given some details on the statement of this result in Jacobi (1837b). A reader coming to his paper would not be aware that the ideas of normal transformations and functional moduli originated with Jacobi. (His reference to Desboves' paper on two theorems of Jacobi (1837b) did not include its title or much detail about its contents.).

[35] It is important to note that the term "canonical" is used just once in Jacobi (1837b) and that this is the only place that this term does appear in his published work prior to 1866. Donkin himself uses the term to characterize the symmetric form of Hamilton's equations. The word appears thirteen times in Donkin (1855) (but not at all in Donkin (1854)).

$x_1, \ldots, x_n, y_1, \ldots, y_n$ and $\xi_1, \ldots, \xi_n, \eta_1, \ldots, \eta_n$ are connected by the relations

$$\frac{dP}{d\xi_i} = \eta_i, \frac{dP}{dy_i} = x_i. \tag{139}$$

In this setting Donkin's Theorem VIII asserts that the transformed variables $\xi_1, \ldots, \xi_n, \eta_1, \ldots, \eta_n$ satisfy the canonical equations

$$\xi_i{}' = \frac{d\Phi}{d\eta_i}, \eta_i{}' = -\frac{d\Phi}{d\xi_i}, \tag{140}$$

where the function $\Phi$ is given as $\Phi = Z - \frac{dP}{dt}$.

If P does not contain $t$ then $\frac{dP}{dt} = 0$ and $\Phi = Z$, the same generating function that had been used by Jacobi and Desboves. (140) then becomes

$$\xi_i{}' = \frac{dZ}{d\eta_i}, \eta_i{}' = -\frac{dZ}{d\xi_i}. \tag{141}$$

We will give Donkin's proof for this case because it is simpler than the corresponding derivation of (140) and provides a clear comparison to the result proved by Jacobi and Desboves.

Donkin undertook a revision of the notation used in Part One so that

$$x_1, \ldots, x_n, y_1, \ldots, y_n, a_1, \ldots, a_n, b_1, \ldots, b_n$$

are now, respectively, designated as

$$\xi_1, \ldots, \xi_n, \eta_1, \ldots, \eta_n, y_1, \ldots, y_n, x_1, \ldots, x_n.$$

Note that the underlying concept of "variable constant" is no longer present and that all of the quantities under consideration are simply variables. Donkin (1855, 313) observed, "It is the object of the present section to consider the general class of transformations of which the method [of the variation of elements] in question is a particular, and not the only useful case."

In term of the new notation, the set of identities (137) are now written

$$\begin{array}{l} \frac{d\xi_i}{dy_j} = -\frac{dx_j}{d\eta_i}, \frac{d\xi_i}{dx_j} = \frac{dy_j}{d\eta_i} \\ \frac{d\eta_i}{dy_j} = \frac{dx_j}{d\xi_i}, \frac{d\eta_i}{dx_j} = -\frac{dy_i}{d\xi_i} \end{array}. \tag{142}$$

Consider next the time derivative of $\xi_i$:

$$\xi_i{}' = \sum_j \left( \frac{d\xi_i}{dx_j} x_j{}' + \frac{d\xi_i}{dy_j} y_j{}' \right). \tag{143}$$

Substituting from the original canonical Eqs. (138), (143) becomes

$$\xi_i{}' = \sum_j \left( -\frac{\mathrm{d}Z}{\mathrm{d}x_j}\frac{\mathrm{d}\xi_i}{\mathrm{d}y_j} + \frac{\mathrm{d}Z}{\mathrm{d}y_j}\frac{\mathrm{d}\xi_i}{\mathrm{d}x_j} \right). \tag{144}$$

Using the identities in the first line of (142) we write (144) in the form

$$\xi_i{}' = \sum_j \left( \frac{\mathrm{d}Z}{\mathrm{d}x_j}\frac{\mathrm{d}x_j}{\mathrm{d}\eta_i} + \frac{\mathrm{d}Z}{\mathrm{d}y_j}\frac{\mathrm{d}y_j}{\mathrm{d}\eta_i} \right). \tag{145}$$

The right side of (145) is $\frac{\mathrm{d}Z}{\mathrm{d}\eta_i}$ and so we have

$$\xi_i{}' = \frac{\mathrm{d}Z}{\mathrm{d}\eta_i}. \tag{146}$$

A similar derivation using the identities in the second line of (142) leads to (141) and the theorem is established.

We outline now how Donkin obtained this result for a more general time-dependent generating function $P$ of $\xi_1, \ldots, \xi_n, y_1, \ldots, y_n$ and $t$. The variable $\xi_i$ is now a function of $x_1, \ldots, x_n, y_1, \ldots, y_n$ and $t$. Hence

$$\xi_i{}' = \frac{\mathrm{d}\xi_i}{\mathrm{d}t} + \sum_j \left( \frac{\mathrm{d}\xi_i}{\mathrm{d}x_j} x_j{}' + \frac{\mathrm{d}\xi_i}{\mathrm{d}y_j} y_j{}' \right). \tag{147}$$

In Part One Donkin had shown that

$$\frac{\mathrm{d}\xi_i}{\mathrm{d}t} = -\frac{\mathrm{d}\left(\frac{\mathrm{d}P}{\mathrm{d}t}\right)}{\mathrm{d}\eta_i}. \tag{148}$$

Hence, we obtain

$$\xi_i{}' = -\frac{\mathrm{d}\left(\frac{\mathrm{d}P}{\mathrm{d}t}\right)}{\mathrm{d}\eta_i} + \frac{\mathrm{d}Z}{\mathrm{d}\eta_j} = \frac{\mathrm{d}\left(Z - \frac{\mathrm{d}P}{\mathrm{d}t}\right)}{\mathrm{d}\eta_i} = \frac{\mathrm{d}\Phi}{\mathrm{d}\eta_i},$$

As was the case with Jacobi, Donkin's proof of the transformation theorem is a direct one about differential equations with no reference to variational theory. In comparing Donkin's proof with Jacobi's, one is struck by its relative simplicity. This feature arises from Donkin's use of identities (142), which carry the main weight of the proof. For a critical appraisal, one is led back to Part One where these identities (as (137)) were originally derived. Donkin's' proof is not found in modern textbooks although it would appear to be a respectable derivation of the result that also includes generating functions which depend on time.

Following his original introduction of the "modulus" or generating function $P$ (expressed in terms of $\xi_1, \ldots, \xi_n, y_1, \ldots, y_n$ and $t$), Donkin (1855, 314) had observed that we could employ instead of $P$ one of the following functions:

$Q$ expressed in terms of $x_1, \ldots, x_n, \xi_1, \ldots, \xi_n$ and $t$,
$R$ expressed in terms of $y_1, \ldots, y_n, \eta_1, \ldots, \eta_n$ and $t$.
$S$ expressed in terms of $x_1, \ldots, x_n, \eta_1, \ldots, \eta_n$ and $t$.
For example, consider the case of $Q$. The partial derivatives of $Q$ with respect to $y_i$ and $\eta_i$ ($i = 1, \ldots, n$) are zero. This condition suggests we consider $Q$ of the form.

$$Q = -P + \sum_i x_i y_i. \tag{149}$$

It is evident that the partial derivative of $Q$ with respect to $\eta_i$ is zero. Also, we have

$$\frac{\mathrm{d}Q}{\mathrm{d}y_i} = -\frac{\mathrm{d}P}{\mathrm{d}y_i} + x_i = -x_i + x_i = 0.$$

In this case, the relations (139) that must be satisfied, expressed in terms of $x_1, \ldots, x_n, \xi_1, \ldots, \xi_n$ become

$$\frac{\mathrm{d}Q}{\mathrm{d}x_i} = y_i, \frac{\mathrm{d}Q}{\mathrm{d}\xi_i} = -\eta_i. \tag{150}$$

In a similar way, we obtain expressions for the moduli $R$ and $S$ in terms of $P$ and relations that correspond to (139) and (150).

### 4.3.3 Coordinate transformations are normal

To get a feel for how Donkin used the transformation theorem, we consider his treatment in Sects. 68–70 (Donkin 1855, 318–321) of what he called coordinate transformations. A coordinate transformation relates the coordinates $x_1, \ldots, x_n$ to the corresponding coordinate variables $\xi_1, \ldots, \xi_n$ and the time. In modern textbooks such a transformation is called a point transformation (see for example Goldstein (1950, 238)). Donkin showed that every coordinate transformation is also a normal (i.e., canonical) transformation. We begin with the canonical system of variables $x_1, \ldots, x_n$, $y_1, \ldots, y_n$ with associated Lagrangian $W( x_1, \ldots, x_n, x_1{}', \ldots, x_n{}', t)$ and Hamiltonian $Z(x_1, \ldots, x_n, y_1, \ldots, y_n, t)$. (Donkin did not use these modern names.) We are given a coordinate transformation from $x_1, \ldots, x_n$ to the variables $\xi_1, \ldots, \xi_n$ defined as

$$x_i = (x_i)(\xi_1, \xi_2, \ldots, \xi_n, t), (i = 1, \ldots, n) \tag{151}$$

with "the brackets indicating that $x_1, \ldots, x_n$ are to be expressed in terms of $\xi_1, \ldots, \xi_n$" (Donkin 1855, 318). Express $W$ in terms of $t, \xi_1, \ldots, \xi_n, \xi_1{}', \ldots, \xi_n{}'$ and define $\eta_i$ as $\eta_i = \frac{\mathrm{d}W}{\mathrm{d}\xi_i'}$ Donkin showed that (151) gives rise to a normal transformation from $x_1, \ldots, x_n, y_1, \ldots, y_n$ to $\xi_1, \ldots, \xi_n, \eta_1, \ldots, \eta_n$. The modulus $P$ for this transformation is the function of $t, \xi_1, \xi_2, \ldots, \xi_n, y_1, y_2, \ldots, y_n$ defined as

$$P = (x_1)y_1 + (x_2)y_2 + \ldots + (x_n)y_n. \tag{152}$$

Evidently, we have

$$\frac{dP}{dy_i} = (x_i) = x_i (i = 1, \ldots, n). \tag{153}$$

It is necessary to calculate $\frac{dP}{d\xi_i}$ and show that it is equal to $\eta_i$. We have

$$\frac{dP}{d\xi_i} = y_1 \frac{d(x_1)}{d\xi_i} + y_2 \frac{d(x_2)}{d\xi_i} + \ldots y_n \frac{d(x_n)}{d\xi_i}. \tag{154}$$

Now

$$W = -(Z) + x_1{}' y_1 + x_2{}' y_2 \ldots + x_n{}' y_n. \tag{155}$$

Donkin (1855, 319) wrote "observing that $(Z)$ becomes a function of $\xi_1{}'$, &c., only through $y_1$ &c., we have" from (155)

$$\begin{aligned} \frac{dW}{d\xi_i{}'} = &- \frac{dZ}{dy_1}\frac{dy_1}{d\xi_i{}'} - \frac{dZ}{dy_2}\frac{dy_2}{d\xi_i{}'} - \ldots - \frac{dZ}{dy_n}\frac{dy_n}{d\xi_i{}'} \\ &+ x_1{}' \frac{dy_1}{d\xi_i{}'} + x_2{}' \frac{dy_2}{d\xi_i{}'} + \ldots + x_n{}' \frac{dy_n}{d\xi_i{}'} \\ &+ y_1 \frac{dx_1{}'}{d\xi_i{}'} + y_2 \frac{dx_2{}'}{d\xi_i{}'} + \ldots + x_n{}' \frac{dx_n{}'}{d\xi_i{}'}. \end{aligned} \tag{156}$$

From the canonical equation $x_j{}' = \frac{dZ}{dy_j}$ $(j = 1,\ldots n)$ it follows that the first and second lines on the right side of (156) cancel and so the latter becomes

$$\frac{dW}{d\xi_i'} = y_1 \frac{dx_1'}{d\xi_i'} + y_2 \frac{dx_2'}{d\xi_i'} + \cdots + x_n' \frac{dx_n'}{d\xi_i'}. \tag{157}$$

Because $x_j$ $(j = 1, \ldots n)$ is a function of $t$ and $\xi_1, \ldots, \xi_n$ there follows

$$x_j{}' = \frac{d(x_j)}{dt} + \frac{d(x_j)}{d\xi_1}\xi_1{}' + \frac{d(x_j)}{d\xi_2}\xi_2{}' + \ldots + \frac{d(x_j)}{d\xi_n}\xi_n{}'. \tag{158}$$

From (158) we obtain

$$\frac{dx_j'}{d\xi_i'} = \frac{d(x_j)}{d\xi_i} (j = 1, \ldots n, i = 1, \ldots n). \tag{159}$$

Combining (157) and (159) we have

$$\eta_i = \frac{dW}{d\xi_i'} = y_1 \frac{d(x_1)}{d\xi_i} + y_2 \frac{d(x_2)}{d\xi_i} + \cdots + y_n' \frac{d(x_n)}{d\xi_i}. \tag{160}$$

The right side of (160) is by (154) equal to $\frac{\mathrm{d}P}{\mathrm{d}\xi_i}$ and so we have finally

$$\eta_i = \frac{\mathrm{d}P}{\mathrm{d}\xi_i}. \tag{161}$$

Because (153) and (161) hold the modulus $P$ satisfies conditions (139). Hence $P$ gives rise to a normal transformation from $x_1, \ldots, x_n, y_1, \ldots, y_n$ to $\xi_1, , \ldots, \xi_n, \eta_1, , \ldots, \eta_n$.

### 4.3.4 Conclusion

Donkin's investigation took place against the background of a pronounced English interest in operational calculi and symbolic algebra, a fact that is apparent in his organization and presentation of the mathematical material. His approach differed in this respect from the analytical style of Continental perturbation theory and the variation of constants. Donkin's focus on Jacobi's theorem on canonical transformations also indicated an interest in the notion of mathematical invariance which is at the base of this result and was of central concern in contemporary English algebra.[36]

Given that Jacobi had died in 1851 and his *Vorlesungen and Ueber diejenigen Probleme* were not published until 1866, Jacobi and Donkin could not have known directly of each other's work. Nevertheless, there are some curious similarities in their development of the subject. Donkin's (1854/1855) overall approach was systematic, as was Jacobi's (1866a; b). The term "canonical" is a standard one in Donkin (1855), and it appears frequently in Jacobi (1866a; b), although Jacobi generally used it for variable elements rather than dynamical variables. (See also note 35 above.) Nevertheless, Jacobi (as did Donkin) recognized the scope of the notion of canonical form. In terms of general dynamical theory both Donkin and Jacobi (1866) put the canonical equations of motion at the base of their development of the subject, a fact that no doubt indicated the influence of Hamilton on each of them. Finally, Jacobi (1866a, b) differed from French writers in using square brackets for Poisson brackets. Donkin (1854/1855) also used square brackets.[37]

Following the publication of his dynamical papers, Donkin went on to write articles on mathematical astronomy and acoustics. He suffered from poor health and died

[36] On British symbolic algebra in the first part of the nineteenth century see Pycior (1981). On the emergence of invariant theory in British algebra in the 1840s see Wolfson (2008).

[37] Carl Borchardt was in possession of notes for Jacobi's Königsberg 1842–43 lectures, notes that were published in edited form as Jacobi (1866a). Wilhelm Scheibner had taken notes of Jacobi's 1847–48 Berlin lectures; his transcript would eventually be edited and published by Helmut Pulte as Jacobi (1996). Of course, Jacobi's contributions to canonical transformations were contained not in these works but in the set of notes that were published as Jacobi (1866b), but were probably written sometime around 1840. However, these notes do not appear to have undergone any detailed posthumous editing and apparently did not circulate. See also note 26 above.There are not copious biographical sources on Donkin. We do know that he displayed an early talent for languages (O'Connor and Robertson, 2021). It is worth noting that the term "canonical" does not appear at all in Donkin (1854), but is an established term in Donkin (1855). It would certainly be possible that there was some contact in (say) 1855 between Donkin and German mathematicians who had attended Jacobi's lectures or had some familiarity with his unpublished writings. But there is no record of such contact in print or apparently in any of the biographical sources we have.

in 1866 in Oxford at age 56. Were it not for the publication of Jacobi's *Vorlesungen* and *Ueber diejenigen Probleme* in 1866, Donkin's writings would likely have become a primary source for Hamilton–Jacobi methods. As reported earlier, his two papers were described in some detail by Cayley in his 1858 report. It is unclear to what degree Donkin's work was disseminated, given the more relaxed conventions for citation that seem to have prevailed in the nineteenth century. In 1873 Ernst Schering published an article for the Göttingen Royal Society of Sciences titled "Hamilton-Jacobische Theorie für Krafte" where he referred to Donkin's work. In 1900 Eduard von Weber wrote a substantial overview article on partial differential equations in the *Enzyklopädie der mathematischen Wissenschaften* where he included a section "Die Hamilton–Jacobi'sche Theorie". He referred to both Jacobi and Donkin and cited their work on canonical transformations, which he called "canonical substitutions." Von Weber mentioned the two men again in a subsequent section on the integration of first-order partial differential equations. In 1913 von Weber's essay was translated (with additional references) under the guidance of Édouard Goursat in the French edition of the *Enzyklopädie* volume on analysis. In the second edition of Whittaker's *Analytical Dynamics* Donkin is cited for his extension of the "kinematic potential" or Lagrangian considered by Hamilton to include time (Whittaker 1917, 264). (Whittaker did not mention that was also done by Jacobi (1837a, 1866a; b).)

## 5 Binet (1841) and Delaunay (1860), Tisserand (1868) and miscellaneous researches to 1890: celestial mechanics

### 5.1 Background

The subject of interest in the following researches is primarily lunar theory and to a lesser extent planetary theory. Lunar theory is an area of some complexity and one today that in its traditional form is generally not well known.[38] Our goal here is to give only a simplified outline of the subject as it was envisaged by Charles Delaunay and Tisserand (further details and history may be found in Brown (1896a, b), Brouwer and Clemence (1961), Fraser and Nakane (2003) and Wilson (2008)).

The earth is at the centre of the coordinate system and the motion of the moon about the earth is described by three second-order differential equations that measure the gravitational forces acting between the two bodies. The integration of these equations gives rise to six arbitrary constants or elements, that are normally expressed in terms of standard astronomical parameters. The action of the sun on the moon constitutes a perturbing force. According to the theory of variation, the arbitrary elements are now regarded as variable functions of time given in terms of the partial derivatives of a function measuring the sun's perturbing force. The solution of these equations, usually done by a method of successive approximation, leads to the desired solution of the lunar problem and provides the basis for the computation of lunar tables.

[38] In 1873 in an address to the Royal Astronomical Society Cayley noted that "everything in the Lunar theory is laborious" (Crilly 2006, 240).

## 5.2 Binet and Delaunay

Working in the theory of variation of elements Jacques Binet in an 1841 article in the *Journal de l'École Polytechnique* investigated perturbed elliptical motion. He referred to Hamilton (1834, 1835) and to Jacobi (1837a) and stated that Jacobi had elaborated Hamilton's results to arrive at his integration theorem. However, Binet did not refer to Jacobi's (1837b) note in the *Comptes rendus*, where Jacobi stated his theorem on canonical elements and theorem on canonical transformations. In contrast to Desboves and Donkin, Binet was not concerned with the general theory of transformations.

Binet showed that the Poisson-Lagrange canonical equations for the variable constants could be modified to include a more flexible selection of initial orbital elements. His solution for Keplerian unperturbed motion contained arbitrary constants $C$, $G$, $H$, $c$, $g$, $h$. These quantities would become variable in the perturbed system. Binet succeeded in deriving the canonical perturbation equations

$$\begin{aligned} \frac{\mathrm{d}C}{\mathrm{d}t} &= \frac{\mathrm{d}R}{\mathrm{d}c}, \frac{\mathrm{d}c}{\mathrm{d}t} = -\frac{\mathrm{d}R}{\mathrm{d}C} \\ \frac{\mathrm{d}G}{\mathrm{d}t} &= \frac{\mathrm{d}R}{\mathrm{d}g}, \frac{\mathrm{d}g}{\mathrm{d}t} = -\frac{\mathrm{d}R}{\mathrm{d}G}, \\ \frac{\mathrm{d}H}{\mathrm{d}t} &= \frac{\mathrm{d}R}{\mathrm{d}h}, \frac{\mathrm{d}h}{\mathrm{d}t} = -\frac{\mathrm{d}R}{\mathrm{d}H} \end{aligned} \tag{162}$$

where $R$ is the perturbation function and the variable elements determine the position of the orbit of the body.

Binet's approach was adopted by the Paris mathematician and astronomer Charles Delaunay in his investigation of the lunar theory. Delaunay wrote a memoir on this subject in 1846. Over the next two decades, he devoted himself entirely to the investigation of the moon's motion, publishing the results of his Herculean labours in 1860 and 1867 in two lengthy treatises in the memoirs of the Paris Academy of Sciences. He based his investigation on the Newtonian equations of motion involving the perturbing function. Delaunay modified Binet's equations to remove secular terms that increase indefinitely in magnitude with time while preserving at the same time the canonical form of the equations. He obtained the new equations

$$\begin{aligned} \frac{\mathrm{d}L}{\mathrm{d}t} &= \frac{\mathrm{d}F}{\mathrm{d}l}, \frac{\mathrm{d}l}{\mathrm{d}t} = -\frac{\mathrm{d}F}{\mathrm{d}L} \\ \frac{\mathrm{d}G}{\mathrm{d}t} &= \frac{\mathrm{d}F}{\mathrm{d}g}, \frac{\mathrm{d}g}{\mathrm{d}t} = -\frac{\mathrm{d}F}{\mathrm{d}G}, \\ \frac{\mathrm{d}H}{\mathrm{d}t} &= \frac{\mathrm{d}F}{\mathrm{d}h}, \frac{\mathrm{d}h}{\mathrm{d}t} = -\frac{\mathrm{d}F}{\mathrm{d}H} \end{aligned} \tag{163}$$

where $F$ is the perturbation function and the elements $L$, $G$, $H$, $l$, $g$, $h$ are known in the modern subject as the Delaunay canonical elements or variables.[39] The perturbation

[39] The meaning of these variables in terms of the lunar orbital parameters is as follows. Let $a$ be the semi-major axis, $e$ be the eccentricity, $i$ be the inclination of the orbit to a fixed plane, $\mu$ be the sum of the masses of the earth and moon, $l$ be the mean anomaly, $g$ be the angular distance of the lower apsis from the ascending node, $h$ be the longitude of the ascending node, then $L = \sqrt{\mu a}$, $G = L\sqrt{1-e^2}$, $H = Gcosi$.

function $F$ is equal to $R + \mu^2/2L^2$ ($\mu$ being the sum of the masses of the earth and moon) and was expanded by Delaunay as an infinite series.

Delaunay developed the perturbation function $F$ up to the seventh order of small parameters. The resulting expansion contained 320 terms, including terms multiplied by the cosine of a linear function. He employed a special procedure to eliminate secular terms from $F$. Delaunay repeated the procedure 57 times term by term, finally eliminating the important secular terms and integrating the resulting equations. The coordinates of the moon, its longitude, latitude and parallax, were then expressed as infinite series in which the time occurred only in the arguments of periodic terms. Although the practical use of his result was limited by the slow convergence of the series, it was remarkable that secular terms could be avoided by means of his method.

Following Delaunay's accidental death in 1873, his lunar theory received accolades from other astronomers. In 1876 the American George W. Hill (1876, 65) stated "among the innumerable set of canonical elements it does not appear that a better can be selected." Poincaré admired Delaunay's canonical elements and considered their introduction to be a significant contribution to celestial mechanics. The Delaunay variables and canonical transformations of these variables would be a subject taken up by later researchers, including Tisserand, Poincaré, and Charlier.[40]

### 5.3 Tisserand

François Tisserand's goal in his 1868 dissertation (published both as a free-standing work and as an article in Liouville's journal) was to derive and improve on Delaunay's main result using the methods of Hamilton and Jacobi.[41] He believed this theory provided effective tools of analysis and also led to new transformations of the elements. He emerged as a confirmed advocate of the Hamilton–Jacobi approach to dynamics and recommended its use in preference to the earlier methods of Lagrange and Poisson.

The only researchers mentioned by name in Tisserand's dissertation were Hamilton, Jacobi and Delaunay. There were no actual references to the literature. In the first section, Tisserand stated without proof the main theorems of Jacobi concerning canonical equations and transformations. The first theorem is simply the result that the equations of motion can be put in Hamiltonian canonical form. The second theorem is Jacobi's integration theorem, the third theorem is Jacobi's theorem on canonical elements, and the fourth theorem is Jacobi's transformation theorem. Jacobi is referred to by name only for the transformation theorem. With respect to the latter Tisserand (1868, 258) wrote, "There are infinitely many systems of canonical arbitraries; when one such system is known one may obtain from it infinitely many others by means of

[40] A limitation of Delaunay's theory was the slow convergence of his series. The standard lunar ephemerides published in Britain and United States in the 1920s were based on the different lunar theory of Hill and Ernest W. Brown (see Wilson (2010)). Delaunay's methods have, however, enjoyed a new lease on life with the advent since the 1950s of computer-assisted computation in applied celestial mechanics (artificial satellites, planetary perturbations), a development that is reviewed by Cook (1988, Chapter 9).

[41] The committee for Tisserand's doctoral examination was Delaunay as president with Joseph Serret and Charles Briot as examiners. The first thesis was the work published in Liouville's journal; the second thesis he defended was on astronomical refraction and did not apparently issue as a publication.

the following theorem of Jacobi.” Jacobi (1837a) had published a proof of his integration theorem; the theorem on canonical elements and the transformation theorem itself had been proved by Desboves and Donkin; these developments were documented by Cayley (1858); and Jacobi’s *Ueber diejenigen Probleme* itself appeared with full proofs of all these results in 1866. In fact, Tisserand was not at all concerned with the more theoretical mathematical aspects of Hamilton–Jacobi methods and devoted his efforts primarily to their application to lunar and planetary theory.

Tisserand proceeded in the next sections to derive canonical equations for the lunar elements. He began with the three Newtonian force equations in Cartesian coordinates $x$, $y$, $z$ and changed to a spherical system of coordinates $r$, $\psi$, $\varphi$. (See Jacobi (1866a, lecture 24) and our discussion in Sect. 3.6). Writing down the Hamilton–Jacobi equation for the problem, its integral $S$ was obtained in the form $S = -Ct + H\psi + f(r, \varphi, C, H, G)$, where $C$, $H$ and $G$ are constants of integration. An application of Jacobi’s theorem on canonical elements gives rise to a set of canonical variable elements $C$, $G$, $H$, $c$, $g$, $h$ and the associated set of canonical equations. The latter take the form

$$\begin{aligned} \frac{dC}{dt} &= \frac{dR}{dl}, \frac{dc}{dt} = -\frac{dR}{dL} \\ \frac{dG}{dt} &= \frac{dR}{dg}, \frac{dg}{dt} = -\frac{dR}{dG} \\ \frac{dH}{dt} &= \frac{dR}{dh}, \frac{dh}{dt} = -\frac{dR}{dH} \end{aligned}, \qquad (164)$$

where $R$ is the perturbation function. Hence Tisserand had derived the Delaunay variable elements and associated canonical equations in a new way using the methods of Hamilton and Jacobi.

Tisserand’s (1868, 283–285) next step in what was a very technical investigation involving shifting notation was to transform the elements $C$, $G$, $H$, $c$, $g$, $h$ to a new set of elements by means of Jacobi’s transformation theorem. He first expressed (164) in a revised form using elements designated as $C$, $(G)$, $(H)$, $\tau$, $(g)$, $(h)$. These variables and the associated canonical equations are transformed into a new set of variable $\Lambda'$, $G'$, $H'$; $\lambda$, $\kappa$, $\eta$ with associated canonical equations using a suitably selected generating function. The latter was somewhat complicated in form, and no explanation was given for how it was obtained. The canonical equations which resulted became the basis for eliminating secular terms in the series expressions for the orbital elements.

In the second part of the memoir Tisserand turned to a study of planetary three-body problems, of which the most prominent example was the sun-jupiter-saturn system. Such a system is different from the earth-moon-sun system, where the mass of one of the bodies is negligible in comparison with the other two. Nevertheless, the methods of Hamilton and Jacobi may be applied to planetary systems and lead to canonical equations involving twelve variable orbital elements. This analysis replicated the steps taken in the lunar theory, including the introduction (Tisserand 1868, 298) of a generating function and a canonical transformation to obtain the orbital elements in a form free of secular terms.

### 5.4 Miscellaneous researches up to the 1890s

A perennial theme in work on the three-body problem in the nineteenth century was the reduction of the degree of the system, that is lowering the number of variables required to characterize the motion of the bodies. Jacobi (1843) in an important paper had shown that the degree could be reduced from 18 to 6; however, he did not employ canonical methods in his analysis.

In his impressively documented 1899 survey of work on the three-body problem Edmund Whittaker identified several researchers in the second half of the century who wrote the dynamical equations for the orbital variables in canonical form. In doing so these researchers made use of integrals for quantities such as angular momentum and energy that were conserved. Their work showed an engagement at a concrete level with finding an analytical description of the system in canonical form. Bour (1856), Scheibner (1868), Radau (1870) and Vernier (1894), among others, contributed to this program of research. Whittkaker also drew attention to Delaunay (1860, 1867) and to Lie (1875, 282–286). Oddly enough he made no reference to Tisserand (1868) or to the Jacobi (1837b, 1866b) transformation theory itself, despite the attention the latter had received in Cayley's 1858 report.

Whittaker's (1899, 123) singled out Bour (1856) as a noteworthy early advance and provided a summary of his result. 18 equations in canonical form are reduced for the three-body system in the plane to 6, with coordinates $q_i, p_i (i = 1, 2, 3)$. The system involves the motion of two fictitious masses $m_1$ and $m_2$ whose potential energy depends on the lengths of the lines joining the masses to each other and the distances of the masses to the origin. $q_1, q_2$ are the distances of $m_1, m_{2,}$ respectively, to the origin. $q_3$ is the angle between $q_1$ and $q_2$, $p_1 = \frac{dq_1}{dt}$, $p_2 = \frac{dq_2}{dt}$, and $p_3$ is the is the difference of the angular momenta of $m_1$and $m_2$ about the origin. The equations that describe the motion are

$$\frac{dp_i}{dt} = \frac{\delta H}{\delta q_i}, \frac{dq_i}{dt} = -\frac{\delta H}{\delta p_i}(i = 1, 2, 3), \tag{165}$$

where $H$ is a function of the $q_i$, $p_i$ and $\delta$ denotes partial differentiation.

As we shall see in the next section, Poincaré's study of transformations originally developed out of work on the three-body problem. This is also true for Charlier (1902, 1907) and to some extent of Whittaker (1904).

## 6 Poincaré (1899, 1905, 1892): "changes of variables" and a new theory of transformations

### 6.1 Introduction

Henri Poincaré was a major mathematical figure at the end of the nineteenth century and the only mathematical scientist among the primary figures considered in this article for whom a modern biography has been written (see Gray (2014)).[42] His three-volume *Les Méthodes nouvelles de la mécanique céleste* stands out as a major contribution to mathematical science during the 1890s. It was complemented by his magisterial *Leçons de mécanique céleste*, published over several years of the new century. Volumes 1 and 3 of *Les Méthodes nouvelles* and volume 1 of *Leçons* included opening chapters that explored seminal mathematical methods in celestial mechanics, particularly ones related to Hamilton–Jacobi theory. Indeed, Poincaré was the key figure in the transmission of this subject from its historical origins in the writings of Hamilton and Jacobi to the widespread embrace of the theory by physicists and mathematicians in the first part of the twentieth century.

Poincaré composed his treatises in a more discursive way than one would write an article in a journal. The work was also more uneven than a traditional French *Cours d'analyse* (see for example Jordan (1896)). A sense of Poincaré's style is conveyed by some remarks of the mathematical astronomer Ernest Brown. In a letter to his friend George H. Darwin in 1901 Brown reported on his experience of reading Poincaré's *Méthodes nouvelles*:

> My respect for him grows with every chapter I read and annoyance at the same time. He is very careless—sometimes proofs are faulty—sometimes incomprehensible, and the number of misprints—some of them misleading—is fearful. I have occasionally spent three or four hours on a single page and then found that the difficulty arose from some misprint or misstatement. But it reads beautifully if one doesn't try to go too much into details. (Wilson 2010, 226)

*Methodes Nouvelles* developed out of courses Poincare conducted at the Sorbonne and he did not always attend to detailed textual references and documentation of sources. (On the provenance of this work see Goroff (1993, 113).) In volumes 1 and 3 of the *Méthodes nouvelles* in the opening chapters on dynamical theory there are few references to earlier authors. Textual citations do appear later in each work in relation to specific results in planetary theory. Even in the case of his own work he sometimes neglected to note that the particular subject at hand had been considered (sometimes in detail) by him in some earlier publication. One significant source was Tisserand's writings, especially his *Traité de Mécanique Cél*este, published just three years before the *Méthodes nouvelles*. Poincaré does not explicitly refer to Tisserand in the parts on dynamical theory, although his name comes up several times later in each volume.

Poincaré was familiar with Jacobi's results in dynamical analysis, but there is no evidence that he had consulted Jacobi's original writings. Rather he seems to have learned about the subject from Tisserand, particularly the latter's 1868 doctoral dissertation and subsequent writings in celestial mechanics. Insofar as Hamilton is concerned, little reference is made by Poincaré to the Irish mathematician, despite the great significance attributed to Hamilton by Jacobi throughout his dynamical writings.

[42] Leo Koenigsberger (1904) (then in his mid-60 s) wrote a very mathematical "life and letters" biography of Jacobi but no major biographical study has appeared since then. Also, there are modern biographies for secondary figures (who did not work on canonical transformations) such as Hamilton and Cayley. See Hankins (1980) for Hamilton and Crilly (2006). There are no biographies older or modern for Desboves, Donkin, Tisserand, Charlier and Whittaker.

In discussing Poincaré's work we shall follow his original notation. In particular, we use the same symbol '$d$' for both the derivative and the partial derivative. Although this had been the custom of earlier French and English authors (see Sect. 3.3 above), by the end of the century Poincaré was somewhat out of step with current notation. For example, the partial derivative symbol was used in Tisserand's 1892 *Traité* and appeared in mathematical works such as Camille Jordan's textbooks on analysis, published in the 1890s at the same time Poincaré was writing.

## 6.2 Volume one of the *Méthodes nouvelles* (1892)[43]

In his famous prize memoir on the three-body problem Poincaré (1890, 170–171) introduced examples of two-dimensional transformations that preserved the canonical form of the equations for the problem, although he did not use the term "canonical transformation" nor did he go into detail about the nature of such transformations.[44] In his later writings he referred to a "canonical change of variables" rather than "canonical transformation." It should be noted that the word "transformation" was used in earlier literature of the period to refer to a change of variables. For example, the title of Radau (1868) is "On a transformation of the differential equations of dynamics."

In Jacobi's dynamics and in modern mechanics one begins with the general Hamiltonian which is a function both of the time and the coordinate variables. The case where time does not appear in the Hamiltonian is then developed as a special case, one encountered in problems in astronomy and physics. By contrast, Poincaré from the outset took as primary Hamiltonians and equations of transformation that are independent of time. Although in the *Leçons* he did consider the more general case of time-dependent Hamiltonians, this subject occupied a limited place in his writings.

### 6.2.1 Jacobi's integration theorem and Jacobi's transformation theorem in *Les Méthodes Nouvelles*

Section one of *Les Méthodes nouvelles* (1892) begins with two sets of time-dependent variables $x_1, x_2, \ldots, x_p$ and $y_1, y_2, \ldots, y_p$. Poincaré (1892, 8) stated that the Hamiltonian or canonical equations have the form

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \frac{\mathrm{d}F}{\mathrm{d}y_i}, \frac{\mathrm{d}y_i}{\mathrm{d}t} = -\frac{\mathrm{d}F}{\mathrm{d}x_i}. \tag{166}$$

The Hamiltonian $F$ is a function of the $x_1, x_2, \ldots, x_p$ and $y_1, y_2, \ldots, y_p$. (Poincaré did not use the term "Hamiltonian" nor indeed make any reference to Hamilton in his

[43] English translations of passages in the *Méthodes nouvelles* are from the 1967 NASA English edition (Poincaré 1967). This translation is also used with some emendations in the American Institute of Physics edition of volume one (Poincaré 1993).

[44] Poincaré (1890) introduced such transformations on p.119, pp.170–171 and p.175; English translation (Poincaré 2017), p.104, pp.151–152 and p.155. For the history of this memoir see Barrow Green (1997).

account.) According to Poincaré, the variables $x_i$and$y_i$ are said to form a conjugate pair.[45]

In Sect. 3 Poincaré (1892, 13–14) turned to Jacobi's integration theorem, what he called "Jacobi's first theorem." As in the other results he presented, no proof is given. Consider a quantity $S$ that is a function of $x_1, x_2, \ldots, x_p$, where each of the $x_i$ depend on time. Poincaré asserted that the integration of the ordinary differential Eq. (166) reduces to solving the following partial differential equation for $S$:

$$F\left(x_1, x_2, \ldots, x_p, \ \frac{\mathrm{d}S}{\partial x_1}, \ldots, \ \frac{\mathrm{d}S}{\mathrm{d}x_p}\right) = h_1, \tag{167}$$

Here the variables $y_i$ in $F$ have been replaced by $y_i = \frac{\mathrm{d}S}{\mathrm{d}x_i}$. $h_1$ is a constant. In modern terminology (167) is the Hamilton–Jacobi partial differential equation for the problem. Let

$$S = S\big(x_1, x_2, \ldots, x_p, h_1, \ h_2, \ldots, h_p\big)$$

be a complete solution of (167) containing, in addition to $h_1$, the arbitrary constants $h_2, \ldots, h_p$. Poincaré (1892, 14) wrote "Jacobi has shown that the general integral of equations (166) can be written:

$$\frac{\mathrm{d}S}{\mathrm{d}x_i} = y_i(i = 1, 2, \ldots, p), \quad \frac{\mathrm{d}S}{\mathrm{d}h_i} = h_i'(i = 2, \ldots, p), \frac{\mathrm{d}S}{\mathrm{d}h_1} = t + h_1'. \tag{168}$$

The quantities $h_i'$ are a new set of $p$ arbitrary constants.

The theorem Poincaré presented here is not what one finds in lecture 20 of Jacobi's *Vorlesungen*, the latter being what is typically called Jacobi's theorem in modern textbooks. It is, in fact, the result formulated for conservative systems that Jacobi had presented in lecture 21 (see Sect. 3.6 above).

In Sect. 4, Poincaré (1892, 15) presented without proof Jacobi's transformation theorem, which he called "Jacobi's second theorem." He wrote.

Let

$$S\big(y_1, y_2, \ldots, y_p; h_1, h_2, \ldots, h_p\big)$$

be an arbitrary function of $p$ variables $y$ and of the new $p$ variables $h_i$. Let us now set

$$x_i = \frac{\partial S}{\partial y_i}, h_i{}' = \frac{\partial S}{\partial h_i} \tag{169}$$

45 In modern dynamics the conjugate $y_i$ to $x_i$ is defined as the partial derivative of the Lagrangian $L$ with respect to the time derivative of $x_i$: $y_i = \frac{\partial L}{\partial \dot{x}_i}$, where $\dot{x}_i = \frac{\mathrm{d}x}{\mathrm{d}t}$. Poincaré's terminology is consistent with this definition.

Equations (169) are regarded as defining the relation which connect the old variables[46]

$$x_1, x_2, \ldots, x_q,$$

$$y_1, y_2, \ldots, y_q,$$

to the new variables

$$h_1, h_2, \ldots, h_q,$$

$$h_1{}', h_2{}', \ldots, h_q{}'.$$

Jacobi has demonstrated that if we make this change of variables, the equations will remain canonical and do so whatever the function $S$ may be.

Although Poincaré had announced at the outset that he was considering results from Jacobi's *Vorlesungen*, the subject of transformations was not covered in these lectures. Rather it was taken up in Jacobi's supplementary treatise *Ueber diejenigen Probleme* published in the same volume as the *Vorlesungen.*[47] Furthermore, Jacobi had introduced transformations in his investigation of the constants of integration arising in the solution to the dynamical equations, whereas Poincaré was considering transformations of the dynamical variables themselves.

The function $S$ is here an arbitrary function and is unrelated to the function in the previous theorem, the latter being a solution of a partial differential equation. Later in Sect. 7, Poincaré (1892, 18) considered a function $S$ of $x_i$ and $h_i$. He wrote

> If, therefore, $S$ is an arbitrary function of $\left(x_1, x_2, \ldots, x_p, h_1, h_2, \ldots, h_p\right)$ and if we set
>
> $$y_i = \frac{\mathrm{d}S}{\mathrm{d}x_i}, h_i{}' = \frac{\partial S}{\partial h_i}, \tag{170}$$
>
> the canonical form of the equation will not be altered when we take $h_i$ and $h_i{}'$ as new variables, and when we change $H$ into –$H$ at the same time."

Although Poincaré did not give proof of these results, it is possible to reconstruct the likely reasoning that led to them. A significant clue is provided in Sect. 5 where he considered a canonical transformation that is a linear transformation from the variables $x_i$, $y_i$ to the variables $x_i{}'$, $y_i{}'$ (He abandoned here the notation $h_i$, $h_i{}'$ used elsewhere

[46] Poincaré switches the upper range of the variables from $p$ to $q$, an apparent misprint.

[47] The 1866 volume in which the *Vorlesungen* as well as *Ueber diejenigen Probleme* appeared was itself titled *Vorlesungen über Dynamik*. It is possible that Poincaré for this reason referred to the source of Jacobi's transformation theorem in this way. See also the discussion at the end of §6.5.3.

for the transformed variables. Also, the subscript $i$ now ranges from 1 to $n$.) The transformation is given as

$$\begin{cases} x_i = \alpha_{1.i}x_1{}' + \alpha_{2.i}x_2{}' + \cdots + \alpha_{n.i}x_n{}', \\ y_i = \beta_{1.i}y_1{}' + \beta_{2.i}y_2{}' + \cdots + \beta_{n.i}y_n{}'. \end{cases} \tag{171}$$

The question is how to select the constants $\alpha$ and $\beta$ so that the transformation is canonical.

To investigate this question Poincaré began with a general result that would be true for any transformation. He introduced what he called virtual increments and what were commonly known as variations. Given $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$ the increments or variations of these variables are denoted $\delta x_1, \delta x_2, \ldots, \delta x_n$ and $\delta y_1, \delta y_2, \ldots, \delta y_n$. We have the following expression for the variation of $F$:

$$\sum\left(\frac{\mathrm{d}F}{\mathrm{d}x_i}\delta x_i + \frac{\mathrm{d}F}{\mathrm{d}y_i}\delta y_i\right) = \delta F. \tag{172}$$

Using the canonical Eq. (166) Poincaré (1892, 16) rewrote (172) in the form

$$\sum\left(\frac{\mathrm{d}x_i}{\mathrm{d}t}\delta y_i - \frac{\mathrm{d}y_i}{\mathrm{d}t}\delta x_i\right) = \delta F. \tag{173}$$

He asserted that a necessary and sufficient condition for the transformation from $x_i, y_i$ to $x_i', y_i'$ be canonical is the validity of the following identity:

$$\sum\left(\frac{\mathrm{d}x_i}{\mathrm{d}t}\delta y_i - \frac{\mathrm{d}y_i}{\mathrm{d}t}\delta x_i\right) = \sum\left(\frac{\mathrm{d}x_i'}{\mathrm{d}t}\delta y_i' - \frac{\mathrm{d}y_i'}{\mathrm{d}t}\delta x_i'\right). \tag{174}$$

The proof (which Poincaré omitted) goes as follows. Since $F(x_i, y_i) = F(x_i', y_i')$ it follows from (173) and (174) that

$$\sum\left(\frac{\mathrm{d}x_i{}'}{\mathrm{d}t}\delta y_i{}' - \frac{\mathrm{d}y_i{}'}{\mathrm{d}t}\delta x_i{}'\right) = \delta F \tag{175}$$

But for $F = F(x_i', y_i')$ we have

$$\sum\left(\frac{\mathrm{d}F}{\mathrm{d}x_i{}'}\delta x_i{}' + \frac{\mathrm{d}F}{\mathrm{d}y_i{}'}\delta y_i{}'\right) = \delta F \tag{176}$$

Comparing (175) and (176) we obtain

$$\frac{\mathrm{d}x_i{}'}{\mathrm{d}t} = \frac{\mathrm{d}F}{\mathrm{d}y_i{}'}, \; \frac{\mathrm{d}y_i{}'}{\mathrm{d}t} = -\frac{\mathrm{d}F}{\mathrm{d}x_i{}'}, \tag{177}$$

and it follows that the transformation is canonical. The converse is straightforward.

Poincaré proceeded to apply this general result to the canonical linear transformation (171). We infer from (171) that $dx_i$ depends only on $dx_i'$, $\delta x_i$ only on $\delta x_i'$, $dy_i$ only on $dy_i'$ and $\delta y_i$ only on $\delta y_i'$. Hence, we conclude from (174) that

$$\sum \mathrm{d}x_i \delta y_i = \sum \mathrm{d}x_i' \delta y_i', \sum \mathrm{d}y_i \delta x_i = \sum \mathrm{d}y_i' \delta x_i', \tag{178}$$

Because the relations (171) are linear, it follows that $dx_i$ is to $dx_i'$ as $x_i$ is to $x_i'$, $\delta x_i$ is to $\delta x_i'$ as $x_i$ is to $x_i'$, $dy_i$ is to $dy_i'$ as $y_i$ is to $y_i$' and $\delta y_i$ is to $\delta y_i'$ as $y_i$ is to $y_i'$. Equation (178) therefore remains valid if $dx_i$ and $\delta x_i$ are replaced by $x_i$; $dy_i$ and $\delta y_i$ by $y_i$; $dx_i'$ and $\delta x_i'$ by $x_i'$; and $dy_i'$ and $\delta y_i'$ by $y_i'$. In this case (178) reduces to

$$\sum x_i y_i = \sum x_i{}' y_i{}'. \tag{179}$$

Poincaré inferred that (179) is a necessary and sufficient condition for the transformation (171) to be canonical.

Poincaré next considered a narrower set of linear transformations. We first suppose that $\alpha_{k.i} = \beta_{k.i}$ in (171) so that $x_i$ and $y_i$ are subjected to the same transformation. Furthermore, the coefficients are required satisfy the following relations:

$$\sum_{i=1}^{i=n} \alpha_{ki}^2 = 1, \sum_{i=1}^{i=n} \alpha_{ki}\alpha_{hi} = 0 (k \neq h). \tag{180}$$

A transformation with these properties is said to be orthogonal. (A rotation is an example of such a transformation.) It is not difficult to see that such a transformation (179) holds and the transformation is therefore canonical.

In Sect. 6, Poincaré presented another example. Assume that we have two variables $x_1$ and $y_1$ and a transformation to $x_1'$ and $y_1'$ given by:

$$x_1 = \varphi\left(x_1{}', y_1'\right), y_1 = \psi\left(x_1{}', y_1'\right) \tag{181}$$

Poincaré (1892, 18) wrote "the equations will remain canonical, I say, provided that the functional determinant, or Jacobian, of $x_1$ and $y_1$, with respect to $x_1'$ and $y_1'$ is equal one." He did not give any proof, but it would be necessary to show that the identity (179) holds. He also indicated that a transformation from the variables $q$ and $p$ to $\omega$and$\rho$ defined by

$$q = \sqrt{2\rho} \cos \omega, p = \sqrt{2\rho} \sin \omega \tag{182}$$

has a Jacobian equal to 1 and the canonical form of the equations is preserved. (In this example it would be necessary to show that (174) holds.)

Although Poincaré did not do so, it is possible to go from the general identity (174) to a proof of Jacobi's theorem on canonical transformations stated in Sects. 4 and 7. Consider the theorem in Sect. 4. We have $S = S\left(y_1, y_2, \ldots, y_p; h_1, h_2, \ldots, h_p\right)$ and the transformation from $x_i$, $y_i$ to $h_i$, $h_i'$ is given by (169). One begins with a general identity involving the interchangeability of the differential and variational operations:

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta S = \delta \frac{\mathrm{d}S}{\mathrm{d}t}. \tag{183}$$

Expanding the left side of (183) gives

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\delta S &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}S}{\mathrm{d}y_1}\delta y_1 + \ldots + \frac{\mathrm{d}S}{\mathrm{d}y_p}\delta y_p + \frac{\mathrm{d}S}{\mathrm{d}h_1}\delta h_1 + \cdots + \frac{\mathrm{d}S}{\mathrm{d}h_p}\delta h_p\right)\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\Big(x_1\delta y_1 + \ldots + x_p\delta y_p + h_1'\delta h_1 + \ldots + h_p'\delta h_p\Big),\end{aligned} \tag{184}$$

where we have used (169). We now carry out the differentiation of the right side of this equation and use the canonical Eq. (166) to obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\delta S = &\frac{\mathrm{d}x_1}{\mathrm{d}t}\delta y_1 + x_1\delta\frac{\mathrm{d}y_1}{\mathrm{d}t} + \ldots + \frac{\mathrm{d}x_p}{\mathrm{d}t}\delta y_p + x_p\delta\frac{\mathrm{d}y_p}{\mathrm{d}t}\\ &+ \frac{\mathrm{d}h_1'}{\mathrm{d}t}\delta h_1 + h_1'\delta\frac{\mathrm{d}h_1}{\mathrm{d}t} + \ldots + \frac{\mathrm{d}h_p'}{\mathrm{d}t}\delta h_p + h_p'\delta\frac{\mathrm{d}h_p}{\mathrm{d}t}.\end{aligned} \tag{185}$$

Expanding next to the right side of (183) and using (169) again we have

$$\begin{aligned}\delta\frac{\mathrm{d}S}{\mathrm{d}t} &= \delta\left(\frac{\mathrm{d}S}{\mathrm{d}y_1}\frac{\mathrm{d}y_1}{\mathrm{d}t} + \cdots + \frac{\mathrm{d}S}{\mathrm{d}y_p}\frac{\mathrm{d}y_p}{\mathrm{d}t} + \frac{\mathrm{d}S}{\mathrm{d}h_1}\frac{\mathrm{d}h_1}{\mathrm{d}t} + \cdots + \frac{\mathrm{d}S}{\mathrm{d}h_p}\frac{\mathrm{d}h_p}{\mathrm{d}t}\right)\\ &= \delta\left(x_1\frac{\mathrm{d}y_1}{\mathrm{d}t} + \cdots + x_p\frac{\mathrm{d}y_p}{\mathrm{d}t} + h_1'\frac{\mathrm{d}h_1}{\mathrm{d}t} + \cdots + h_p'\frac{\mathrm{d}h_p}{\mathrm{d}t}\right)\\ &= \frac{\mathrm{d}y_1}{\mathrm{d}t}\delta x_1 + x_1\delta\frac{\mathrm{d}y_1}{\mathrm{d}t} + \cdots + \frac{\mathrm{d}y_p}{\mathrm{d}t}\delta x_p + x_p\delta\frac{\mathrm{d}y_p}{\mathrm{d}t}.\end{aligned} \tag{186}$$

Equating (185) and (186) gives rise after some simplification to

$$\begin{aligned}&\frac{\mathrm{d}x_1}{\mathrm{d}t}\delta y_1 - \frac{\mathrm{d}y_1}{\mathrm{d}t}\delta x_1 + \cdots + \frac{\mathrm{d}x_p}{\mathrm{d}t}\delta y_p - \frac{\mathrm{d}y_p}{\mathrm{d}t}\delta x_p\\ &\quad= \frac{\mathrm{d}h_1}{\mathrm{d}t}\delta h_1' - \frac{\mathrm{d}h_1'}{\mathrm{d}t}\delta h_1 + \cdots + \frac{\mathrm{d}h_p}{\mathrm{d}t}\delta h_p' - \frac{\mathrm{d}h_p'}{\mathrm{d}t}\delta h_p.\end{aligned} \tag{187}$$

But (187) is simply the fundamental identity (174). Hence it follows that the transformation is canonical.

The transformation theorem involving the generating function $S(x_i, h_i)$ in Sect. 7 may be proved in the same way. Later, in the *Leçons*, Poincaré (1905) would give derivations of these results along similar lines, but use methods that did not make use of the symbolic $\delta$ notation. We will return to this subject in Sect. 6.5.1 below.

### 6.2.2 Keplerian motion

In Sect. 8 Poincaré (1892, 19–21) considered what he called "Keplerian motion" involving a mass $M$ that moves under the action of an inverse-square force located at the origin. His account in both notation and substance followed some earlier work of Tisserand ((1868) and (1889, 123–127)); Tissserand in turn followed (with some changes in notation) Jacobi's (1866a, 183–186) formulation and treatment of the problem in lecture 24 of the *Vorlesungen*. (Tisserand made no reference to Jacobi's lecture and Poincaré made no reference to either Jacobi or Tisserand.[48]) The equations of motion are given in canonical form with the Hamiltonian:

$$F = \frac{y_1^2 + y_2^2 + y_3^2}{2} - \frac{2M}{\sqrt{x_1^2 + x_2^2 + x_3^2}}, \tag{188}$$

where $x_1, x_2, x_3$ are the coordinates of $M$, $y_1, y_2, y_3$ are its velocity components and the units are selected in such a way that the Gaussian constant is 1.[49] The partial differential Eq. (167) is

$$\left(\frac{\mathrm{d}S}{\mathrm{d}x_1}\right)^2 + \left(\frac{\mathrm{d}S}{\mathrm{d}x_2}\right)^2 + \left(\frac{\mathrm{d}S}{\mathrm{d}x_3}\right)^2 - \frac{2M}{\sqrt{x_1^2 + x_2^2 + x_3^2}} = 2h, \tag{189}$$

where $h$ is an arbitrary constant. Poincaré introduced spherical coordinates

$$x_1 = r \sin\omega \cos\varphi, x = r \sin\omega \sin\varphi, x_3 = r \cos\omega, \tag{190}$$

and (189) becomes

$$\left(\frac{\mathrm{d}S}{\mathrm{d}r}\right)^2 + \frac{1}{r^2}\left(\frac{\mathrm{d}S}{\mathrm{d}\omega}\right)^2 + \frac{1}{r^2 \sin^2\omega}\left(\frac{\mathrm{d}S}{\mathrm{d}\varphi}\right)^2 = \frac{2M}{\sqrt{x_1^2 + x_2^2 + x_3^2}} + 2h \tag{191}$$

Poincaré's separated the variables to obtain a complete solution of (191) of the form $S = S(r, \omega, \varphi, G, \Theta, h)$ or $S = S(x_1, x_2, x_3, G, \Theta, h)$, where $G$, $\Theta$, $h$ are arbitrary constants. Jacobi's first theorem then gives the general solution of the canonical equations as

$$y_i = \frac{\mathrm{d}S}{\mathrm{d}x_i}\ (i = 1, 2, 3), \quad h' + t = \frac{\mathrm{d}S}{\mathrm{d}h}, \quad g = \frac{\mathrm{d}S}{\mathrm{d}G}, \quad \theta = \frac{\mathrm{d}S}{\mathrm{d}\Theta} \tag{192}$$

48 The chapter in which the problem appears in Tisserand (1889) is Chapter VII, "Intégration des équations différentielles du mouvment elliptique par la méthode de Jacobi". While the method is attributed to Jacobi the solution itself is Tisserand's own invention.

49 According to Brown (1896a, b, 1) the Gaussian constant of attraction is defined as follows. For any two masses $m$ and $m$' a distance $r$ apart the gravitational force between them is $F = k\frac{mm'}{r^2}$. $k$ is the Gaussian constant of attraction. The units may be chosen so that $k = 1$.

where $g, \theta, h'$ are new arbitrary constants. Poincaré provided more astronomically convenient expressions involving these constants and related them to the orbital elements which were used in celestial mechanics.

Poincaré (1892, 20–21) next extended the analysis to more general situations. He wrote, "If the moving mass, instead of being subject to the attraction of the mass $M$, were subject to other forces, we could, nevertheless, construct the function $S$ and define six new variables $L$, $G$, $\Theta$, $l$, $g$, $\theta$.

as a function of $x_i$ and $y_i$ by equations.

$$y_i = \frac{dS}{dx_i} \ (i = 1, 2, 3), \quad \frac{dS}{dL} = l, \quad \frac{dS}{dG} = g, \quad \frac{dS}{d\Theta} = \theta, \tag{193}$$

$L$, $G$, $\Theta$, $l$, $g$, and $\theta$ would no longer be constants." Poincaré named these six quantities Keplerian variables.

In Sect. 10, Poincaré (1892, 24–26) considered the central-force motion of a mass $m$ with coordinates $x_1, x_2, x_3$ and velocity components $y_1, y_2, y_3$. The force need not be restricted to an inverse-square force and is given in terms of a general potential function $V$. Poincaré began by setting down the canonical differential equations for the problem, where the Hamiltonian $F$ is given as

$$F = \frac{1}{2}\left(p_1^2 + p_2^2 + p_3^2\right) + V, \tag{194}$$

with $p_i = my_i$. As before, the function $S = S(x_1, x_2, x_3, G, \Theta, L)$ is a complete solution of the partial differential equation for the problem. Equation (193) give rise to a change of variables from $x_1, x_2, x_3, y_1, y_2, y_3$ to the Keplerian variables $L$, $G$, $\Theta$, $l$, $g$, $\theta$. Referring to Sect. 7, Poincaré asserted that the transformed equations will also be canonical, and have the form

$$\begin{aligned} &\tfrac{dL}{dt} = -\tfrac{dF}{dL}, \ \tfrac{dG}{dt} = -\tfrac{dF}{dg}, \ \tfrac{d\Theta}{dt} = -\tfrac{dF}{d\theta}, \\ &\tfrac{dl}{dt} = \tfrac{dF}{dL}, \ \tfrac{dg}{dt} = \tfrac{dF}{dG}, \ \tfrac{d\theta}{dt} = \tfrac{dF}{d\Theta}. \end{aligned} \tag{195}$$

Equation (195) are simply Delaunay's equations, the established basis of the theory of lunar and planetary perturbations, obtained here from the canonical equations of motion by means of a canonical transformation.[50]

In the present problem $S$ is regarded as a generating function (to use a modern term) that gives a canonical transformation from the original dynamical coordinates and momenta $x_1, x_2, x_3, y_1, y_2, y_3$ to the variable constants $L$, $G$, $\Theta$, $l$, $g$, $\theta$. It should be noted that Poincaré's development is very different from Jacobi's procedure. Jacobi (1866b) had obtained canonical equations for the variable constants from his theorem on canonical elements (Theorem IX of Jacobi (1866b)) before transformations were even introduced. By contrast, Poincaré in the case at hand derived these equations using $S$ as a generating function in an application of Jacobi's theorem on canonical transformations (Theorem X of Jacobi (1866b)).

[50] In terms of the Delaunay elements identified in note 39, Poincaré is setting $\mu = 1$.

The idea of using a solution of a partial differential equation as the generating function of a canonical transformation was first introduced in published form by Desboves in 1848. It was also successfully deployed in a fairly modern way in the final sections of Jacobi's *Ueber diejenigen Probleme* (1866b), as we saw in Sect. 3.7.3 above. The idea is fundamental to the modern subject. Although Poincaré in the analysis above used the complete solution $S$ as a generating function, his approach is different from the modern one. In the latter, this move is part of a derivation of Jacobi's integration theorem. By contrast, Poincaré's goal was to effect a change of variables that lead to canonical equations for the Keplerian elements. A more general difference is that Poincaré was applying the method of variation of arbitrary constants, whereas the modern theory is developed independently of this method. As we shall see in Sect. 6.5.3, in the *Leçons* Poincaré would use a canonical transformation to prove Jacobi's integration theorem in a way that is similar to modern derivations.

### 6.3 Poincaré's 1897 paper on the three-body problem

In 1896 and 1897 Poincaré published papers with the same title, "Sur une forme nouvelle des équations du problème des trois corps." Both contained results on canonical transformations. The 1897 paper was the more complete of the two and is the one we shall focus on in this account.

Following other celestial mechanicians of the period, Poincaré sought a transformation of variables that would reduce the degrees of freedom in a three-body problem. He tried to find transformations that were linear and canonical, and for which the equation of areas (angular momentum) was preserved. He formulated the three-body problem as follows.

Let $A$, $B$, and $C$ be three bodies whose coordinate are: $A$: $(x_1, x_2, x_3)$, $B$: $(x_4, x_5, x_6)$, $C$: $(x_7, x_8, x_9)$. The mass of $A$ is denoted $m_1 = m_2 = m_3$; the mass of $B$, $m_4 = m_5 = m_6$; and the mass of $C$, $m_7 = m_8 = m_9$. The momentum components are then $p_i = m_i \frac{\mathrm{d}x_i}{\mathrm{d}t}$. The canonical Eq. (166) consists of 18 equations. Consider a transformation from the variables $x_i$, $y_i$ to the new variables $x_i{}'$, $y_i{}'$. Poincaré asserted that the canonical form of (166) will be preserved in the new variables if and only if the following differential form is exact:

$$\sum x_i' dy_i' - \sum x_i \mathrm{d}y_i. \tag{196}$$

If the integrating function for this form is $S = S\left(y_i, y_i'\right)$ then the exactness condition implies that,

$$x_i = -\frac{\mathrm{d}S}{\mathrm{d}y_i}, x_i{}' = \frac{\mathrm{d}S}{\mathrm{d}y_i'}, \tag{197}$$

which is the same condition (169) he had given in 1892.

It is not clear why Poincaré here adopted the terminology of exact forms rather than the generating functions he had used in 1892. A possible advantage of expressing the condition that the transformation is canonical using an exact form is that the coefficients in the form show the relations involving partial derivatives which connect the old and new variables. There also seems to have been a tendency among researchers to develop dynamics using ideas from differential geometry, and Poincaré's approach was in keeping with this approach.

Poincaré in the paper went on to examine in some detail canonical transformations with particular properties related to angular momentum and the centre of gravity of three bodies, but we will not follow him in these investigations.

## 6.4 The mature theory: volume three of the *Méthodes nouvelles* (1899)

### 6.4.1 Proof (1899) of Jacobi's transformation theorem

In 1896 Poincaré published a short paper that analyzed the properties of periodic solutions of the three-body problem using the principle of least action. The paper examined the existence of such solutions when the force law is an inverse cube or higher power of the mutual distances. Poincaré introduced what he called the Hamiltonian action

$$J = \int_{t_0}^{t} (T + U) dt \tag{198}$$

where $T$ is the kinetic energy and $U$ is minus the potential energy ($T + U$ is the Lagrangian, a term not used by Poincaré). The assertion that $J$ is a minimum is what Poincaré called the principle of least action and what modern writers call Hamilton's principle.

Chapter 29 of Volume 3 of *Les Méthodes nouvelles* (1899) contains a derivation of Jacobi's theorem on canonical transformations from Hamilton's principle. Although some of the ideas had arisen several years earlier in volume one, the approach adopted here is more rigorous and better grounded in basic variational theory.

We have given the variables $x_i$ and $y_i$ ($i = 1, \ldots, n$) and a function $F$ of these variables. Poincaré introduced the action integral $J$,

$$J = \int_{t_0}^{t_1} \left( -F + \sum y_i \frac{\mathrm{d}x_i}{\mathrm{d}t} \right) dt \tag{199}$$

(In the usual dynamical case the integrand in the integral on the right of (199) is simply the Lagrangian $T + U$.) The variation of $J$ is

$$\delta J = \int_{t_0}^{t_1} \left( -\delta F - \sum \delta y_i \frac{\mathrm{d}x_i}{\mathrm{d}t} - \sum y_i \frac{\mathrm{d}\delta x_i}{\mathrm{d}t} \right) dt. \tag{200}$$

Poincaré asserted that in order for the variation of $J$ to vanish it is necessary that the canonical Eq. (166) hold:

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \frac{\mathrm{d}F}{\mathrm{d}y_i}, \frac{\mathrm{d}y_i}{\mathrm{d}t} = -\frac{\mathrm{d}F}{\mathrm{d}x_i}. \tag{201}$$

In fact, the latter are simply the standard Euler–Lagrange variational equations corresponding to the integrand in (199), regarded as a function of the $2n$ variables $x_i$ and $y_i$ and their derivatives. It is a basic theorem of the calculus of variations that (166) follow from the condition $\delta J = 0$. It is not clear whether Poincaré expected the reader to derive these equations directly from (200), or if he was assuming that the reader was familiar with the variational theorem. (Indeed, it is not entirely clear if he himself was thinking of the theorem.) Rather surprisingly, this seems to have been the first time (201) were derived in this simple and natural way.

Poincaré proceeded to give proof of Jacobi's transformation theorem. One has a change of variables from $x_i$ and $y_i$ to $x_i'$ and $y_i'$. The new and old variables are connected using the generating function $S = S(x_i, x_i')$ by means of the exact differential identity

$$\sum y_i' dx_i' - \sum y_i \mathrm{d}x_i = dS. \tag{202}$$

Poincaré set

$$J' = \int_{t_0}^{t_1} \left( -F + \sum y_i' \frac{\mathrm{d}x_i'}{\mathrm{d}t} \right) dt. \tag{203}$$

From (199), (203) and (202) we deduce that

$$J' - J = \int \frac{\mathrm{d}S}{\mathrm{d}t} dt = S_{t_1} - S_{t_0}. \tag{204}$$

Hence there follows

$$\delta J' = \delta J + [\delta S]_{t=t_0}^{t=t_1} \tag{205}$$

It is supposed in the variational process that $\delta x_i = 0$ at $t_0$ and $t_1$ and so we have $[\delta S]_{t=t_1}^{t=t_0} = 0$. Hence it follows from (205) and $\delta J = 0$ that $\delta J' = 0$. The Euler–Lagrange variational equations corresponding to $\delta J' = 0$ are evidently:

$$\frac{\mathrm{d}x_i'}{\mathrm{d}t} = \frac{\mathrm{d}H}{\mathrm{d}y_i'}, \quad \frac{\mathrm{d}y_i'}{\mathrm{d}t} = -\frac{\mathrm{d}F}{\mathrm{d}x_i} \tag{206}$$

and the transformation is canonical.

After some discussion of the theory when the limits $t_0$ and $t_1$ are variable, Poincaré turned to the primary dynamical case where $F = T - U$. $T$ is a second-degree homogenous function in the variables $y_\mathrm{i}$ and $U$ is a function of $x_i$. Poincaré stated that from Hamilton's principle it follows that the integral

$$\int_{t_0}^{t_1} (T + U) dt \tag{207}$$

is a minimum. Poincaré denoted the Lagrangian $T + U$ by $H$ and wrote Lagrange's equations in the form[51]

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial H}{\partial x_i'} = \frac{\partial H}{\partial x_i}. \tag{208}$$

In this context, Poincaré introduced the idea of a kinetic focus and applied it to the investigation of periodic solutions using ideas from his papers of 1896 and 1897. In the calculus of variations, a conjugate point is the point at which the second variation is equal to zero and marks the limiting value at which there is no longer an actual extremum, but simply a stationary solution to the variational equations. "Kinetic focus" was the name for the conjugate point in problems in variational dynamics. Although not directly related to the theory of transformations as such, Poincaré's interest in this subject is indicative of the prominent place at this time of the calculus of variations in his dynamical researches.

### 6.4.2 Comparison of Jacobi and Poincaré

It is of interest to compare Jacobi's proof of the transformation theorem (set out above in Sect. 3.7.2) with Poincaré's 1899 derivation of the same result. Jacobi did not explain how he came up with the method of generating functions for producing canonical transformations; the genesis of his ideas is unknown. In his account, the theorem is a result about differential equations and involves no appeal to anything in the calculus of variations, whether the latter is viewed as the study of optimization or simply as a formal subject involving the analytical comparison of families of functions. Jacobi's proof deploys the operations of ordinary and partial differentiation and nothing more than these. A restriction on Jacobi's result was the time independence of the generating function, required in the proof. (In 1907 Charlier would make the necessary additions to Jacobi's proof to extend it to time-dependent generating functions.)

By contrast, Poincaré's derivation uses variational ideas and laws at a fundamental level. There is an underlying family of comparison functions and a process of continuous variation from one functional curve to another. The motion of the system is governed by a variational principle. The structural character of the proof gives it a deceptive simplicity that is lacking in Jacobi's algorithmic and direct approach. Poincaré's proof also allowed for time-dependent generating functions. It is the simplicity of Poincaré's proof that accounts for its popularity in modern textbook literature, where it typically appears as something that has arisen from the general contemplation of the subject; there is never any indication that it was Poincaré's invention.[52]

---

51 It should be noted that the symbol "'∂" as it was understood by Poincaré is not quite the partial derivative in the modern sense. Assume we have the function *f(x,y)* and the function $g(x,x')$, where $x' = \frac{\mathrm{d}x}{\mathrm{d}t}$ is the time derivative of $x$. Poincaré wrote the partial derivative of $f$ with respect to $x$ as $\frac{\mathrm{d}f}{\mathrm{d}x}$ and the partial derivative of $g$ with respect to x as $\frac{\partial g}{\partial x}$. The "∂ rond" is used when $x$ occurs in an expression that also contains *x'* but the "*d*" is used otherwise.

52 In his excelllent introduction to the revised English translation of Volume One of *Les méthodes nouvelles* (that also extends to later aspects of Poincaré's work), Daniel Goroff (1993, I26-27) presents Poincaré's

### 6.5 Poincaré's *Leçons de mécanique céleste* (*1905*)

In his three-volume *Leçons de mécanique céleste* (1905–1909) Poincaré published another systematic account from a mathematical viewpoint of celestial mechanics. In the introduction to this work he compared it to his earlier *Méthodes nouvelles* and to Félix Tisserand's multi-volume *Traité de mécanique céleste* (1889–1896). He stated that in comparison to the *Méthodes nouvelles*, the *Leçons* would be more expository and less focused on a detailed study of technical questions such as convergence. However, the work still included subjects primarily of mathematical interest. For readers who were concerned with astronomical details he recommended Tisserand's work.[53] He also emphasized that he would not attend to historical aspects of the development of celestial mechanics, an aspect again said to be covered by Tisserand.

#### 6.5.1 Jacobi's transformation theorem

Poincaré's account of mechanical principles and equations is set out in the first chapter of volume one (1905) of the *Leçons*. Here he provided a more detailed investigation of the theory related to Jacobi's transformation theorem that he had presented in 1892 at the beginning of volume one of his *Méthodes nouvelles*. He neglected his influential development of the theory in 1899 in volume 3, as if the latter had never existed. The proof he develops of the transformation theorem is different from the one six years earlier.

Poincaré first considered $2n$ variables $x_1, \ldots, x_n,\ y_1, \ldots, y_n$ and a function $F$ of these variables. We are given the canonical equations

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \frac{\mathrm{d}F}{\mathrm{d}y_i}, \quad \frac{\mathrm{d}y_i}{\mathrm{d}t} = -\frac{\mathrm{d}F}{\mathrm{d}x_i}. \tag{209}$$

The solution of these equations gives $x_i$ and $y_i (i = 1, \ldots n)$ as functions of $t$ and $2n$ constants of integration $\alpha_1, \alpha_2, \ldots, \alpha_{2n}$. In the derivation, he would make use of differentiation with respect to the constants of integration $\alpha_k$. It should be noted that differentiation with respect to $\alpha_k$ corresponds to the $\delta$ operation Poincaré had employed in 1892 in section five of *Méthodes nouvelles*, discussed above in Sect. 6.2.1.[54]

At this point, Poincaré began to write in terms of $x$ and $y$, where these variables (without subscripts) stand in for $x_i$ and $y_i$ $(i = 1, \ldots, n)$. He set down the identity

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum x\frac{\mathrm{d}y}{\mathrm{d}\alpha_k} - \frac{\mathrm{d}}{\mathrm{d}\alpha_k}\sum x\frac{\mathrm{d}y}{\mathrm{d}t} = \sum \frac{\mathrm{d}x}{\mathrm{d}t}\frac{\mathrm{d}y}{\mathrm{d}\alpha_k} - \sum \frac{\mathrm{d}y}{\mathrm{d}t}\frac{\mathrm{d}x}{\mathrm{d}\alpha_k}, \tag{210}$$

Footnote 52 continued
1899 derivation of Jacobi's theorem on canonical transformations. However, Goroff does not point out that the proof is seldom if ever attributed to Poincaré in modern textbooks.

53 For a modern account of one of Tisserand's (1896) results in his investigation of the three-body problem see Murray and Dermott (2012, 71–73).

54 For example, Eq. (6.7) from *Méthodes Nouvelles* becomes Eq. (6.45) in the *Leçons*.

where the summation is taken over $x$ and $y$ and $k = 1, 2, \ldots, 2n$. (210) follows from the identity

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}y}{\mathrm{d}\alpha_k} = \frac{\mathrm{d}}{\mathrm{d}\alpha_k}\frac{\mathrm{d}y}{\mathrm{d}t}. \tag{211}$$

Now

$$\frac{\mathrm{d}F}{\mathrm{d}\alpha_k} = \sum \frac{\mathrm{d}F}{\mathrm{d}y}\frac{\mathrm{d}y}{\mathrm{d}\alpha_k} + \sum \frac{\mathrm{d}F}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}\alpha_k}. \tag{212}$$

From (212) and (227) it follows that

$$\frac{\mathrm{d}F}{\mathrm{d}\alpha_k} = \sum \frac{\mathrm{d}x}{\mathrm{d}t}\frac{\mathrm{d}y}{\mathrm{d}\alpha_k} + \sum -\left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)\frac{\mathrm{d}x}{\mathrm{d}\alpha_k}. \tag{213}$$

From (210) and (213) we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum x\frac{\mathrm{d}y}{\mathrm{d}\alpha_k} - \frac{\mathrm{d}}{\mathrm{d}\alpha_k}\sum x\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\mathrm{d}F}{\mathrm{d}\alpha_k}. \tag{214}$$

The $2n$ Eq. (214) follow from the $2n$ canonical Eq. (166).

Conversely, assume the system of Eq. (214) hold. Then it follows from (211) and (212) that

$$\sum\left(\frac{\mathrm{d}x}{\mathrm{d}t} - \frac{\mathrm{d}F}{\mathrm{d}y}\right)\frac{\mathrm{d}y}{\mathrm{d}\alpha_k} - \sum\left(\frac{\mathrm{d}y}{\mathrm{d}t} + \frac{\mathrm{d}F}{\mathrm{d}x}\right)\frac{\mathrm{d}x}{\mathrm{d}\alpha_k} = 0, \tag{215}$$

for $k = 1, \ldots, 2n$. These $2n$ equations are linear with respect to the $2n$ unknowns

$$\frac{\mathrm{d}x}{\mathrm{d}t} - \frac{\mathrm{d}F}{\mathrm{d}y}, -\left(\frac{\mathrm{d}y}{\mathrm{d}t} + \frac{\mathrm{d}F}{\mathrm{d}x}\right). \tag{216}$$

Poincaré (1905, 3) asserted "the determinant of these $2n$ equations (215) which is nothing but the functional determinant of $x$ and $y$ with respect to $\alpha$ cannot be identically zero." Hence the system of canonical Eq. (227) follows from (215). Poincaré concluded that the system of Eq. (227) is fully equivalent to the system of Eq. (214).

Poincaré proceeded to state and prove Jacobi's transformation theorem. We consider a transformation from $x$, $y$ to the new variables $x'$, $y'$, defined in terms of the exact form[55]

$$\sum x'dy' - \sum xdy = dS. \tag{217}$$

55 (6.50) evidently expresses the fact that the generating function $S$ is a function of $y$ and $y$' and that the following relations hold.

$$\frac{\partial S}{\partial y} = -x, \frac{\partial S}{\partial y'} = x'.$$

Poincaré asserted that the Eq. (227) expressed in terms of the new variables will take the canonical form

$$\frac{\mathrm{d}x_i'}{\mathrm{d}t} = \frac{\mathrm{d}F}{\mathrm{d}y_i'}, \frac{\mathrm{d}y_i'}{\mathrm{d}t} = -\frac{\mathrm{d}F}{\mathrm{d}x_i'}. \tag{218}$$

Hence the transformation from $x$, $y$ to $x'$, $y'$ preserves the canonical form of (227). To prove this result Poincaré derived from (217) the two equations

$$\sum x'\frac{\mathrm{d}y'}{\mathrm{d}\alpha_k} - \sum x\frac{\mathrm{d}y}{\mathrm{d}\alpha_k} = \frac{\mathrm{d}S}{\mathrm{d}\alpha_k}, \tag{219}$$

$$\sum x'\frac{\mathrm{d}y'}{\mathrm{d}t} - \sum x\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\mathrm{d}S}{\mathrm{d}t}.$$

Using the equality $\frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}S}{\mathrm{d}\alpha_k} = \frac{\mathrm{d}}{\mathrm{d}\alpha_k}\frac{\mathrm{d}S}{\mathrm{d}t}$ we obtain from (219) the identity

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum x\frac{\mathrm{d}y}{\mathrm{d}\alpha_k} - \frac{\mathrm{d}}{\mathrm{d}\alpha_k}\sum x\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\sum x'\frac{\mathrm{d}y'}{\mathrm{d}\alpha_k} - \frac{\mathrm{d}}{\mathrm{d}\alpha_k}\sum x'\frac{\mathrm{d}y'}{\mathrm{d}t}. \tag{220}$$

From (214) and (220) there follows

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum x'\frac{\mathrm{d}y'}{\mathrm{d}\alpha_k} - \frac{\mathrm{d}}{\mathrm{d}\alpha_k}\sum x'\frac{\mathrm{d}y'}{\mathrm{d}t} = \frac{\mathrm{d}F}{\mathrm{d}\alpha_k}. \tag{221}$$

According to the result Poincaré had proved earlier, the validity of (221) is equivalent to the validity of the canonical Eq. (218).

As we noted above, Poincaré's proof here of the transformation theorem is different from his proof in 1899 (see Sect. 6.4.1 above). In the latter, variational laws played a fundamental role and the canonical equations in the transformed variables were obtained from Hamilton's principle, where they appeared as the Euler–Lagrange variational equations. The 1905 result, first presented in outline form in 1892, evidently expresses Poincaré's original line of thinking about the transformation theorem. It is direct if somewhat involved proof with no reference to variational theory, and in this respect is similar to the ones of Jacobi and Donkin.[56]

### 6.5.2 Point transformations are canonical

In this section and the next two, we look at the different ways that Poincaré deployed the transformation theorem to arrive at significant dynamical results. The derivations are original and not altogether typical of how the same results are obtained today and are noteworthy for this reason.

56 Modern textbooks standardly derive Jacobi's transformation theorem using the 1899 proof. However, an exception is ter Haar (1964, 99–100), where the derivation follows the ideas of Poincaré (1905), although ter Haar's account is less rigorous and complete than Poincaré's. Another exception is Brouwer and Clemence (1961, 531–533) whose proof bears some similarities to those of Jacobi (1866b) and Charlier (1907).

In Sect. 7 Poincaré considered a free system in space of $n/3$ particles, and designated the coordinates of the bodies of the system by $x_i$, so that $i$ goes from 1 to $n$. We have $m_1 = m_2 = m_3$, $m_4 = m_5 = m_6$ and so on. The kinetic energy $T$ and the potential energy $U$ are introduced (Poincaré's terms!), and one obtains the total energy $F = T + U$. (The notation has changed here from the *Méthodes nouvelles* (1899), where $U$ was minus the potential energy and the total energy was $F = T - U$.) We introduce the variable $y_i$ defined as $y_i = m_i \frac{\mathrm{d}x_i}{\mathrm{d}t}$. The kinetic energy $T$ is given as

$$T = \frac{1}{2} \sum m_i \left( \frac{\mathrm{d}x_i}{\mathrm{d}t} \right)^2. \tag{222}$$

Written in term of $y_i$ (222) becomes

$$T = \frac{1}{2} \sum \frac{y_i^2}{m_i}. \tag{223}$$

It follows that

$$\frac{\mathrm{d}T}{\mathrm{d}y_i} = \frac{y_i}{m_i} = \frac{\mathrm{d}x_i}{\mathrm{d}t}. \tag{224}$$

Also we have the force equation

$$m_i \frac{\mathrm{d}^2 x_i}{\mathrm{d}t^2} = -\frac{\mathrm{d}U}{\mathrm{d}x_i}, \tag{225}$$

which may be rewritten as

$$\frac{\mathrm{d}y_i}{\mathrm{d}t} = -\frac{\mathrm{d}U}{\mathrm{d}x_i}. \tag{226}$$

In the total energy $F = T + U$, $T$ is a function of $x_i'$ alone and $U$ is a function of $x_i$ alone. Hence $\frac{\mathrm{d}F}{\mathrm{d}x_i} = \frac{\mathrm{d}U}{\mathrm{d}x_i}$ and $\frac{\mathrm{d}F}{\mathrm{d}y_i} = \frac{\mathrm{d}T}{\mathrm{d}y_i}$. Thus Eq. (224) and (170) become

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \frac{\mathrm{d}F}{\mathrm{d}y_i}, \frac{\mathrm{d}y_i}{\mathrm{d}t} = -\frac{\mathrm{d}F}{\mathrm{d}x_i}, \tag{227}$$

and the differential equations of the system are canonical.

In Sect. 8 Poincaré examined a change of variables from $x_1, x_2, \ldots, x_n$ to a new set of coordinates $q_1, q_2, \ldots, q_n$, where the old and new coordinates are related by the equations

$$x_i = \varphi_i(q_1, q_2, \ldots, q_n). \tag{228}$$

Such a transformation is known in the modern subject as a point transformation. The derivatives $\frac{\mathrm{d}x_i}{\mathrm{d}t}$ and $\frac{\mathrm{d}q_i}{\mathrm{d}t}$ of $x_i$ and $q_i$ are now designated as $\frac{\mathrm{d}x_i}{\mathrm{d}t} = x_i'$ and $\frac{\mathrm{d}q_i}{\mathrm{d}t} = q_i'$.

(This is a departure from Poincaré's practice elsewhere in the *Leçons* and his other works, where the prime notation is used to denote the transform of a variable.) The conjugate variable $p_i$ is defined as

$$\frac{\mathrm{d}T}{\mathrm{d}q_i'} = p_i. \tag{229}$$

Poincaré showed that Eq. (166) expressed in terms of the new variables $q_i$ and $p_i$ remain canonical. It follows that Hamilton's equations hold for any set of coordinate variables $q_i$ related to $x_i$ by Eq. (228). Put another way, point transformations are canonical. This result had been proved by Donkin (1855, Sects. 69–70) in a somewhat different way fifty years earlier (see our account above in Sect. 4.3.2).

Poincaré's proof begins with the identity

$$\mathrm{d}x_i = \sum \frac{\mathrm{d}\varphi_i}{\mathrm{d}q_k}\mathrm{d}q_k, \tag{230}$$

or

$$x_i' = \sum \frac{\mathrm{d}\varphi_i}{\mathrm{d}q_k} q_k'. \tag{231}$$

Poincaré posited an increment $\mathrm{d}q'$ in which $q'$ is varied but $q$ remains constant and reasoned that $x$ (which depends only on $q$) will not vary, but $x'$ will vary and there will result an increment $\mathrm{d}x'$. Hence we obtain from (231) the equation

$$\mathrm{d}x_i{}' = \sum \frac{\mathrm{d}\varphi_i}{\mathrm{d}q_k}\mathrm{d}q_k{}'. \tag{232}$$

There is the general identity

$$dT = \sum \frac{\mathrm{d}T}{\mathrm{d}q'}dq' = \sum \frac{\mathrm{d}T}{\mathrm{d}x'}dx', \tag{233}$$

or

$$\sum pdq' = \sum ydx'. \tag{234}$$

Comparing (230) and (232), Poincaré inferred that $\mathrm{d}x'$ has the same relation to $\mathrm{d}q'$ as $\mathrm{d}x$ has to $\mathrm{d}q$. Hence in the identity (234) we can replace $\mathrm{d}x'$ by $\mathrm{d}x$ and $\mathrm{d}q'$ by $\mathrm{d}q$ and obtain

$$\sum pdq = \sum ydx. \tag{235}$$

This equation may in turn be re-expressed in the form

$$\sum q\mathrm{d}p - \sum x\mathrm{d}y = \mathrm{d}\Big(\sum pq - \sum yx\Big). \tag{236}$$

Equation (236) is an exact differential with generating function given in terms of $y$ and $p$. Poincaré concluded that the change of variables from $x_i, y_i$ to $q_i, p_i$ is canonical and we have

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = \frac{\mathrm{d}F}{\mathrm{d}p_i}, \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\mathrm{d}F}{\mathrm{d}q_i}. \tag{237}$$

### 6.5.3 Jacobi's integration theorem obtained from a canonical transformation

In the three volumes of the *Méthodes nouvelle*s Poincaré never actually proved Jacobi's integration theorem ("Jacobi's first theorem"), despite the prominent place of this result in his development of the theory. In Sect. 10 of the *Leçons*, he derived it for conservative systems by means of a canonical transformation. This form of proof would become standard in later textbooks, particularly among physicists, although one also finds in the modern literature the older derivation going back to Jacobi's (1837a) early published findings in the theory.[57]

We are given the function $F = F(x_i, y_i)$. Let $S$ be an unknown function, implicitly taken to be a function of $x_1, x_2, \ldots, x_n$. Replace $y_i$ by $\frac{\mathrm{d}S}{\mathrm{d}x_i}$ in $F$. We set the result equal to a constant and obtain the Hamilton–Jacobi partial differential equation in $S$ and $x_i$:

$$F\left(x_i, \frac{\mathrm{d}S}{\mathrm{d}x_i}\right) = \text{const}, \tag{238}$$

which Poincaré called simply "Jacobi's equation."

Poincaré considered what he called a "particular" solution $S$ to (238) containing the $n$ arbitrary constants $\beta_1, \ldots, \beta_n$. ($S$ is today more commonly called a complete solution.) The constant on the right in (238) is then a function $\varphi$ of these $n$ quantities and this equation becomes

$$F\left(x_i, \frac{\mathrm{d}S}{\mathrm{d}x_i}\right) = \varphi(\beta_1, \beta_2, \ldots, \beta_n). \tag{239}$$

$S$ is a function of $x_i$ and $\beta_i$. Hence, we have

$$dS = \sum \frac{\mathrm{d}S}{\mathrm{d}x} dx + \sum \frac{\mathrm{d}S}{\mathrm{d}\beta} d\beta. \tag{240}$$

We now posit the following relations:

$$\frac{\mathrm{d}S}{\mathrm{d}x_i} = y_i, \frac{\mathrm{d}S}{\mathrm{d}\beta_i} = \gamma_i. \tag{241}$$

[57] Courant and Hilbert (1962, 107–109, 129–131) give both proofs of Jacobi's integration theorem. Gelfand and Fomin (1963, 91–93) also give both proofs.

Equation (241) gives $2n$ relations among the four variables $x_i$, $y_i$, $\beta_i$, $\gamma_i$. They allow us to express $\beta_i$ and $\gamma_i$ as functions of $x_i$, $y_i$, and conversely, to express $x_i$ and $y_i$ as functions of $\beta_i$ and $\gamma_i$.

Combining (240) and (241) gives

$$dS = \sum ydx + \sum \gamma d\beta, \tag{242}$$

or,

$$\sum \gamma d\beta - \sum xdy = d\Big(S - \sum xy\Big), \tag{243}$$

which is an exact differential. Equation (243) gives rise to a change of variables from $x, y$ to $\beta, \gamma$. By Jacobi's transformation theorem this change of variables is canonical. Hence we have the equations

$$\frac{\mathrm{d}\gamma_i}{\mathrm{d}t} = \frac{\mathrm{d}F}{\mathrm{d}\beta_i}, \frac{\mathrm{d}\beta_i}{\mathrm{d}t} = -\frac{\mathrm{d}F}{\mathrm{d}\gamma_i}. \tag{244}$$

From (239), Eq. (244) reduce to

$$\frac{\mathrm{d}\gamma_i}{\mathrm{d}t} = \frac{\mathrm{d}\varphi}{\mathrm{d}\beta_i}, \frac{\mathrm{d}\beta_i}{\mathrm{d}t} = 0. \tag{245}$$

Hence, we find that the $\beta_i$ are constant and so the $\frac{\mathrm{d}\varphi}{\mathrm{d}\beta_i}$ are also constant. It follows that (245) integrates to

$$\beta_i = \text{constant}, \gamma_i = \frac{\mathrm{d}\varphi}{\mathrm{d}\beta_i} t + \omega_i, \tag{246}$$

where the $\omega_i$ are $n$ new constants of integration. A solution to the original canonical Eq. (227) is obtained by substituting (246) into (241) which leads to the desired expressions for $x_i$ and $y_i$.

Poincaré's solution is (setting aside matters of presentation) the same as the one given in Jacobi (1866b, 467–468) in Sect. 41 of his *Ueber diejenigen Probleme* (discussed by us in Sect. 3.7.2). There is no indication that Poincaré had read Jacobi; he appears to have learned about the theorem from its statement in Tisserand (1868) and Tisserand's later writings; the proof was apparently his own invention. It should be noted that from the viewpoint of Hamilton–Jacobi theory the proof is a fairly natural one.

### 6.5.4 Changes of variables that involve the time

In his writings on mechanics up to 1905 Poincaré restricted his attention to functions $F = F(x_i, y_i)$ that are independent of time. (He did not in any of his writings use the modern name Hamiltonian for $F$.) However, in Sect. 12 of the *Leçons* he generalized his analysis to the case where time is an explicit variable in $F$, a problem that leads for transformations to the consideration of generating functions that contain the time. Poincaré's idea was to reduce the case of time-dependent functions $F$ to the theory developed earlier by regarding time formally as another dependent variable in the problem.

The function $F$ under consideration is a function of $x_i$, $y_i$ and the time $t$: $F(x_i, y_i; t)$. We introduce the new variable $u = t$. Consider the modified function $F$':

$$F' = F(x_i, y_i; u) + v. \tag{247}$$

Here $v$ may be regarded as a variable that is conjugate to $u$, in the same way that the $x_i$ are conjugate to $y_i$ (Poincaré does not use the word "conjugate.") In addition to $x_i, y_i$ we now have the two auxiliary dependent variables $u$ and $v$. The canonical equations corresponding to $x_i, y_i, u, v$ are

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}t} &= \frac{\mathrm{d}F'}{\mathrm{d}y}, \frac{\mathrm{d}y}{\mathrm{d}t} = -\frac{\mathrm{d}F'}{\mathrm{d}x}, \\ \frac{\mathrm{d}u}{\mathrm{d}t} &= \frac{\mathrm{d}F'}{\mathrm{d}v}, \frac{\mathrm{d}v}{\mathrm{d}t} = -\frac{\mathrm{d}F'}{\mathrm{d}u}. \end{aligned} \tag{248}$$

Because $\frac{\mathrm{d}F}{\mathrm{d}x} = \frac{\mathrm{d}F'}{\mathrm{d}x}$ and $\frac{\mathrm{d}F}{\mathrm{d}y} = \frac{\mathrm{d}F'}{\mathrm{d}y}$ Eq. (248) are simply the original canonical Eq. (227). The first equation in (248) gives $\frac{\mathrm{d}u}{\mathrm{d}t} = 1$ and so $u = t$, as assumed. The second equation in (248) leads to the equation

$$F(x_i, y_i; u) + v = \text{constant}. \tag{249}$$

Consider now a change of variables from $x_i, y_i, t$ to $x_i{}', y_i{}', t$. (Poincaré has returned to his normal practice of using the prime symbol to denote the transform of a variable.) Suppose that the following exact differential identify holds:

$$\sum x'dy' - \sum xdy = dS + Wdt, \tag{250}$$

where $dS$ is an exact differential and $W$ is a function of $x$, $y$ and $t$ (or alternatively of $x'$, $y'$ and $t$). Setting $v' = v + W$ and replacing $t$ by $u$ in (250) we obtain

$$\left(\sum x'\mathrm{d}y' + u\mathrm{d}v'\right) - \left(\sum x\mathrm{d}y + u\mathrm{d}v\right) = \mathrm{d}(S + uW). \tag{251}$$

Consider now a change from the variables $x_i$, $y_i$, $u$, $v$ to $x_i'$, $y_i'$, $u$, $v'$. By Jacobi's theorem on canonical transformations, the validity of (251) implies that the canonical Eq. (248) expressed in terms of $x_i'$, $y_i'$, $u$, $v'$ remain canonical. Now $F' = F + v = F + v' - W$. Hence

$$\frac{\mathrm{d}F'}{\mathrm{d}x'} = \frac{\mathrm{d}(F-W)}{\mathrm{d}x'}, \frac{\mathrm{d}F'}{\mathrm{d}y'} = -\frac{\mathrm{d}(F-W)}{\mathrm{d}y'}, \tag{252}$$

where we have used the fact that the partial derivatives $\frac{\mathrm{d}(v')}{\mathrm{d}x'}$ and $\frac{\mathrm{d}(v')}{\mathrm{d}y'}$ are zero. The canonical Eq. (6.80a) are therefore given as

$$\frac{\mathrm{d}x'}{\mathrm{d}t} = \frac{\mathrm{d}(F-W)}{\mathrm{d}y'}, \frac{\mathrm{d}y'}{\mathrm{d}t} = -\frac{\mathrm{d}(F-W)}{\mathrm{d}x'}. \tag{253}$$

(Note that from (250) we have (in modern notation) $W = -\frac{\partial S}{\partial t}$, and so the transformed Hamiltonian (to use the modern term) takes the form $F + \frac{\partial S}{\partial t}$.)

Poincaré concluded the section by showing how the Hamilton–Jacobi equation may be introduced using the original canonical Eq. (248). As noted above, Eq. (249) follows from (6.80b). We now set

$$\frac{\mathrm{d}S}{\mathrm{d}x_i} = y_i, \frac{\mathrm{d}S}{\mathrm{d}u} = v, \tag{254}$$

and substitute into (249), obtaining

$$F\left(x_i, \frac{\mathrm{d}S}{\mathrm{d}x_i}, u\right) + \frac{\mathrm{d}S}{\mathrm{d}u} = \text{constant}. \tag{255}$$

Replacing $u$ by $t$ in (255) we arrive at Jacobi's equation (the *H–J* equation) for time-dependent functions $F$:

$$F\left(x_i, \frac{\mathrm{d}S}{\mathrm{d}x_i}, t\right) + \frac{\mathrm{d}S}{\mathrm{d}t} = \text{constant}. \tag{256}$$

## 7 Charlier: bracket methods (1902) and generating functions (1907)

Carl Charlier's two-volume work *Die Mechanik der Himmels* (1902, 1907) contributed substantially to the theory of canonical transformations. Working partly from Jacobi (1866a, b), Charlier also drew on Poincaré's writings and those of other researchers. A notable aspect of his research that distinguished it from his contemporaries was the use of Hamilton–Jacobi theory to investigate conditionally periodic motions. In particular, Charlier employed separability methods to integrate the Hamilton–Jacobi partial differential equation, drawing on some results of Paul Stäckel (1893). German physicists in the decade following 1907 took note of this part of Charlier's work, a development that contributed to the adoption of Hamilton–Jacobi methods in the old quantum theory. (See Hund (1974, 86–88), Darrigol (1992, 110–111), Mehra and Rechenberg (1982, 228) and Shore (2003).)

### 7.1 Canonical Substitutions (1902)

In 1900 the theory of the variation of constants remained a prominent subject in the analysis of planetary motion. The theory was taken up and investigated in some detail by Charlier in Chapter 6 "Perturbation theory" of *Die Mechanik der Himmels* (1902). Here he introduced canonical transformations in a way that was different from Jacobi. At this point, Charlier used the word "substitution" rather than "transformation" and employed the bracket methods originally developed in the French school. Jacobi's notion of a generating function did not appear in the theory that he developed.

In the first section of chapter 6, Charlier (1902, 289—296) investigated substitutions or transformations which preserve the canonical form of the dynamical equations. The latter are given in the form

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \frac{\partial F}{\partial y_i}, \frac{\mathrm{d}y_i}{\mathrm{d}t} = -\frac{\partial F}{\partial x_i}(i = 1, 2, \ldots, n) \tag{257}$$

where $F = F(x_1, \ldots, x_n, y_1, \ldots, y_n, t)$. The variables $x_i$, $y_i$ are connected to a second set of variables $\xi_i$, $\eta_i$ by the relations

$$x_i = f_i(\xi_1, \ldots, \xi_n; \eta_1, \ldots, \eta_n), y_i = g_i(\xi_1, \ldots, \xi_n; \eta_1, \ldots, \eta_n). \tag{258}$$

The goal is to find conditions so that the transformed equations are also canonical:

$$\frac{\mathrm{d}\xi_i}{\mathrm{d}t} = \frac{\partial F}{\partial \eta_i}, \frac{\mathrm{d}\eta_i}{\mathrm{d}t} = -\frac{\partial F}{\partial \xi_i}(i = 1, 2, \ldots, n) \tag{259}$$

Charlier introduced the bracket designation

$$[a, b] = \sum_{s=1}^{n}\left(\frac{\partial a}{\partial x_s}\frac{\partial b}{\partial y_s} - \frac{\partial a}{\partial y_s}\frac{\partial b}{\partial x_s}\right) \tag{260}$$

which is the Poisson bracket of $a$ and $b$, although Charlier did not use this term. He first presented the identity

$$\frac{\mathrm{d}\xi_i}{\mathrm{d}t} = \sum_{s=1}^{n}\left(\frac{\partial \xi_i}{\partial x_s}\frac{\mathrm{d}x_s}{\mathrm{d}t} + \frac{\partial \xi_i}{\partial y_s}\frac{\mathrm{d}y_s}{\mathrm{d}t}\right) = \sum_{s=1}^{n}\left(\frac{\partial \xi_i}{\partial x_s}\frac{\partial F}{\partial y_s} - \frac{\partial \xi_i}{\partial y_s}\frac{\partial F}{\partial x_s}\right) = [\xi_i, F]. \tag{261}$$

Since

$$\frac{\partial F}{\partial y_s} = \sum_{k=1}^{n}\left(\frac{\partial F}{\partial \xi_k}\frac{\partial \xi_k}{\partial y_s} + \frac{\partial F}{\partial \eta_k}\frac{\partial \eta_k}{\partial y_s}\right), \frac{\partial F}{\partial x_s} = \sum_{k=1}^{n}\left(\frac{\partial F}{\partial \xi_k}\frac{\partial \xi_k}{\partial x_s} + \frac{\partial F}{\partial \eta_k}\frac{\partial \eta_k}{\partial x_s}\right) \tag{262}$$

we have the identity

$$\frac{\partial \xi_i}{\partial x_s}\frac{\partial F}{\partial y_s} - \frac{\partial \xi_i}{\partial y_s}\frac{\partial F}{\partial x_s} = \sum_{k=1}^{n}\frac{\partial F}{\partial \xi_k}\left(\frac{\partial \xi_i}{\partial x_s}\frac{\partial \xi_k}{\partial y_s} - \frac{\partial \xi_i}{\partial y_s}\frac{\partial \xi_k}{\partial x_s}\right) + \sum_{k=1}^{n}\frac{\partial F}{\partial \eta_k}\left(\frac{\partial \xi_i}{\partial x_s}\frac{\partial \eta_k}{\partial y_s} - \frac{\partial \xi_i}{\partial y_s}\frac{\partial \eta_k}{\partial x_s}\right) \tag{263}$$

Then, from (261) one obtains

$$\frac{\mathrm{d}\xi_i}{\mathrm{d}t} = [\xi_i, F] = [\xi_i, \xi_1]\frac{\partial F}{\partial \xi_1} + \cdots + [\xi_i, \xi_n]\frac{\partial F}{\partial \xi_n} + [\xi_i, \eta_1]\frac{\partial F}{\partial \eta_1} + \cdots + [\xi_i, \eta_n]\frac{\partial F}{\partial \eta_n}. \tag{264}$$

Similarly we find that

$$\frac{\mathrm{d}\eta_i}{\mathrm{d}t} = [\eta_i, F] = [\eta_i, \xi_1]\frac{\partial F}{\partial \xi_1} + \cdots + [\eta_i, \xi_n]\frac{\partial F}{\partial \xi_n} + [\eta_i, \eta_1]\frac{\partial F}{\partial \eta_1} + \cdots + [\eta_i, \eta_n]\frac{\partial F}{\partial \eta_n}. \tag{265}$$

It is evident from definition (260) that for all $i$, $[\xi_i, \xi_i] = [\eta_i, \eta_i] = 0$. Suppose also that the following conditions are satisfied:

$$[\xi_i, \xi_r] = 0, [\eta_i, \eta_r] = 0, [\xi_i, \eta_r] = 0 (i \neq r), [\xi_i, \eta_i] = +1 (i, r = 1, 2, \ldots, n). \tag{266}$$

Then it follows that Eqs. (264) and (265) reduce to (259) and the given change of variables is canonical. Equation (266) are conditions that give rise to a canonical transformation.

Brackets were employed in this way by Desboves (1848), and Jacobi (1866b), although these authors considered Lagrange rather than Poisson brackets. Jacobi had used bracket methods in his proof of Lagrange's fundamental identity and in his theorem on canonical elements. (See Sect. 3.7.1 above). The derivation of canonical transformations by means of bracket methods would be taken up by some later authors, including Rademacher and Iglisch (1925) and Brouwer and Clemence (1961, Chapter XVII). The disadvantage of this approach is that it requires looking at a given transformation and then checking if it satisfies the requisite conditions to confirm that it is canonical. By contrast, the method of generating functions enables one to obtain a canonical transformation by taking any function whatsoever as the generator.

Charlier proceeded to give some examples of canonical substitutions that were useful in the theory of perturbations. The first example is the simplest and involves the Delaunay variable elements $(L, G, H, l, g, h)$ (see Sect. 5.2). A change of variables beginning with these elements was reasonably standard in the literature and appeared for example in Brown (1896a, b, 134).[58] The Delaunay elements satisfy the canonical equations

$$\begin{aligned} \frac{\mathrm{d}L}{\mathrm{d}t} &= \frac{\mathrm{d}F}{\mathrm{d}l}, \frac{\mathrm{d}l}{\mathrm{d}t} = \frac{\mathrm{d}F}{\mathrm{d}L} \\ \frac{\mathrm{d}G}{\mathrm{d}t} &= \frac{\mathrm{d}F}{\mathrm{d}g}, \frac{\mathrm{d}g}{\mathrm{d}t} = \frac{\mathrm{d}F}{\mathrm{d}G} \\ \frac{\mathrm{d}H}{\mathrm{d}t} &= \frac{\mathrm{d}F}{\mathrm{d}h}, \frac{\mathrm{d}h}{\mathrm{d}t} = \frac{\mathrm{d}F}{\mathrm{d}H} \end{aligned} . \tag{267}$$

58 Charlier did not give any references to the literature.

Charlier transformed these elements to a new set $(\Lambda, \Gamma, Z, \lambda, \gamma, z)$ by means of the relations

$$\begin{array}{c}\Lambda = L, \Gamma = L - G, Z = G - H \\ \lambda = l + g + h, \gamma = -g - h, z = -h\end{array}. \tag{268}$$

In terms of the notation above, we have $x_1 = L$, $x_2 = G$, $x_3 = H$, $y_1 = l$, $y_2 = g$, $y_3 = h$ and $\xi_1 = \Lambda, \xi_2 = \Gamma, \xi_3 = Z, \eta_1 = \lambda, \eta_2 = \gamma, \eta_3 =$ z. Then $[\xi_i, \xi_r] = 0$ because $\xi_i, \xi_r$ are given only in terms of the $x_j$ and so $\frac{\partial \xi_j}{\partial y_k} = 0$ for all $j$ and $k$. Similarly, $[\eta_i, \eta_r] = 0$ because $\eta_i, \eta_r$ are given only in term of $y_j$ and so $\frac{\partial \eta_j}{\partial x_k} = 0$ for all $j$ and $k$. A simple calculation shows further that $\left[\xi_i, \eta_j\right] = 0$ if $i \neq j$ and $[\xi_i, \eta_i] = 1$ for all $i$. Hence conditions (266) are satisfied and the transformation from $(L, G, H, l, g, h)$ to $(\Lambda, \Gamma, Z, \lambda, \gamma, z)$ is canonical. Charlier concluded his discussion by giving expressions for $(\Lambda, \Gamma, Z, \lambda, \gamma, z)$ in terms of the elliptical orbital elements.[59]

Charlier and Poincaré before him recognized that canonical transformations provide an effective tool for generating new sets of variable elements in a form suitable for orbital analysis. Charlier next transformed the elements $(\Lambda, \Gamma, Z, \lambda, \gamma, z)$ to a new set $(\Lambda, \xi, p, \lambda, \eta, q)$ given as

$$\begin{array}{c}\Lambda = \Lambda, \xi = \sqrt{2\Gamma}\cos\gamma, p = \sqrt{2Z}\cos z \\ \lambda = \lambda, \eta = \sqrt{2\Gamma}\sin\gamma, q = \sqrt{2Z}\sin z\end{array}. \tag{269}$$

Poincaré (1892, 18) had presented a similar example (see Eq. (183) above). A calculation reveals that $[\xi, \xi] = 0$, $[\eta, \eta] = 0$. Furthermore

$$\begin{aligned}[\xi, \eta] &= \frac{\partial \xi}{\partial \Gamma}\frac{\partial \eta}{\partial \gamma} - \frac{\partial \xi}{\partial \gamma}\frac{\partial \eta}{\partial \Gamma} \\ &= \frac{1}{\sqrt{2\Gamma}} cos\gamma \bullet \sqrt{2\Gamma} cos\gamma - \sqrt{2\Gamma}(-sin\gamma) \bullet \frac{1}{\sqrt{2\Gamma}} sin\gamma \\ &= cos^2\gamma + sin^2\gamma = 1\end{aligned} \tag{270}$$

Similarly, we find that $[p,p] = [q,q] = 0$ and $[p,q] = 1$. Hence conditions (266) are satisfied and the transformation (269) is canonical.

Charlier referred to $(\Lambda, \xi, p, \lambda, \eta, q)$ as Poincaré elements because they had been first introduced by Poincaré (1892, Sect. 12); they are known by this name in the modern subject (Brouwer and Clemence 1961, 540). Goroff (1993, 125) suggests that Poincaré selected these elements because they are advantageous when the eccentricities and inclinations of the planets are small. Poincaré had transformed the elements using trigonometric functions and inferred that the resulting transformations were canonical

[59] The expressions for the Delaunay elements in terms of the elliptical orbital parameters were given in note 39 above. Charlier provided a similar set of expressions for the new element set; we will not follow him in his subsequent detailed investigation of perturbed planetary motion. This particular canonical transformation of the Delaunay elements appears in modern textbooks, for example Brouwer and Clemence (1961, 539–540)..

based on his discussion in Chapter 1 Sect. 6, which, however, did not include a proof. Having proved that $(\Lambda, \xi, p, \lambda, \eta, q)$ are canonical Charlier expressed these quantities in terms of the elliptical orbital elements and used the resulting expressions in further investigations of perturbation theory.

## 7.2 Charlier's development of the Jacobi transformation theory (1907)

### 7.2.1 Jacobi's theorem on canonical transformations

In the second volume *Mechanik des Himmels* (1907), Charlier again took up the topic of changes of variables that preserve the canonical form of the dynamical equations. Here he explicitly used the term "transformation" rather than "substitution" and developed the theory in terms of generating functions, what he called "transformation functions" (Charlier 1907, 335). He referred to Jacobi's (1866b) *Ueber diejenigen Probleme* and cited the republication of this treatise in Jacobi's *Werke* (Volume 5).[60] He even produced within quotation marks what is ostensibly the original statement of Jacobi's Theorem X from that treatise. Nevertheless, in "quoting" Jacobi he altered slightly what was written, replacing Jacobi's variable constants by dynamical variables and replacing terms such as "Element" by "Grösse". (Jacobi wrote "… drückt die Störungsfonction $H_1$ durch $t$ und diese neuen Elemente $\alpha_1', \alpha_2', \ldots, \alpha_m', \beta_1', \beta_2', \ldots, \beta_m'$ aus", which Charlier quoted as "… drückt $H$ durch $t$ und diese neuen Grössen $\xi_1, \xi_2, \ldots, \xi_m, \beta_1, \beta_2, \ldots, \beta_m$ aus".)

In Charlier's notation, we have the variables $x_i$, $y_i$ and the canonical equations

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \frac{\partial H}{\partial y_i}, \frac{\mathrm{d}y_i}{\mathrm{d}t} = -\frac{\partial H}{\partial x_i}(i = 1, 2, \ldots, m). \tag{271}$$

A second set of variables $\xi_i$, $\eta_i$ is connected to $x_i$, $y_i$ by the relations

$$\frac{\partial \psi}{\partial x_i} = y_i; \frac{\partial \psi}{\partial \xi_i} = -\eta_i (i = 1, 2 \ldots, m), \tag{272}$$

where the generating function $\psi$ is given as

$$\psi = \psi(x_1, x_2, \ldots, x_m; \xi_1, \xi_2, \ldots, \xi_m). \tag{273}$$

Expressing $H$ as a function of $\xi_i$, $\eta_i$ it follows that we have

$$\frac{\mathrm{d}\xi_i}{\mathrm{d}t} = \frac{\partial H}{\partial \eta_i}, \frac{\mathrm{d}\eta_i}{\mathrm{d}t} = -\frac{\partial H}{\partial \xi_i}(i = 1, 2, \ldots, m), \tag{274}$$

and the transformation is canonical.

[60] It is noteworthy that references by authors from around 1900 to Jacobi's *Vorlesungen* and *Ueber diejenigen Probleme* are to the editions of these works published in Jacobi's *Gesammelte Werke* in 1886 and 1890 respectively, rather than to the original volume published in 1866. It suggests that the publication of the collected works may have led in the 1890s to a renewed interest in Jacobi's work.

Charlier (1907, 334–338) followed Donkin (who is not cited)[61] in extending Jacobi's result to generating functions that contain the time:

$$\psi = \psi(x_1, x_2, \ldots, x_m; \xi_1, \xi_2, \ldots, \xi_m, t). \tag{275}$$

The Hamiltonian $H$ is transformed to $R$ where

$$R = H + \frac{\partial \psi}{\partial t}, \tag{276}$$

and the canonical Eq. (274) become

$$\frac{\mathrm{d}\xi_i}{\mathrm{d}t} = \frac{\partial R}{\partial \eta_i}, \frac{\mathrm{d}\eta_i}{\mathrm{d}t} = -\frac{\partial R}{\partial \xi_i} (i = 1, 2, \ldots, m). \tag{277}$$

The proof, which we will not give, was the same as Jacobi's described by us above in Sect. 3.7.2 suitably extended to encompass this more general case. As we noted in Sect. 3.7.2, Jacobi's proof was a natural one, formulated as a result about differential equations. Charlier's *Mechanik* was in the twentieth century one of the standard reference works for mathematical methods of celestial mechanics. His adoption and extension of Jacobi's proof ensured its place in the heritage of modern dynamical analysis."

### 7.2.2 Intermediate orbits and Jacobi's theorem on canonical elements

Charlier (1907, 339–340) next showed how canonical transformations may be used to go from an approximate solution to a problem of planetary motion to a solution of the orbit itself. We have a primary orbit with Hamiltonian $H$ and an intermediate orbit with Hamiltonian $H_1$. The notion of an intermediate orbit as a formal concept had been introduced by Hugo Gyldén (1885), and was adopted by celestial mechanicians in the decades which followed ((Poincaré 1893, 223–224) and Whittaker (1899, 142–143)), but appears to have fallen out of use in the modern subject. The intermediate orbit is in an approximation to the actual orbit, the latter being obtained from the former through a process of perturbation. The key idea is that the solution to the intermediate orbit is more tractable. Taking a solution to the differential equations for this orbit and applying a canonical transformation one obtains the solution for the actual orbit.

The Hamilton–Jacobi equation for the intermediate orbit is

$$0 = \frac{\partial V}{\partial t} + H_1\left(x_1, x_2., \ldots, x_m; \frac{\partial V}{\partial x_1}, \frac{\partial V}{\partial x_2}, \ldots, \frac{\partial V}{\partial x_m}, t\right), \tag{278}$$

[61] There are some points of similarity between Charlier's work and the memoirs from the 1850s of Donkin – in choice of notation, the extension of generating functions to time-dependent functions, the use of Poisson square brackets and the division of generating functions into four types. Donkin's memoirs had appeared in the *Philosophical Transactions*, the premier English-language scientific journal in the nineteenth century. However, Charlier did not mention Donkin and may not have read his papers. Charlier's proof of the theorem on canonical transformations followed Jacobi rather than Donkin.

where the principal function is

$$V = V(x_1, x_2, \ldots, x_m; t), \tag{279}$$

and $H_1$ is the Hamiltonian for the intermediate orbit. A complete solution of (278) is given in the form

$$V = V(x_1, x_2, \ldots, x_m; \alpha_1, \alpha_2, \ldots, \alpha_m; t) \tag{280}$$

with arbitrary constants $\alpha_1, \alpha_2, \ldots, \alpha_m$. Referring tacitly to Jacobi's integration theorem, Charlier (1907, 338) wrote that it is known that $x_i$ and $y_i$ are given by the equations

$$\frac{\partial V}{\partial x_i} = y_i, \quad \frac{\partial V}{\partial \alpha_i} = -\beta_i, (i = 1, \ldots, m) \tag{281}$$

involving the constants $\alpha_1, \alpha_2, \ldots, \alpha_m$ and the $m$ additional arbitrary constants $\beta_1, \beta_2, \ldots, \beta_m$. Charlier had in fact proved this theorem in Volume 1 of *Die Mechanik die Himmels* (1902, 62–66) in essentially the same way as Jacobi (1837a).

Although $V$ was derived for the intermediate orbit, Charlier's next step was to use it as a generating function that defines a canonical transformation for the primary orbit with Hamiltonian $H$, where the quantities $\alpha_i$ and $\beta_i$ that appear in (281) are now regarded as variable functions of $t$. Equations (281) are the relations involving the generating function that relate the old variables $x_i$, $y_i$ and $t$ to the new variables $\alpha_i$, $\beta_i$ and $t$. By Jacobi's transformation theorem we obtain

$$\frac{\mathrm{d}\alpha_i}{\mathrm{d}t} = \frac{\partial R}{\partial \beta_i}, \frac{\mathrm{d}\beta_i}{\mathrm{d}t} = -\frac{\partial R}{\partial \alpha_i}, (i = 1, 2, \ldots, m), \tag{282}$$

where

$$R = H + \frac{\partial V}{\partial t}. \tag{283}$$

But $0 = \frac{\partial V}{\partial t} + H_1$ and so $R = H - H_1$ and $H = H_1 + R$. Thus, with (282) we have arrived precisely at Jacobi's theorem on canonical elements (see Sect. 3.7.1 above). Equations (7.26) combined with (7.25) for the intermediate orbit lead to the actual orbit and the solution to the problem.

Charlier referred to the theory of variation of constants, which he contributed to in this elegant derivation using canonical transformations. It should be noted that he had proved the theorem on canonical elements earlier in volume one of *Mechanik des Himmels* (Charlier 1902, 69–74) along the same lines as Jacobi's (1866b, 355–358) proof in Theorem IX of *Ueber diejenige Probleme*. (See our account in Sect. 3.7.1) The proof we have just presented from volume two of Charlier's work in turn harkened back to Jacobi's (1866b, 464–468) account in Sect. 41 where he showed that the theorem on canonical elements could also be derived from Theorem X on canonical transformations. Because Charlier was using a time-dependent generating function obtained from a solution to the *H–J* equation he was able to simplify Jacobi's procedure. (This method of proof was not available to Jacobi (1866b) or Desboves (1848) because they only considered time-independent generating functions.)

Some perspective on Charlier's derivation is provided by a comparison with standard dynamical theory as is found, for example, in mechanics textbooks of Goldstein (1950) or Landau and Lifshitz (1969). Charlier is beginning with Jacobi's integration theorem *as known*. (This theorem was in fact proved in volume one (Charlier 1902, 62–66).) A transformation then leads to Jacobi's theorem on canonical elements with the desired Eq. (280) for these elements. By contrast, in Landau and Lifshitz (1969) or Goldstein (1950) a canonical transformation obtained from a complete solution to the *H–J* equation is used *to prove* Jacobi's integration theorem. Furthermore, these authors do not include the theory of variation of constants as a subject in their books.

However, Charlier's derivation embodies some of the characteristics that are in principle at the core of Hamilton–Jacobi methods in modern celestial mechanics. Consider the following passage from Vinti (1998, 57):

> At this point the question may arise: Since we have already solved the Kepler problem, why solve it again with such a complicated piece of machinery as the *HJ* procedure? The answer is this: The *HJ* solution will yield a canonical transformation of the Cartesian *q's* and *p's* or the spherical coordinate *q's* and *p's* to the $\alpha$*'s* and $\beta$*'s*, which are so closely related to the Keplerian elements. Most problems in orbital mechanics and celestial mechanics are solved by a method of perturbations, beginning with a solution of a problem already solved, such as the Kepler problem. If we begin with the Keplerian solution, we use the Keplerian elements as variables in the perturbed problem. Once we have solved the perturbed problem by finding the variable Keplerian elements as functions of time, we can write down the solution for the position vector $\boldsymbol{r}$ and the velocity $\dot{\boldsymbol{r}}$ …".

### 7.2.3 Conservative systems and time-independent generating functions

Charlier (1907, 339) remarked that the above result was of limited practical use in calculating orbits because terms appear in Eq. (276) involving time $t$ as a multiplier, and this introduces serious complications into the integration. In a three-body problem, it is necessary to employ other means. In fact, he suggested, there is no general method beginning with an intermediate orbit for obtaining a suitable transformation.

Charlier (1907, 339–342) turned to the case where the Hamiltonian is independent of time and obtained results that were more productive. Here he derived equations corresponding to (282) for the variable elements. The Hamilton–Jacobi equation for the intermediate orbit is

$$0 = \frac{\partial V}{\partial t} + H_1\left(x_1, x_2., \ldots, x_m; \frac{\partial V}{\partial x_1}, \frac{\partial V}{\partial x_2}, \ldots, \frac{\partial V}{\partial x_m}\right) \tag{284}$$

where $H_1$ does not involve time. Jacobi's integration theorem is next presented in the following form. We consider an integral of (284) of the form

$$V = -Ct + W, \tag{285}$$

where $C$ is a constant and $W = W(x_1, x_2, \ldots, x_m)$ does not contain $t$. Because $\frac{\partial V}{\partial x_i} = \frac{\partial W}{\partial x_i}$, (284) and (285) give rise to the equation

$$C = H_1\left(x_1, x_2, \ldots, x_m; \frac{\partial W}{\partial x_1}, \frac{\partial W}{\partial x_2}, \ldots, \frac{\partial W}{\partial x_m}\right). \tag{286}$$

A complete integral of (286) has the form

$$W = W(x_1, x_2, \ldots, x_m; \alpha_1, \alpha_2, \ldots, \alpha_m). \tag{287}$$

The constant $C$ in (286) is taken to be a function of $\alpha_1, \alpha_2, \ldots, \alpha_m$:

$$C = C(\alpha_1, \alpha_2, \ldots, \alpha_m). \tag{288}$$

In Jacobi's original statement of Jacobi's integration theorem for conservative systems, the constant $C$ was made equal to $\alpha_m$ (see Sect. 3.6 above). In Charlier's formulation, $C$ is given more generally by (288). From (281) and (285) we obtain

$$\frac{\partial W}{\partial x_i} = y_i, \quad \frac{\partial W}{\partial \alpha_i} = \frac{\partial C}{\partial \alpha_i} t - \beta_i, (i = 1, \ldots, m). \tag{289}$$

Let $\omega_i = -\frac{\partial C}{\partial \alpha_i} t + \beta_i$. Returning to the original Hamiltonian $H$, we now take $\alpha_i$ to be a variable function of time while $\omega_i$ is a linear function of time. The generating function $W$ and Eq. (289) give rise to a canonical transformation from $x_i$, $y_i$ to $\alpha_i$, $\omega_i$. Hence we have the canonical equations

$$\begin{aligned} \frac{d\alpha_i}{dt} &= \frac{\partial H}{\partial \omega_i}, \\ & \qquad (i = 1, 2, \ldots, m) \\ \frac{d\omega_i}{dt} &= -\frac{\partial H}{\partial \alpha_i}, \end{aligned} \tag{290}$$

for the variable elements $\alpha_i$ and $\omega_i$.

Charlier continued with an investigation of intermediate orbits, focusing on *H–J* partial differential equations that could be solved by the method of separation of variables. A key object of interest was periodic orbits. The concept of an action-angle variable developed out of this analysis (see Nakane (2015)). Although these researchers were not closely related to the subject of canonical transformations, the resulting methods and solutions played an important role in the further development of *H–J* theory.

### 7.3 Classification of generating functions

In considering canonical transformations from $x_1, \ldots, x_m; y_1, \ldots, y_m$ to $\xi_1, \ldots, \xi_m; \eta_1, \ldots, \eta_m$ it should be noted that there are generating functions other than $\psi(x_1, \ldots, x_m; \xi_1, \ldots, \xi_m)$. This fact had been noted and discussed by Donkin (1855, 314) and Poincaré had considered generating functions of more than one form. At the end of section one of the eleventh chapter Charlier (1907, 355–356) presented without any further commentary four types of generating functions and the relations each generating function must satisfy to produce a canonical transformation. He used the same symbol $\psi$ in each case; we add subscripts to distinguish between the different cases.

1. $\psi_1 = \psi_1(x_1, \ldots, x_m; \xi_1, \ldots, \xi_m)$ with the relations

$$\frac{\partial \psi_1}{\partial x_i} = y_i; \ \frac{\partial \psi_1}{\partial \xi_i} = -\eta_i. \tag{291}$$

2. $\psi_2 = \psi_2(y_1, \ldots, y_m; \eta_1, \ldots, \eta_m)$ with the relations

$$\frac{\partial \psi_2}{\partial y_i} = x_i; \ \frac{\partial \psi_2}{\partial \eta_i} = -\xi_i. \tag{292}$$

3. $\psi_3 = \psi_3(x_1, \ldots, x_m; \eta_1, \ldots, \eta_m)$ with the relations

$$\frac{\partial \psi_3}{\partial x_i} = y_i; \ \frac{\partial \psi_3}{\partial \eta_i} = \xi_i. \tag{293}$$

4. $\psi_4 = \psi_4(y_1, \ldots, y_m; \xi, \ldots, \xi_m)$ with the relations

$$\frac{\partial \psi_4}{\partial y_i} = x_i; \ \frac{\partial \psi_2}{\partial \xi_i} = \eta_i. \tag{294}$$

The transformation theorem had been proved for the first case $\psi_1$ with the associated relations $\frac{\partial \psi_1}{\partial x_i} = y_i$; $\frac{\partial \psi_1}{\partial \xi_i} = -\eta_i$. Charlier provided no proof that the second, third and fourth functions $\psi_k (k = 2, 3, 4)$ and the associated relations give rise to canonical transformations. Of course, It might seem clear that the original proof for the first case could be suitably revised for each of the other cases. However, it seems more likely that Charlier's was thinking along the following elementary lines. We show how this works for the third case; the others follow in a similar way. We have $\psi_3 = \psi_3(x_1, \ldots, x_m; \eta_1, \ldots, \eta_m)$ with the relations as prescribed in the third case, $\frac{\partial \psi_3}{\partial x_i} = y_i$; $\frac{\partial \psi_3}{\partial \eta_i} = \xi_i$. Since $\psi_3$ does not involve $y_i$ or $\xi_i$ we have $\frac{\partial \psi_3}{\partial y_i} = 0$ and $\frac{\partial \psi_3}{\partial \xi_i} = 0$. These last conditions suggest that we try $\psi_1$ given as

$$\psi_1 = \psi_3 - \sum_1^m \xi_i \eta_i. \tag{295}$$

For $\psi_1$ so defined we have $\frac{\partial\psi_1}{\partial x_i} = \frac{\partial\psi_3}{\partial x_i} = y_i$. Furthermore, we find that

$$\frac{\partial\psi_1}{\partial\xi_i} = -\eta_i.$$

Hence for $\psi_3$ with the conditions $\frac{\partial\psi_3}{\partial x_i} = y_i$; $\frac{\partial\psi_3}{\partial\eta_i} = \xi_i$ we are led to $\psi_1$ with conditions $\frac{\partial\psi_1}{\partial x_i} = y_i$; $\frac{\partial\psi_1}{\partial\xi_i} = -\eta_i$ and so the transformation is canonical.

In the cases of $\psi_2$ and $\psi_4$ we are led directly to the following expressions for $\psi_1$:

$$\psi_1 = \sum_1^m (x_i y_i - \xi_i\eta_i) - \psi_2 \tag{296}$$

$$\psi_1 = \sum_1^m x_i y_i - \psi_4 \tag{297}$$

In each case, the prescribed relations imply that $\frac{\partial\psi_1}{\partial x_i} = y_i$; $\frac{\partial\psi_1}{\partial\xi_i} = -\eta_i$ and thus the transformation given by these relations is canonical.[62]

### 7.4 Charlier: concluding comments

Although Charlier was primarily an astronomer, he contributed to making canonical transformations a standard mathematical tool in the analysis of perturbations. The level of this work was high, with lucid derivations of mathematical methods and their various applications to problems of perturbed motion. He preserved Jacobi's theory largely in its original form with some innovations and minor rearrangements. As we saw in Sect. 7.2, he showed in solving dynamical problems how solutions to the *H–J* partial differential equation could be effectively used as generating functions for canonical transformations. However, unlike Poincaré (1905) (Sect. 6.5.3) and Whittaker (1904) (Sect. 8.3) he did not use transformations obtained in this way to prove Jacobi's integration theorem and instead derived this result directly.

In the years following the publication of *Die Mechanik des Himmels* Charlier became a pioneer in the statistical study of stellar motion (see Trumpler 1933). This work was formative in the development of mathematical statistics. He was also interested in cosmology and invented something known as the hierarchic theory of the universe. It was one of the last grand cosmological schemes prior to the advent in the 1930s of the expanding-universe paradigm of cosmology. (See North (1965, 18–22) for details.)

[62] Charlier does not give the expressions (7.37) and (7.38) for $\psi_3$, $\psi_2$, $\psi_4$. They were given by Donkin (1855) and can be found in modern textbooks; see for example, Goldstein (1950, 240–243).

## 8 Whittaker (1904): new directions

### 8.1 Introduction

Edmund Whittaker's (1899) survey of research on the three-body problem established him as a knowledgeable writer about celestial mechanics. In 1904 he published *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, a substantial theoretical study of mathematical mechanics. Chapters X, XI and XII of Whittaker's book were devoted to Hamilton–Jacobi theory, including Chapter XI on transformation theory. His approach to the subject differed somewhat from Poincaré's or Charlier's, with a stronger emphasis on differential geometry and a more abstract approach to the subject matter. He also drew on some of Sophus Lie's ideas about transformations and their properties. In his 1899 survey, he had included nothing on canonical or contact transformations. However, in his *Analytical Dynamics* he elevated the subject to a whole area of investigation with its own chapter. Indeed, the concept of a contact or canonical transformation was at the base of his organization of the theory.

We will examine two results reported on by Whittaker: his formulation and proof of Jacobi's transformation theorem in Chapter XI, and his proof of Jacobi's integration theorem in Chapter XII.

### 8.2 Jacobi's transformation theorem

In modern mathematical dynamics circa 1960, the term "contact transformation" is more or less equivalent in meaning to "canonical transformation," that is, a transformation that preserves the canonical form of the canonical equations.[63] Whittaker's conception of a "contact transformation" was somewhat different, and originated in a notion introduced by Sophus Lie.[64] Whittaker's book contains statements such as the following: "The whole course of a dynamical system can thus be regarded as the gradual self-unfolding of a contact-transformation." In the second edition (1917, 290)

[63] In Corben and Stehle's *Classical Mechanics* (1950) Sect. 68 the term contact transformation is used. In the index of the book under the entry "Canonical transformations" one finds "see Contact transformations.".In Sect. 67 (p. 218) of Corben and Stehle's book a transformation is said to be a contact transformation from $q,p$ to $Q,P$ if there exists a function $S = S(q,Q)$ such that $pdq\text{-}PdQ = dS$. The authors then give Poincaré's proof that such a transformation preserves the canonical form of the canonical equations (with no reference to Poincaré.) In Sect. 68 (p. 221) the authors write, "The essential reason for the introduction of contact transformations is the property exhibited by (67.3) [the canonical Eqs. (8.1) in our account] that they leave the form of the canonical equations invariant.".In Goldstein's *Classical Mechanics* (1950) the term "canonical transformation" is used in the same way. The terms contact transformation and canonical transformation mean the same thing in the two books. "Canonical transformation" has become the standard in today's mechanics.(It should be noted that Carathéodory (1935, 78-121) some years earlier wrote in detail about contact and canonical transformations and the relation between them. He proved a theorem (Satz 1) in Sect. 121 on pp. 107-108 that asserts that the two concepts—general homogeneous canonical transformation and general contact transformation—are the same (except for matters of notation).).

[64] Whittaker (1904) doesn't identify where in Lie's work the relevant mathematical theory was developed. Presumably Lie and Georg Scheffers' *Geometrie der Berührungstransformationen* (Leipzig, 1896) was an important source. In the second (1917) edition Whittaker adds a new section at the beginning of Chapter XI where he uses ideas from Hamilton's optics to introduce transformations in a geometric way, and he here refers to Lie's contact transformation, again with no references to sources.

he writes, "From the point of view of Pure Mathematics, we regard the change from the set of variables (*x*, *y*, *z*, *l*, *m*, *n*) to the set of variables (*x'*, *y'*, *z'*, *l'*, *m'*, *n'*) or (to express it geometrically) from the surfaces $\sigma$ to the surfaces $\Sigma$, as a *transformation*. The function *V* is thus to be regarded as determining a transformation of space which changes any surface $\sigma$ into a new surface $\Sigma$."

Nevertheless, in analytical terms, Whittaker definition seems to have reduced to a transformation given in terms of a generating function. This fact is apparent in the definition given at the beginning of Chapter XI. In Sect. 126 (1904, 284) one finds:

> … the transformation from the variables $(q_1, q_2, \ldots, q_n, p_1, \ldots, p_n)$ to variables $(Q_1, Q, \ldots, Q_n, P_1, \ldots, P_n)$ is a contact transformation, for the expression
>
> $$\sum_{r=1}^{n}(P_r dQ_r - p_r \mathrm{d}q_r)$$
>
> becomes, in virtue of these equations, $dW$, and so is a perfect differential.

Whittaker stated Jacobi's transformation theorem in Sect. 136, "On the transformation of a given dynamical system into another dynamical system." In subsequent editions the title is changed to "Jacobi's integration theorem on the transformation of a given dynamical system into another dynamical system." In these later editions a reference is added to Jacobi (1837b), although no mention is made of Jacobi's (1866b) more noteworthy *Ueber diejenigen Probleme* that contained a detailed discussion and full proof of the result. (Furthermore, the result Whittaker presented is incomplete, and the full result appears only in Sect. 138.) We here provide only an outline of the derivation.

Whittaker's proof is spread over two chapters and three sections. It involves three parts: a result in Sect. 116 (Chapter X) about "relative integral invariants"; a partial proof in Sect. 136 (Chapter XI) of the theorem; and the completion of the proof in Sect. 138 (Chapter XI) giving the form of the transformed Hamiltonian. The proof is altogether a very elaborate construction. We begin with the canonical system

$$\frac{\mathrm{d}q_r}{\mathrm{d}t} = \frac{\partial H}{\partial p_i}, \frac{\mathrm{d}p_r}{\mathrm{d}t} = -\frac{\partial H}{\partial q_r}, \quad (r = 1, 2, \ldots, n), \tag{298}$$

where $H = H(q_r, p_r, t)$ is a given function of $q_r, p_r, t$. Let $(Q_1, Q_2, \ldots, Q_n, P_1, \ldots, P_n)$ be $2n$ variables related to $(q_1, q_2, \ldots, q_n, p_1, \ldots, p_n)$ by means of a contact transformation. Analytically this means we have a function $W = W(q_r, Q_r)$ such that

$$\sum_{r=1}^{n} P_r dQ_r - \sum_{r=1}^{n} p_r \mathrm{d}q_r = dW. \tag{299}$$

We now consider a process of variation denoted by $\delta$. Under this variation we have

$$\sum_{r=1}^{n} P_r \delta Q_r - \sum_{r=1}^{n} p_r \delta q_r = \delta W. \tag{300}$$

In the preceding chapter "Hamiltonian systems and their-integral invariants" Whittaker developed a body of results on invariants associated with transformations defined with respect to a class of differential equations. Despite the title, the theory was more general than transformations defined for Hamiltonian or canonical systems of differential equations. The approach was formalistic and rather abstract. Whittaker was stimulated by researches of Poincaré and Lie to develop the subject in the spirit of what he called "Pure Mathematics."

Whittaker concluded from (300) that $\sum_{r=1}^{n} P_r \delta Q_r$ is something called a "relative integral invariant." Appealing to results presented in the preceding chapter, Whittaker reasoned that one is able to infer the existence of a function $K = K(Q_r, P_r, t)$ such that

$$\frac{\mathrm{d}Q_r}{\mathrm{d}t} = \frac{\partial K}{\partial p_i}, \frac{\mathrm{d}P_r}{\mathrm{d}t} = -\frac{\partial K}{\partial P_r}, \quad (r = 1, 2, \ldots, n) \tag{301}$$

Whittaker's proof hinged on this existence result. In Sect. 138 (Chapter XI) he proceeded to deduce a somewhat complicated expression for $K$ that reduces in the conventional problem to the form

$$K = H + \frac{\partial W}{\partial t}. \tag{302}$$

His reasoning involved the differential form (299) and some facts about differential equations. One concludes finally that the transformation defined by (300) preserves the canonical form of the Eq. (298), with $H$ being replaced by $K$ in the transformed equations.

Whittaker's proof of Jacobi's transformation theorem is different from those of Desboves (1848), Donkin (1855), Jacobi (1866b), and Poincaré (1892/1905 and 1899), and is challenging to follow and to understand. In contrast to Poincaré, his proof made no use of variational laws and was a result of the theory of differential equations developed from the viewpoint of differential geometry. The derivation was not taken up by later physicists and mathematical astronomers who otherwise found great value in the theory of canonical transformations. Whittaker's proof remains a curious piece of analytical exotica from a formative period in the development of the theory of transformations.

## 8.3 Proof of Jacobi's integration theorem

In Jacobi's *Vorlesungen* the Hamilton–Jacobi differential equation is derived in lecture 8 in terms of Hamilton's principal function. In his statement about the integration theorem in lecture 20 Jacobi takes the *H–J* equation as given from the outset. A complete integral of this equation then leads to a solution of the dynamical equations.

In Sect. 142 of his *Analytical Dynamics* Whittaker proceeded in a different way. His proof of the integration theorem is the one given by Jacobi (1866b, Sect. 41) in his *Ueber diejenigen Probleme*, although nowhere in the *Analytical Dynamics* (in any of its editions) does Whittaker refer to this source. We have the Hamiltonian $H = H(q_1, q_2, \ldots, q_n, p_1, \ldots, p_n, t)$. Whittaker begins with a function $W$ of the variables $q_1, q_2, \ldots, q_n, Q_1, \ Q_2, \ldots, Q_n, t$:

$$W = W(q_1, q_2, \ldots, q_n, Q_1, \ Q_2, \ldots, Q_n, t) \tag{303}$$

We introduce another set of variables $P_1, \ P_2, \ldots, P_n$ and consider a contact or canonical transformation from $q_1, q_2, \ldots, q_n, p_1, \ldots, p_n$ to $Q_1, Q_2, \ldots, Q_n, P_1, P_2, \ldots, P_n$ defined using $W$ by the relations

$$p_r = \frac{\partial W}{\partial q_r}, P_r = -\frac{\partial W}{\partial Q_r}(r = 1, 2, \ldots, n). \tag{304}$$

The transformed Hamiltonian K is

$$K = H + \frac{\partial W}{\partial t}. \tag{305}$$

The transformation is canonical and so we have

$$\frac{\mathrm{d}Q_r}{\mathrm{d}t} = \frac{\partial K}{\partial P_r}, \frac{\mathrm{d}P_r}{\mathrm{d}t} = -\frac{\partial K}{\partial Q_r}. \tag{306}$$

The next step is to let $K = 0$, obtaining in this way the Hamilton–Jacobi partial differential equation:

$$H + \frac{\partial W}{\partial t} = 0. \tag{307}$$

Let $W$ be a solution of (307). Then $K = 0$ and Eq. (306) become

$$Q_r = \alpha_r, P_r = \beta_r, (r = 1, 2, \ldots, n) \tag{308}$$

where $\alpha_r$ and $\beta_r$ ($r = 1, \ldots, n$) are constants. Equations (304) may then be used to express the $q_r$ and $p_r$ as functions of $t$ and the arbitrary constants $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$. Jacobi's integration theorem has been proved. In the second edition, Whittaker would refer to Jacobi (1837a), without noting that the theorem is proved in a different way there.

In Chapter 13, Whittaker reviewed work that had been done on reducing the degree of the three-body problem. In his 1899 report, he had independently examined each mathematician's procedure. By contrast, the treatment in Chapter 13 is unified around the use of transformations and generating functions by different researchers (including himself) to achieve the given reductions.

### 8.4 Historical reflections on Whittaker's *Analytical Dynamics*

In reference to Whittaker's *Analytical Dynamics*, Herbert Goldstein (1950, 269) writes at the end of the chapter in his *Classical Mechanics* on canonical transformations, "Much of the material given in the Nordheim and Fues article is also discussed by Whittaker in Chapters IX and X, more from the viewpoint of the mathematician, and it is interesting to contrast the two treatments."[65] This note and several others are omitted in the second (1980) edition of the book, replaced by newer references to the literature. In the second edition Goldstein (1980, p. viii) states in the preface, "Occasionally, a line of reasoning presented in the first edition has been supplemented by a different way of looking at the problem. The most important example is the introduction of the symplectic approach to canonical transformations, in parallel with the older technique of generating functions."

Beginning in the 1950s, symplectic geometry was adopted by some investigators in abstract theoretical mechanics. A book on the new theory was Ralph Abraham's 1967 *Foundations of Mechanics*. On p. 100, Abraham writes, "Symplectic diffeomorphisms are known classically as **homogeneous canonical** (or **contact**) **transformations**" [emphasis in the original].

The symplectic approach to mechanics only got underway in the 1950s and 1960s. However, S. C. Coutinho in a study of Whittaker's *Analytical Dynamics* suggests that this approach was foreshadowed in that book. Coutinho writes (p. 356).

> Although the word symplectic was only coined by H. Weyl in 1939 … the subject has its roots in the work of Lagrange and Poisson … with important contributions by G. Darboux and S. Lie. The present approach, in terms of differential forms, goes back to Lie's theory of contact transformations (Berührungstransformation) …

On the next page he continues,

> Even a cursory look at the table of contents of *Analytical Dynamics* suggests why it is still useful: although couched in an older language one finds there most of the topics dealt with in modern tracts, including several that we commonly associate with the symplectic approach that became prevalent since the 1960s.

Although Whittaker's development of Hamilton–Jacobi's theory was not entirely successful it at least embodied tendencies that would germinate half a century later in a formal approach to dynamics cultivated by one school of modern researchers.

## 9 Conclusion

In the nineteenth century, a major historical shift took place in the investigation of canonical equations and transformations, from an emphasis before 1850 on the theory of perturbations and the detailed investigation of the time rate of change of the orbital elements, to an embrace during the second half of the century of general dynamical

[65] Goldstein erred here, the relevant Chapters in Whittaker were X, XI and XII.

theory. Lagrange, Poisson, Jacobi and Desboves were focused on the variation of constants arising in the integration of the differential equations of gravitating systems of bodies. Jacobi's theorem on canonical elements was a central result in his exposition of the theory, and it was also at the centre of Desboves' investigation. Of course, Jacobi also pursued theoretical lines of research and recognized more general possibilities implicit in the theory, although these possibilities were not fully explored in his researches. It is also true that astronomers such as Tisserand and Charlier in the decades around 1900 continued to investigate the variation of constants and perturbed orbits. Nevertheless, the more mathematical-minded investigators, Donkin, Poincaré, and Whittaker among them, concentrated on general dynamical analysis and Hamilton–Jacobi theory insofar as they could be applied to perturbation problems. In modern physics textbooks, Jacobi's theorem on canonical elements is nowhere to be found, and the subject of perturbations is a specialized if an important field of investigation.[66] Jacobi's theorem on canonical elements remains a standard result in celestial mechanics.[67]

The idea of transformations that preserve the canonical form of a system of differential equations originated with Jacobi in 1837. A method for generating such transformations was a signal contribution of his inscrutable creative genius.[68] In the next sixty years, such transformations were the subject of isolated investigation by researchers in France and England, culminating at the end of the century with Poincaré's systematic exposition. In the latter canonical transformations occupy a much more prominent place than they did in Jacobi's dynamical corpus. In the first decades of the new century the theory of canonical transformations emerged as a subject in its own right, a change that is evident in the prominence it received in the books of Whittaker (1904) and Charlier (1907).

Jacobi invented the idea of the canonical transformation, a mathematical advance not anticipated in the earlier work of Hamilton or others. Jacobi's strength was his unparalleled grasp of algorithmic mathematics and his ability to find operational, formal solutions and methods.[69] Bell (1986, 327) in his book *Men of Mathematics*

66 For an informative account of perturbation theory in atomic physics in the 1920s see Fues (1927), in which the subject receives its own chapter in the *Handbuch der Physik*. Fues includes some remarks on the relation of perturbations in physics to traditional celestial mechanics. He writes (p. 132), "In the past, physics had little reason to be interested in this mode of calculation, until the establishment of Bohr's atomic model suddenly created a close relationship between atomic theory and cosmic astronomy." For an historical study of this subject see Shore (2003).

67 Vinit's book on orbital and celestial mechanics contains a chapter on Hamilton–Jacobi theory that features Jacobi's theorem on canonical elements. See Vinti (1998, Chapter7).

68 In his biography of Jacobi, Koenigsberger (1904, 247) states Jacobi's transformation theorem in the part dealing with Jacobi's researches in Konigsberg at the end of the 1830s. He doesn't indicate where in Jacobi's writings the theorem is to be found or give many details about it but does comment that Jacobi applied it to the problem of attraction to two fixed centres. This problem is taken up in lecture 29 of the 1866 *Vorlesungen* where it is solved using Jacobi's integration theorem that had been presented in lecture 20. (Koenigsberger also doesn't give any of this information. The problem of attraction to two fixed centres is cited by Nakane and Fraser (2002, 215).) The transformation theorem itself does not appear in lecture 29 or anywhere else in the *Vorlesungen*. Koenigsberger does not appear to be aware that the transformation theorem was first stated in Jacobi's Paris note of 1837 and proved as Theorem X in the 1866 supplemental work *Ueber diejenigen Probleme*. (See also the discussion in Sect. 3.7.2 above.).

69 It should be noted that despite Jacobi's strong algorithmic and formal tendencies, his work possessed a more developed sense of mathematical deduction and proof than was the case for such older figures as

fittingly titled his chapter on Jacobi, "The great algorist."[70] In Jacobi's approach to the calculus of variations the formal, algorithmic aspect was paramount, although it is true that some of his innovations, for example the conjugate point, had geometric implications. While the subject of canonical transformations developed out of a part of dynamical analysis with strong roots in the calculus of variations, Jacobi's study of transformations was analytical and did not draw on notions involving continuous variational processes.[71]

By contrast, variational processes were critical to Poincaré's approach and imparted a "topological" element to the theory. A key contribution consisted of his treatment in 1899 of Jacobi's transformation theorem, where a variational law provided a powerful structural tool in the derivation of this result. The contributions of Poincaré in this respect are highlighted by Gaston Darboux in his 1916 éloge:

> … Jacobi had established a theory which appeared to be one of the most complete in dynamics. For fifty years we lived on the theorems of the illustrious German mathematician, applying them and studying them from all angles, but without adding anything essential. It was Poincaré who first shattered these rigid frames in which the theory seemed to be encased and contrived for it vistas and new windows on the external world. He introduced or used, in the study of dynamical problems, different notions: the first, which had been given before and which, moreover, is applicable not solely to mechanics, is that of variational equations, namely, linear differential equations that determine solutions of a problem infinitely near to a given solution; the second, that of integral invariants, which belong entirely to him and play a capital part in these researches.[72]

From its origins up until the early years of the twentieth century Hamilton–Jacobi theory was primarily of interest to mathematical analysts and researchers in celestial mechanics, the latter being active throughout this period. After around 1915 the methods of Hamilton–Jacobi theory were adopted by atomic physicists in Germany. Prominent in this group were Arnold Sommerfeld in Munich and Max Born in Göttingen. Sommerfeld (1919, 485) wrote: "Up to a few years ago it was possible to consider that the methods of mechanics of Hamilton and Jacobi could be dispensed with for physics and to regard it as serving only the requirements of the calculus of astronomic

Footnote 69 continued
Lagrange. Jacobi's thinking about the foundations of mechanics and mathematics in relation to Lagrange is examined by Pulte (1996, 1997, 1998). Pulte draws attention to Jacobi's Berlin lectures (Jacobi 1996), which were not published during the period although they appear (as lectures themselves or as lecture notes) to have influenced Bernhard Riemann and Carl Neumann. In terms of the work considered in the present study the most important influence on Jacobi was Hamilton rather than Lagrange.

70 Jacobi was a remarkable figure and a leading mathematician of the nineteenth century; in Klein's (1926) history of nineteenth-century mathematics he is accorded the same prominence as Gauss, Riemann and Weierstrass. However, Jacobi does not seem to enjoy a correspondingly high profile in today's mathematical culture. For example, he does not appear in Victor Katz's *History of Mathematics*, a widely read textbook today.

71 A lack of a systematic geometric perspective affected not only Jacobi's contributions to mathematical dynamics but also his work in pure analysis. Demidov (1982, 339–340) states that this limitation prevented Jacobi from creating a general theory of first-order partial differential equations, something that would be achieved by Sophus Lie in the 1870s.

72 English translation in Bell (1986, 544–545).

perturbations and the interests of mathematics.” As a result of the rapid development of quantum theory, the situation had changed dramatically. Sommerfeld continued: “ … it seems [today] almost as if Hamilton’s method was expressly created for treating the most important problems of physical mechanics.”[73]

In the late fall of 1922 Born led a private Poincaré reading circle in his cramped apartment in Göttingen, which included the young Werner Heisenberg (Cassidy 1992, 147). Poincaré was a major mathematical inspiration. Cassidy (1992, 147) writes that “Werner also studied Poincaré with Born “with every ounce of energy.”” (The quoted words are from a letter Heisenberg wrote to his parents.) A focus of their work was action-angle variables and periodic atomic orbits treated using the mathematical methods Poincaré had developed for celestial mechanics. Another source of inspiration at Göttingen was the mathematician David Hilbert’s (2009) lectures on quantum mechanics (winter semester 1922–1923), which covered the subject of canonical transformations and their properties.

Insofar as the treatment of canonical transformations in modern textbooks is concerned, a key source became Born’s 1925 *Atommechanik*. In the first chapter, there is an account of Hamilton–Jacobi theory. Canonical transformations and the Hamilton—Jacobi partial differential equation are introduced along the lines set out by Poincaré. In particular, Born took his proof of Jacobi’s theorem on canonical transformations from Poincaré (1899, Chapter 29), although no specific citation is made to Poincaré. (Born elsewhere in the book does cite Poincaré (1899, Chapters 22–24) as well as Poincaré (1893).)

In subsequent history, the Poincaré derivation of canonical transformations became an integral part of mathematical mechanics, although Poincaré’s name is seldom If ever mentioned. Lothar Nordheim was Born’s student and worked with him closely in the 1920s. He was also Hilbert’s assistant and was heavily involved in writing Hilbert’s lectures on quantum mechanics.[74] A seminal contribution was Nordheim and Fues (1927) essay in the *Handbuch der Physik* (discussed by us above in Sect. 1). Textbooks on mechanics from the middle of the century draw mainly on Born (1925)

[73] English translation in Sommerfeld (1923, 555–556).

[74] Nordheim was Hilbert’s assistant from 1922–1927, although his primary mentor was Born (Nordheim completed his physics PhD under Born in 1923). At this time Hilbert was entering retirement-age years while Nordheim was in his early to mid-20 s. The more mathematical character of Nordheim’s writings in the 1920s compared to other physicists was no doubt influenced by his contact with Hilbert. In Hilbert (2009) the editors state that Nordheim and fellow Born student Gustav Heckmann “worked out” Hilbert’s lectures on quantum mechanics from 1922–23 and Nordheim “worked out” Hilbert’s lectures on the same subject in 1926–1927. It is not clear precisely what “worked out” means in this context but it may simply have been that Nordheim and Heckmann wrote the lectures and Hilbert provided editorial emendations. Fifty years later Nordheim would reflect on his time with Hilbert in less than positive terms. In a 1977 interview with historian of physics Bruce Wheaton, he recounted: “First I must say, during the time I was his assistant, he [Hilbert] was very sick. And not the genius he had been. He lived very much in the past in a way. His mathematical interest was logic, which was not terribly appealing to me. But he had the conviction that the best thing for a young man was to work with him. That was a reward in itself. And everything else, financial and family considerations, would be way down in importance.” (From the website of the American Institute of Physics: https://www.aip.org/history-programs/niels-bohr-library/oral-histories/5074.) Of course, personal memories of events long in the past may be selective. From a mathematical viewpoint there are some interesting foundational aspects to Nordheim and Fues’ 1927 article that were likely influenced by Nordheim’s association with Hilbert (Nordheim was the first author on this article).

and Nordheim and Fues (1927) in their presentation of Hamilton–Jacobi theory.[75] Goldstein's (1950, 269–270) chapter on canonical transformations concludes with a survey of the literature, all of which followed (without citation) the Poincaré 1899 proof of Jacobi's theorem on canonical transformations. A somewhat more abstract development of the theory (using terminology from tensor analysis and differential geometry) is given by Yusuke Hagihara (1970, Chapter 1).

The subject of canonical transformations was also investigated by mathematicians, a notable contributor being Constantin Carathéodory. In 1925 he wrote an essay on the calculus of variations for Frank and von Mises' *Die Differential- und Integralgleichungen der Mechanik und Physik* that included a section on canonical transformations. Although framed in a more general and formalistic way than the corresponding work of the physicists, the basic idea of the proof of the theorem on canonical transformations seems to have originated in ideas from Poincaré. This subject was also examined in Carathéodory's 1935 *Variationsrechnung und partielle Differentialgleichingen erster Ordnung* (Chapter 6). A prominent concern here was the role of canonical transformations in the investigation of partial differential equations.

Not all later textbooks on the calculus of variations include canonical transformations as a subject, but several do, including Weinstock (1952), Gelfand and Fomin (1963) and Akheizer (1988). In these writings, the basic theorem on canonical transformations using generating functions is from Poincaré (1899), with no mention and probably no awareness of this source.

Note: Except where otherwise noted, all translations from French and German in this paper are by the authors.

**Acknowledgements** The authors would like to thank Jesper Lützen and Helmut Pulte for reading and advising on the present paper. We are also grateful for the feedback that we have received at academic conferences (Asian History of Astronomy Group, Physical Society of Japan, American Mathematical Society, Canadian Mathematical Society and Canadian Society for the History and Philosophy of Mathematics) at various stages in the preparation of the present study. Research for this article was financially supported by the Social Sciences and Humanities Research Council of Canada, Japanese Grants-in-Aid for Scientific Research (JP17500686), and the Eleanor P. May Funds of the University of Toronto.

## Declarations

**Conflict of interest** On behalf of all authors, the corresponding author states that there is no conflict of interest.

## Appendix: Notation for Lagrange and Poisson Brackets

We have two sets of variables $q_i$, $p_j$ and $a_i$, $b_j$ $(i, j = 1, 2, \ldots, n)$. The two sets of variables are functionally connected.

[75] Physicists Herbert Corben and Philips Stehle (1950, v) write, "The parts of classical mechanics which are of present-day interest to the physicist are not those which were of paramount interest in the nineteenth century. At present fundamental physics is the physics of particles and fields, whereas nineteenth-century mechanics was the study of the n-body problem. There are many points in common between these problems, but the points of view are vastly different." See also Fues' comment in note 66 above.

The Lagrange bracket as originally introduced by Lagrange (1809a, b) is defined as

$$[a_i, b_j] = \sum_{k=1}^{n} \left( \frac{\partial q_k}{\partial a_i} \frac{\partial p_k}{\partial b_j} - \frac{\partial q_k}{\partial b_j} \frac{\partial p_k}{\partial a_i} \right).$$

The Poisson bracket as originally introduced by Poisson (1809a, b) is defined as

$$(a_i, b_j) = \sum_{k=1}^{n} \left( \frac{\partial a_i}{\partial q_k} \frac{\partial b_j}{\partial p_k} - \frac{\partial a_i}{\partial p_k} \frac{\partial b_j}{\partial a_i} \right).$$

Some later authors adopted square brackets for Lagrange brackets and parentheses for Poisson brackets. Other authors adopted the reverse. The table below gives the convention followed by each author.

| | Lagrange bracket | Poisson bracket |
|---|---|---|
| Lagrange (1809a, 1811a, b) | Square [xx] | Round (xx) |
| Poisson (1809a, b) | Square [xx] | Round (xx) |
| Hamilton (1835) | – | Brace {xx} |
| Desboves (1848) | Square [xx] | – |
| Donkin (1854) | – | Square [xx] |
| Cayley (1858)(Cayley used the term "Coefficient" rather than "bracket" and employed the same round parentheses to denote both Lagrange and Poisson brackets. In the edition of the report that was published in Cayley's collected papers (1890) the following sentence is added in Sect. 13: "It may be noticed that throughout the Report, I speak of the Lagrange's Coefficients *(a, b)*, and Poisson's Coefficients *(a, b),* distinguishing them in this manner, and not by any difference of notation.".) | Round (xx) | Round (xx) |
| Jacobi (1866a) | Round (xx) | Square [xx] |
| Poincaré (1905) | Round (xx) | Square [xx] |
| Brown (1903) | Round (xx) and square [xx] | – |
| Charlier (1902) | Round (xx) | Square [xx] |
| Whittaker (1904) | Square [xx] | Round (xx) |
| Carathéodory (1935) | Square [xx] | Round (xx) |

| | Lagrange bracket | Poisson bracket |
|---|---|---|
| Nordheim and Fues (1927) | Square [xx] | Round (xx) |
| Carathéodory (1935) | Square [xx] | Round (xx) |
| Lanczos (1949) | Square [xx] | Round (xx) |
| Goldstein (1950) | Brace {xx} | Square [xx] |
| Corben and Stehle (1950) | Square [xx] | Round (xx) |
| Brouwer and Clemence (1961) | Square [xx] | – |
| Landau and Lifshitz (1969) | – | Square [xx] |
| Akhiezer (1988) | Square [xx] | – |
| Vinti (1998) | Square [xx] | Round (xx) |

## References


Abraham, Ralph. 1967. *Foundations of mechanics*. New York: W. A. Benjamin, Inc.

Arthurs, A.M. 1975. *Calculus of variations*. London: Routledge & Kegan Paul.

Akhiezer, N. I. 1988. *The Calculus of Variations*. Translated from the Russian edition of 1981 by Michael E. Alferieff. Harwood Academic Publishers. London etc. The author was born in 1901 and the references to the literature are mainly from before 1970. Chapter 4 is titled "Canonical Variables" and deals with Hamilton-Jacobi theory and canonical transformations.

Barrow Green, June. 1997. *Poincaré and the Three Body Problem*. Volume 11 in the series "History of Mathematics", American and London Mathematical Societies.

Bell, Eric T. 1986. *Men of Mathematics*. Touchstone Press. New York et al. Reprint of the original 1937 edition.

Binet, Jacques Phillipe Marie. 1841. "Mémoire sur la variation des constantes arbitraires dans les formules générales de la dynamique." *Journal de l'École Polytechnique*, (28. cahier), 1–94.

Bliss, Gilbert. 1946. *Lectures on the calculus of variations*. Chicago: University of Chicago Press.

Bolza, Oscar. 1909. *Vorlesungen über Variationsrechnung*. Leipzig: Teubner.

Born, Max. 1925. *Vorlesungen über Atommechanik Erster Band*. Verlag von Julius Springer. Berlin. From lectures delivered in 1923-1924. English translation 1927, The *Mechanics of the Atom*, G. Bell and Sons, Ltd. London.

Bour, Edmond. 1856. Mémoire sur le problème des trois corps. *Journal De L'école Polytechnique* 21: 35–38.

Brouwer, Dirk, and Gerald M. Clemence. 1961. *Methods of celestial mechanics*. New York: Academic Press.

Brown, Ernest W. 1896a. *An introductory treatise on the lunar theory*. Cambridge: Cambridge University Press.

Brown, Ernest W. 1896b. On the application of Jacobi's dynamical method to the general problem of three bodies. *Proceedings of the London Mathematical Society* 28: 130–142.

Brown, Ernest W. 1903. On the variation of the arbitrary and given constants in dynamical equations. *Transactions of the American Mathematical Society* 4: 333–350.

Cajori, Florian. 1929. *A history of mathematical notations volume ii notations mainly in higher mathematics*. Chicago: The Open Court Publishing Company.

Carathéodory, Constantin. 1925. “Variationsrechnung.” Chapter 5 of Frank and von Mises (1925).

Carathéodory, Constantin. 1935. *Variationsrechnung und partielle Differentialgleichungen erster Ordnung*. B. G. Teubner. Leipzig and Berlin.

Cassidy, David C. (1992). *Uncertainty: the life and science of Werner Heisenberg*. W. H. Freeman. New York. Cassidy, David C. (1992).

Cauchy, Augustin L. 1819. “Note sur l’intégration des équations aux différences partielles du premier ordre à un nombre quelconque de variables.” *Bulletin des Sciences, par la Société Philomatique de Paris*, 101–21.

Cayley, Arthur. 1858. “Report on the recent progress of theoretical dynamics.” *Report of the Twenty-Seventh Meeting of the British Association for the Advancement of Science*, pp. 1–42. John Murray, London. The meeting was held in Dublin in August and September of 1857. Republished in *The Collected Mathematical Papers of Arthur Cayley*, V. 3 (Cambridge University Press, 1890), pp. 156–204. It should be noted that supplementary comments are inserted in the edition of the report that appears in the collected papers.

Charlier, Carl L. 1902. *Die Mechanik des Himmels*. Leipzig: Erster Band. Verlag von Veit & Comp.

Charlier, Carl L. 1907. *Die Mechanik des Himmels*. Leipzig: Zweiter Band. Verlag von Veit & Comp.

Corben, Herbert C., and Philip Stehle. 1950. *Classical mechanics*. New York: Wiley.

Cook, Alan. 1988. *The motion of the moon*. Bristol-Philadelphia.

Costabel, Pierre. 1981. Siméon Denis Poisson. In *Dictionary of scientific biography Volume 15, Supplement 1*, ed. Charles C. Gillispie, 480–490. New York: Charles Scribner’s Sons.

Courant, Richard. 1953. *Calculus of variations lectures given during 1945–1946*. New York: New York University, Institute of Mathematical Sciences.

Courant, Richard, and David Hilbert. 1962. *Methods of mathematical physics volume 2 partial differential equations*. Revised and supplemented English translation (translator not specified) of the 1937 German edition. Interscience Publishers, a division of John Wiley & Sons. New York and London.

Coutinho, S.C. 2014. Whittaker’s analytical dynamics: A biography. *Archive for History of Exact Sciences* 68: 355–407.

Crilly, Tony. 2006. *Arthur Cayley. Mathematician Laureate of the Victorian Age*. Baltimore. Johns Hopkins University Press. Baltimore, Md.

Darboux, Gaston. 1916. “Éloge historique de Henri Poincaré.” In Poincaré (1916), pp. vii-lxxi.

Darrigol, Olivier. 1992. *From C-numbers to Q-numbers: the classical analogy in the history of quantum theory*. Berkeley: University of California Press.

Delaunay, Charles-Eugène. 1846. Mémoire sur une nouvelle méthode pour la détermination du mouvement de la lune. *Comptes Rendus De L’académie Des Sciences* 22: 32–37.

Delaunay, Charles-Eugène. 1860. Théorie du mouvement de la lune I. *Mémoires De L’académie Des Sciences* 28: 1–883.

Delaunay, Charles-Eugène. 1867. Théorie du mouvement de la lune II. *Mémoires De L’académie Des Sciences* 29: 1–931.

Demidov, Sergey S. 1982. The study of partial differential equations of the first order in the $18^{th}$ and $19^{th}$ centuries. *Archive for History of Exact Sciences* 26: 325–350.

Desboves, Adolphe H. 1848. Démonstration de deux théorèmes de M. Jacobi—Application au problème des perturbations planétaires. *Journal de mathématiques pures et appliquées* Tome 13, 397—411. Thesis presented to the Paris University Faculty of Sciences 3 April 1848.

Donkin, William F. 1850. On certain theorems in the calculus of operations. *The Cambridge and Dublin Mathematical Journal* 5 (1850): 10–18.

Donkin, William F. 1854. On a class of differential equations, including those which occur in dynamical problems.—Part I. *Philosophical Transactions of the Royal Society of London* 144: 71–113.

Donkin, William F. 1855. On a class of differential equations, including those which occur in dynamical problems. Part II. *Philosophical Transactions of the Royal Society of London* 145: 299–358.

Fox, Charles. 1954. *An introduction to the calculus of variations*. Oxford: Oxford University Press.

Frank, Philipp, and Richard von Mises (Eds.). 1925. *Die Differential- und Integralgleichungen der Mechanik und Physik Erster Mathematische Teil*. Fridr. Vieweg & Sohn. Braunschweig. Second edition published 1930.

Fraser, Craig. 1983. J.L. Lagrange’s early contributions to the principles and methods of mechanics. *Archive for History of Exact Sciences* 28, 197–241.

Fraser, Craig. 1996. Jacobi's result (1837) in the calculus of variations and its reformulation by Otto Hesse (1857). A study in the changing interpretation of mathematical theorems. In *History of mathematics and education*, Eds. H. N Jahnke and N. Knoche. Volume 11 of the series "Studien zur Wissenschafts-, Sozial und Bildungsgeschichte der Mathematik". Vandenhoeck. Göttingen.
Fraser, Craig and Michiyo Nakane. 2003, La meccanica celeste dopo Laplace: la teoria di Hamilton-Jacobi. In *Storia della scienza*, editor-in-chief Sandro Petruccioli, Rome, Istituto della Enciclopedia Italiana, V. 7, 234–243.
Fues, Erwin. 1927. Störungsrechnung. In Geiger and Scheel (1927), 131–177.
Geiger, Hans, and Karl Scheel (Eds.). 1927. *Handbuch der Physik Volume 5*. Editor R. Grammel. Julius Springer. Berlin.
Gelfand, Israel M. and Sergey V. Fomin. 1963. *Calculus of Variations*. Revised and translated by Richard A. Silverman. Prentice-Hall Inc. Englewood Cliffs, New Jersey. Based on lectures given by Gelfand at Moscow State University.
Goldstein, Herbert. 1950. *Classical mechanics*. Reading: Addison-Wesley Publishing Company Inc.
Goldstein, Herbert. 1980. *Classical mechanics*. Second edition. Addison-Wesley Publishing Company, Inc., Reading, Massachusetts etc. The second edition is greatly expanded, almost twice the length of the first edition.
Goroff, Daniel L. 1993. Henri Poincaré and the birth of chaos theory: An introduction to the English translation of *Méthodes nouvelles de la Mécanique celeste*. In Poincaré (1993, I1-I107).
Gray, Jeremy. 2014. *Henri Poincaré: a scientific biography*. Princeton: Princeton University Press.
Gyldén, Hugo. 1885. Die Intermediäte Bahn des Mondes. *Acta Mathematica* 7: 125–172.
Hagihara, Yusuke. 1970. *Celestial mechanics volume 1 dynamical principles and transformation theory*. Cambridge: MIT Press.
Hamilton, William Rowan. 1834. On a general method in dynamics. *Philosophical Transactions of the Royal Society*, 247–308.
Hamilton, William Rowan. 1835. Second essay on a general method in dynamics. *Philosophical Transactions of the Royal Society*, 95–144.
Hankins, Thomas L. 1980. *Sir William Rowan Hamilton*. Baltimore: The Johns Hopkins University Press.
Hawkins, Thomas. 1991. Jacobi and the birth of Lie's theory of groups. *Archive for History of Exact Sciences* 42: 187–278.
Hilbert, David. 2009. *Mathematische Grundlagen der Quantentheorie*. Text of lectures delivered in the winter semester 1922–1923 at the University of Göttingen. In Tilman Sauer and Ulrich Majer (Eds.), *David Hilbert's lectures of the foundations of physics 1915–1927 relativity, quantum theory and epistemology in collaboration with Arne Schirrmacher and Heniz-Jürgen* " Dordrecht, Heidelberg, London, New York.
Hill, George W. 1876. Demonstration of the Differential Equations Employed by Delaunay in the Lunar Theory. *The Analyst* 3: 65–70.
Hund, Friedrich. 1974. *The history of quantum theory*. London: Harrap.
Jacobi, Carl Gustav J. 1837a. Ueber die Reduction der Integration der partielle Differentialgleichungen erster ordnung zwischen irgend einer Zahl Variabeln auf die Integration eines einzigen Systemes gewöhnlicher Differentialgleichungen. *Journal für die Reine und Angewandte Mathematik* 17, 97–162. Reprinted in Jacobi's *Werke* volume 4, 57–127.
Jacobi, Carl Gustav J.1837b. Note sur l'intégration des équations différentielles de la dynamique. *Comptes rendus de l'Académie des Sciences* 5, 61–67. Reprinted in Jacobi's *Werke* volume 4, 129–136.
Jacobi, Carl Gustav J.1837c. Zur Theorie der Variationsrechnung und der Differentialgleichungen. *Journal für die Reine und Angewandte Mathematik* 17, 68–82. Reprinted in Jacobi's *Werke* volume 4, 39–55.
Jacobi, Carl Gustav J. 1838. Sur la réduction de l'intégration des équations différentielles du premier ordre entre un nombre quelconque de variables à l'intégration d'un seul système d'équations différentielles ordinaires. *Journal de mathématiques pures et appliquées* Tome 3, 60–96. French translation of Jacobi (1837a).
Jacobi, Carl Gustav J. 1841. De Determinantibus functionalibus. *Journal für die reine and angewandte Mathematik* 22, 319–359. Reprinted in Jacobi's *Werke* volume 3, 393–438.
Jacobi, Carl Gustav J. 1843. Sur l'élimination des noeuds dans le problème des trois corps. *Journal für die reine and angewandte Mathematik* 26, 115–131.
Jacobi, Carl Gustav J. 1846. Neues Theorem der analytischen Mechanik. *Journal für die reine and angewandte Mathematik* 30, 137–142. Originally published in 1838 in *Bericht über die Verhandlungen*

*der die zur Bekanntmachung geeigneten Verhandlungen der Königlich-Preussischen Akademie der Wissenshaften zu Berlin*, December, 178–182. Reprinted in Jacobi's *Werke* volume 4, 137–142.

Jacobi, Carl Gustav J. 1866a. *Vorlesungen über Dynamik von C. G. J. Jacobi nebst Fünf Hinterlassenen Abhandlungen Desselben*. Verlag Georg Reimer. Berlin. Text of lectures delivered in the winter semester 1842–1843 at the University of Königsberg as well as five supplementary treatises. The *Vorlesungen* was published without the appendices in 1884 as a supplemental volume to Jacobi's *Werke*. The re-publication is referred to as "the second revised edition" although the two editions are virtually identical.

Jacobi, Carl Gustav J. 1866b. *Ueber diejenigen Probleme der Mechanik, in welchem eine Kräftefunction Existiert und über die Theorie der Störungen.* The first of the five supplementary works that are published in the same volume as Jacobi (1866a), pp. 303–470. Reprinted with minor editorial revision in Volume 5 of Jacobi's *Gesammelte Werke* (1890), pp. 217–395. Examples of the revision are on p. 225 (1890) (compare p. 309 (1866)) and p. 235 (1890) (compare p. 319 (1866)).

Jacobi, Carl Gustav J. 1996. Vorlesungen *über analytische Mechanik* Berlin 1847/48 Nach einer Mitschrift von Wilhelm Scheibner. Editor Helmut Pulte. Part of the Dokumente zur Geschichte der Mathematik series (Volume 8). Springer.

Jacobi, Carl Gustav J. 2009. *Jacobi's Lectures on Dynamics Delivered at the University of Königsberg in the winter semester 1842–1843 and according to the notes prepared by C. W. Brokhardt [sic]*. Translated by K. Balagangadharan, translation edited by Biswarup Banerjee. In the series Texts and Readings in Mathematics 51. Hindustan Book Agency, Gurgaon.

Jordan, Camille. 1896. *Cours d'analyse de l'École polytechnique, Tome 3: Équations Différentielles*. Deuxième édition. Gauthier-Villars. Paris.

Klein, Fleix. 1926. *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert. Teil I*. Für den Druck bearbeitet von R. Courant und O. Neugebauer. Verlag von Julius Springer. Berlin.

Klein, Fleix. 1979. *Development of mathematics in the 19th century*. English translation of Klein (1926) by M. Ackerman. Also includes an appendix by Robert Hermann titled "Kleinian mathematics from an advanced standpoint," 363–630. Math Sci Press. Brookline, Massachusetts.

Koenigsberger, Leo. 1904. *Carl Gustav Jacob Jacobi Festschrift zur Feier der hundertsten Wiederkehr seines Geburtstages*. Leipzig: B. G. Teubner.

Lagrange, Joseph Louis. 1783. "Théorie des variations séculaires des élémens des Planetes (Première Partie contenant les principes & les formules générales pour déterminer ces variations)." *Nouveaux Mémoires de l'Académie Royale des Sciences année 1781* 2e pag., 199–276. *Œuvres* V.12, 125–207.

Lagrange, Joseph Louis. 1784. "Théorie des variations séculaires des élémens des Planetes (Seconde Partie contenant la détermination de ces variations pour chacune des Planetes principales." *Nouveaux Mémoires de l'Académie Royale des Sciences année 1782* 2e pag., 169–292 ; *Œuvres* V. 5, 211–344.

Lagrange, Joseph Louis. 1788. *Méchanique analitique*. La Veuve Desaint. Paris.

Lagrange, Joseph Louis. 1809a. "Mémoire sur la théorie des variations des éléments des planètes et en particulier des variations des grands axes de leurs orbites." *Mémoires de la première Classe de l'Institut de France, Année 1808a* Volume 6 (read on 22 August 1808a) = *Oeuvres*, vol. 6, 711–768.

Lagrange, Joseph Louis. 1809b. "Mémoire sur la théorie générale de la variation des constantes arbitraires dans tous les problèmes de la mécanique." *Mémoires de la première Classe de l'Institut de France, Année 1808b*, Volume 6, (read on 13 March 1809) = *Oeuvre*s vol.6, 769–805.

Lagrange, Joseph Louis. 1810. Second mémoire sur la théorie de la variation des constantes arbitraires dans les problèmes de Mécanique. *Mémoires de la première Classe de l'Institut de France, Année 1809*, Vol. 10, (read on 19 February 1810) = *Oeuvres*, vol. 6, 807–816.

Lagrange, Joseph Louis. 1811. *Mécanique analytique*. 2nd ed., Volume 1. Ve. Courcier. Paris.

Lagrange, Joseph Louis. 1813. *Mécanique analytique*. Paris, 2nd ed., vol. 2. Ve. Courcier. Paris.

Lanczos, Cornelius. 1949. *The variational principles of mechanics*. Toronto: University of Toronto Press.

Landau, L. D., and E. M. Lifshitz. 1969. *Mechanics*. Second edition. First edition appeared in 1960. Translated by J. B. Sykes and J. S. Bell. Oxford: Pergamon Press.

Lie, Sophus. 1872. Zur Theorie partieller Differentialgleichungen. *Nachrichten von der Königlichen Gesellschaft der Wissenschaften und der Georg-Augusts-Universität zu Göttingen*, 473–489.

Lie, Sophus. 1875. Begründung einer Invarianten-Theorie der Berührungs-Transformationen. *Mathematische Annalen* 8, 215-303.

Lie, Sophus, and Georg Scheffers. 1896. *Geometrie der Berührungstransformationen*. Leipzig.

Liouville, Joseph. 1847. Sur quelques cas particuliers où les équations du mouvement d'un point matériel peuvent s'intégrer. *Journal Des Mathématiques Pures Et Appliquées* 12: 410–444.

Lützen, Jesper. 1990. *Joseph Liouville 1809–1882*. Berlin: Springer.
Mehra, Jagdish, and Helmut Rechenberg. 1982. *The historical development of quantum mechanics, vol. 1., part 1, the quantum theory of Planck, Einstein, Bohr, and Sommerfeld: Its foundation and the rise of its difficulties 1900–1925*. Springer.
Murray, Carl D., and Stanley F. Dermott. 2012. *Solar system dynamics*. Cambridge: Cambridge University Press.
Nakane, Michiyo. 2015. The origins of action angle variables and Bohr's introduction of them in a 1918 paper. In *One hundred years of the Bohr Atom: Proceedings from a conference*, ed. Finn Aaserud and Helge Kragh, 290–309. The Royal Danish Academy of Science and Letters. Copenhagen.
Nakane, Michiyo, and Craig Fraser. 2002. The early history of Hamilton-Jacobi theory. *Centaurus* 44: 161–227.
Nordheim, Lothar, and Erwin Fues. 1927. "Die Prinzipe der Dynamik. Die Hamilton-Jacobische Theorie der Dynamik. " Chapter 3 of Geiger and Scheel (1927), 91–130.
North, John. 1965. *Measure of the universe*. Oxford University Press.
O'Connor, John J., and Edmund F. Robertson. 2021. William Fishburn Donkin. Online in *MacTutor History of Mathematics Archive* at https://mathshistory.st-andrews.ac.uk/Biographies/Donkin/.
Poincaré, Henri. 1890. Sur le problème des trois corps et les équations de la dynamique. *Acta Mathematica* 13: 1–270.
Poincaré, Henri. 1892. *Les Méthodes nouvelles de la mécanique céleste*, vol. 1. Paris: Gauthier-Villars et fils.
Poincaré, Henri. 1893. *Les méthodes nouvelles de la mécanique céleste*, vol. 2. Paris: Gauthier-Villars et fils.
Poincaré, Henri. 1896. Sur les solutions périodiques et le principe de moindre action. *Comptes Rendus* 123: 915–918.
Poincaré, Henri. 1897a. Sur une forme nouvelle des équations du problème des trois corps. *Acta Mathematica* 21, 83–97.
Poincaré, Henri. 1897b. Les solutions périodiques et le principe de moindre action. *Comptes rendus* 124, 713–716.
Poincaré, Henri. 1899. *Les méthodes nouvelles de la mécanique céleste*. Volume 3. Paris : Gauthier-Villars.
Poincaré, Henri. 1905. *Leçons de mécanique céleste*. Volume 1. Paris : Gauthier-Villars et fils.
Poincaré, Henri. 1916. *Œuvres de Poincaré*. Volume 2. Edited by Gaston Darboux in collaboration with N. E. Nörlund and Ernest Lebon. Gauthier Villars et Cie. Paris.
Poincaré, Henri. 1967. *New methods of celestial mechanics*. English translation of Poincaré (1892, 1893, 1899). Translator not specified. National Aeronautics and Space Administration. NASA technical translation: F-450–52.
Poincaré, Henri. 1993. *New methods of celestial mechanics volume one*. Edited reprint of the 1967 English translation of Poincaré (1892), with a new introduction by Daniel L. Goroff. In the series *History of Modern Physics and Astronomy*, v. 13. Los Angeles, Calif.: Tomash Publishers; Woodbury, NY: American Institute of Physics.
Poincaré, Henri. 2017. *The three-body problem and the equations of dynamics*. Springer International Publishing. English translation with introduction by Bruce D. Popp of Poincaré (1890).
Poisson, Siméon Denis. 1809a. Mémoire sur les inégalités séculaires des moyens mouvements des planètes. *Journal de l'École Polytechnique*, vol.8 (15. cahier), 1–56. (Presented to the Academy on June 20, 1808).
Poisson, Siméon Denis. 1809b. Mémoire sur la variation des constantes arbitraires dans les questions de la mécanique. *Journal de l'École Polytechnique*, vol.8, (15. cahier), 266–344. (Presented to the Academy on October 16, 1809b.)
Pulte, Helmut. 1996. "Einleitung." Editor's introduction to Jacobi (1996), xviii–lviii.
Pulte, Helmut. 1997. After 150 years: news from Jacobi about Lagrange's analytical mechanics. *Mathematical Intelligencer* 19, 48–54.
Pulte, Helmut. 1998. Jacobi's criticism of Lagrange: the changing role of mathematics in the foundations of classical mechanics. *Historia Mathematica* 25, 154–184.
Pycior, Helena. 1981. George Peacock and the British origins of symbolical algebra. *Historia Mathematica* 8: 23–45.
Radau, Rodolph. 1868. Sur une transformation des équations différentielles de la dynamique. *Annales scientifiques de l'École Normale Supérieure 1re série, tome 5*, 311–375.

Radau, Rodolph. 1870. Ueber gewisse Eigenschaften der Differentialgleichungen der Dynamik. *Mathematische Annalen* 2, 167–181.

Rademacher, Hans, and Robert Iglisch. 1925. “Anfangswertprobleme.” Chapter 15 of Frank and von Mises (1925). In the section on Partielle Differentialgleichungen, 627–737 of the 1930 edition of Frank and Rademacher.

Sagan, Hans. 1969. *An introduction to the calculus of variations*. McGraw Hill. New York. Scheibner, Wilhelm. 1868. “Ueber das Problem der drei Körper,” *Journal für die Reine und Angewandte Mathematik* 68, 390–392.

Scheibner, Wilhelm. 1868. Ueber das Problem der drei Körper. (Auszug aus einem Schreiben an den Herausgeber). *Journal für die reine und angewandte Mathematik* 65: 390–392.

Schering, Ernst. 1873. Hamilton-Jacobische Theorie für Kräfte, deren Maass von der Bewegung der Körper abhängt. *Abhandlungen der Königichen Gesellschaft der Wissenschaften zu Göttigen* V. 18, *Mathematischen Classe* 3–54.

Shore, Steven N. 2003. Macrocosms/microcosms: Celestial mechanics and old quantum theory. *Historia Mathematica* 30: 494–513.

Sommerfeld, Arnold. 1919. *Atombau und Spektrallinien.* Fridr. Vieweg & Sohn. Braunschweig.

Sommerfeld, Arnold. 1923. *Atomic structure and spectral lines*. Methuen & Co. Ltd. London. English translation of the 1922 third edition (1922) of *Atombau und Spektrallinien*,

Stäckel, Paul. 1893. Ueber die Bewegung eines Punktes in einer n-fachen Mannigfaltigkeit. *Mathematische Annalen* 42: 537–563.

Tisserand, François. 1868. Exposition, d'après les principes de Jacobi, de la méthode suivie par M. Delaunay dans sa Théorie du Mouvement de la lune autour de la Terre; Extension de la méthode. *Journal de Mathématiques Pures et Appliquées*, vol.13 (2nd series), 255–303. This article also appeared in the same year as the free-standing publication of Tisserand's doctoral dissertation by Gauthier-Villars.

Tisserand, François.1889. *Traité de Mécanique Céleste Volume* 1. Volumes 2, 3 and 4 appeared respectively in 1891, 1894 and 1896. Gauthier Villars. Paris.

Ter Haar, Dirk. 1964. *Elements of Hamiltonian dynamics*, 2nd ed. Oxford: Pergamon Press.

Truesdell, Clifford. 1968. Whence the law of moment of momentum? In *Essays in the history of mechanics*, 239–271. Springer.

Trumpler, Robert J. 1933. Address of the retiring president of the society in awarding the Bruce gold medal to professor C.V.L. Charlier. *Publications of the Astronomical Society of the Pacific* 45: 5–14.

Vernier, Paul. 1894. Sur la transformation des équations canoniques du problème des trois corps. *Comptes Rendus De L'académie Des Sciences* 119: 452–454.

Vinti, John. 1998. *Orbital and celestial mechanics*, ed. Gim J. Der, Nino L. Bonavito. American Institute and Astronautics, Inc. Based on notes from lectures given by Vinti in 1966.

Von Weber, Edouard. 1900. Partielle Differentialgleichungen. *Encyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen Volume 2–5 Analysis*, 294–399. Leipzig: B. G. Teubner.

Von Weber, Edouard. 1913. Équations non linéaires du premier ordre. Équations d'ordre plus grand que un. *Équations aux derivées partielles. Encyclopédie des sciences mathématiques pures et appliquées Tome 2 Analyse Volume 4*, 56–160. Gauthier-Villars: B. G. Teubner.

Weinstock, Robert. 1952. *Calculus of variations with applications to physics and engineering.* McGraw Hill. New York. In 1974 Dover published “an unabridged and corrected republication” of the 1952 edition.

Whittaker, Edmund T. 1899. Report on the progress of the solution of the problem of three bodies. In *Report of the sixty-ninth meeting of the British Association for the advancement of science*, 121–159. London: John Murray.

Whittaker, Edmund T. 1904. *A treatise on the analytical dynamics of particles and rigid bodies: With an introduction to the problem of three bodies*. Cambridge: Cambridge University Press.

Whittaker, Edmund T. 1917. *A treatise on the analytical dynamics of particles and rigid bodies: With an introduction to the problem of three bodies second edition*. Cambridge University Press. Cambridge. The second edition is a revised and enlarged edition of the first edition, with citations and historical remarks added. The second edition was translated into German and published in 1924, and is cited by Nordheim and Fues (1927, 92, note 1).

Wilson, Curtis. 2010. *The Hill-Brown Theory of the Moon's Motion: Its Coming to Be and Short-lived Ascendancy (1877–1984)*. New York: Springer.

Wolfson, Paul. 2008. George Boole and the origins of invariant theory. *Historia Mathematica* 35: 37–46.