ENCYKLOPÄDIE DER MATHEMATISCHEN WISSENSCHAFTEN MIT EINSCHLUSS IHRER ANWENDUNGEN

VOL. IV, PART 2: MECHANICS

EDITED BY

FELIX KLEIN † AND **CONRAD MÜLLER**

IN GÖTTINGEN IN HANNOVER

THE GENERAL METHODS OF INTEGRATION IN ANALYTICAL MECHANICS

BY

GEORG PRANGE

IN HANNOVER

TRANSLATED BY

D. H. DELPHENICH

LEIPZIG

PUBLISHED AND PRINTED BY B. G. TEUBNER

1904-1935

TABLE OF CONTENTS

———————

BIBLIOGRAPHY

The theories that are presented in this treatise are treated, in part, in the numerous works on elementary analytical mechanics, and the ones that appeared up to 1906 are summarized in the bibliography to Article **6** (in this Encyclopedia). Therefore, only the recent presentations of general mechanics that appeared since 1906 will be listed here under **1**. The presentations in handbooks, etc., are given in **2.** By contrast, in **3**, the best works that are essential for all of the matter that is treated in this treatise are summarized (while repeating the ones that were listed before in Article **6**), while in **4**, the papers of a more monographic character are listed.

1. Elementary presentations of analytical mechanics since 1906.

A. Brill, *Vorlesungen sur Einführung in die Mechanik raumerfüllender Massen*, Leipzig and Berlin 1906.

A. Brill, *Vorlesungen über allgemeine Mechanik*, Munich and Berlin 1928.

G. Bouligand, *Précis de mécanique rationelle*, t. 1, Paris 1926.

P. Burgatti, *Lezioni di meccanica razionale*, 1st ed., Bologna, 1916, 2nd ed. 1919.

C. L. Charlier, *Mechanik des Himmels*, 2 vols. Leipzig 1902 and 1907.

I. Chazy, *Cours de mécanique rationelle*, 2 vols. Paris 1933.

A. E. Haas, *Die Grundgleichungen der Mechanik, dargestellt auf Grund der geschischtlichen Entwicklung*, Leipzig 1914.

G. Hamel, *Elementare Mechanik*, Leipzig and Berlin 1912.

H. Lamb, *Higher Mechanics*, Cambridge 1920.

T. Levi-Civita and **U. Amaldi**, *Compendio di meccanica razionale*, 2 vols. Bologna 1928/29.

C. H. Müller and **G. Prange**, *Allgemeine Mechanik*, Hannover 1923.

W. Müller, *Dynamik*, 2 books, Leipzig 1925.

P. Painlevé, *Cours de mécanique* I, Paris 1930.

M. Planck, *Einführung in die allgemeine Mechanik*, Leipzig 1916, 4th ed. 1928.

H. Poincaré, *Leçons de mécanique celeste*, 3 vols. Paris 1905/10.

Ch. J. de la Vallée-Poussin, *Leçons de mécanique analytique*, 2 vols. Löwen and Paris 1924/25.

2. Presentations in handbooks, etc.

Handbuch der Physik, ed. by **R. Geiger** and **K. Scheel**, Bd. V, "Grundlagen der Mechanik, Mechanik der Punkte und starren Körper," ed. by **R. Grammel**, Berlin 1927.

Handbuch der physikalischen und technischen Mechanik, ed. by **F. Auerbach** and **W. Hart**, 2 vols. Leipzig 1929 and 1930.

Die Kultur der Gegenwart III, 3, **1**, Physik, ed. by **E. Lecher**, Leipzig and Berlin 1915. In particular, art. 1 by **E. Wiechert** and art. 34 by **M. Planck** come into question.

Repertorium der Physik by **R. H. Weber** and **R. Gans** 1^1 and 1^2, Leipzig and Berlin 1915/16.

Die Differential- und Integralgleichungen der Mechanik und Physik (as the 7th ed. of **Riemann-Weber**'s *Partiellen Differentialgleichungen der mathematischen Physik*), ed. by **Ph. Frank** and **R. von Mises**, Bd. II, Braunschweig 1927, esp. Section 1.

Cl. Schaefer, *Einführung in die theoretische Physik*, Bd. I, Leipzig 1914, 3rd ed. 1929.

3. Main works cited in the present treatise

P. Appell, *Traité de mécanique rationelle*, 5 vols., Paris, 1st ed., 1893, 4th ed. 1919.

L. Boltzmann, *Vorlesungen über die Prinzipe der Mechanik*, 2 vols., Leipzig 1897 and 1904 (cited as "Prinzipe").

E. Delassus, *Leçons sur la dynamique des systèmes matériels*, Paris 1913.

H. Hertz, *Die Prinzipien der Mechanik in neuem Zusammenhange dargestellt*, Leipzig 1894 (*Ges. Werke*, Bd. III, cited as "Prinzipien").

K. Heun, *Lehrbuch der Mechanik, I Teil: Kinematik*, Leipzig 1906 (cited as "Kinematik").

C. G. J. Jacobi, *Vorlesungen über Dynamik, gehalten 1845/46*, ed. by **A. Clebsch**, Berlin 1866 (*Ges. Werke, Supplementband*, cited as "Vorlesungen," from the Supplementary Volume to the *Werke*).

J. L. Lagrange, *Mécanique analytique*, Paris 1788, *Mécanique analytique*, 2nd ed., 2 vols. 1811/14 (*Œuvres completes*, t. 11/12, cited as "Mécanique" from the *Œuvres*).

P. S. Laplace, *Traité de mécanique celeste*, 5 vols., 1799-1825 (*Œuvres completes*,t. 1-5).

T. Levi-Civita and **U. Amaldi**, *Lezioni di meccanica razionale*, vol. I, II^1, and II^2, Bologna 1923, 1926, and 1927, 2nd ed. 1930 (cited as "Lezioni").

P. Painlevé, *Leçons sur l'intégration des équations de la mécanique* (lithographed), Paris 1895 (cited as "L'intégr. d. équ. de la méc.").

H. Poincaré, *Les méthodes nouvelles de la mécanique celeste*, 3 vols., Paris 1892 to 1899 (cited as "Méthod. nouv.").

W. Thomson (Lord Kelvin) and **P. G. Tait**, *Treatise on natural philosophy*, Part I, Cambridge 1879 (cited as "Nat. Philos." from the ed. of 1923).

E. T. Whittaker, *A treatise on the analytical dynamics of particles and rigid bodies*, Cambridge 1904, 3rd ed. 1927 (cited as "dynamics"); German from the 2nd ed. of **F.** and **K. Mittelsten**. Scheid, Berlin 1924.

4. Papers of a monographic character.

P. Appell, "Les mouvements de roulement en dynamique," Scientia 4, Paris 1899.

P. Appel, "Sur une forme générale des équations de la dynamique," Mémorial des sciences math., fasc. 1, Paris 1925.

G. D. Birkhoff, *Dynamical systems*, Amer. Math. Soc. Colloq. Publ., vol. 9, New York 1927.

E. Husson, "Les trajectoires de la dynamiques," Mém. des sc. math., fsc. **55**, Paris 1932.

E. Kasner, *Differential-geometric Aspects in Dynamics*, Amer. Math. Soc. Colloq. Publ.,vol. **3** (the Princeton collection), New York 1913.

C. G. J. Jacobi, "Über diejenigen Probleme der Mechanik, in welchen eine Kräftefunktion existiert und über die Theorie der Störungen," ed. by **A. Clebsch**, Berlin 1866 (*Ges. Werke*, Bd. V, pp. 219-395, cited as "Probleme der Mech." from Bd. V of *Werke*).

Ph. E. B. Jourdain, *The principle of least action*, Chicago 1913.

Cl. Schaefer, *Die Prinzipe der Mechanik*, Berlin and Leipzig 1919.

INTRODUCTION

1. Scope of the book. Historical preliminaries. – In the plan of this *Encyklopädie*, Article 12, together with Article 13, shall define a continuation of Article 6 (**P. Stäckel**) in a certain sense. Whereas Article 6 was concerned with dynamics, to the extent that it dealt with the "elementary" methods of analysis, here, we shall report on the "advanced" methods that lead to an Ansatz for the equations of motion of mechanical systems with finite degrees of freedom and are otherwise meaningful in their integration.

In order to arrive at an Ansatz for the equations of motion, one must obviously go back to the principles of mechanics that were summarized in Article 1 (**A. Voss**), and here they will be given only a formal reshaping of their details. In that way, *generalized coordinates* [cf., IV 1 (**A. Voss**), no. **37**] will be introduced from the outset in order to characterize the configuration of a mechanical system, in which it will prove convenient to interpret the motion of a mechanical system with n degrees of freedom as the motion of the individual mass-points (the so-called "representative" points) in a space of n dimensions [cf., IV 6 (**P. Stäckel**), no. **2**]. Of course, it is simplest for one to regard such an R_n as a general (in the simplest case, **Riemannian**) manifold [cf., III D 11 (**L. Berwald**), no. **17**] whose arc-length element is inferred from the kinetic energy of the systems by means of ([1]):

$$ds^2 = 2T\,dt^2 .$$

We do not need to go further into the details of how one makes that convenient choice of general coordinates for the individual systems at this point, since that is treated in Article 11 (**K. Heun**), and on the same grounds, despite their importance for all applications, the concomitant determination of the reaction forces of the constraints can also be left out of consideration.

If one initially makes no restricting assumptions on the constraints of the system but lets them be holonomic or non-holonomic constraints [cf., IV 1 (**A. Voss**), nos. **37** and **38**] then the *differential principles* will prove to be the most far-reaching ones for defining an Ansatz for the equations of motion. In the normal cases, it would be simplest to start from the **Lagrange** formulation of **d'Alembert**'s principle [cf., IV 1 (**A. Voss**), no. **36**], and one needs to appeal to **Gauss**'s principle of least constraint [cf., IV 1 (**A. Voss**), no. **39**] only in those few cases in which the **Lagrange** formula is not sufficient.

If one could arrive at the *Ansatz* for the equations of motion by differential principles alone in that way then it would be, however, exceptionally significant that variational principles were introduced into mechanics on the basis of teleological-philosophical speculations, even if the original teleological basis for those principles is generally borrowed from the exact natural sciences today ([2]). That is because the fact that the equations of motion arise as the **Euler** equations of a variational problem proves to be fundamental in their integration, as **W. R. Hamilton** first recognized. Namely, when he started from the variational problem for **Fermat**'s principle for the

([1]) Cf., on this, **P. Stäckel**, "Bericht über die Mechanik mehrfacher Mannigfaltigkeiten," Jahresber. d. Deutsch. Math.-Vereinig. **12** (1903), pp. 469, as well as **J. L. Synge**, "On the geometry of dynamics," Trans. London Phil. Soc. (A) **226** (1927), pp. 31.

([2]) Cf., e.g., **M. Planck**, in "Die Kultur der Gegenwart," III 3 **1** (Physik), pp. 698.

shortest light path in his investigations into ray optics, he thought it would be obvious for him to focus upon the entire bundle of light rays that are emitted from a luminous point, rather than an individual light ray. He only needed to observe that in the sense of the wave theory of light, the propagation of light for the light path along the individual light ray that is taken, one can see that the *wave surfaces* of the bundle are nothing but those surfaces that are bounded by all individual rays of the bundle with *the same light path*, so the relationships between wave surfaces and rays that optics develops are just the relationships between the rays (viz., the extremals of the variational problem) and the surfaces of constant light path that are associated with the bundle by the variational problem. **Hamilton** likewise saw that this could be adapted to any variational problem, and in particular, applied his methods to mechanics by combining the trajectories (space-time lines, resp.) of the motion into suitable bundles and associated such bundles with the *characteristic function* (*principal function*, resp.) with the help of the extremal integral of the variational problem. In that way, he arrived at a representation of the integral of the equations of motion that proved to be so significant that since then that *principle of varying action* has competed with the *principle of least* (or better yet, *stationary) action* as an Ansatz for the integration of the equations of motion.

The representation of the integrals of the equations of motion that **Hamilton** arrived at with the help of the functions that he introduced made it possible for him to present the perturbation calculations [cf., VI[2], 15 (**Karl F. Sundmann**), no. **4**] for the so-called many-body problem (in particular, in the three-body problem), which he constructed in an exceptionally clear way using the process of **L. Euler** by **J. L. Lagrange**, **P. S. Laplace**, and **S. D. Poisson**, and to also extend them. At the same time, that consideration led him to give the equations of motion the general form of the *canonical system*, which is a form that had already been arrived at in some special cases of the perturbation calculations, but its proper meaning was still not appreciated.

The research of **C. G. J. Jacobi** was connected with *first-order partial differential equations* that the characteristic (principle, resp.) function satisfied, and he constructed a systematic theory of the integration of the canonical system from that. It is especially significant in it that he drew attention to the so-called ***Poisson*** *theorem* that **Poisson** had observed before in perturbation theory and that he could make it the center of the theory of integration, which is a result that found its coronation in the *theory of function groups that* **S. Lie** addressed. Meanwhile, the formulation of that theory of integration by the **Jacobi** school threatened to degenerate into pure formalism. However, its close contact with the computational astronomy always hindered the overgrowth of formulas.

When one starts from the boundary formulas of the variational problem, which already defined the starting points for **Hamilton**'s arguments, one will further arrive, with no further analysis, at the *theory of integral invariants* that **H. Poincaré** addressed. On the other hand, the boundary formula can be interpreted as a ***Pfaffian*** *expression*, which makes the canonical system take the form of the **Pfaffian** system of that expression and the **Hamilton-Jacobi** partial differential equation will take the form of the partial differential equation of the **Pfaff** problem. In that way, the analytical tools of the theory of the **Pfaff** problem will be beneficial in the definition of the Ansatz for the equations of motion, as well as their integration.

Another starting point for the systematic integration of the equations of motion is the concept of *cyclic coordinates*, which was first employed by **W. Thomson** (Lord **Kelvin**) and then, above all, by **H. von Helmholtz**. Now, the associated impulse coordinate is constant for a cyclic

coordinate, while the equations of motion will go to themselves when a cyclic coordinate is present when one performs a "parallel displacement" in the direction of that coordinate. One correspondingly casts a glance to the connection between an integral of the equation of motion and an associated one-parameter group of transformations that take the system of equations of motion into itself, which is connected with **S. Lie**'s ideas about the integration of differential equations.

The variation of constants that is used in perturbation theory makes contact with the ideas of **S. Lie** in a different way. The introduction of the constants of the "unperturbed problem" as the coordinates of the "perturbed problem" can, in fact, be interpreted as the introduction of new coordinates into the canonical system. When one now seeks to choose the constants (which was first achieved by **W. R. Hamilton** by appealing to his characteristic function) in such a way that the perturbation equations will again have the canonical form, one will arrive at a simple generalization of the idea of the general *canonical transformation*, whose systematic formulation was first taken up by **E. Schering**, and which **S. Lie** could then identify as a special case of his general *theory of contact transformations*.

Finally, **Lie**'s theory of transformation groups led to the question of the *equivalence of two mechanical problems*, which was taken up from various angles and posed by **P. Painlevé** especially.

However, the meaning of those general theories for the treatment of the individual problems was unfortunately restricted by the fact that they were, as a rule, applicable to only integration "in the small." That is because, for example, the integrals of the equations of motion are, as a rule, infinitely multi-valued, such that they will say something essential about the motion only in the neighborhood of an isolated space-time point. Thus, **H. Poincaré** showed for the three-body problem that, aside from the ten elementary integrals (viz., center of mass integrals, area integrals, and energy integral), there were no other single-valued analytic integrals [cf., IV^2, 12 (**E. T. Whittaker**), no. **4**]. Moreover, analogous statements are true for the equivalence of two mechanical problems. They can be equivalent "in the small," without needing to be correspondingly equivalent "in the large."

For the treatment of individual problems, everything will then come down to what goal one has in mind. If one restricts oneself to following the evolution of an individual process of motion then one can exploit the general theory, and one will go down that path especially when one tries to adapt the method of variation of constants that was introduced into astronomy to "terrestrial" mechanics [cf., IV 11 (**K. Heun**), no. **17**]. By contrast, if one would like to survey the totality of all possible forms of motion (i.e., integration "in the large") then deeper problems will present themselves. Here, as a preparation, one must look for the singularities of the equations of motion (cf., no. **7**) and investigate the motion more closely in the neighborhood of those special places. Above all, what has been treated more precisely in that sense is the motion in the neighborhood of the equilibrium point, which gave rise to the *theory of small oscillations* and is connected to the examination of stability. As a generalization, one must then take up the investigation of *periodic solutions*, etc., which represent "the first breach" (to use **H. Poincaré**'s term), through which one can find access to the integration "in the large." In fact, with those results (which will, however, no longer be discussed here), one seeks to gain a glimpse of the totality of all possible forms of motion and the total course of a particular one of those forms of motion. It is in the nature of things that one can next take up only entirely specialized problems that one must choose to be as simple

as possible. One thus confines oneself to problems of motion with two degrees of freedom and chooses the simplest such problem that is not trivial to be the so-called *restricted three-body problem* [cf., VI^2, 12 (**E. T. Whittaker**), no. **2**]. Along with that, the problem of force-free motion on a two-dimensional surface (at rest or in motion) in ordinary Euclidian space will also appear, which will come down to the determination of the *geodetic lines* [cf., III D 3 (**R. von Lilienthal**), nos. **14-18**] for surfaces at rest, in particular. The work of **G. D. Birkhoff** ([3]) and his school is connected with those two problem statements, which employed a continuation of **H. Poincaré**'s ideas, and in particular, the topological methods. In order to lay the foundation for such a theoretical investigation of the restricted three-body problem, numerous periodic orbits of that problem were calculated numerically at the instigation of **E. Strömgren**. The next problem was to employ that material for the construction of the theory, for which individual Ansätze were proposed in recent times ([4]).

([3]) **G. D. Birkhoff**, *Dynamical Systems*.

([4]) That work is compiled under the title: Publikationer og mindre Meddelelser fra Köbenhavns Observatorium.

CHAPTER I

THE DIFFERENTIAL EQUATIONS OF MOTION AND THEIR DEFINITION IN TERMS OF DIFFERENTIAL PRINCIPLES.

2. Concept of a mechanical system with finitely-many degrees of freedom. – Celestial and terrestrial mechanics can describe the simplest mechanical processes (falling and throwing, relative motion of a planet) by developing the equations of motion for an isolated mass-point. If one turns one's gaze from the motion of isolated planets to the total system of the Sun and its planets, along with their satellites, then one will be dealing with a *system of mass-points* whose individual points interact with each other according to *Newton's* law of gravitation. On the other hand, such a simple problem as the mathematical pendulum yields an example of the motion of a mass-point with a constraint on the motion, and indeed, one that can be given very simply in terms of mathematical formulas. One needs only to combine the two together in order to arrive at the concept of a *system of material points* on whose individual points forces act, while on the other hand, those points are coupled with each other by some sort of links that can be expressed mathematically by a number of equations (or inequalities) for the evolution of certain quantities (position and velocity) ([5]). That concept became all the more necessary due to the fact that the results of celestial mechanics, in conjunction with the resurrection of the atomistics of antiquity, brought the concept into physics that one must reduce all natural phenomena to the motion of material points that interact with each other by central forces [cf., IV 6 (**P. Stäckel**), no. **2**]. According to that conception of things, solid bodies and fluids can also be regarded as *systems of mass-points*, and indeed as systems of *infinitely-many* mass-points. For example, the *rigid body* can be regarded a system of infinitely-many mass-points whose individual points are continually kept at the same distance from each other by some sort of coupling. Similarly, by introducing somewhat-more-complicated constraints, the elastic solid body and the fluid can be called systems of infinitely-many mass-points [cf., IV 6 (**P. Stäckel**), no. **22**]. However, whereas a system of *finitely-many* mass-points can naturally always have finitely-many degrees of freedom, for systems that consist of infinitely-many points, either infinitely-many degrees can appear, as with fluids and solid (elastic) bodies, or one will arrive at a restriction to finitely-many degrees of freedom by introducing sufficiently-many constraints, as in the case of the rigid body or the chain that is constructed from rigid bodies. The latter are then associated with the mechanics of systems with *finitely-many degrees of freedom* that are treated here, while the mechanics of elastic solid bodies fall within the scope of the *mechanics of continua* [cf., IV 6 (**P. Stäckel**), no. **22**]. The bridge between the two domains is defined by *statistical mechanics* [IV 32, **P** and **T. Ehrenfest**], for which the methods that are discussed in this article will naturally take on their special meaning.

The *spatial position of a system that consists of finitely-many* (say r) *mass-points* with masses $m_1, m_2, \ldots, m_r$ is established, perhaps relative to a three-axis coordinate system, when one knows

([5]) The example of the physical pendulum was of fundamental significance in the development of dynamics. Cf., **J. L. Lagrange**, *Mécanique* 2, sect. I, nos. 7-9, (*Œuvres*, 11, pp. 248).

the 3*r* coordinates of the individual x_i, y_i, z_i ($i = 1, 2, \ldots, r$), and the motion of the system will be known when one has determined those 3*r* coordinates as functions of *time t*:

$$(1) \qquad x_i = x_i(t), \qquad y_i = y_i(t), \qquad z_i = z_i(t) .$$

In so doing, those functions must be continuous functions of time that are assumed to be differentiable sufficiently often, which is obviously physically plausible.

The general couplings in such a system will be represented by a number of equations (and possibly inequalities, but they will not be considered in what follows) between the position coordinates x_i, y_i, z_i at time *t* and the velocity components $\dot{x}_i = \frac{dx_i}{dt}$, $\dot{y}_i = \frac{dy_i}{dt}$, $\dot{z}_i = \frac{dz_i}{dt}$ of the *r* mass-points of the system. Let there be, say, *l* such equations:

$$(2) \qquad \begin{cases} f_1(x_i, y_i, z_i, \dot{x}_i, \dot{y}_i, \dot{z}_i, t) = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f_1(x_i, y_i, z_i, \dot{x}_i, \dot{y}_i, \dot{z}_i, t) = 0 \end{cases} \qquad (l < 3r) .$$

According to **H. Hertz**, one must then distinguish whether the velocity components do or do not appear in those constraint equations. **H. Hertz** called the constraint equation *holonomic* when the velocity components are missing and *non-holonomic* in the other case [cf., IV 1 (**A. Voss**), no. **37**, as well as IV 6 (**P. Stäckel**), no. **4**]. *Holonomic* and *non-holonomic* systems are very different from each other accordingly. According to **L. Boltzmann** ([5.a]), each of those two classes can be further subdivided when one decides whether the time *t* does or does not appear explicitly in the constraint equations. If the variable *t* is absent then the constraints will be called *scleronomic*, while it if does appear then one will have *rheonomic* constraints.

One arrives at the equations of motion for such constrained systems in such a way that one expresses the influence of the constraints by suitable reaction forces and then treats the mass-points of the system as free mass-points [cf., IV 6 (**P. Stäckel**), no. **4** and no. **7**]. The Ansatz of the reaction forces is thus required by the special nature of the constraints. Meanwhile, analytical mechanics is generally restricted to the case of *frictionless* constraints, for which the reaction forces are subject to the condition that they must do no work under a virtual displacement (cf., no. **3**). Therefore, the Ansatz for reaction forces with holonomic constraints is given immediately, and likewise for non-holonomic constraints, when the constraint equations are linear in the components of the velocity, which is almost the only case that appears in the applications, moreover. In the scleronomic cases of such linear non-holonomic constraints, the constraint equations will always have the form:

$$\sum_{\kappa=1}^{n} [a_{\lambda\kappa}(x_i, y_i, z_i)\,\dot{x}_\kappa + b_{\lambda\kappa}(x_i, y_i, z_i)\,\dot{y}_\kappa + c_{\lambda\kappa}(x_i, y_i, z_i)\,\dot{z}_\kappa] = 0$$

([5.a]) Cf., **L. Boltzmann**, *Prinzipe* II, § **4**, pp. 16.

in applications, i.e., they are *homogeneous in the velocity components*. As **C. Carathéodory** had remarked, that homogeneity is necessary in the spirit of vanishing friction since the reaction forces can also do no work under actual motions in the scleronomic case. As a generalization of that argument, it also seemed to **C. Carathéodory** that non-holonomic constraints that were not linear in the velocity components would also be required in order for the constraint equations:

$$f_\lambda(x_i, y_i, z_i, \dot{x}_i, \dot{y}_i, \dot{z}_i) = 0 \qquad (\lambda = 1, \ldots, l)$$

to be *homogeneous* in the velocity components in the scleronomic case, in contrast to the Ansätzen that one finds in the literature ([6]). The dependency of the functions (2) on the velocity components is also subject to certain restrictions in rheonomic cases, even though those functions no longer need to be homogeneous.

a) *Holonomic systems*. – For holonomic constraints, equations (2) will simplify to:

$$\text{(3.a)} \qquad \begin{cases} f_1(x_i, y_i, z_i) = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ f_l(x_i, y_i, z_i) = 0 \end{cases}$$

in the holonomic case and to:

$$\text{(3.b)} \qquad \begin{cases} f_1(x_i, y_i, z_i, t) = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ f_l(x_i, y_i, z_i, t) = 0 \end{cases}$$

in the case of rheonomic constraints, resp. The number of degrees of freedom is restricted to:

$$\text{(4)} \qquad n = 3r - l$$

by those constraint equations (3.a) [(3.b), resp.].

One does not necessarily need to begin with the determination of the reaction forces in order to derive the equations of motion. That is because in generalizing a process that has proved very

([6]) For practical examples of such general constraints, cf., **P. Appell**, "Exemple de movement d'un point assujetti à une liaison exprimée par une relation non linéaire entre les composantes de la Vitesse," Palermo Rend. **32** (1911), pp. 48, as well as **P. Appell**, "Sur les liaisons non linéaire par rapport aux vitesses," Palermo Rend. **33** (1912), pp. 259. Cf., also **É. Delassus**, "Sur la realization matérielle des liaisons," C. R. Acad. Sci. Paris **152** (1911), pp. 1739, as well as "Sur les liaisons non linéaire," C. R. Acad. Sci. Paris **152** (1911), pp. 626; "Sur les liaisons non linéaire et les mouvements étudiés par *M. Appell*," C. R. Acad. Sci. Paris **152** (1911), pp. 707, and furthermore "Sur les liaisons liaisons d'ordre quelconque des systèmes matériels," An. Éc. Norm. (3) **29** (1912), pp. 305, cf., also **É. Delassus**, *Dynamique des systèmes matériels*, Paris, 1913, esp. pp. 26, *et seq.*

For the Ansatz of the equations of motion when completely-general constraints are present, in which acceleration components can also appear, cf., **A. Przeborski**, "Die allgemeinsten Gleichungen der klassischen Dynamik," Math. Zeit. **36** (1933), pp. 184.

fruitful in the special cases, **J. L. Lagrange** ([7]) has shown that, in general, in the case of *scleronomic conditions* (3.a), one can dispense with those forces in such a way that the $x_1, \ldots, z_r$ are expressed as functions of n parameters $q_1, \ldots, q_n$, viz., the *generalized* coordinates, in such a way that the latter are chosen in such a way that equations (3.a) will be fulfilled identically when one introduces them. One then finds that:

$$\text{(5.a)} \qquad \begin{cases} x_i = \varphi_i(q_1, \ldots, q_n), \\ y_i = \psi_i(q_1, \ldots, q_n), \\ z_i = \chi_i(q_1, \ldots, q_n), \end{cases}$$

and the motion will be known when one has determined the $q_1, \ldots, q_n$ as functions of time [cf., IV 1 (**A. Voss**), no. **37**]. If one regards the functions:

$$\text{(6)} \qquad q_1 = q_1(t), \qquad q_2 = q_2(t), \qquad \ldots, \qquad q_n = q_n(t)$$

as representing a curve in the n-dimensional $(q_1, q_2, \ldots, q_n)$-space then one can interpret the motion of the system as a motion of *one* material point in n-dimensional space, and one must refer to the curve (6) as the trajectory of the material "representative point." If one would also like to geometrically represent the temporal course of the motion then one would have to interpret equations (6) as an analytical representation of a curve in the $(n + 1)$-dimensional $(q_1, \ldots, q_n, t)$-manifold that one can also call the $(q_1, \ldots, q_n, t)$-*world*, as a generalization of the usual terminology in the theory of relativity [cf., also IV 6 (**P. Stäckel**), no. **2**].

In general, one is also dealing with the case of holonomic constraints when a system of infinitely-many points is reduced to finitely-many degrees of freedom by infinitely-many constraints. One example is, say, the free rigid body, which possesses six degrees of freedom. As a generalization of that special case (e.g., by going to a chain of rigid bodies), one can also imagine that a system with n degrees of freedom is given here in general. The coordinates of each point of the system are then expressed by relations of the form (5) as functions of the generalized coordinates and possible time. In that way, the isolated points, as they would naturally appear in continuously-extended bodies, will be characterized by suitable parameters instead of one number. In rigid bodies, one cares to introduce, e.g., a coordinate system that is fixed in the body, in which the individual points of the rigid body might perhaps have the coordinates (a, b, c). One can then imagine that the the orientation of that comoving coordinate system with respect to the (x, y, z)-system is given by, e.g., the coordinates of the initial point and the **Euler** angles. If one denotes the former by q_1, q_2, q_3, and the latter by q_4, q_5, q_6 then the coordinates x, y, z of any point (a, b, c) of the system relative to the given reference system will be given by formulas of the form:

([7]) **J. L. Lagrange**, *Mécanique* 2, sect. IV (*Œuvres* 11, pp. 325).

$$(7)\qquad \begin{cases} x = \varphi(a,b,c,q_1,\ldots,q_6), \\ y = \psi(a,b,c,q_1,\ldots,q_6), \\ z = \chi(a,b,c,q_1,\ldots,q_6). \end{cases}$$

The motion of the rigid body is known when one knows the q_1, …, q_6 as functions of time [cf. IV 11 (**K. Heun**), no. **2**].

Moreover, the replacement of the original rectangular coordinates by general coordinates is in no way the only convenient one when there are a number of holonomic constraint equations to eliminate. Many times, the introduction of general coordinates (whose number naturally amounts to 3*r* again) offers certain advantages in the treatment of a system that is not restricted by constraint equations when one wishes to achieve an analytic representation that is suited to the special structure of the system. In that case, the relations between the x_i, y_i, z_i, and the q_1, …, q_6 will possess the form (5.a), in which $n = 3r$.

b) *Non-holonomic systems*. – If the constraint equations (2) include a number of holonomic constraints along with the non-holonomic ones then one can eliminate the former, as was just done. In order to describe the system, one might then once more appeal to *n* generalized coordinates q_1, …, q_6 again, but between which *k* non-holonomic constraints exist:

$$(8)\qquad \begin{cases} g_1(\dot{q}_1,\ldots,\dot{q}_n,q_1,\ldots,q_n,t) = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ g_k(\dot{q}_1,\ldots,\dot{q}_n,q_1,\ldots,q_n,t) = 0, \end{cases} \qquad (k < n)$$

and *none* of which should be representable as equations that are free of the $\dot{q}_1$, …, $\dot{q}_n$. Corresponding to **C. Carathéodory**'s remark, those functions must be homogeneous in the $\dot{q}_\rho$ in the scleronomic case and must also satisfy certain conditions with regard to the dependency of the $\dot{q}_\rho$ in the rheonomic case, in general.

The problem in which the constraint equations (2) are *linear in the velocity components*, which is the most important for applications, and is essentially the only problem that has been considered up to now, will also be linear in the $\dot{q}_\rho$ then, and can then be written in the form of **Pfaff** equations:

$$(9)\qquad \begin{cases} a_{11}\,dq_1 + a_{12}\,dq_2 + \cdots + a_{1n}\,dq_n + a_{1,n+1}\,dt = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{k1}\,dq_1 + a_{k2}\,dq_2 + \cdots + a_{kn}\,dq_n + a_{k,n+1}\,dt = 0 \end{cases}$$

$$a_{\sigma\rho} = a_{\sigma\rho}\,(q_1,\ \ldots,\ q_n,\ t)\ .$$

In the scleronomic case, in which time *t* does not appear explicitly, one must have $a_{\rho,\,n+1} \equiv 0$, such that the system of **Pfaff** equations will possess the form:

$$(9.a)\qquad \begin{cases} a_{11}\,dq_1 + a_{12}\,dq_2 + \cdots + a_{1n}\,dq_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{k1}\,dq_1 + a_{k2}\,dq_2 + \cdots + a_{kn}\,dq_n = 0, \end{cases}$$

$$a_{\sigma\rho} = a_{\sigma\rho}\,(q_1, \ldots, q_n)\ .$$

Such non-holonomic constraint equations give a restriction on the freedom of motion only *infinitesimally*. Indeed, $q_1, \ldots, q_n, t$ can be chosen freely, while only the choice of the velocity components $\dot{q}_\rho$ is restricted. That is a fundamental difference from the holonomic constraints, because it is therefore impossible to employ non-holonomic constraints for a further reduction of the number of coordinates. Nonetheless, one can also drop the auxiliary conditions, in a certain sense, in the definition of the equations of motion by further generalizing the concept of coordinates. Namely, since velocity components appear in the constraint equations (9), as well as the position coordinates, one can introduce the velocity components $\dot{q}_1, \ldots, \dot{q}_n$ as independent variables, along with the $q_1, \ldots, q_n, t$. The $\dot{q}_1, \ldots, \dot{q}_n$ can then be replaced with new variables – say $\omega_1, \ldots, \omega_n$ – without altering the $q_1, \ldots, q_n$, and in that way one can choose the new variables in such a way that the equations (9) [(9.a), resp.] will take on the simple form:

$$(10)\qquad \omega_{n-k+1} = 0\ , \ldots, \omega_n = 0\ .$$

In that way, the ω_ρ will be represented as linear functions of the $\dot{q}_1, \ldots, \dot{q}_n$, say in the form:

$$(11)\qquad \begin{cases} \omega_1 = \beta_{11}\,\dot{q}_1 + \cdots + \beta_{1n}\,\dot{q}_n + \beta_{1,n+1}, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \omega_n = \beta_{n1}\,\dot{q}_1 + \cdots + \beta_{nn}\,\dot{q}_n + \beta_{n,n+1}. \end{cases}$$

Moreover, one often conveniently introduces such expressions in place of the $\dot{q}_\rho$ without the appearance of non-holonomic constraints, e.g., in the motion of a rigid body, when one prefers to operate with the components of the instantaneous rotation vector, instead of the time derivatives of the **Euler** angles [cf. IV 6 (**P. Stäckel**), no. **16**]. In that case, it frequently happens that t does not appear in the $\beta_{\rho\sigma}$, and one has $\beta_{\rho,\,n+1} \equiv 0$, such that the ω_ρ will be homogeneous in the $\dot{q}_1, \ldots, \dot{q}_n$.

In many situations, infinitesimal quantities $d\kappa_\rho = \omega_\rho\,dt$ are also introduced in place of the ω_ρ that are coupled with the differentials of the position coordinates and time by relations of the form:

$$(11.a)\qquad d\kappa_\rho = \beta_{\rho 1}\,dq_1 + \ldots + \beta_{\rho n}\,dq_n + \beta_{\rho,\,n+1}\,dt.$$

Naturally, one cannot determine, say, the quantities κ_ρ as functions of the $q_1, \ldots, q_n, t$ from that. Nevertheless, the formulation of the argument will occasionally simplify when one speaks of the

κ_ρ as if they were coordinates. One then refers to them as *quasi-coordinates* ([8]). The solution of equations (11) for the $\dot{q}_\lambda$ might be:

(11.b) $$\dot{q}_\lambda = B_{1\lambda}\,\omega_1 + \ldots + B_{n\lambda}\,\omega_n + B_{n+1,\,\lambda}\,,$$

$$\sum_{\rho=1}^{n} \beta_{\rho\mu}\,B_{\rho\nu} = \begin{cases} 0 & \mu \neq \nu, \\ 1 & \mu = \nu, \end{cases} \qquad B_{n+1,\,\lambda} = -\sum_{\rho=1}^{n} \beta_{\rho,n+1}\,B_{\rho\nu}\,,$$

from which it naturally emerges that:

(11.c) $$dq_\lambda = B_{1\lambda}\,d\kappa_1 + \ldots + B_{n\lambda}\,d\kappa_n + B_{n+1,\,\lambda}\,dt\,.$$

With the introduction of quasi-coordinates, as in (11.a), the non-holonomic constraint equations (9) will take on precisely the form:

(10.a) $$d\kappa_{n-k+1} = 0\,, \quad \ldots, \quad d\kappa_n = 0\,.$$

If one then further sets:

$$\omega_1 = \dot{q}_1\,, \qquad \ldots, \qquad \omega_{n-k} = \dot{q}_{n-k}$$

then one can keep the $q_1 = \kappa_1, \ldots, q_{n-k} = \kappa_{n-k}$ as ordinary coordinates.

3. Virtual displacements in generalized coordinates. – If one would like to arrive at an Ansatz for the equations of motion of a mechanical system without having previously determined the reaction forces of the constraints then one must start from the **Lagrange** formulation of **d'Alembert**'s principle and operate with the concept of the virtual displacement of the system [cf., IV 1 (**A. Voss**), nos. **30** and **36**]. One understands a *virtual displacement of the system* to mean a conversion of its individual mass-points at their positions that are given by x_i, y_i, z_i to $x_i + \delta x_i, y_i + \delta y_i, z_i + \delta y_i$, in which the infinitesimal quantities $\delta x_i, \delta y_i, \delta z_i$ are chosen in such a way that the neighboring positions in the case of scleronomic conditions likewise satisfy the condition equations, while in the rheonomic case, they must fulfill the condition equations for fixed values of t, i.e., for the virtual displacements, rheonomic constraints must become scleronomic with the form that they have at the instant t considered ([8.a]). For *holonomic* constraints, the infinitesimal coordinate variations $\delta x_i, \delta y_i, \delta z_i$ must satisfy the condition equations:

([8]) Cf., **E. T. Whittaker**, *Dynamics*, pp. 41. **G. Hamel** employed the less-concise term "non-holonomic (generalized) coordinates." **G. Hamel**, "Über die virtuellen Verschiebungen in der Mechanik," Math. Ann. **59** (1904), pp. 416, esp. pp. 421.

([8.a]) One correspondingly cares to define the virtual displacements as "compatible with the condition equations" and "timeless." Geometrically, one can regard the virtual displacement for holonomic constraints in such a way that one interprets equations (3.a) [(3.b), resp.] as an $(n + 1 = 3r - l + 1)$-dimensional manifold in the $(3r + 1)$-dimensional space-time manifold. In the scleronomic case of equations (3.a), that M_{n+1} is, in particular, a "cylindrical" manifold

$$(12)\qquad \sum_{i=1}^{n}\left(\frac{\partial f_\rho}{\partial x_i}\,\delta x_i+\frac{\partial f_\rho}{\partial y_i}\,\delta y_i+\frac{\partial f_\rho}{\partial z_i}\,\delta z_i\right)=0 \qquad (\rho=1,\ldots,l),$$

regardless of whether one is dealing with scleronomic constraints (3) or rheonomic ones (3.a). Obviously, those equations will be nonetheless fulfilled when one chooses the infinitesimal variations $\delta q_1, \ldots, \delta q_n$ of the general coordinates (5.a) [(5.b), resp.] completely arbitrarily and then sets:

$$(13)\qquad \begin{cases} \delta x_i=\dfrac{\partial \varphi_i}{\partial q_1}\delta q_1+\dfrac{\partial \varphi_i}{\partial q_2}\delta q_2+\cdots+\dfrac{\partial \varphi_i}{\partial q_n}\delta q_n, \\ \delta y_i=\dfrac{\partial \psi_i}{\partial q_1}\delta q_1+\dfrac{\partial \psi_i}{\partial q_2}\delta q_2+\cdots+\dfrac{\partial \psi_i}{\partial q_n}\delta q_n, \\ \delta z_i=\dfrac{\partial \chi_i}{\partial q_1}\delta q_1+\dfrac{\partial \chi_i}{\partial q_2}\delta q_2+\cdots+\dfrac{\partial \chi_i}{\partial q_n}\delta q_n. \end{cases}$$

For *non-holonomic* constraints, a virtual displacement can be defined only when the condition equations are ***Pfaff*** *equations.* In that case, the infinitesimal quantities $\delta q_1, \ldots, \delta q_n$ will represent a virtual displacement of the system with the non-holonomic conditions in the form (9), as well as (9.a), when they satisfy the equations:

$$(14)\qquad \begin{cases} a_{11}\,\delta q_1+a_{12}\,\delta q_2+\cdots+a_{1n}\,\delta q_n=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{k1}\,\delta q_1+a_{k2}\,\delta q_2+\cdots+a_{kn}\,\delta q_n=0. \end{cases}$$

The explanation of virtual displacements that is then given for all cases was completely sufficient for one to apply the **Lagrange** formulation of **d'Alembert**'s principle. However, if one would later like to convert that formulation into ***Lagrange****'s central equation* and then go on to ***Hamilton****'s principle* then one must not consider an *isolated* state of motion of the system at a *fixed* time, but a *motion* of the system *during its temporal evolution:*

$$(15)\qquad q_1=\bar{q}_1(t),\qquad q_2=\bar{q}_2(t),\qquad \ldots,\qquad q_n=\bar{q}_n(t),$$

which is a representation that one can interpret as either the *trajectory of the system* in M_n with the parametric representation $q_1, \ldots, q_n$ or also as a *space-time line of the system* in the M_{n+1} of (q_ρ, t). If one now chooses the components of the virtual displacement $\delta q_1, \ldots, \delta q_n$ to be arbitrary

with generators that are parallel to the *t*-axis. In the case of rheonomic constraints, one intersects the M_{n+1} with the manifold t = const. and lays a cylindrical M_{n+1} with generators parallel to the *t*-axis through the section M_n. A neighboring position of the system that is generated by a virtual displacement must belong to the cylindrical M_{n+1} in both cases.

functions of time that are naturally continuous and differentiable sufficiently often then one will have constructed a neighboring curve in the M_n of q_ρ by way of:

$$(15.a) \qquad q_1 = \bar{q}_1(t) + \delta q_1(t), \quad q_2 = \bar{q}_2(t) + \delta q_2(t), \quad \ldots, \qquad q_n = \bar{q}_n(t) + \delta q_n(t)$$

that relates to the trajectory (15) point-wise. On the other hand, any point of the new curve is also assigned a time, namely, since:

$$(15.b) \qquad \delta t = 0,$$

it will be the same time as the associated point on the initial trajectory (15), such that (15.a) can also be regarded as the representation of a neighboring space-time curve that likewise relates to the space-time curve (15) point-wise.

For *holonomic constraints*, the neighboring curve (15.a) that is created by the virtual displacement will determine a motion of the system in its own right, under which the values of the coordinates $\bar{x}_\rho + \delta x_\rho$, $\bar{y}_\rho + \delta y_\rho$, $\bar{z}_\rho + \delta z_\rho$ of the individual system points and the associated value of the time t will likewise satisfy the condition equations, as in the case of the motion that was represented by (15) itself. The curve (15.a) then represents a motion of the system that is compatible with the conditions that can be compelled by the application of suitable supplementary forces, or as one says, the neighboring curve (15.a) is a *kinematically-possible trajectory* of the system ([9]).

Things are different for *non-holonomic constraints.* One then constructs a neighboring curve to the initial curve (15) with the help of the virtual displacements in such a way that one chooses the $\delta q_1, \ldots, \delta q_n$ to be functions of time that fulfill the conditions (14) identically and each point of the neighboring curve is assigned the same time that the associated point initial curve (15) has attained, moreover. However, if one would now like to imagine that the motion that is given by the temporal association is realized along that neighboring curve then that "neighboring motion" would neither satisfy the conditions equations (9) in the rheonomic case or the conditions (9.a) in the scleronomic case. The neighboring motion is *not kinematically possible* ([9.a]). Namely (and this will be true in the holonomic, as well as the non-holonomic case), when one denotes the variation in the velocity vector of the neighboring motion in comparison to its magnitude at the corresponding point of the initial motion by $\delta \dot{q}_1, \ldots, \delta \dot{q}_n$ then the relation will follow immediately from the given definition of the neighboring motion that:

$$(16) \qquad \delta \dot{q}_\rho = \frac{d\,\delta q_\rho}{dt},$$

which one ordinarily writes in the form:

([9]) Cf., e.g., **Müller-Prange**, *Allg. Mech.*, Chap. VI, no. **9**, pp. 520.

([9.a]) Cf., **L. Boltzmann**, *Prinzipe II*, pp. 31, or also **Müller-Prange**, *Allg. Mech.*, Chap. VI, no. **10**, pp. 531.

H. Hertz was the first to refer to that distinction in *Prinzipien*, *Ges. Werke III*, pp. 23. A further clarification was then achieved by **O. Hölder**, "Über die Prinzipien von Hamilton und Maupertuis," Gött. Nachr. (1896), pp. 122.

(16.a) $$\delta\, dq_\rho = d\ \delta q_\rho\ ,$$

and which one cares to refer to as the *commutation of the operator symbols d and* δ. Now, if the neighboring motion is kinematically possible in the non-holonomic case then since equations (14) are fulfilled for every point of the neighboring curve, one must also have:

$$\begin{aligned} &d\ (a_{11}\ \delta q_1 + \ldots + a_{1n}\ \delta q_n) = 0\ , \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &d\ (a_{k1}\ \delta q_1 + \ldots + a_{kn}\ \delta q_n) = 0\ . \end{aligned}$$

On the other hand, varying (9) will give the equations:

$$\begin{aligned} &\delta\,(a_{11}\ dq_1 + \ldots + a_{1n}\ dq_n + a_{1,n+1}\ dt) = 0\ , \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &\delta\,(a_{k1}\ dq_1 + \ldots + a_{kn}\ dq_n + a_{k,n+1}\ dt) = 0\ . \end{aligned}$$

Those two systems of equations must be compatible with each other, and it will then follow from this that the left-hand sides of (9) must be complete differentials. The same conclusions will be true in the case of equations (9.a).

The commutation relations (16) between d and δ, which arise immediately from the definition of the δq_λ as functions of time, will be true only as long as one has actual coordinates. When one is dealing with *quasi*-coordinates (cf., no. **2**), $d\,(\delta\kappa_\rho) - \delta\,(d\kappa_\rho)$ will no longer be equal to zero, but rather one will get from (11.a) that:

$$\delta\kappa_\rho = \beta_{\rho 1}\ \delta q_1 + \ldots + \beta_{\rho n}\ \delta q_n\ ,$$

and therefore:

(17) $$d\,(\delta\kappa_\rho) - \delta\,(d\kappa_\rho) = \sum_{\lambda,\mu=1}^{n}\left(\frac{\partial\beta_{\rho\lambda}}{\partial q_\mu} - \frac{\partial\beta_{\rho\mu}}{\partial q_\lambda}\right)(dq_\mu\,\delta q_\lambda - \delta q_\mu\,dq_\lambda) + dt\sum_{\lambda=1}^{n}\left(\frac{\partial\beta_{\rho\lambda}}{\partial t} - \frac{\partial\beta_{\rho,n+1}}{\partial q_\lambda}\right)\delta q_\lambda\ ,$$

in which the right-hand side vanishes only when the right-hand sides of (11.a) are complete differentials. From (11.c), one can also write (17) as:

(17.a) $$d\,(\delta\kappa_\rho) - \delta\,(d\kappa_\rho)$$

$$\begin{aligned} &= \sum_{\sigma,\tau}\left[\sum_{\lambda,\mu}\left(\frac{\partial\beta_{\rho\lambda}}{\partial q_\mu} - \frac{\partial\beta_{\rho\mu}}{\partial q_\lambda}\right)B_{\sigma\mu}B_{\tau\lambda}\right](d\kappa_\mu\,\delta\kappa_\lambda - \delta\kappa_\mu\,d\kappa_\lambda) \\ &+\ dt\sum_{\sigma}\left[\sum_{\lambda,\mu}\left(\frac{\partial\beta_{\rho\lambda}}{\partial q_\mu} - \frac{\partial\beta_{\rho\mu}}{\partial q_\lambda}\right)(B_{n+1,\mu}\,B_{\sigma\lambda} - B_{n+1,\lambda}\,B_{\sigma\mu}) + \sum_{\lambda}\left(\frac{\partial\beta_{\rho\lambda}}{\partial t} - \frac{\partial\beta_{\rho,n+1}}{\partial q_\lambda}\right)\right]\delta\kappa_\sigma \end{aligned}$$

$$= \sum_{\sigma,\tau} \gamma_{\sigma\tau}^{\rho} (d\kappa_\sigma\, \delta\kappa_\tau - \delta\kappa_\sigma\, d\kappa_\tau) + dt \sum_{\sigma} \gamma_{n+1,\sigma}^{\rho}\, \delta\kappa_\sigma ,$$

in which the second term will be missing when the term with *dt* does not appear in the definition (11.a) of quasi-coordinates. **K. Heun** ([10]) referred to those relations as the *transitivity equations* (transition equations) ([10.a]).

Moreover, one will obviously have:

(17.b) $$\gamma_{\sigma\tau}^{\rho} = -\,\gamma_{\tau\sigma}^{\rho}, \qquad \text{so} \qquad \gamma_{\sigma\sigma}^{\rho} = 0 .$$

In particular, as **G. Hamel** showed ([11]), the $\gamma_{\sigma\tau}^{\rho}$ can be related to the theory of continuous groups [cf., II A 6 (**L. Maurer** and **H. Burkhardt**)]. Namely, if one regards each of the formulas (11.c) (in which *t* might not appear explicitly) as an infinitesimal transformation and writes:

$$X_\lambda f = B_{1\lambda} \frac{\partial f}{\partial q_1} + \cdots + B_{n\lambda} \frac{\partial f}{\partial q_n}$$

correspondingly then the formulas for the bracket expression [cf., II A 6 (**L. Maurer** and **H. Burkhardt**), no. **5**] will be:

$$(X_\mu X_\nu) = \sum_{\rho} \gamma_{\mu\nu}^{\rho} X_\rho .$$

The $\gamma_{\mu\nu}^{\rho}$ are constant if and only if the *n* infinitesimal transformations generate an *n*-parameter continuous group.

Along with the virtual displacements, as they were explained here, *variations of a different type* are also applied to the trajectories (space-time curves, resp.) of the motion in mechanics, and the problem then remains of *recognizing the connection between those variations and the components of the virtual displacement.* In general, a space-time line of the motion will go to a neighboring one when one shifts each space-time point x_i, y_i, z_i, t to a neighboring one $x_i + \Delta x_i$, $y_i + \Delta y_i$, $z_i + \Delta z_i$, $t + \Delta t$.

([10]) **K. Heun**, "Die Bedeutung des *d'Alembert*schen Prinzips für starre Systeme und Gelenkmechanismen," Arch. Math. Phys. (3) **2** (1901), pp. 57 and pp. 298, esp., pp. 300. **Lagrange** had already treated the case of a rigid body that can rotate about a fixed point in the second edition of *Mécanique analytique*, t. II, 2, sect. IX, Chap. 1, § **1** (end) = *Œuvres* 12, pp. 216.

([10.a]) They were expressed in **G. Hamel**, "Die Lagrange-Eulerschen Gleichungen der Mechanik," Zeit. Math. Phys. **50** (1904), pp. 1, esp. pp 10, in which one should observe that one can set $dt = dq_{n+1} = d\kappa_{n+1}$. Cf., also **V. Volterra**, "Sopra una classe di equazioni dinamiche," Turin Atti dell'acc. delle scienze **33** (1897), pp. 451.

([11]) **G. Hamel**, Zeit. Math. Phys. **50** (1904), pp. 1, esp. pp. 10, also **G. Hamel**, "Über die virt. Versch. in der Mech.," Math. Ann. **59** (1904), pp. 416. The parallel between the virtual displacements and the infinitesimal transformations of a transitive continuous group was also employed by **H. Poincaré** in a passing remark about the way that one presents the equations of motion. **H. Poincaré**, "Sur une forme nouvelle des équations de la méc.," C. R. Acad. Sci. Paris **132** (1901), pp. 369.

In so doing, for holonomic conditions, one chooses the variations Δx_i, Δy_i, Δz_i, Δt in such a way that the varied space-time coordinates also satisfy the condition equations for the system (3.a) [(3.b), resp.]. In the case of scleronomic conditions, the system of Δx_i, Δy_i, Δz_i represents a virtual displacement of the system for each individual position of the system, with no further discussion, such that the set of all systems of values Δx_i, Δy_i, Δz_i will yield a variation of the *trajectory* that satisfies the demand that it should be a virtual displacement at each individual point. However, if one regards the time evolution of the motion on the neighboring trajectory that is determined by Δt then as a result of the variation of time, the course of the motion on the varied trajectory will be different from what it must be if the neighboring motion is to satisfy the condition $\Delta t = 0$ of the virtual displacement.

For rheonomic constraints, the Δx_i, Δy_i, Δz_i that one obtains, when considered by themselves, would not represent virtual displacements either. If one would like to obtain virtual displacements in the form of the Δx_i, Δy_i, Δz_i then one would not have to preserve the condition for the variations Δx_i, Δy_i, Δz_i, Δt in the original form:

$$f(x_i + \Delta x_i, y_i + \Delta y_i, z_i + \Delta z_i, t + \Delta t) = \Delta f(x_i, y_i, z_i, t) = 0 ,$$

but must be replaced with:

$$\Delta f(x_i, y_i, z_i, t) - \frac{\partial f}{\partial t}\Delta t = 0 . \tag{18}$$

O. Hölder ([12]) had, in fact, prescribed that condition for the variation.

Correspondingly, for non-holonomic conditions (which are linear in the differentials), one will not demand that the values $x_i + \Delta x_i$, $y_i + \Delta y_i$, $z_i + \Delta z_i$, $t + \Delta t$ must satisfy the prescribed condition equations either, but rather one prescribes equations (14) for the spatial coordinates instead of them and then imagines that Δt is added arbitrarily. As a result of that temporal ordering, one will get:

$$\Delta\left(\frac{dx_i}{dt}\right) = \frac{d(\Delta x_i)}{dt} - \dot{x}_i \frac{d(\Delta t)}{dt} , \text{ etc.} \tag{19}$$

for the change in the velocity components $\dot{x}_i$, $\dot{y}_i$, $\dot{z}_i$.

Another way of looking at things is the following one, which goes back to **E. J. Routh** ([13]) in the literature. One considers the Δx_i, Δy_i, Δz_i, Δt to be variations that make the varied space-time curves satisfy the condition equations, and one then asks whether the points of the varied space-time curve that belong to the same value of the time coordinate belong to the starting point of the actual space-time curve considered. The transition to that point will be mediated by ([14]):

([12]) **O. Hölder**, "Über die Prinzipien von *Hamilton* and *Maupertuis*," Gött. Nachr. (1826). pp. 122. Naturally, the Δx_i, Δy_i, Δz_i thus-defined, together with an arbitrary Δt, do not fulfill the condition equations, but only when $\Delta t = 0$.

([13]) **E. J. Routh**, *Dynamics of a system of rigid bodies*, German translation by **A. Schepp**, Leipzig 1898, Bd., II, Chap. 10, pp. 327.

([14]) Correspondingly, for the difference between the velocity of the neighboring curve and the initial curve at points with the same value of t, one will have:

(20) $$\Delta x_i - \dot{x}_i \Delta t , \quad \Delta y_i - \dot{y}_i \Delta t , \quad \Delta z_i - \dot{z}_i \Delta t .$$

For holonomic constraints, those are certainly virtual displacements, because in the scleronomic, as well as rheonomic, case, it likewise follows from $\Delta f = 0$ that:

(20.a) $$\sum_{i=1}^{n} \left[\frac{\partial f}{\partial x_i} (\Delta x_i - \dot{x}_i \Delta t) + \frac{\partial f}{\partial y_i} (\Delta y_i - \dot{y}_i \Delta t) + \frac{\partial f}{\partial z_i} (\Delta z_i - \dot{z}_i \Delta t) \right] = 0 .$$

If one produces the variations Δx_i, Δy_i, Δz_i, Δt in such a way that one gives the variations Δq_1, …, Δq_n ([15]), Δt to the parameters q_1, …, q_n, and t then for scleronomic conditions, the:

(21.a) $$\Delta x_i = \sum_{\rho=1}^{n} \frac{\partial \varphi_i}{\partial q_\rho} \Delta q_\rho , \qquad \text{etc.,}$$

that arise from (7.a) will naturally be virtual displacements. By contrast, for rheonomic constraints, one will get virtual displacements from ([16]):

(21.b) $$\Delta x_i - \dot{x}_i \Delta t = \sum_{\rho=1}^{n} \frac{\partial \varphi_i}{\partial q_\rho} (\Delta q_\rho - \dot{q}_\rho \Delta t) .$$

On the other hand, for non-holonomic conditions, it is never possible to produce variations such that the varied trajectories (space-time lines, resp.) will satisfy the prescribed conditions and go to the virtual displacements in the sense of the definition above in a simple way ([17]).

---

(20) $$\Delta \dot{x}_i - \ddot{x} \Delta t_i = \frac{d \Delta x_i}{dt} - \dot{x}_i \frac{d \Delta t}{dt} - \ddot{x}_i \Delta t = \frac{d}{dt} (\Delta x_i - \dot{x} \Delta t_i) ,$$

such that the change in the velocity components will be precisely the time derivative of the change in the position coordinates.

([15]) Naturally, the Δq_i can always be regarded as virtual displacements (just like the Δx_i, Δy_i, Δz_i in the scleronomic case) since one indeed says nothing at all about a time ordering for them. For example, one can always take $\Delta t = 0$.

([16]) When one has variations Δx_i, Δy_i, Δz_i, Δt such that the rheonomic conditions are fulfilled, one will get the **Hölder** variations from them in the form:

$$\sum_{\rho=1}^{n} \frac{\partial \varphi_i}{\partial q_\rho} \Delta q_\rho = \Delta \varphi_i - \frac{\partial \varphi_i}{\partial t} \Delta t = \Delta x_i - \frac{\partial x_i}{\partial t} \Delta t ,$$

which should not be confused with (21.b). Cf., **Ph. E. B. Jourdain**, "On those principles of mechanics which depend upon processes of variation," Math. Ann. **65** (1908), pp. 513. That paper was the last word in a discussion between **M. Rethy** and **Ph. E. B. Jourdain** in Math. Ann. **62** and **64** (1906 and 1907).

([17]) Here, one tries to introduce variations Δq_i, Δt that will satisfy the condition equations (9) when dq_i is replaced with Δq_i and dt is replaced with Δt. Cf., **A. Voss**, "Über die Prinzipe von Hamilton und Maupertuis," Gött. Nachr. (1900), pp. 322. However, the varied curve will naturally by no means satisfy the prescribed conditions equations in that way.

The viewpoint that **G. Hamel** ([18]) introduced rests upon a different basis. Here, as well, the variations of the position coordinates are initially determined in such a way that they are virtual displacements, but then, along with those variations δq_ρ of the position coordinates, arbitrary variations of the velocity components $\delta \dot{q}_\rho$ will be imposed that are completely independent of the latter.

The individual "elements" q_ρ, $\dot{q}_\rho$ of the initial trajectory will go to neighboring elements $q_\rho + \delta q_\rho$, $\dot{q}_\rho + \delta \dot{q}_\rho$, each of which will determine its own trajectory in its own right, moreover. However, only the individual element itself will come into question here. On the other hand, if one now focusses on only the variations $\delta q_\rho(t)$ then they will yield a neighboring curve in the parametric representation (but not the trajectory of the system that **G. Hamel** spoke of), and the tangent direction to that curve is determined by $\left(\dot{q}_\rho + \frac{d\,\delta q_\rho}{dt} \right)$. Since the increase has nothing at all to do with the variation of the velocity $\frac{d\,\delta q}{dt}$, one will naturally have:

$$\delta \dot{q}_\rho \neq \frac{d\,\delta q}{dt}, \tag{22}$$

in general, or as one can also write ([19]):

$$\delta\, dq_\rho \neq d\,\delta q_\rho\,. \tag{22.a}$$

For the x_i, y_i, z_i, it will then follow accordingly that $\delta\, dx_i \neq d\,\delta x_i$, etc., and indeed, for scleronomic conditions, one will have:

$$d\,\delta x_i - \delta\, dx_i = \sum_{\rho=1}^{n} \frac{\partial \varphi_i}{\partial q_\rho} (d\,\delta q_\rho - \delta\, dq_\rho)\ . \tag{22.b}$$

Hamel considered the rheonomic case by setting $t = q_{n+1}$, $dt = dq_{n+1}$, and $\delta t = 0$ ([20]).

4. Defining the equations of motion on the basis of Lagrange's formulation of d'Alembert's principle and Lagrange's central equation. – If forces are applied to the r mass-points whose resultant possesses the components X_i, Y_i, Z_i for the individual mass-points then the

([18]) **G. Hamel**, "Über die virt. Versch. in der Mech.," Math. Ann. **59** (1904), pp. 416. Cf., also **K. Heun**, *Kinematik*, Leipzig, 1906, pp. 94.

([19]) Since **Hamel**'s variation of the elements δq_ρ, $\delta \dot{q}_\rho$ does not produce a trajectory (a varied space-time line, either, when one adds $\delta t = 0$, resp.), it would be useless as long as one applies it to the variational problems. It is only applicable to **Lagrange**'s central equation since only the variation of the velocity components enter into it above and beyond the virtual displacements.

([20]) One must have $dt = 0$ for the virtual variation, so when one advances along the initial curve, one must also have $d\,\delta t = 0$.

virtual work that is performed by the forces under a virtual displacement of the system will be equal to:

$$\sum_{i=1}^{r}(X_i\,\delta x_i + Y_i\,\delta y_i + Z_i\,\delta z_i)\,. \tag{23}$$

If one introduces the general coordinates into that while eliminating all holonomic conditions then that will give the expression for the virtual work in the form:

$$Q_1\,\delta q_1 + \ldots + Q_n\,\delta q_n\,, \tag{23.a}$$

in which:

$$Q_\rho = \sum_{i=1}^{r}\left(X_i\frac{\partial\varphi_i}{\partial q_\rho} + Y_i\frac{\partial\psi_i}{\partial q_\rho} + Z_i\frac{\partial\chi_i}{\partial q_\rho}\right)\,, \tag{23.b}$$

regardless of whether the φ_i, ψ_i, χ_i do or do not include t explicitly. One refers to those Q_ρ as the *generalized forces* that act on the system. [Cf., IV 1 (**A. Voss**), no. **37**] When one represents the motion of the system with the picture of the motion of a mass-point in an n-dimensional space, one will have to speak of the $Q_1, \ldots, Q_n$ as the *components of the applied force* that acts on that mass-point. If one assumes that the force components X_i, Y_i, Z_i are completely-general functions of time t, the position coordinates x_i, y_i, z_i, and the velocity components $\dot{x}_i$, $\dot{y}_i$, $\dot{z}_i$ [cf., IV 6 (**P. Stäckel**), no. **4**] then the force components Q_ρ will take the form of functions of t, $q_1, \ldots, q_n$, $\dot{q}_1, \ldots, \dot{q}_n$. It is important to observe that when one goes from one system of general coordinates to another:

$$\begin{cases} \overline{q}_1 = \varphi_1(q_1,\ldots,q_n), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \overline{q}_n = \varphi_n(q_1,\ldots,q_n), \end{cases} \tag{24}$$

the virtual work, as a scalar quantity, will be invariant:

$$Q_1\,\delta q_1 + \ldots + Q_n\,\delta q_n = \overline{Q}_1\,\delta\overline{q}_1 + \cdots + \overline{Q}_n\,\delta\overline{q}_n\,,$$

so the displacement components and the force components must transform contragrediently to each other [cf., III D 10 (**R. Weitzenböck**), Part 2, no. **11**]. The transformation formulas for the *(contravariant) displacement vector:*

$$\delta\overline{q}_\rho = \frac{\partial\varphi_\rho}{\partial q_1}\delta q_1 + \cdots + \frac{\partial\varphi_\rho}{\partial q_n}\delta q_n \qquad (\rho = 1, \ldots, n) \tag{24.a}$$

will then imply the conversion formula for the components of the *covariant force vector:*

(24.b) $$\bar{Q}_1 \frac{\partial \varphi_1}{\partial q_\sigma} + \cdots + \bar{Q}_n \frac{\partial \varphi_n}{\partial q_\sigma} \delta q_n = Q_\rho \qquad (\sigma = 1, \ldots, n).$$

In order to exhibit the equations of motion of the system, one can now start from *the Lagrange formulation of d'Alembert's principle* [cf., IV 1 (**A. Voss**), no. **36**] and demand that:

(25) $$\sum \{(X_i - m_i \ddot{x}_i)\delta x_i + (Y_i - m_i \ddot{y}_i)\delta y_i + (Z_i - m_i \ddot{z}_i)\delta z_i\} = 0$$

for all virtual displacements of the system. For holonomic constraints, as well as for non-holonomic constraints in the form of **Pfaff** expressions, one can generally obtain the equations of motion in that way ([21]). Moreover, it is not convenient to introduce the general coordinates q_ρ into formula (25) directly since one will have to perform the laborious calculation of the second derivatives. Therefore, **J. L. Lagrange** ([22]) had already put that formula into another form to which **K. Heun** gave the name of the *central equation.* Namely, if one regards the virtual displacements δx_i, etc., as functions of time that satisfy conditions $\delta \dot{x}_i = d\,\delta x_i / dt$, etc., that correspond to (16) then one will get from (25) that:

$$\sum_{i=1}^{r} \{(X_i\,\delta x_i + Y_i\,\delta y_i + Z_i\,\delta z_i) + m_i\,(\dot{x}_i\,\delta \dot{x}_i + \dot{y}_i\,\delta \dot{y}_i + \dot{z}_i\,\delta \dot{z}_i) - \frac{d}{dt}[m_i(\dot{x}_i\,\delta x_i + \dot{y}_i\,\delta y_i + \dot{z}_i\,\delta z_i)]\} = 0\,,$$

or when one introduces the *kinetic energy* of the system:

(26) $$T = \sum_{i=1}^{r} \tfrac{1}{2} m_i\,(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2)\,,$$

that

(27) $$\delta T + \sum_{i=1}^{r} (X_i\,\delta x_i + Y_i\,\delta y_i + Z_i\,\delta z_i) = \frac{d}{dt}\left[\sum_{i=1}^{r} m_i(\dot{x}_i\,\delta x_i + \dot{y}_i\,\delta y_i + \dot{z}_i\,\delta z_i)\right],$$

in which the expression to the right of the time derivative represents the *virtual work done by impulse*. (It is work done by impulse, since it emerges from the expression for the work done by the applied forces when one replaces the force components X_i, etc., with the impulse components $m_i\,\dot{x}_i$, etc.)

If one now introduces the general coordinates into the kinetic energy (26) by means of (5.a) [(5.b), resp.] then *in the scleronomic case, T will become a quadratic form in* $\dot{q}_\rho$, say:

([21]) **J. Hadamard**, "Sur la mise en équation des problèmes de mécanique," Nouvelles annales de math. (4) **6** (1906), pp. 97 pointed out that when one knows the kinematical mobility of constrained systems, one can see immediately whether a system of equations of motion whose number of equations equals the number of degrees of freedom consists of independent equations and is then sufficient to determine the motion.

([22]) **J. L. Lagrange**, *Mécanique*, 2 part., 4 sect. = *Œuvres* 11, pp. 325.

(26.a) $$T = \tfrac{1}{2}\sum_{\lambda,\mu=1}^{n} g_{\lambda\mu}\,\dot{q}_\lambda\,\dot{q}_\mu\,, \qquad g_{\lambda\mu} = g_{\lambda\mu}\,(q_1, \ldots, q_n)\,,$$

while, *in the rheonomic case, it will be a quadratic function*, in which, along with the terms that are quadratic in $\dot{q}_\rho$, terms that are linear in them and a term that is free of the $\dot{q}_\rho$ will appear:

(26.b) $$T = \tfrac{1}{2}g_{00} + \sum_{\lambda=1}^{n} g_{0\lambda}\,\dot{q}_\lambda + \tfrac{1}{2}\sum_{\lambda,\mu=1}^{n} g_{\lambda\mu}\,\dot{q}_\lambda\,\dot{q}_\mu = T_0 + T_1 + T_2\,,$$

in which the g_{00}, $g_{0\lambda}$, $g_{\lambda\mu}$ can include time t explicitly, in addition to q_ρ ([23]):

(26.b*) $$g_{\sigma\tau} = g_{\sigma\tau}\,(q_1, \ldots, q_n, t)\,.$$

The introduction of the general coordinates into the virtual work done by impulse yields:

(28) $$\sum_{i=1}^{r} m_i(\dot{x}_i\,\delta x_i + \dot{y}_i\,\delta y_i + \dot{z}_i\,\delta z_i) = \sum_{i=1}^{r}\left(\frac{\partial T}{\partial \dot{x}_i}\delta x_i + \frac{\partial T}{\partial \dot{y}_i}\delta y_i + \frac{\partial T}{\partial \dot{z}_i}\delta z_i\right)$$

$$= \sum_{\rho=1}^{n}\left[\sum_{i=1}^{r}\left(\frac{\partial T}{\partial \dot{x}_i}\frac{\partial x_i}{\partial q_\rho} + \frac{\partial T}{\partial \dot{y}_i}\frac{\partial y_i}{\partial q_\rho} + \frac{\partial T}{\partial \dot{z}_i}\frac{\partial z_i}{\partial q_\rho}\right)\delta q_\rho\right] = \sum_{\rho=1}^{n}\frac{\partial T}{\partial \dot{q}_\rho}\,\delta q_\rho = \sum_{\rho=1}^{n} p_\rho\,\delta q_\rho\,,$$

in which the relations $\dfrac{\partial x_i}{\partial q_\rho} = \dfrac{\partial \dot{x}_i}{\partial \dot{q}_\rho}$ were used. [Cf., IV 1 (**A. Voss**), no. **37**]. The quantities:

(29) $$p_\rho = \frac{\partial T}{\partial \dot{q}_\rho}$$

in that are the *general impulse components*. Since the work done by impulse is also invariant under the introduction of new coordinates, the general impulse coordinates will transform in the same way as the general force components:

(29.a) $$\bar{p}_1\frac{\partial \varphi_1}{\partial q_\rho} + \cdots + \bar{p}_n\frac{\partial \varphi_n}{\partial q_\rho} = p_\sigma \qquad (\sigma = 1, \ldots, n),$$

i.e., the *impulse is a covariant vector*, like the applied force.

([23]) That form was given before, e.g., **R. Lehmann-Filhès**, "Über einige Fundamentalsätze der Dynamik," Astr. Nachr. **125** (1890), pp. 49.

If one appeals to (23.a) and (28) then *Lagrange's central equation* (27) will take the form ([24]):

(27.a) $$\delta T + (Q_1\,\delta q_1 + \ldots + Q_n\,\delta q_n) = \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_1}\delta q_1 + \cdots + \frac{\partial T}{\partial \dot{q}_n}\delta q_n\right)$$

or

(27.b) $$\frac{d}{dt}\left(\left(\frac{\partial T}{\partial \dot{q}_1}\right) - \frac{\partial T}{\partial q_n} - Q_1\right)\delta q_1 + \cdots + \frac{d}{dt}\left(\left(\frac{\partial T}{\partial \dot{q}_n}\right) - \frac{\partial T}{\partial q_n} - Q_n\right)\delta q_n = 0\,.$$

If there are no non-holonomic constraints present, and if the *holonomic* constraints are eliminated by introducing the general coordinates then the δq_ρ will be completely arbitrary, and one can then conclude the *Lagrange equations* (art. **2**) of motion:

(30) $$\left(\frac{\partial T}{\partial \dot{q}_1}\right) - \frac{\partial T}{\partial q_n} = Q_\rho \qquad (\rho = 1, \ldots, n)$$

from the central equation. [Cf., IV 1 (**A. Voss**), no. **37** and IV 6 (**P. Stäckel**), no. **7**]. The left-hand sides of these equations transform in the same way as the components Q_ρ of the covariant force vector on the right-hand sides under the introduction of new coordinates. One then has the transformation formulas:

(30.a) $$\left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_1}\right) - \frac{\partial T}{\partial q_n}\right)\frac{\partial \varphi_1}{\partial q_\sigma} + \cdots + \left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_n}\right) - \frac{\partial T}{\partial q_n}\right)\frac{\partial \varphi_n}{\partial q_\sigma} = \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\sigma}\right) - \frac{\partial T}{\partial q_\sigma},$$

as one also easily confirms by direct calculation. Therefore, the left-hand sides of (30) are the components of a covariant vector, namely, the *general acceleration vector*. The great significance of the **Lagrange** equations (30) is based in that invariance under the introduction of new coordinates.

If *non-holonomic* constraints in the form of **Pfaff** equations then one can consider the conditions (14) with the help of **Lagrange** factors and obtain the equations of motion in the form:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\sigma}\right) - \frac{\partial T}{\partial q_\sigma} = Q_\rho + \sum_{\sigma=1}^{k} \lambda_\sigma\, a_{\sigma\rho} \qquad (\rho = 1, \ldots, n),$$

([24]) Under the assumption that $d\,\delta q_\rho \neq \delta\, dq_\rho$, **G. Hamel** wrote the so-called *central equation* in its place, which reads:

(27.c) $$\delta T + \sum_\rho Q_\rho\,\delta q_\rho + \sum_\rho \frac{\partial T}{\partial \dot{q}_\rho}\left(\frac{d\,\delta q_\rho}{dt} - \delta\frac{dq_\rho}{dt}\right) = \frac{d}{dt}\left(\sum_\rho \frac{\partial T}{\partial \dot{q}_\rho}\delta q_\rho\right).$$

Cf., **G. Hamel**, "Über die virt. Verschieb. i. d. Mechanik," Math. Ann. **59** (1904), pp. 416, see esp. pp. 424.

which determine the $(n + k)$ unknowns $q_1, \ldots, q_n, \lambda_1, \ldots, \lambda_k$, together with the k equations (9) [(9.a), resp.].

On the other hand, for *non-holonomic* constraints, one can avoid the appearance of **Lagrange** factors in the equations of motion completely by *introducing quasi-coordinates* and exploiting the equations of motion in them in such a way that the number of equations of motion will take the least-possible value. That is because, from (11.c), the virtual displacements δq_λ are expressed in terms of the $\delta\kappa_\rho$ of the quasi-coordinates in the form:

$$\delta q_\lambda = B_{1\lambda}\,\delta\kappa_1 + \ldots + B_{n\lambda}\,\delta\kappa_n \,. \tag{31}$$

If one then introduces the quasi-coordinates in such a way that the non-holonomic constraints impose simply the conditions:

$$\delta\kappa_{n-k+1} = 0 \,, \quad \ldots, \quad \delta\kappa_n = 0$$

on the virtual displacements $\delta\kappa_\rho$ then one will get:

$$\delta q_\lambda = B_{1\lambda}\,\delta\kappa_1 + \ldots + B_{n-k,\,\lambda}\,\delta\kappa_{n-k} \,,$$

in which the $\delta\kappa_1, \ldots, \delta\kappa_{n-k}$ are quantities that can be chosen arbitrarily. If one then substitutes those values in the form (27.b) of **Lagrange**'s central equation then that will give:

$$\sum_{\rho=1}^{n-k}\left[\sum_{\lambda=1}^{n} B_{\rho\lambda}\left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot q_\lambda}\right) - \frac{\partial T}{\partial q_\lambda} - Q_\lambda\right)\right]\delta\kappa_\rho = 0 \,,$$

and one finds the equations of motion from that in the form:

$$\sum_{\lambda=1}^{n} B_{\rho\lambda}\left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot q_\lambda}\right) - \frac{\partial T}{\partial q_\lambda}\right) = \sum_{\lambda=1}^{n} B_{\rho\lambda} Q_\lambda \qquad (\rho = 1, \ldots, n-k)$$

that **G. A. Maggi** gave ([25]), which determine the coordinates $q_1, \ldots, q_k$ as functions of t, together with the k equations (9) [(9.a), resp.], but likewise possess an even-more complicated form.

One will come to an essentially clearer form for the equations of motion when one introduces the quasi-coordinates into the central equation (27.a) immediately. However, since the use of quasi-coordinates in the equations of motion can also be of value when no non-holonomic constraints are present (as the ***Euler** equations of motion of the top* show), the introduction of quasi-coordinates into **Lagrange**'s central equation will be possible in full generality. From (11.b), the kinetic energy T will be a function of degree two in the ω_σ :

([25]) **G. A. Maggi**, "Di alcune nuove forme delle equ. della dinam. applicabili ai sistemi anolonomi," Rom. Acc. Linc. Rend. (5) 10^2 (1901), pp. 287.

(32) $$T^* = \tfrac{1}{2} c_{00} + \sum_{\rho=1}^{n} c_{0\rho}\,\omega_\rho + \tfrac{1}{2}\sum_{\lambda,\mu=1}^{n} c_{\lambda\mu}\,\omega_\lambda\,\omega_\mu\,,$$

in which the $c_{\rho\sigma}$ are once more functions of the $q_1, q_2, \ldots, q_n, t$ ([26]). The recalculation of the virtual work done by impulse in quasi-coordinates then gives:

(33) $$\sum_\lambda \frac{\partial T}{\partial \dot q_\lambda}\delta q_\lambda = \sum_\rho \frac{\partial T^*}{\partial \omega_\rho}\left(\sum_\lambda \frac{\partial \omega_\rho}{\partial \dot q_\lambda}\delta q_\lambda\right) = \sum_\rho \frac{\partial T^*}{\partial \omega_\rho}\left(\sum_\lambda \beta_{\rho\lambda}\,\delta q_\lambda\right) = \sum_\rho \frac{\partial T^*}{\partial \omega_\rho}\delta\kappa_\rho\,.$$

Corresponding to the interpretation of the $p_\rho = \partial T^* / \partial \dot q_\rho$ as the impulse components that belong to the general coordinates q_ρ, the quantities:

(33.a) $$J_\rho = \frac{\partial T^*}{\partial \omega_\rho}$$

will be referred to as the *quasi-impulses* of the system. In the same way that the general impulses p_ρ are linear functions of the general velocity components $\dot q_1, \ldots, \dot q_n$, the quasi-impulses J_λ are linear functions of the *quasi-components of the velocity* $\omega_1, \ldots, \omega_n$.

Analogously, one will get the recalculation of the *virtual work done by the applied forces:*

(34) $$\begin{aligned} Q_1\,\delta q_1 + \ldots + Q_n\,\delta q_n &= \sum_{\rho=1}^{n} Q_\rho\left(\sum_\sigma B_{\sigma\rho}\,\delta\kappa_\sigma\right) = \sum_{\sigma=1}^{n}\left(\sum_{\rho=1}^{n} Q_\rho\,B_{\sigma\rho}\right)\delta\kappa_\sigma \\ &= K_1\,\delta\kappa_1 + \ldots + K_n\,\delta\kappa_n\,, \end{aligned}$$

in which one sets:

(34.a) $$K_\sigma = B_{\sigma 1}\,Q_1 + B_{\sigma 2}\,Q_2 + \ldots + B_{\sigma n}\,Q_n\,,$$

or

([26]) In particular, one has:

(32.a) $$\begin{cases} c_{00} = g_{00} + \sum_\lambda g_{0\lambda}\,B_{n+1,\lambda} + \sum_{\lambda,\mu} g_{\lambda\mu}\,B_{n+1,\mu}, \\ c_{0\rho} = \sum_\lambda g_{0\lambda}\,B_{\rho\lambda} + \sum_{\lambda,\mu} g_{\lambda\mu}(B_{\rho\lambda}B_{n+1,\mu} + B_{n+1,\lambda}B_{\rho\mu}), \\ c_{\rho\sigma} = \sum_{\lambda,\mu} g_{\lambda\mu}\,B_{\rho\lambda}B_{\sigma\mu}\,. \end{cases}$$

If T were originally a quadratic form in $\dot q_\rho$ and one expresses the ω_σ as linear forms in the $\dot q_\rho$ then T^* will become a quadratic form in the ω_σ.

(34.b) $$Q_\rho = \beta_{1\rho} K_1 + \beta_{2\rho} K_2 + \ldots + \beta_{n\rho} K_n ,$$

resp. ([27]), and the K_1, …, K_n can perhaps be referred to as the *quasi-components of the applied force*.

The ***Lagrange*** *central equation* will then take on the following form *for quasi-coordinates* ([28]):

(35) $$\delta T^* + (K_1 \delta\kappa_1 + \cdots + K_n \delta\kappa_n) = \frac{d}{dt}\left(\frac{\partial T^*}{\partial \omega_1}\delta\kappa_1 + \cdots + \frac{\partial T^*}{\partial \omega_n}\delta\kappa_n \right),$$

which can be regarded as a generalization of (27.a).

In order to arrive at the equations of motion from that, one replaces the δq_ρ in δT^* with $\delta\kappa_\rho$ by means of:

$$\sum_\rho \frac{\partial T^*}{\partial q_\rho}\delta q_\rho = \sum_\rho \left[\frac{\partial T^*}{\partial q_\rho}\left(\sum_\sigma B_{\sigma\rho}\delta\kappa_\sigma \right) \right] = \sum_\sigma \left(\sum_\rho \frac{\partial T^*}{\partial q_\rho} B_{\sigma\rho} \right)\delta\kappa_\sigma .$$

Here, one can symbolically denote ([29]):

(36) $$\sum_\rho \frac{\partial T^*}{\partial q_\rho} B_{\sigma\rho} \quad \text{by} \quad \left(\frac{\partial T^*}{\partial \kappa_\sigma} \right),$$

which will make:

$$\delta T^* = \sum_\rho \left[\frac{\partial T^*}{\partial \omega_\rho}\delta\omega_\rho + \left(\frac{\partial T^*}{\partial \kappa_\rho} \right)\delta\kappa_\rho \right] = \sum_\rho \left[J_\rho \, \delta\omega_\rho + \left(\frac{\partial T^*}{\partial \kappa_\rho} \right)\delta\kappa_\rho \right].$$

With that, the central equation (35) will then take on the form:

$$\sum_{\rho=1}^{n} \left[\left(\frac{d}{dt}\left(\frac{\partial T^*}{\partial \omega_\rho} \right) - \left(\frac{\partial T^*}{\partial \kappa_\rho} \right) - K_\rho \right)\delta\kappa_\rho + \frac{\partial T^*}{\partial \omega_\rho}\left(\frac{d\,\delta\kappa_\rho}{dt} - \delta\frac{d\kappa_\rho}{dt} \right) \right] = 0$$

([27]) Since the work is a scalar, the force components must be substituted contragrediently to the displacement components.

([28]) Cf., **V. Volterra**, "Sopra una classe di equ. din.," Turin Atti **33** (1897), pp. 451, as well as **G. Hamel**, "Die Lagrange-Eulerschen Gleichungen der Mechanik," Zeit. Math. Phys. **50** (1904), pp. 1, esp. pp. 15.

([29]) If the components of the virtual displacements define the infinitesimal transformations of a transitive continuous group then **H. Poincaré** [cf., ([11])] obtained the equations of motion in the form:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \omega_s} \right) = \sum_i c_{ski} \frac{\partial T}{\partial \omega_i}\omega_k + \sum_i \left(\frac{\partial T}{\partial q_i} + Q_i \right) B_{si} ,$$

which are included in the equation above as a special case. **N. Četajew**, C. R. Acad. Sci. Paris **185** (1927), pp. 1577 has rewritten it by introducing the quasi-impulses.

in which the second terms in the individual summands are converted with the help of the transitivity equations (17.a). That will ultimately give:

$$(37) \qquad \sum_{\rho=1}^{n}\left\{\frac{d}{dt}\left(\frac{\partial T^*}{\partial \omega_\rho}\right)+\sum_{\lambda,\mu=1}^{n}\gamma_{\mu\rho}^{\lambda}\frac{\partial T^*}{\partial \omega_\lambda}\omega_\mu+\sum_{\lambda=1}^{n}\gamma_{n+1,\rho}^{\lambda}\frac{\partial T^*}{\partial \omega_\lambda}-\left(\frac{\partial T^*}{\partial \kappa_\rho}\right)-K_\rho\right\}\delta\kappa_\rho=0$$

then. Now, when *there are no auxiliary conditions*, the $\delta\kappa_\rho$ will be completely arbitrary quantities, and that will give the following *n* relations as the *equations of motion for the quasi-coordinates:*

$$(38) \qquad \frac{d}{dt}\left(\frac{\partial T^*}{\partial \omega_\rho}\right)+\sum_{\lambda,\mu=1}^{n}\gamma_{\mu\rho}^{\lambda}\frac{\partial T^*}{\partial \omega_\lambda}\omega_\mu+\sum_{\lambda=1}^{n}\gamma_{n+1,\rho}^{\lambda}\frac{\partial T^*}{\partial \omega_\lambda}-\left(\frac{\partial T^*}{\partial \kappa_\rho}\right)-K_\rho=0\,,$$

and together with the *n* equations (11):

$$\omega_\rho=\beta_{\rho 1}\dot{q}_1+\cdots+\beta_{\rho n}\dot{q}_n+\beta_{\rho,n+1}$$

that represents a system of 2*n* first-order differential equations for the 2*n* unknown functions q_1, …, q_n, ω_1, …, ω_n . Those equations will be even more useful in the applications when one introduces the quasi-impulses J_1, …, J_n in place of the ω_1, …, ω_n in them by means of (33.a), in which one naturally has to replace the relations (11) with relations between the J_ρ and $\dot{q}_\rho$. One will then get 2*n* first-order differential equations in the 2*n* unknowns J_1, …, J_n, q_1, …, q_n . Since they were given before by **Euler** in the case of rigid bodies and further analyzed by **Lagrange** [cf., IV 6 (**P. Stäckel**), nos. **29** and **30**], **G. Hamel** gave them the name of the *Lagrange-Euler equations of mechanics.*

On the grounds of the form (37) of **Lagrange**'s central equation, *non-holonomic* constraints that are represented by **Pfaff** equations in the form (9) can be exploited to reduce the number of equations of motion. Corresponding to the system (9), one introduces:

$$(39) \qquad \omega_1=\dot{q}_1\,, \qquad \ldots, \qquad \omega_{n-k}=\dot{q}_{n-k}$$

and further sets:

$$(39.a) \qquad \begin{cases} \omega_{n-k+1} & =a_{11}\dot{q}_1+\cdots+a_{1n}\dot{q}_n+a_{1,n+1}, \\ \ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \omega_n & =a_{k1}\dot{q}_1+\cdots+a_{kn}\dot{q}_n+a_{k,n+1}. \end{cases}$$

The auxiliary conditions will then be simply:

$$(39.b) \qquad \omega_{n-k+1}=0\,, \qquad \ldots, \qquad \omega_n=0\,,$$

and the conditions for the virtual displacement will read:

$$\text{(39.c)} \qquad \delta\kappa_{n-k+1} = 0\,, \quad \ldots, \quad \delta\kappa_n = 0$$

in the associated quasi-coordinates, such that the sum runs from only 1 to $n - k$ in the variational formula (37). Since the remaining quantities $\delta\kappa_1, \ldots, \delta\kappa_{n-k}$ remain completely arbitrary, one will get the $(n - k)$ equations:

$$\text{(40)} \qquad \frac{d}{dt}\left(\frac{\partial T^*}{\partial \omega_\rho}\right) + \sum_{\lambda=1}^{n}\left(\sum_{\mu=1}^{n-k}\gamma_{\mu\rho}^{\lambda}\,\omega_\mu\right)\frac{\partial T^*}{\partial \omega_\lambda} + \sum_{\lambda=1}^{n}\gamma_{n+1,\rho}^{\lambda}\frac{\partial T^*}{\partial \omega_\lambda} - \left(\frac{\partial T^*}{\partial \kappa_\rho}\right) = K_\rho \qquad (\rho = 1, \ldots, n-k)$$

(in which the $\omega_{n-k+1}, \ldots, \omega_n$ are set equal to zero in the differentiation), which determine the $2n - k$ unknowns $\omega_1, \ldots, \omega_{n-k}, q_1, \ldots, q_n$ along with the n equations (39) and (9).

However, when one now recalls (39), the $\kappa_1 = q_1, \ldots, \kappa_{n-k} = q_{n-k}$ are not quasi-coordinates at all, but actual coordinates. One then sees from (17.a) that one has:

$$\text{(39.d)} \qquad \gamma_{\mu\rho}^{\lambda} \equiv 0\,, \qquad \gamma_{n+1,\rho}^{\lambda} \equiv 0 \qquad (\lambda = 1, 2, \ldots, n-k),$$

along with (39.b), i.e., that the *equations of motion* (40) read, more simply ([30]):

$$\text{(40.a)} \qquad \frac{d}{dt}\left(\frac{\partial T^*}{\partial \omega_\rho}\right) + \sum_{\lambda=n-k+1}^{n}\left(\sum_{\mu=1}^{n-k}\gamma_{\mu\rho}^{\lambda}\,\omega_\mu\right)\frac{\partial T^*}{\partial \omega_\lambda} + \sum_{\lambda=n-k+1}^{n}\gamma_{n+1,\rho}^{\lambda}\frac{\partial T^*}{\partial \omega_\lambda} - \left(\frac{\partial T^*}{\partial \kappa_\rho}\right) = K_\rho \qquad (\rho = 1, \ldots, n-k),$$

([30]) Now, in that way, from (34.b), the K_ρ are coupled with $Q_1, \ldots, Q_n$ by the equations:

$$\text{(40.b)} \qquad Q_\rho = K_\rho + a_{1\rho}K_{n-k+1} + a_{2\rho}K_{n-k+2} + \ldots + a_{k\rho}K_n \qquad (\rho = 1, \ldots, n-k),$$

and analogously, from (36), the derivatives $\dfrac{\partial T^*}{\partial q_\rho}$ and $\left(\dfrac{\partial T^*}{\partial \kappa_\rho}\right)$ are coupled by:

$$\text{(40.c)} \qquad \frac{\partial T^*}{\partial q_\rho} = \left(\frac{\partial T^*}{\partial \kappa_\rho}\right) + a_{1\rho}\left(\frac{\partial T^*}{\partial \kappa_{n-k+1}}\right) + \cdots + a_{k\rho}\left(\frac{\partial T^*}{\partial \kappa_n}\right) \qquad (\rho = 1, \ldots, n-k)\,,$$

while for $r > n - k$, one has:

$$\text{(40.b*)} \qquad Q_\rho = a_{1\rho}K_{n-k+1} + \ldots + a_{k\rho}K_n\,,$$

$$\text{(40.c*)} \qquad \frac{\partial T^*}{\partial q_\rho} = a_{1\rho}\left(\frac{\partial T^*}{\partial \kappa_{n-k+1}}\right) + \cdots + a_{k\rho}\left(\frac{\partial T^*}{\partial \kappa_n}\right)\,.$$

The $\gamma_{\mu\rho}^{\lambda}$ and $\gamma_{n+1,\rho}^{\lambda}$ are also greatly simplified in comparison to the general formula (17.a). Cf., **G. Hamel**, "Die Lagrange-Eulerschen Gleichungen der Mechanik," Zeit. Math. Phys. **50** (1904), esp., pp. 21.

in which one sets $\omega_1 = \dot{q}_1$, …, $\omega_{n-k} = \dot{q}_{n-k}$, $\omega_{n-k+1} = 0$, …, $\omega_n = 0$ after the differentiation. They are $(n - k)$ second-order differential equations for the $q_1, \ldots, q_n$. The $q_1, \ldots, q_n$ can be determined as functions of time t from them and the k first-order differential equations that the conditions (9) produce. The equations of motion will have been reduced to the least-possible number then.

All sorts of flawed arguments were then made in regard to exhibiting the least-possible number of equations of motion for non-holonomic constraints. The correct form was then achieved almost simultaneously by several parties. The representation that is given here is closely connected with **G. Hamel** ([31]). **Hamel**'s general representation was already the basis for the arguments of **P. Woronetz** ([32]), who solved the k non-holonomic conditions (9) for $\dot{q}_{n-k+1}, \ldots, \dot{q}_n$:

$$(41)\qquad \left\{\begin{array}{ll} \dot{q}_{n-k+1} & = b_{11}\,\dot{q}_1 + \cdots + b_{1,n-k}\,\dot{q}_{n-k} + b_1, \\ \dot{q}_{n-k+1} & = b_{21}\,\dot{q}_1 + \cdots + b_{2,n-k}\,\dot{q}_{n-k} + b_2, \\ \ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \dot{q}_n & = b_{k1}\,\dot{q}_1 + \cdots + b_{k,n-k}\,\dot{q}_{n-k} + b_n, \end{array}\right.$$

and eliminated the $\dot{q}_{n-k+1}, \ldots, \dot{q}_n$ from the kinetic energy T corresponding to the ideas that were expressed in (39), so from (32), it then took the form:

$$(42)\qquad T^* = \tfrac{1}{2}\,c_{00} + \sum_{\rho=1}^{n-k} c_{0\rho}\,\dot{q}_\rho + \tfrac{1}{2}\sum_{\lambda,\mu=1}^{n-k} c_{\lambda\mu}\,\dot{q}_\lambda\,\dot{q}_\mu\,,$$

in which the c_{00}, $c_{0\rho}$, and $c_{\lambda\mu}$ are functions of $q_1, \ldots, q_n, t$. The variational formula (35) then reads:

$$(43)\quad \delta T^* + (Q_1\,\delta q_1 + \cdots + Q_{n-k}\,\delta q_{n-k} + Q_{n-k+1}\,\delta q_{n-k+1} + \cdots + Q_n\,\delta q_n) = \frac{d}{dt}\left(\frac{\partial T^*}{\partial \dot{q}_1}\delta q_1 + \cdots + \frac{\partial T^*}{\partial \dot{q}_{n-k}}\delta q_{n-k}\right),$$

in which:

([31]) **L. Boltzmann**, *Prinzipe II*, **§ 27**, pp. 104 replaced the rectangular coordinates with quasi-coordinates and interpreted the additional terms that arose from the transitivity equation in rectangular coordinates. Cf., also **L. Boltzmann**, "Über die Form der Lagrangeschen Gleichungen für nichtholonome generalisierte Koordinaten," Wien Sitzungsber. **111**, II[a] (1903), pp. 1603 – *Ges. Abh. III*, pp. 682.

For the older literature, which was initially oriented towards individual problems, cf., IV 1 (**A. Voss**), no. **88**. A summary of the results of the older work can be found in **J. Quanjel**, "Les équat. gén. de la mécanique dans le cas des liaisons non holonomes," Palermo Rend. **22** (1906), pp. 293. A direct recalculation of the **Lagrange** equations of the first kind was given by **P. Burgatti**, Rom. Acc. Linc. Rend. (5) 18^2 (1909), pp. 135 and 340. Direct calculation was also employed by **J. Tzenoff**, J. de math. (8) **3** (1920), pp. 245 and Math. Ann. **91** (1924), pp. 161, without achieving anything essentially new. On that, cf., **G. Hamel**, Math. **92** (1924), pp. 33.

([32]) **P. Woronetz**, Moscow Math. Collection **22** (1901), pp. 659 (Russian). A representation of the process is also in **P. Woronetz**, "Über die Bewegung eines starren Körpers, der ohne Gleitung auf einer beliebigen Fläche rollt," Math. Ann. **70** (1911), pp. 410, esp., in Chap. II, §§ **5** – **8**.

$$\delta T^* = \frac{\partial T^*}{\partial \dot{q}_1}\delta \dot{q}_1 + \cdots + \frac{\partial T^*}{\partial \dot{q}_{n-k}}\delta \dot{q}_{n-k} + \frac{\partial T^*}{\partial q_1}\delta q_1 + \cdots + \frac{\partial T^*}{\partial q_{n-k}}\delta q_{n-k} + \frac{\partial T^*}{\partial q_{n-k+1}}\delta q_{n-k+1} + \cdots + \frac{\partial T^*}{\partial q_n}\delta q_n \ .$$

From (41), one has to set:

(41.a)
$$\left\{ \begin{array}{ll} \delta q_{n-k+1} & = b_{11}\,\delta q_1 + \cdots + b_{1,n-k}\,\delta q_{n-k}, \\ \ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \delta q_n & = b_{k1}\,\delta q_1 + \cdots + b_{k,n-k}\,\delta q_{n-k} \end{array} \right.$$

in that, and one will then get the $(n - k)$ equations of motion that determine the q_1, q_2, …, q_n as functions of time, along with the k equations (41) ([33]).

One must investigate whether the $(n - k)$ equations of motion that one gets from (43) [the $(n - k)$ equations of motion (40.a), resp.] include ones that take the form of the ordinary **Lagrange** equations:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}}\right) - \frac{\partial T}{\partial q} = Q \ ,$$

or *how many* of the equations of motion can assume that simple form for the most favorable choice of parameter, resp. To **P. Appell**, the least-possible number of equations that cannot read that way determines the *order of non-holonomity* of the system. **E. B. Schieldrop** has developed a method that allows one to decide whether the simple **Lagrange** equation will occur for a general coordinate on the basis of kinematical arguments in which he introduced a new concept, namely, the so-called "non-holonomic deviation," and with whose help one can likewise arrive at a kinematical interpretation of the additional terms ([34]).

5. The principle of least constraint and the definition of the equations of motion. – For *non-holonomic constraints* in the general *nonlinear form* (2), one cannot arrive at an Ansatz for the equations of motion from the principle of virtual displacements (**Lagrange**'s central equation, resp.). Ordinarily, one then cares to appeal to the *principle of least constraint* [cf., IV 1 (**A. Voss**), no. **39**]. **C. F. Gauss** defined the *constraint* by the expression:

(44)
$$Zw = \sum_{i=1}^{r} \frac{1}{m_i}\left[(X_i - m_i\,\ddot{x}_i)^2 + (Y_i - m_i\,\ddot{y}_i)^2 + (Z_i - m_i\,\ddot{z}_i)^2\right] \ ,$$

([33]) **P. Woronetz** is connected with: **A. Bilimowitsch**, C. R. Acad. Sci. Paris **156** (1913), pp. 381 and *ibid.*, pp. 1216.

([34]) Cf., **P. Appell**, "Sur une forme gén. des équ. de la dyn.," Mém. des scienc. math. fasc. **1**, Paris 1925), pp. 13, as well as **P. Appell**, "Sur l'ordre d'un système non holonome," C. R. Acad. Sci. Paris **179** (1924), pp. 549 and furthermore **E. B. Schieldrop**, "Sur une notion de déviation non holonome," Skand. Matematiker Kongressen **6** (Kopenhagen) Beretning (1925), pp. 281.

which can be regarded as a function of the acceleration components $\ddot{x}_i$, $\ddot{y}_i$, $\ddot{z}_i$ (i.e., the position coordinates x_i, y_i, z_i the velocity components $\dot{x}_i$, $\dot{y}_i$, $\dot{z}_i$, and time t are regarded as constant) The rectangular coordinates in that expression shall next be replaced with the general coordinates ([35]), in which the condition equations are ignored completely for the time being, i.e., the 3r rectangular coordinates x_i, y_i, z_i are replaced with 3r general parameters by means of the formulas:

$$(45) \qquad \begin{cases} x_i = \varphi_i(q_1, \ldots, q_{3r}), \\ y_i = \psi_i(q_1, \ldots, q_{3r}), \\ z_i = \chi_i(q_1, \ldots, q_{3r}). \end{cases}$$

Moreover, one can even assume (without needing to the change the argument essentially) that the transformation formulas for the introduction of general parameter include time explicitly, so they will read:

$$(46) \qquad \begin{cases} x_i = \varphi_i(q_1, \ldots, q_{3r}, t), \\ y_i = \psi_i(q_1, \ldots, q_{3r}, t), \\ z_i = \chi_i(q_1, \ldots, q_{3r}, t). \end{cases}$$

Since the direct recalculation of the expression (44) into the new coordinates would be quite tiresome, due to the appearance of the second derivatives, along with the constraint (44), **R. Lipschitz** represented the kinetic energy as:

$$(47) \qquad T = \sum_{i=1}^{r} \tfrac{1}{2} m_i (\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2) ,$$

which will go to ([36]):

$$(47.a) \qquad T = T_0 + T_1 + T_2 = \tfrac{1}{2} g_{00} + \sum_{\lambda=1}^{3r} g_{0\lambda} \dot{q}_\lambda + \tfrac{1}{2} \sum_{\rho,\sigma=1}^{3r} g_{\rho\sigma} \dot{q}_\rho \dot{q}_\sigma ,$$

$$g_{\lambda\mu} = g_{\lambda\mu}(q_1, \ldots, q_{3r}, t)$$

under the transformation (46), and based the recalculation of (44) on the argument that an expression that is covariant under the transformation (46) with the kinetic energy (47) must be the desired recalculation of the constraint (44) in the event that it goes to the expression (44) for $q_1 = x_1$, $q_2 = y_1$, …, $q_{3r} = z_r$.

([35]) **R. Lipschitz**, "Bemerkungen zu dem Prinzip des kleinsten Zwanges," J. f. Math. **82** (1877), pp. 316.

([36]) In particular, T would yield the quadratic form:

$$T = T_2 = \tfrac{1}{2} \sum_{\rho,\sigma=1}^{3r} g_{\rho\sigma}(q_1, \ldots, q_n) \dot{q}_\rho \dot{q}_\sigma$$

under the transformation (45).

Now, the expressions:

$$\left[Q_\rho - \left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\rho} \right) - \frac{\partial T}{\partial q_\rho} \right) \right]$$

will go to $(X_i - m_i \ddot{x}_i)$, etc., resp., for $q_1 = x_1, \ldots, q_{3r} = z_r$, but otherwise they will transform contragrediently to the coordinate differentials $dq_1, \ldots, dq_{3r}$ according to (24.b) and (30.a). If one then introduces the associated quantities $g^{\rho\sigma}$, which transform cogrediently to the coordinate differentials, to the coefficients $g_{\rho\sigma}$ in the quadratic form T_2 in (47.a), which transform contragrediently to the coordinate differentials ([37]), then one will get a *covariant of the kinetic energy* (47.a) in the expression:

$$\sum_{\rho,\sigma=1}^{3r} g^{\rho\sigma} \left\{ Q_\rho - \left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\rho} \right) - \frac{\partial T}{\partial q_\rho} \right) \right\} \left\{ Q_\sigma - \left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\sigma} \right) - \frac{\partial T}{\partial q_\sigma} \right) \right\} .$$

Since the $g_{\rho\sigma}$ assume the values:

$$g_{\rho\sigma} = 0 \qquad (\rho \neq \sigma),$$

$$g_{3i-2,\,3i-2} = g_{3i-1,\,3i-1} = g_{3i,\,3i} = m_i ,$$

moreover, for $q_1 = x_1, \ldots, q_{3r} = z_r$, the $g^{\rho\sigma}$ will take the values:

$$g^{\rho\sigma} = 0 \qquad (\rho \neq \sigma),$$

$$g^{3i-2,3i-2} = g^{3i-1,3i-1} = g^{3i,3i} = \frac{1}{m_i},$$

in this case, i.e., precisely the values that appear in (44). It then follows from this that the *constraint in general coordinates* will be represented by the expression ([38]):

(44.a) $$Zw = \sum_{\rho,\sigma=1}^{3r} g^{\rho\sigma} \left\{ Q_\rho - \left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\rho} \right) - \frac{\partial T}{\partial q_\rho} \right) \right\} \left\{ Q_\sigma - \left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\sigma} \right) - \frac{\partial T}{\partial q_\sigma} \right) \right\} ,$$

([37]) In particular, T will be given as the quadratic form:

$$T = T_2 = \tfrac{1}{2} \sum_{\rho,\sigma=1}^{3r} g_{\rho\sigma}(q_1, \ldots, q_n)\, \dot{q}_\rho\, \dot{q}_\sigma$$

under the transformation (45).

([38]) **R. Lipschitz**, *loc. cit.* ([35]), gave a more general expression, in that he based the arc-length element *ds*, not on the square root of a quadratic form in the coordinate differentials, but the p^{th} root of a form of order p.

which can now be regarded as a function of the second derivatives $\ddot{q}_1, \ldots, \ddot{q}_{3r}$, in the sense that time *t*, the position coordinates q_ρ, and the velocity components $\dot{q}_\rho$ are thought of as fixed ([38.a]).

The differential of the constraint, when regarded in that way, is ([39]):

$$\delta Zw = \sum_{\rho=1}^{3r} \left\{ Q_\rho - \left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\rho} \right) - \frac{\partial T}{\partial q_\rho} \right) \right\} \delta \ddot{q}_\rho \ , \tag{48}$$

which is a formula that takes the form of the variational formula for (44):

$$\delta Zw = \sum_{i=1}^{r} [(m_i \ddot{x}_i - X_i)\delta \ddot{x}_i + (m_i \ddot{y}_i - Y_i)\delta \ddot{y}_i + (m_i \ddot{z}_i - Z_i)\delta \ddot{z}_i] \tag{49}$$

([38.a]) That can also be shown by direct calculation: **A. Voss**, "Bemerkungen über die Prinzipien der Mechanik," **31** (1901), pp. 167. **Voss** emphasized that such a calculation is possible only when the number of parameters $q_1, \ldots, q_{3r}$ is *the same* as the number of rectangular coordinates. In that way, he was of the opinion that the constraint could be equal to zero for only free systems. By contrast, if one also wants to speak of the constraint for a system in a space with a general metric then one must define it by expressions of the form (44.a) [(44.b), resp.] in that case.

For the calculations in the case where the number of general parameters is less than the number of rectangular coordinates, cf., **A. Wassmuth**, "Über die Transformation des Zwanges in allgemeine Koordinaten," Wien Sitzungsber. **104**, II[a] (1895), pp. 281, as well as **M. Radakovic**, "Über die analyt. Darst. des Zwanges eines mater. Syst. in allg. Koordin.," Monatsh. Math. Phys. **7** (1896), pp. 27. In that case, it will be:

$$Zw = \sum_{\rho,\sigma=1}^{3r} g^{\rho\sigma} \left\{ Q_\rho - \left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\rho} \right) - \frac{\partial T}{\partial q_\rho} \right) \right\} \left\{ Q_\sigma - \left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\sigma} \right) - \frac{\partial T}{\partial q_\sigma} \right) \right\} + \Phi(t, q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n) \ ,$$

in which the $g^{\rho\sigma}$ refer to the quadratic form T_2 in *n* variables. The form of the function Φ was determined by **A. Wassmuth**, "Das Restglied bei der Transformation des Zwanges in allgemeine Koordinaten," Wien Sitzungsber. **110**, II[a] (1901), pp. 387. A calculation of the constraint that started from the **Lagrange** equations in general coordinates was given by **R. Leitinger**, "Über die Ableitung des Gaußschen Prinzips des kleinsten Zwanges aus den allgemeinsten Lagrangschen Gleichungen zweiter Art," Wien Sitzungsber. **116**, II[a2] (1907), pp. 1321.

([39]) That is because, from (44.a), one has:

$$\frac{\partial Zw}{\partial \ddot{q}_\nu} = \sum_{\sigma=1}^{3r} \left[\sum_{\rho=1}^{3r} g^{\rho\sigma} \left\{ Q_\rho - \left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\rho} \right) - \frac{\partial T}{\partial q_\rho} \right) \right\} \frac{\partial}{\partial \ddot{q}_\nu} \left\{ Q_\sigma - \left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\sigma} \right) - \frac{\partial T}{\partial q_\sigma} \right) \right\} \right] ,$$

or since:

$$\frac{\partial}{\partial \ddot{q}_\nu} \left\{ Q_\sigma - \left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\sigma} \right) - \frac{\partial T}{\partial q_\sigma} \right) \right\} = - \frac{\partial}{\partial \ddot{q}_\nu} \left\{ \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\sigma} \right) \right\} = - g_{\nu\sigma} \, ,$$

and because of (47.c), one has:

$$\frac{\partial Zw}{\partial \ddot{q}_\nu} = \sum_{\rho=1}^{3r} \left[\left\{ \left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\rho} \right) - \frac{\partial T}{\partial q_\rho} \right) - Q_\rho \right\} \sum_{\sigma=1}^{3r} g^{\rho\sigma} g_{\nu\sigma} \right] = \left\{ \left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\nu} \right) - \frac{\partial T}{\partial q_\nu} \right) - Q_\nu \right\} .$$

in rectangular coordinates, in particular.

From **Gauss**'s principle of least constraint, one must have:

(50) $$\delta Zw = 0$$

for all variations of the acceleration components that are compatible with the condition equations ([40]). Now, if *holonomic constraints* are present then, regardless of whether they are scleronomic or rheonomic, one can introduce the general coordinates in such a way that the associated condition equations will take the form:

(51) $$q_{n+1} = 0\ , \qquad \ldots, \qquad q_{3r} = 0\ .$$

Equations (51) imply the relations:

(51*) $$\delta\ddot{q}_{n+1} = 0\ , \qquad \ldots, \qquad \delta\ddot{q}_{3r} = 0$$

as conditions for the acceleration components, such that the variational formula (48) will lead to

$$\sum_{\nu=1}^{n}\left\{\left(\frac{d}{dt}\left(\frac{\partial T}{\partial\dot{q}_\nu}\right)-\frac{\partial T}{\partial q_\nu}\right)-Q_\nu\right\}\delta\ddot{q}_\nu = 0\ ,$$

and since the $\delta\ddot{q}_1, \ldots, \delta\ddot{q}_n$ are arbitrary infinitesimal quantities, the equations:

$$\left(\frac{d}{dt}\left(\frac{\partial T}{\partial\dot{q}_\nu}\right)-\frac{\partial T}{\partial q_\nu}\right)-Q_\nu = 0 \qquad (\nu = 1, \ldots, n)$$

will yield equations of motion after introducing (51).

If non-holonomic constraints appear (along with any holonomic ones that might possible exist) then one can imagine that they are represented by equations of the form [possibly by appealing to (51)]:

(51.a) $$f_\rho(\dot{q}_1,\ldots,\dot{q}_n,q_1,\ldots,q_n,t) = 0 \qquad (\rho = 1, \ldots, k),$$

from which, the conditions for the $\delta\ddot{q}_\nu$ will emerge:

$$\frac{\partial f_\rho}{\partial\dot{q}_1}\delta\ddot{q}_1+\cdots+\frac{\partial f_\rho}{\partial\dot{q}_n}\delta\ddot{q}_n = 0 \qquad (\rho = 1, \ldots, k)\ .$$

([40]) Those conditions can possibly be first converted into conditions on the acceleration components by differentiation.

The variational formula (50.a) then implies the equations of motion in the form:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\nu}\right) - \frac{\partial T}{\partial q_\nu} = Q_\nu + \sum_{\rho=1}^{k} \lambda_\rho \frac{\partial f_\rho}{\partial \dot{q}_\nu} \ .$$

If there are no applied forces then the principle of least constraint can be connected with the *principle of the straightest path* that **H. Hertz** exhibited [cf., IV 1 (**A. Voss**), no. **28** and no. **39**]. The expression (44.a) for the constraint will then, in fact, become:

$$Zw = \sum_{\rho,\sigma=1}^{n} g^{\rho\sigma}\left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\rho}\right) - \frac{\partial T}{\partial q_\rho}\right)\left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\sigma}\right) - \frac{\partial T}{\partial q_\sigma}\right),$$

and when the kinetic energy possesses the form:

$$T = \tfrac{1}{2}\sum_{\rho,\sigma=1}^{n} g_{\rho\sigma}\,\dot{q}_\rho\,\dot{q}_\sigma \qquad\qquad [g_{\rho\sigma} = g_{\rho\sigma}(q_1, \ldots, q_n)],$$

such that the manifold of $q_1, \ldots, q_n$ can be referred to as a **Riemannian** space (cf., no. **6**), it will be precisely the expression for the *geodetic curvature of the trajectory* ([40.a]). The trajectories are then given as the curves whose geodetic curvature has the smallest value that is compatible with the constraints ([40.b]).

As **J. W. Gibbs** ([41]) first emphasized, the variational formula (49) can also be written in the form:

(52) $$\delta\sum_{i=1}^{r} \tfrac{1}{2} m_i(\ddot{x}_i^2 + \ddot{y}_i^2 + \ddot{z}_i^2) - \sum_{i=1}^{r}(X_i\,\delta\ddot{x}_i + Y_i\,\delta\ddot{y}_i + Z_i\,\delta\ddot{z}_i) = 0 \ .$$

([40.a]) With the notation that will be introduced in no. **6**, the expression can be immediately converted into:

$$Zw = \sum_{\rho,\sigma=1}^{n} g_{\rho\sigma}\,L^\rho L^\sigma \ ,$$

which is ordinarily given as a representation of the geodetic curvature. [Cf., ([48.a])].

([40.b]) Let it be mentioned that **Ph. E. B. Jourdain**, "Note on an analogue of Gauss's principle of least constraint," Quart. J. of math. **40** (1909), pp. 153, pointed out that it would suffice to employ the variational formula:

$$\sum_{i=1}^{3r}[(X_i - m_i\,\ddot{x}_i)\,\delta\dot{x}_i + (Y_i - m_i\,\ddot{y}_i)\,\delta\dot{y}_i + (Z_i - m_i\,\ddot{z}_i)\,\delta\dot{z}_i] = 0$$

as the Ansatz the equation of motion for general non-holonomic constraints, instead of the variational formula (49) for the principle of least constraint.

([41]) **J. W. Gibbs**, "On the fundamental formulae of dynamics," Amer. J. of math. **2** (1879), pp. 49 = *Scient. Papers II*, pp. 1.

If one now introduces any new parameters $q_1, \ldots, q_n$ such that the virtual displacements δx_i, δy_i, δz_i can be expressed by linear homogeneous functions of the independent variations $\delta q_1, \ldots, \delta q_n$ then (52) will read:

$$\delta \sum_{i=1}^{r} \tfrac{1}{2} m_i (\ddot{x}_i^2 + \ddot{y}_i^2 + \ddot{z}_i^2) - \sum_{\rho=1}^{r} Q_\rho \, \delta \ddot{q}_\rho = 0 , \tag{52.a}$$

or since the Q_ρ are free of $\ddot{q}_1, \ldots, \ddot{q}_n$:

$$\delta \left[\sum_{i=1}^{r} \tfrac{1}{2} m_i (\ddot{x}_i^2 + \ddot{y}_i^2 + \ddot{z}_i^2) - \sum_{\rho=1}^{r} Q_\rho \, \ddot{q}_\rho \right] = 0 . \tag{52.b}$$

Here, the parameters $q_1, \ldots, q_n$ are also incorporated in the first term in the brackets, which is the so-called *Appell acceleration energy* ([42]):

$$S = \sum_{i=1}^{r} \tfrac{1}{2} m_i (\ddot{x}_i^2 + \ddot{y}_i^2 + \ddot{z}_i^2) \tag{52.c}$$

[cf., IV 1 (**A. Voss**), no. **38**], from which one will get the equations of motion in the form:

$$\frac{\partial S}{\partial \ddot{q}_\rho} - Q_\rho = 0 . \tag{52.d}$$

From that point onward, when *non-holonomic constraints* appear ([43]), one will arrive at the least number of equations of motion, just as when one starts from the general Ansatz for the equations of motion in quasi-coordinates.

([42]) The term *acceleration energy* was proposed by **A. de Saint-Germain**, C. R. Acad. Sci. Paris **130** (1900), pp. 1174, who showed that an analogue of **König**'s theorem on kinetic energy is true for it.

([43]) That is the path that **P. Appell** went down as an Ansatz to the equations of motion in the case of non-holonomic constraints. **P. Appell**, "Sur une forme générale des équations de la dynamique et sur le principe de Gauß," J. f. Math. **122** (1900), pp. 205. There is a survey presentation in **P. Appell**, "Les mouvements de roulement en dynamique," Paris 1899 (Scientia 4) and **P. Appell**, "Sur une forme générale des équations de la dynamique," Paris 1925 (Mémorial des scien. math., fasc. 1).

If one has represented the x_i, y_i, z_i as functions of $q_1, q_2, \ldots, q_n, t$ in the form (5.a) or (5.b) then one defines the expressions dx_i, dy_i, dz_i and employs any **Pfaff** equations (9) that might be present in order to eliminate the $dq_{n-k+1}, \ldots, dq_n$ such that:

$$\begin{aligned} \dot{x}_i &= a_{1i}\dot{q}_1 + a_{21i}\dot{q}_2 + \cdots + a_{n-k,i}\dot{q}_{n-k} + a_i , \\ \dot{y}_i &= b_{1i}\dot{q}_1 + b_{21i}\dot{q}_2 + \cdots + b_{n-k,i}\dot{q}_{n-k} + b_i , \\ \dot{z}_i &= c_{1i}\dot{q}_1 + c_{21i}\dot{q}_2 + \cdots + c_{n-k,i}\dot{q}_{n-k} + c_i . \end{aligned}$$

$\ddot{x}_i$, $\ddot{y}_i$, $\ddot{z}_i$ are then obtained by repeated differentiation and are then substituted in (52.b). Naturally, $q_{n-k+1}, \ldots, q_n$, as well as t, also appear in the coefficients. Further developments, with applications are in **P. Appell**, "Développements sur une forme nouvelle des équations de la dynamique," J. de math. (5) **6** (1900), pp. 5 and **P. Appell**, "Remarques

6. The Lagrange equations and their solutions. – In the case of *holonomic* constraints, the ***Lagrange*** *equations* read:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\rho}\right) - \frac{\partial T}{\partial q_\rho} = Q_\rho \qquad (\rho = 1, 2, \ldots, n). \tag{53}$$

These are second-order differential equations that are linear in the second derivatives and can be easily solved for them. The general solution of such a system will include $2n$ arbitrary constants $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$:

$$q_1 = q_1\,(t, \alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n)\,, \ldots, q_n = q_n\,(t, \alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n)\,. \tag{54}$$

One separates the individual solutions from that when one fixes the constants, perhaps by prescribing the position coordinates and the velocity components a certain time $t = t_0$:

$$t = t_0 \quad \begin{cases} q_1 = q_1^{(0)}, & \ldots, \quad q_n = q_n^{(0)}, \\ \dot{q}_1 = \dot{q}_1^{(0)}, & \ldots, \quad \dot{q}_n = \dot{q}_n^{(0)}, \end{cases} \tag{54.a}$$

and if one interprets such an individual solution as either a *space-time curve* in the $(n + 1)$-dimensional space-time manifold of the $(q_1, \ldots, q_n, t)$ or when one regards time as a parameter, as a *trajectory* in the n-dimensional spatial manifold $(q_1, \ldots, q_n)$. The totality of space-time lines represents a $2n$-parameter family of curves, and the same thing is also generally true for the trajectories since every space-time line belongs to a trajectory and conversely. A bundle of space-time lines goes through every prescribed space-time point since the individual space-time line is determined by its "direction" $dq_1 : \ldots : dq_n : dt$, i.e., by the n components of the initial velocity. The bundle defines an n-parameter family of curves. One and only one space-time line of the family of curves goes through every space-time point that is sufficiently close to the center of the bundle. As one says, the family defines a *field* in a sufficiently close proximity to the center of the bundle and associates each point of the field with the corresponding velocity components $\dot{q}_1, \ldots, \dot{q}_n$. Things are different for the trajectories. A trajectory through a spatial point is still not determined when one gives the direction $dq_1 : \ldots : dq_n$, but only when one is given the velocity components $\dot{q}_1, \ldots, \dot{q}_n$. Therefore ∞^1 trajectories go through a spatial point in a prescribed direction. Moreover, that is the reason why there are ∞^n trajectories.

---

d'ordre analytique sur une nouvelle forme des équations de la dynamique," J. de math. (5) **7** (1901), pp. 5. One also obtains the Ansatz for the equations of motion for constraints of the general form (2) in that way, **P. Appell**, "Sur les liasons exprimées par des relations non linéaires entre les vitesses," C. R. Acad. Sci. Paris **152** (1911), pp. 1197. According to **Appell**, the second terms in the bracket in (52.b) will still remain when a further condition is added. **P. Appell**, "The principe du minimum de l'énergie d'accélérations et la substitution des liaisons aux forces," C. R. Acad. Sci. Paris **159** (1914), pp. 989. The connection between the principle of least constraint and **Appell**'s argument was also treated by **H. Brell**, "Über eine neue Form des Gaußschen Prinzips des kleinsten Zwanges," Wien Sitzungsber. **122**, II^{a2} (1913), pp. 1531. The inverse problem of calculating the acceleration energy from the equations of motion was treated by **E. Cotton**, "A propos des équations de M. Appel," Nouv. Ann. de math. (4) **7** (1907), pp. 539. **Cotton** also treated the kinetostatic problem of calculating the reactions of non-holonomic constraints there.

The number of trajectories will then amount to ∞^{2n-1} only when the force components $Q_1, \ldots, Q_n$, which could generally be functions of $t, q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n$ up to now, *do not* depend upon the velocity components and time, but are functions of only the position coordinates and when, at the same time, the kinetic energy does not include time t explicitly. Namely, a space-time curve will always go to another space-time curve of the system under the one-parameter group of "parallel translations" in the t-direction. All space-time curves of the one-parameter family that is generated by the group are the "parallel curves" in a cylindrical M_2 with generators that lie parallel to the t-direction, but they belong to one and the same trajectory, which is obtained from a section of the cylindrical M_2 by the M_n of $(q_1, \ldots, q_n)$. The motion therefore possesses only ∞^{2n-1} trajectories, each of which can be specified in ∞^1 ways, and indeed one will get all other trajectories from the course of motion along one trajectory when one increases the times at which the individual spatial points are attained by that specification by the same constant (cf., also no. **10**).

Far-reaching investigations were then carried out in that case. Let the *kinetic energy* be a quadratic form in the velocity components whose coefficients *do not contained time explicitly:*

$$(55) \qquad T = \tfrac{1}{2}\sum_{\rho,\sigma} g_{\rho\sigma}\,\dot{q}_\rho\,\dot{q}_\sigma\,, \qquad g_{\rho\sigma} = g_{\rho\sigma}(q_1, \ldots, q_n)\,.$$

Now, using the kinetic energy, one can associate an *arc-length element* ([44]):

$$(55.a) \qquad ds^2 = 2T\,dt^2 = \sum_{\rho,\sigma=1}^{n} g_{\rho\sigma}\,dq_\rho\,dq_\sigma\,,$$

such that this M_n will be regarded as an n-dimensional *Riemannian space* [cf., III D 11 (**L. Berwald**), no. **17**]. The distinction between contravariant and covariant vectors will disappear in a **Riemannian** space. There are only vectors, *per se*. However, there are two types of components for every vector, which are distinguished by referring to them as the contravariant and covariant components ([45]). Thus, along with the n contravariant components $\dot{q}_1, \ldots, \dot{q}_n$ that were employed up to now, the *velocity vector* also has n covariant components that are calculated from the contravariant components by means of:

([44]) The convention generalizes the relation:

$$T = \tfrac{1}{2}\dot{s}^2 = \tfrac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)\,,$$

which couples the kinetic energy of an isolated mass-point of mass 1 with the arc-length element $ds^2 = dx^2 + dy^2 + dz^2$ in three-dimensional Euclidian space.

([45]) In the notation of **Ricci**'s absolute differential calculus, one cares to distinguish the covariant and contravariant components of a vector by lower (upper, resp.) indices. It should be observed that the contravariant components of the velocity vectors are therefore "falsely" denoted as $\dot{q}_1, \ldots, \dot{q}_n$.

(56) $$\sum_{\sigma=1}^{n} g_{\rho\sigma}\,\dot{q}_\sigma = \frac{\partial T}{\partial \dot{q}_\rho} = p_\rho\,.$$

The covariant components of the velocity vector are then precisely the *impulse components* (no. **4**). In that sense, the $Q_1, \ldots, Q_n$ are the covariant components of the *applied force vector*, moreover, and one will get their contravariant components by means of:

(56.a) $$\sum_{\lambda=1}^{n} g^{\rho\lambda}\,Q_\lambda = Q^\rho\,,$$

in which the $g^{\rho\lambda}$ are explained correspondingly ([37]). If one defines the following differential equations from the Q^ρ:

$$\frac{dq_1}{Q^1} = \frac{dq_2}{Q^2} = \ldots = \frac{dq_n}{Q^n}$$

then they will determine an $(n-1)$-parameter family of curves in the R_n of $q_1, \ldots, q_n$, namely, the *lines of force* of the field of the applied forces that point in the direction of the applied force at the individual points of the R_n.

The left-hand sides of the **Lagrange** equations of motion (53) ([46]):

(57) $$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\lambda}\right) - \frac{\partial T}{\partial q_\lambda} = \sum_{\rho=1}^{n} g_{\lambda\rho}\,\ddot{q}_\rho + \sum_{\sigma,\tau=1}^{n} \begin{bmatrix} \sigma\,\tau \\ \lambda \end{bmatrix} \dot{q}_\sigma\,\dot{q}_\tau$$

are referred to as the covariant components of the *acceleration vector*. If one goes from the covariant components of both sides of the **Lagrange** equations, which now read:

(57.a) $$\sum_{\rho=1}^{n} g_{\lambda\rho}\,\ddot{q}_\rho + \sum_{\sigma,\tau=1}^{n} \begin{bmatrix} \sigma\,\tau \\ \lambda \end{bmatrix} \dot{q}_\sigma\,\dot{q}_\tau = Q_\lambda \qquad (\lambda = 1, 2, \ldots, n)$$

to the covariant ones then one will ultimately obtain the *equations of motion* in the form ([47]):

([46]) The *square Christoffel three-index symbols* [cf., III D 10 (**R. Weitzenböck**), Part 2, no. **19**] were introduced here:

$$\begin{bmatrix} \sigma\,\tau \\ \lambda \end{bmatrix} = \frac{1}{2}\left(\frac{\partial g_{\sigma\lambda}}{\partial q_\tau} + \frac{\partial g_{\tau\lambda}}{\partial q_\sigma} - \frac{\partial g_{\sigma\tau}}{\partial q_\lambda}\right).$$

([47]) The *curly Christoffel three-index symbols* that are employed here are defined by:

$$\begin{Bmatrix} \sigma\,\tau \\ \lambda \end{Bmatrix} = \sum_{\lambda=1}^{n} g^{\rho\lambda} \begin{bmatrix} \sigma\,\tau \\ \lambda \end{bmatrix}$$

[cf., *loc. cit.* ([46])].

(57.b) $$\ddot{q}_\rho + \sum_{\sigma,\tau=1}^{n} \left\{ \begin{matrix} \sigma\,\tau \\ \rho \end{matrix} \right\} \dot{q}_\sigma \, \dot{q}_\tau \;=\; Q^\rho \qquad\qquad (\rho = 1, 2, \ldots, n),$$

when one likewise solves for the second derivatives $\ddot{q}_\rho$. The contravariant components of the acceleration vector that appears on the left-hand side are the so-called *covariant derivatives* [cf., III D 10 (**R. Weitzenböck**), Part 2, no. **19**] of the contravariant components $\dot{q}_1$, …, $\dot{q}_n$ of the velocity vector, in the sense of **Ricci**'s absolute differential calculus.

In order to integrate those equations (57.b), it is obvious that one must first determine the *trajectories* in the **Riemannian** space of the q_1, …, q_n, and determine the course of the motion along the individual trajectories afterwards. However, instead of establishing the individual members among those ∞^{2n-1} trajectories by way of the *initial position* $q_1^{(0)}$, …, $q_n^{(0)}$ and the components of the *initial velocity* $\dot{q}_1^{(0)}$, …, $\dot{q}_n^{(0)}$, one would rather appeal to their characterization by geometric quantities. If one now introduces the components of the initial direction ([48]) $\left(\frac{dq_1}{ds}\right)_0$, …, $\left(\frac{dq_n}{ds}\right)_0$, in place of the initial velocity, and the magnitude of the initial velocity $v_0 = (\dot{s})_0$ then the trajectory will appear to be determined for a prescribed starting point and a prescribed direction by magnitude of the initial velocity v_0. Now, when v_0 changes, the (geodetic) curvature ([48.a]) of the trajectory at the starting point will change, such that the individual trajectories that go through

([48]) For the direction components dq_1 / ds, …, dq_n / ds, one has:

$$\sum_{\lambda,\mu=1}^{n} g_{\lambda\mu} \frac{dq_\lambda}{ds} \frac{dq_\mu}{ds} = 1 .$$

Furthermore, one has:

$$\dot{s}^2 = \sum_{\lambda,\mu=1}^{n} g_{\lambda\mu} \, \dot{q}_\lambda \, \dot{q}_\mu ,$$

from which it follows:

$$\dot{s}\ddot{s} = \sum_{\mu=1}^{n} (g_{1\mu}\,\dot{q}_1 + \cdots + g_{n\mu}\,\dot{q}_\mu) \left(\ddot{q}_\mu + \sum_{\sigma,\tau=1}^{n} \left\{ \begin{matrix} \sigma\,\tau \\ \mu \end{matrix} \right\} \dot{q}_\sigma \, \dot{q}_\tau \right) .$$

([48.a]) The covariant derivatives of the direction components:

$$L^\rho = \frac{d^2 q_\rho}{ds^2} + \sum_{\sigma,\tau=1}^{n} \left\{ \begin{matrix} \sigma\,\rho \\ \rho \end{matrix} \right\} \frac{dq_\sigma}{ds} \frac{dq_\tau}{ds}$$

are the contravariant components of a vector that possesses the direction of the principal normal to the curve $q_1 = q_1(s)$, …, $q_n = q_n(s)$. If K_g means the *geodetic curvature* of the curve then one will have:

$$K_g^2 = \sum_{\rho,\sigma=1}^{n} g_{\rho\sigma} \, L^\rho L^\sigma .$$

[Cf., III D 11 (**L. Berwald**), no. **18**].

a given point in a given direction can also be characterized by giving the geodetic curvature at the starting point.

In order to express that in formulas, one writes the equations of motion (57.b) in the form:

$$\frac{d^2 q_\rho}{ds^2}\dot{s}^2 + \frac{dq_\rho}{ds}\ddot{s} + \dot{s}^2 \sum_{\sigma,\tau} \left\{ \begin{matrix} \sigma\ \tau \\ \rho \end{matrix} \right\} \frac{dq_\sigma}{ds}\frac{dq_\tau}{ds} = Q^\rho$$

or

$$L^\rho = \frac{Q^\rho - \frac{dq_\rho}{ds}\ddot{s}}{\dot{s}^2} .$$

On the other hand ([48]), when the contravariant force vector is introduced using (57.b), $\ddot{s}$ will take the form ([49]):

$$\ddot{s} = \sum_{\mu=1}^{n} \left(g_{1\mu}\frac{dq_1}{ds} + \cdots + g_{n\mu}\frac{dq_n}{ds} \right) Q^\mu$$

or ([49.a]):

$$\ddot{s} = Q_1\frac{dq_1}{ds} + \cdots + Q_n\frac{dq_n}{ds} = W . \tag{58}$$

Therefore, the equations of motion ultimately read ([50]):

$$L^\rho = \frac{Q^\rho - \frac{dq_\rho}{ds}W}{\dot{s}^2}, \tag{59}$$

which implies the connection between the geodetic curvature and the magnitude of the velocity in the form of the relation ([51]):

([49]) The expressions in the parentheses are the *n* covariant components of the direction vector whose contravariant components are dq_1 / ds, …, dq_n / ds.

([49.a]) W is the projection of the force vector Q_1, …, Q_n onto the tangent to the trajectory. Equation (58) is then nothing but the theorem of the tangential acceleration [IV 6 (**P. Stäckel**), no. **5**].

([50]) On this subject, cf., **J. Lipka**, "On the geometry of motion in a curved *n*-space," J. of math. and phys. **1** (1922), pp. 21. Similar arguments are also given by **L. Berwald** and **Ph. Frank**, "Über eine kovariante Gestalt der Differentialgleichungen der Bahnkurven allgemeiner mechanischer Systeme," Math. Zeit. **21** (1924), pp. 154. Cf., moreover, **J. L. Synge**, "On the geometry of dynamics," Trans. London Phil. Soc. (A) **226** (1927), pp. 31.

Furthermore, **C. G. J. Jacobi** had already appealed to a geometric interpretation in his formulation of the principle of least action (cf., no. **10**). Moreover, **H. Hertz** made use of an interpretation is higher-dimensional spaces in his *Prinzipien*.

([51]) The sum in parentheses is the length of the force vector, and equation (60) is the generalization of the theorem of the normal acceleration [IV 6 (**P. Stäckel**), no. **5**]. That is because since the vector L^ρ / K_g is the unit vector in the direction of the principal normals of the trajectory [cf., III D 1 (**L. Berwald**), no. **18**]:

(60) $$K_g^2 = \frac{1}{\dot{s}^4}\left(\sum_{\rho,\sigma=1}^{n} g_{\rho\sigma} Q^\rho Q^\sigma - W^2 \right).$$

One sees from this that for those mechanical problems, the curves $K_g = 0$, i.e., the *geodetic lines* [cf., III D 11 (**L. Berwald**), no. **18**] of the **Riemannian** space:

$$\frac{d^2 q_\rho}{ds^2} + \sum_{\sigma,\tau=1}^{n} \left\{ \begin{matrix} \sigma\ \rho \\ \rho \end{matrix} \right\} \frac{dq_\sigma}{ds}\frac{dq_\tau}{ds} = 0 \qquad (\rho = 1, \ldots, n)$$

are included in the trajectories, and indeed, they belong to the ones with $\dot{s} = \infty$, i.e., they will be assigned infinitely-large velocities ([52]).

Naturally, that is true only as long as not all Q^ρ vanish. If all Q^ρ are equal to zero, i.e., if the force vector is identically zero then, from (60), the geodetic curvature of the trajectories will be equal to zero, so the trajectories will become *geodetic lines* in **Riemannian** space. Therefore, the mechanical problem *without* applied forces might also be briefly referred to as the *geodetic problem* that belongs to the kinetic energy ([52.a]). Since the magnitude of the acceleration $\ddot{s}$ on a

---

(61) $$V = \frac{1}{K_g}\sum_{\rho=1}^{n} Q_\rho L^\rho$$

will be the projection of the applied force vector onto the principal normal of the trajectory. On the other hand, if one multiplies equations (59) by Q_ρ and sums over ρ then it will follow that:

$$\frac{1}{K_g}\sum_{\rho=1}^{n} Q_\rho L^\rho = \frac{\sum_{\rho=1}^{n} Q_\rho Q^\rho - W^2}{\dot{s}^2} = \frac{\sum_{\sigma,\tau=1}^{n} g_{\sigma\tau} Q^\sigma Q^\tau - W^2}{\dot{s}^2} = \dot{s}^2 K_g^2$$

such that (61) will go to:

(61.a) $$V = K_g \dot{s}^2 = \frac{\dot{s}^2}{\rho_g},$$

in which ρ_g is understood to mean the radius of geodetic curvature of the trajectory.

([52]) More precisely: The geodetic line through a point appears to be the limiting curve that the trajectories will approach when one lets the magnitude of the initial velocity go to ∞. On that subject, cf., **P. Painlevé**, "Sur les mouvements et les trajectoires réels des systèmes," Bull. Soc. math. Fr. **22** (1894), pp. 136, as well as **J. Andrade**, "Sur une propriété mécanique des lignes géodésiques," Bull. Soc. math. Fr. **22** (1894), pp. 186. Any mechanical problem will then have a family of ∞^{2n-2} trajectories in common with the geodetic problem that belongs to the same expression for the kinetic energy T. One can ask whether it might be correspondingly possible for two mechanical problems that have the same expression for kinetic energy but different (non-zero) force vectors to have a family of ∞^{2n-2} trajectories in common. According to **P. Painlevé**, "Sur la transformation des équations de la dynamique," J. de math. (4) **10** (1894), pp. 5, as long as $n > 2$, that would be impossible. For $n = 2$, it can happen, as **Painlevé** showed in an example (cf., pp. 32).

([52.a]) Furthermore, from (59) the geodetic lines will also prove to be trajectories when one has:

$$Q^\rho = \ddot{s}\frac{dq_\rho}{ds} \qquad (\rho = 1, \ldots, n).$$

geodetic line is zero, from (58), the geodetic lines will be assigned a velocity with constant magnitude. However, since the numerical value of the magnitude of the velocity remains undetermined, the geodetic line can be traversed with any arbitrary velocity, i.e., one will always get the same trajectory once the starting point and starting direction have been prescribed, independently of the initial velocity. In that case, one therefore has only ∞^{2n-2} trajectories that are each associated with a double infinitude of ways of traversing them. In geometric language: Each geodetic line belongs to ∞^2 space-time curves. Namely, if, for the sake of geometric interpretation, one lays parallels to the *t*-axis through the individual points of the geodetic line in the space-time manifold of the (q_1, …, q_n, t), and in so doing generates a cylindrical M_2 then a pencil of ∞^1 space-time lines will go through every space-time point of that M_2 that correspond to the possible values that the magnitude of the velocity might have. However, since one can displace any space-time curve parallel to itself on the cylindrical M_2, one will get, in total, ∞^2 space-time curves on the cylindrical M_2, each of which represents a possible motion along the geodetic line.

In the case of the general *forced motion*, after eliminating $\dot{s}$ from (59) by means of (61.a), that will yield the following *n* second-order differential equations as the differential equations for the trajectories ([53]):

$$L^{\rho} = \left(Q^{\rho} - \frac{dq_{\rho}}{ds} W \right) \frac{K_g}{V} \ ,$$

to which yet another differential equation will be added that determined K_g as a function of *s*. One gets it from equation (61), which will that next imply that:

$$2\ddot{s} = \frac{d}{ds}\left(\frac{V}{K_g} \right)$$

when one appeals to (58) in the form:

However, that is possible only when the force components Q_1, …, Q_n depend upon not only the position coordinates, but also the velocity components. **P. Painlevé**, "Sur les mouvements et les trajectoires réels des systèmes," Bull. Soc. math. Fr. **22** (1893), pp. 136.

([53]) When extracting the square root in (60), the sign is chosen in such a way that:

(60.a) $$\dot{s}^2 = \frac{V}{K_g}$$

will be positive. That is because a real motion naturally belongs to a positive value of $\dot{s}^2$. However, one should observe that the equations of the trajectory will remain unchanged when one changes the sign of the force components, i.e., inverts the direction of the force. In that way, $\dot{s}^2$ will take the negative sign, such that no real motion can belong to it. Following **P. Painlevé**, one nonetheless speaks of the *true* and *conjugate motions* in those cases. The trajectories of motion with the forces (Q_1, …, Q_n) and the motion with the forces ($-Q_1$, …, $-Q_n$) are identical then. If one of the trajectories for the one force system is a true trajectory ($\dot{s}^2 > 0$) then it will be a conjugate one for the other problem ($\dot{s}^2 < 0$). There are also trajectories with one part of their arc-length that belongs to a true motion and another that belongs to a conjugate one (cf., *infra*).

$$2W K_g^2 = K_g \frac{dV}{ds} - V \frac{dK_g}{ds}.$$

The equations of the trajectory will then be:

$$(62) \qquad \begin{cases} \dfrac{d^2 q_\rho}{ds^2} = -\sum\limits_{\sigma,\tau=1}^{n} \left\{ \begin{matrix} \sigma\ \tau \\ \rho \end{matrix} \right\} \dfrac{dq_\sigma}{ds} \dfrac{dq_\tau}{ds} + \left(Q^\rho - \dfrac{dq_\rho}{ds} W \right) \dfrac{K_g}{V} \quad (\rho = 1, \ldots, n), \\ \dfrac{dK_g}{ds} = \dfrac{K_g}{V} \left(\dfrac{dV}{ds} - 2W K_g \right), \end{cases}$$

in which the right-hand side of the last equation must be represents as a function of the $q_1, \ldots, q_n$, $\frac{dq_1}{ds}, \ldots, \frac{dq_n}{ds}$ ([54]).

If one has found the trajectory $q_1 = q_1(s), \ldots, q_n = q_n(s)$ then one can succeed in determining the time t when an individual point is reached by a quadrature, namely, by ([54.a]):

$$(63) \qquad dt = \pm \sqrt{\frac{K_g}{V}}\, ds, \qquad t - t_0 = \pm \int \sqrt{\frac{K_g}{V}}\, ds,$$

in which the integral is taken along the trajectory in question. Equation (63) will break down when $V \equiv 0$, i.e., when the force vector $(Q_1, \ldots, Q_n)$ continually falls along the tangent direction to the trajectory. In that case, since one cannot have $\dot{s} = 0$ identically, one must necessarily have $K_g = 0$, i.e., the curve must be a geodetic line, such that this can happen only when the *lines of force are likewise geodetic lines* of the arc-length. If such *remarkable trajectories* [*trajectoires remarquables* according to **P. Painlevé** ([55])] exist then ∞^2 motions will be possible along them, because one would then have to appeal to (59) and obtain ([56]):

([54]) In the literature, as a rule, one does not choose the arc-length to be the independent variable, but one of the variables, e.g., q_1 . Cf., **P. Painlevé**, "Sur la transformation des équations de la dynamique," J. de math. (4) **10** (1894), pp. 5. The differential equations of the trajectories are found on pp. 24. When the energy integral exists, one can eliminate the time t with their help. **P. Stäckel** did this in "Über dyn. Probl., deren Differentialgl. eine infinit. Transform. gestatten," Leipzig Ber. **45** (1893), pp. 331, and he also used q_1 as an independent variable.

([54.a]) The double sign means that every trajectory can be traversed in the sense of increasing arc-length, as well as decreasing. The transition from the one sense of traversal to the opposite one comes about when one replaces t with $(-t)$. One might also refer to such mechanical problems as *reversible* then.

([55]) **P. Painlevé**, "Sur les mouvements et les trajectoires réels des systèmes," Bull. Soc. math. Fr. **22** (1894), pp. 136, esp. pp. 145.

([56]) Moreover, one can also try to interpret the trajectories of the motion under an applied force as the geodetic lines of a **Riemannian** space. **L. P. Eisenhart** succeeded in doing that while restricting himself to the case in which the applied force arises from a potential in "Dynamical trajectories and geodesics," Ann. Math. (2) **30** (1929), pp. 591 (cf., esp., pp. 603), by introducing an $(n + 1)$-dimensional **Riemannian** space.

(63.a)
$$\begin{cases} \frac{1}{2}\dot{s}^2 = \int W\,ds + K = f(s) + K, \\ t - t_0 = \pm\int \dfrac{ds}{\sqrt{2(f(s)+K)}}. \end{cases}$$

If *time t appears explicitly* in the kinetic energy (and possibly in the components of the applied force) ([57]) then one can start the integration with the equations of motion (53) directly. Since the kinetic energy is not a quadratic form now, but a quadratic function:

(64)
$$T = T_0 + T_1 + T_2 = \tfrac{1}{2} g_{00} + \sum_{\rho=1}^{n} g_{0\rho}\,\dot{q}_\rho + \tfrac{1}{2}\sum_{\sigma,\rho=1}^{n} g_{\sigma\rho}\,\dot{q}_\sigma\,\dot{q}_\rho\,,$$

in which g_{00}, $g_{0\rho}$, $g_{\sigma\rho}$ are functions of $q_1, \ldots, q_n$, and t, the **Lagrange** equations of motion (53) will take on the form ([58]):

(64.a)
$$\sum_{\rho=1}^{n} g_{\rho\lambda}\,\ddot{q}_\rho + \sum_{\sigma,\rho=1}^{n} \left[\begin{matrix}\sigma\,\rho\\ \lambda\end{matrix}\right]\dot{q}_\sigma\,\dot{q}_\rho + 2\sum_{\rho=1}^{n}\left[\begin{matrix}0\,\rho\\ \lambda\end{matrix}\right]\dot{q}_\rho + \left[\begin{matrix}0\,0\\ \lambda\end{matrix}\right] = Q_\lambda \qquad (\lambda = 1, \ldots, n),$$

which is analogous to (57.a), and from that, one will further get the following form ([58.a]):

(64.b)
$$\ddot{q}_\lambda + \sum_{\sigma,\rho=1}^{n} \left\{\begin{matrix}\sigma\,\rho\\ \lambda\end{matrix}\right\}\dot{q}_\sigma\,\dot{q}_\rho + 2\sum_{\rho=1}^{n}\left\{\begin{matrix}0\,\rho\\ \lambda\end{matrix}\right\}\dot{q}_\rho + \left\{\begin{matrix}0\,0\\ \lambda\end{matrix}\right\} = Q^\lambda\,.$$

([57]) Since t cannot be switched with $(-t)$ here, as before, one might refer to such problems of motion as *irreversible*. Cf., **J. Lipka**, "On irreversible dynamical systems," J. math. phys, **2** (1923), pp. 73.

([58]) These were probably first given by **T. Levi-Civita**, "Sugli integrali algebrici delle equaz. dinam.," Turin Atti **31** (1895), pp. 816. In it, one has:

$$\left[\begin{matrix}0\,\rho\\ \lambda\end{matrix}\right] = \frac{1}{2}\left(\frac{\partial g_{0\lambda}}{\partial q_\rho} + \frac{\partial g_{\rho\lambda}}{\partial t} - \frac{\partial g_{0\rho}}{\partial q_\lambda}\right),$$

$$\left[\begin{matrix}0\,0\\ \lambda\end{matrix}\right] = \frac{\partial g_{0\lambda}}{\partial t} - \frac{1}{2}\frac{\partial g_{00}}{\partial q_\lambda}\,.$$

([58.a]) The curly three-index symbols are defined in the same way as in ([47]) with the quantities $g^{\rho\sigma}$, which one gets from the quadratic form T_2 in the same way that one gets them from the quadratic form T itself in the case of the **Riemannian** R_n. In particular:

$$\left\{\begin{matrix}0\,\rho\\ \lambda\end{matrix}\right\} = \sum_{\sigma=1}^{n} g^{\sigma\lambda}\left[\begin{matrix}0\,\rho\\ \lambda\end{matrix}\right], \qquad \left\{\begin{matrix}0\,0\\ \lambda\end{matrix}\right\} = \sum_{\sigma=1}^{n} g^{\sigma\lambda}\left[\begin{matrix}0\,0\\ \sigma\end{matrix}\right],$$

and furthermore:

$$Q^\lambda = \sum_{\sigma=1}^{n} g^{\sigma\lambda} Q_\sigma\,,$$

as before.

Of course, the Q^λ here are not referred to as the contravariant components of the force vector, since the R_{n+1} of the $q_1, \ldots, q_n, t$ is not a **Riemannian** space ([58.b]). Similarly, the impulse components, which are defined by:

$$p_\rho = \sum_{\sigma=1}^{n} g_{\rho\sigma}\dot{q}_\sigma + g_{0\rho} \tag{64.c}$$

here ([58.c]), will no longer play the role of covariant components of the velocity vectors.

In the case of *non-holonomic constraints*, one first considers the *scleronomic* case ([59]). The M_n of the $q_1, \ldots, q_n$ is a **Riemannian** space then whose arc-length element is imposed upon it by the kinetic energy T corresponding to (55.a). The motion in that space restricted by the k non-integrable **Pfaffian** equations (9.a), which can now be written in the form ([59.a]):

$$a_{(\mu)1}\, dq_1 + \ldots + a_{(\mu)n}\, dq_n = 0\,, \qquad a_{(\mu)\rho} = a_{(\mu)\rho}\,(q_1, \ldots, q_n) \qquad (\mu = 1, \ldots, k). \tag{65}$$

The k vectors that are determined in that way, along with the contravariant components $(a^1_{(\mu)},, \ldots, a^n_{(\mu)})$, span a k-dimensional element at each point of the **Riemannian** R_n, and the **Pfaff** equations (65) say that the velocity vector (the direction vector, resp.) to the trajectory belongs to the $(n-k)$-dimensional element that is perpendicular to it. The equations of motion that were given in covariant form on pp. 23 can be written in the contravariant form:

$$\ddot{q}_\tau + \sum_{\rho,\sigma=1}^{n} \left\{ \begin{matrix} \rho\,\sigma \\ \tau \end{matrix} \right\} \dot{q}_\rho\, \dot{q}_\sigma = Q^\tau + \sum_{\mu=1}^{k} \lambda_\mu\, a^\tau_{(\mu)}\,, \tag{65.a}$$

([58.b]) According to **L. P. Eisenhart**, *loc. cit.* ([56]), when a potential exists for the applied force, the space-time lines can be interpreted as geodetic lines in an $(n+2)$-dimensional **Riemannian** space.

([58.c]) The solution of those relations for the $\dot{q}_\rho$ reads:

$$\dot{q}_\rho = \sum_{\sigma=1}^{n} g^{\rho\sigma}(p_\sigma - g_{0\sigma})\cdot$$

([59]) **G. Vraneceanu**, "Studio geometrico dei sistemi anholonomi," Ann. di mat. (4) **6** (1929), pp. 6, in which he summarized some earlier scattered little papers, as well as **J. L. Synge**, "Geodesics in non-holonomic geometry," Math. Ann. **99** (1928), pp. 738.

([59.a]) In this, it is suggested that the first index represents simply a number that characterizes the equation, while the second index, by contrast, expresses the idea that the $a_{(\mu)\,1}, \ldots, a_{(\mu)\,n}$ are the covariant components of a vector in the **Riemannian** R_n whose associated contravariant components are then:

$$a^\rho_{(\mu)} = \sum_{\sigma=1}^{n} g^{\rho\sigma} a_{(\mu)\sigma}\,.$$

One can imagine that the **Pfaff** equations are given from the outset in such a way that for every μ, the $a^1_{(\mu)}, \ldots, a^n_{(\mu)}$ are the components of a unit vector and that two different unit vectors among those k will be mutually-perpendicular:

$$\sum_{\sigma=1}^{n} g_{\rho\sigma}\, a^\rho_{(\mu)}\, a^\sigma_{(\nu)} = \begin{cases} 0 & (\mu \neq \nu), \\ 1 & (\mu = \nu). \end{cases}$$

in which the **Lagrange** factors (i.e., the reaction forces) λ_μ take the form:

$$\lambda_\mu = -\left(\sum_{\sigma=1}^{n} a_{(\mu)\sigma} Q^\sigma + \sum_{\rho,\sigma=1}^{n} a_{(\mu)\rho\sigma} \dot{q}_\rho \, \dot{q}_\sigma \right) ,$$

in the event that the $(a^1_{(\mu)},, \ldots, a^n_{(\mu)})$ are mutually-perpendicular unit vectors, and the $a_{(\mu)\rho\sigma}$ the covariant derivative of $a_{(\mu)\rho}$ with respect to q_σ ([59.b]).

The transition from equations (65.a) to the equations of motion with the lowest number of equations can be geometrically illustrated by regarding the totality of all $(n - k)$-dimensional elements that the **Pfaff** equations (65) associate with the individual points of the **Riemannian** space as a non-holonomic manifold M_n^{n-k} that is spanned in the **Riemannian** space ([59.c]). Namely, since the reactions of the constraints in (65.a) are perpendicular to the $(n - k)$-dimensional at each point, if one wishes to eliminate then from the equations of motion, one needs only to project all vectors onto the M_n^{n-k} ([59.d]), and in that way obtain the lowest number of equations of motion. For the analytical representation, it would then be preferable to introduce quasi-coordinates (in the chosen way of looking at things, they are usually referred to as *non-holonomic parameters*) in order to represent the M_n^{n-k} by the simple equations (39.b). The kinetic energy will then take on a form whose coefficients are the projections of the fundamental tensor of the g_{ik} (or arc-length on the **Riemannian** space) onto the M_n^{n-k}, and the coefficients that appear in the transitivity equations are closely connected with the translation quantities for vectors that the **Riemannian** R_n imprints upon the M_n^{n-k} that they span.

That consideration, which initially referred to only scleronomic constraints, can be generalized to rheonomic constraints. That is because it is basically inessential that the R_n of the $q_1, \ldots, q_n$ should be a **Riemannian** space, rather it is essential that the **Christoffel** three-index symbols should define a direction of translation in it and thus make it possible to form covariant derivatives. Therefore, **A. Wundheiler** ([60]) regarded the $(n + 1)$-dimensional manifold of $q_1, \ldots, q_n, t$ as an *affine space* A_{n+1}. One defines a linear translation in that A_{n+1} [cf., III D 11 (**L. Berwald**), no. **28**] with the help of the three-index symbols that appear in the equations of motion (64.b) ([60.a]). If non-holonomic constraints of the form (9) exist, which might now be written:

([59.b]) Cf., **J. L. Synge**, "On the geometry of dynamics," Trans. London Phil. Soc. (A) **226** (1927), pp. 31, esp., pp. 53, *et seq.*

([59.c]) The notation shall suggest that the manifold belongs to *all* points in the **Riemannian** space R_n, but that at each point of it, the only allowable directions of advance are the ones that belong to the $(n - k)$-dimensional elements.

([59.d]) Cf., **J. A. Schouten**, "Über nichtholonome Übertragungen in einer L_n," Math. Zeit. **30** (1929), pp. 149, as well as **J. A. Schouten** and **E. R. van Kampen**, "Zur Einbettungs- und Krümmungstheorie nichtholonomer Gebilde," Math. Ann. **103** (1930), pp. 752. The starting point in those treatises was chosen more generally than was necessary here.

([60]) **A. Wundheiler**, "Über die Variationsgleichungen für affine geodätische Linien und nichtholonome, nichtkonservative dynamische Systeme," Prace matemat.-fis. **38** (1931), pp. 126.

([60.a]) If one chooses the translation quantities according to the formulas:

(66) $$a_{(\mu)1}\, dq_1 + \ldots + a_{(\mu)n}\, dq_n = 0 \qquad (\mu = 1, \ldots, k),$$

corresponding to (65), then they will define a non-holonomic manifold M_{n+1}^{n-k+1} in the affine space A_{n+1} that is embedded in the A_{n+1}, and the linear translation in the A_{n+1} then determines a linear translation in that non-holonomic M_{n+1}^{n-k+1}. One also starts from the equations of motion, which will possess the form ([60.b]):

(66.a) $$\ddot{q}_\tau + \sum_{\rho,\sigma=1}^{n} \left\{ \begin{matrix} \rho\, \sigma \\ \tau \end{matrix} \right\} \dot{q}_\rho\, \dot{q}_\sigma + 2\sum_{\rho=1}^{n} \left\{ \begin{matrix} 0\, \rho \\ \tau \end{matrix} \right\} \dot{q}_\rho + \left\{ \begin{matrix} 0\, 0 \\ \tau \end{matrix} \right\} = Q^\tau + \sum_{\mu=1}^{k} \lambda_\mu\, a_{(\mu)}^\tau$$

here, corresponding to (65.a), and when one projects them onto the M_{n+1}^{n-k+1} (while discarding the reaction forces), one will get the lowest number of equations of motion, in which one once more finds it preferable to appeal to a convenient analytical representation of the non-holonomic M_{n+1}^{n-k+1} quasi-coordinates (non-holonomic parameters). The coefficients in the transitivity equations are connected with the quantities that describe the direction of translation in the M_{n+1}^{n-k+1}.

7. Painlevé's general discussion of the singularities of the trajectories of a system ([61]). – In order to get a glimpse into the singularities of the trajectories, one must next look at the singularities of the functions g_{ik} and those of the force components $Q_1, \ldots, Q_n$ and treat the course of motion in the neighborhood of those singular points according to the general rules [cf., II A 4.a

$$\Gamma_{\rho\sigma}^{\tau} = \left\{ \begin{matrix} \rho\, \sigma \\ \tau \end{matrix} \right\}, \qquad \Gamma_{0\rho}^{\tau} = \left\{ \begin{matrix} 0\, \rho \\ \tau \end{matrix} \right\}, \qquad \Gamma_{00}^{\tau} = \left\{ \begin{matrix} 0\, 0 \\ \tau \end{matrix} \right\} - Q^\tau, \quad \Gamma_{\rho\sigma}^{0} = 0, \quad \Gamma_{00}^{\tau} = 0 \qquad (\rho, \sigma, \tau = 1, \ldots, n)$$

then the space-time lines of the motion are precisely the geodetic lines of the A_{n+1}. In that way, the geodetic lines are imagined to be defined to be the curves whose tangents are parallel to the direction of translation [cf., III D 11 (**L. Berwald**), no. **18**].

([60.b]) The $a_{(\mu)}^\tau$ in that are defined by:

$$a_{(\mu)}^\tau = \sum_{\rho=1}^{n} g^{\tau\rho}\, a_{(\mu)\rho},$$

as in (59.a).

([61]) The investigation is restricted to the case in which the kinetic energy T is a quadratic form in the $\dot{q}_\rho$ whose coefficients do not include time t, just like the components Q_ρ of the applied force.

(**P. Painlevé**) and III D 8 (**H. Liebmann**)] ([61.a]). Moreover, the so-called *equilibrium points* ([62]), i.e., the points at which all n force components $Q_1, \ldots, Q_n$ vanish, are especially significant for the form of the trajectory. Moreover, they are not only singular points for the system of trajectories, but also for the system of lines of force.

The construction that one ordinarily uses to prove the existence of solutions to the differential equations (64) of the trajectories can be interpreted as follows: At the starting point, the direction vector of the trajectory and the contravariant force vector will span a two-dimensional planar manifold that is the osculating M_2 for the desired trajectory. Since one knows the initial value of the curvature, one can construct an arc-length to the curve according to the last of equations (62). One then finds the new tangent direction and the new value of the curvature, and since the new value of the force vector, together with the direction vector, once more implies the osculating M_2, one can repeat the construction. The construction will not even fail when the direction of the force vector and the tangent direction of the trajectory coincide, so the normal components of the force will be $V = 0$. The trajectory can then be continued at the points where one has $K_g = 0$ or $\dot{s} = 0$ ([63]). The latter, which **Painlevé** referred to as *stopping points* ([64]), will become significant on the basis of things that will soon become clear. By contrast, the construction can break down when one reaches an equilibrium point ([65]). In that case, the trajectory must end at the equilibrium point and cannot be continued beyond the equilibrium point, which is otherwise true for only singular points of the g_{ik} and Q_i . A necessary condition for the trajectory to not be continued beyond the equilibrium configuration is that the velocity must go to zero when the system approaches the

([61.a]) At a regular point, the g_{ik} must be twice continuously differentiable and the Q_i must be once continuously differentiable. Furthermore, the *determinant G* of the quadratic form:

$$ds^2 = \sum_{i,k=1}^{n} g_{ik}\, dq_i\, dq_k$$

cannot vanish there. Moreover, one can also see immediately how to proceed when the g_{ik} (Q_i, resp.) are not single-valued. On the subject of singularities, see also **P. Painlevé**, "Leçons sur la théorie analytique des équations différentielles," Paris 1897 (lithogr.), esp. pp. 543-589.

([62]) Cf., in addition to (55), also **P. Painlevé**, *L'intégr. d. équ. de la méc.*, esp. Leçons 16.

([63]) For $\dot{s}^2 = 0$, the curvature K_g indeed assumes the indeterminate form 0 / 0, but it will nonetheless always yield a well-defined limiting value for K_g as long as the $\left\{ \begin{matrix} \rho\, \sigma \\ \lambda \end{matrix} \right\}$ and the Q_λ are twice continuously differentiable. Analogous statements will be true for the value of $\dot{s}^2$ when the curvature is $K_g = 0$.

([64]) For **Painlevé**, the term was *point d'arrét*. Moreover, for $\dot{s} = 0$, i.e., for $\dot{q}_1 = 0, \ldots, \dot{q}_n = 0$, the trajectory must necessarily possess the direction of the line of force, as one sees immediately from equations (59).

([65]) However, the trajectory can also be continued past an equilibrium point. If the moving system is brought into an equilibrium configuration, and the velocity is zero there ($\dot{q}_1 = 0, \ldots, \dot{q}_n = 0$) then the trajectory will consist of the single point $q_1 = a_1, \ldots, q_n = a_n$. When a trajectory moves towards an equilibrium configuration, it can attain the equilibrium configuration with a certain direction and in a finite arc-length (without it then needing to be continuable beyond the equilibrium configuration). However, it can move towards the equilibrium configuration without its tangent direction approaching a certain limiting position in that way.

equilibrium configuration. **Painlevé** then subdivided the equilibrium points into *regular* ones ([66]), i.e., ones at which $\dot{s}$ does not vanish, and *singular* ones, at which $\dot{s}$ does equal zero.

When one now focusses on the *time evolution* of the motion along an individual trajectory, one must decide whether one has:

$$\dot{s}^2 = \frac{V}{K_g} \geq 0 \qquad \text{or} \qquad \dot{s}^2 = \frac{V}{K_g} \leq 0 \,.$$

If $\dot{s}^2 > 0$ for all points of a trajectory then according to **Painlevé**, it will be called a *true trajectory* (trajectory of the true motion). By contrast, if $\dot{s}^2$ is negative everywhere along the trajectory then **Painlevé** spoke of a *conjugate trajectory* ([67]) (trajectory of the conjugate motion). It is implicitly assumed in this that $\dot{s}^2$ always has the same sign along the trajectory, such that the velocity $\dot{s}$ will not be zero. A true trajectory will then be traversed completely in one sense. Every finite point that is reached at a finite time $\bar{t}$ will also be reached with a certain finite velocity, such that the motion can be continued beyond the time-point $\bar{t}$ ([68]). However, the general case is naturally the one in which there are points along a trajectory for which the $\dot{s}$ are equal to zero. If such a *stopping point* $\dot{s} = 0$ is not coincidentally an equilibrium point then the motion will reverse at the stopping point and traverse the same trajectory backwards, such that it will pass through the same point on the trajectory at time $t^* + \tau$ that it passed through at time $t^* - \tau$, in which t^* is understood to mean that time at which the stopping point is reached ([68.a]). In fact, the root in (63) will change sign at such a point since V will vanish simultaneously with $\dot{s}$ ([69]). If one now pursues the value of $\dot{s}^2$ beyond a stopping point then V will change sign along with $\dot{s}^2$. The stopping point then splits the curve into two pieces, for which the motion in question will be real along one of them, while the conjugate motion will be real for the other one. Such trajectories with stopping points then carry the true motion along part of their extent and the conjugate motion along the other. On those grounds, **Painlevé** referred to them as *mixed trajectories* (*trajectoires mixtes*). There are ∞^n such mixed trajectories, because for every point there is a mixed trajectory that possesses that point as

---

([66]) ∞^1 trajectories will go through a regular equilibrium point with a prescribed direction, just like an ordinary regular point, corresponding to the fact that the value of $\dot{s}$ can be freely assigned. However, the difference between a regular equilibrium point and an ordinary regular point is that all of those trajectories will possess the geodetic curvature $K_g = 0$ at the equilibrium point.

([67]) The terminology "conjugate" is understood to mean that it would be a "true" trajectory if one replaced t with (it), where i means the imaginary unit.

([68]) Furthermore, not all q_i need to remain finite for a finite $\bar{t}$ by any means.

If the system approaches an equilibrium configuration and its velocity then goes to zero as a result then that *cannot* happen in a finite length of time; rather, t must become infinite. Conversely, if the system approaches a certain finite configuration when $t \to \infty$ then that configuration must be an equilibrium point, and the system must approach that configuration in such a way that the velocity components go to zero like $1 / t$.

([68.a]) **Ph. Frank** thus referred to those points as *turning points* (*Umkehrpunkte*).

([69]) If one then develops the q_ρ in powers of $t = (t - t^*)$ then only even powers of that quantity can appear in such a development. For singular points of the g_{ik} (the Q_i, resp.), cf., **P. Painlevé**, *loc. cit.* ([55]), esp., pp. 153.

a stopping point ([70]). It must then be true that any point of the mixed trajectory can be a stopping point of a motion. In that case, since ∞^1 motions are possible along the mixed trajectories, they must be "remarkable trajectories," i.e., they must coincide with the lines of force, which are then, at the same time, geodetic lines of the arc-length. It is only in that case that the totality of mixed trajectories will reduce to ∞^{n-1}. Conversely, the remarkable trajectories are naturally always mixed trajectories for which every point is a possible stopping point. If two or more stopping points occur on a mixed trajectory ([71]) then the trajectory will be divided by them into pieces that alternately belong to the true or conjugate motions. The motion then results in such a way that the system (the associated representative point, resp.) will move *periodically* back and forth between the two stopping points.

Of the *singular* trajectories, **Painlevé** investigated the ones along which there were equilibrium points with vanishing velocities. Such an equilibrium point is an asymptotic point for the *motion* of the system in the sense that the system (the representative point, resp.) approaches such an equilibrium point as $t \to \infty$. By contrast, the *trajectory* can go through such a point in a normal way ([72]). Of course, it is also possible that it ends at such a point and cannot be continued beyond it. Indeed, it does not even need to have a well-defined tangent there, nor does the arc-length up to the equilibrium point need to be finite. It is always assumed then that the velocity between an arbitrary point on the trajectory and the equilibrium point possesses only finitely-many zeroes. **Painlevé** gave an example in which infinitely-many zeroes of the velocity could also appear along a trajectory that would accumulate at the equilibrium point ([73]).

([70]) A pencil of mixed trajectories goes through a point, so any one of them possess its stopping point at that point. The direction is that of the line of force that goes through the point, and the value of the geodetic curvature K_g is found to be determined completely by a passage to the limit.

([71]) Aside from the remarkable trajectories, that can occur only in exceptional cases. Only a finite number of stopping points can ever lie between two given finite points of a trajectory that is not a remarkable trajectory.

([72]) In particular, the remarkable trajectories can always be continued beyond any equilibrium point that might occur.

([73]) If the trajectory can be continued beyond an equilibrium point then further equilibrium points can appear on it. There can also be infinitely many of them, and that sequence can possess an accumulation point.

CHAPTER II

THE VARIATIONAL PRINCIPLES.

8. Hamilton's principle. – One will easily arrive at the transition to the variational principles by starting from the differential principles. One starts from the formulas for the ***Lagrange*** *central equation* in the form (27) [(27.a), resp.] of no. **4**. If one regards the virtual displacements δq_1, …, δq_n as functions of time in the sense of no. **3** then one can integrate formula (27.a) between two fixed values of time $t = t_1$ and $t = t_2$ ([74]) and obtain:

$$\int_{t_1}^{t_2} [\delta T + (Q_1 \delta q_1 + \cdots + Q_n \delta q_n)\, dt = \left[\frac{\partial T}{\partial \dot{q}_1} \delta q_1 + \cdots + \frac{\partial T}{\partial \dot{q}_n} \delta q_n \right]_{t_1}^{t_2}, \tag{67}$$

which is a relation that one can refer to as the *integrated central equation* ([75]). If the forces arise from a potential $\Phi(q_1, \ldots, q_n, t)$, so:

$$Q_\rho = - \frac{\partial \Phi}{\partial q_\rho},$$

then (67) will take the form:

$$\int_{t_1}^{t_2} \delta (T - U)\, dt = \left[\frac{\partial T}{\partial \dot{q}_1} \delta q_1 + \cdots + \frac{\partial T}{\partial \dot{q}_n} \delta q_n \right]_{t_1}^{t_2}. \tag{67.a}$$

Now, from no. **3**, the virtual displacements δq_ρ are performed for constant t, such that the variation symbol δ will commute with the integral sign. Moreover, Φ is independent of the $\dot{q}_\rho$, so $\partial \Phi / \partial \dot{q}_\rho = 0$. Therefore, one can convert formula (67.a) into:

([74]) This argument was already applied, in principle, by **J. L. Lagrange** himself, but he derived the principle of least action (cf., no. 10) in that way, see, *Mécanique*, 2. partie, sect. III, **§ 6** = *Œuvr.* 11, pp. 315. Cf., moreover, e.g., **L. Boltzmann**, *Prinzipe* II, **§ 1**.

([75]) One can also integrate the differential formula of least constraint:

$$\int_{t_1}^{t_2} \sum_{i=1}^{r} [(m_i \ddot{x}_i - X_i)\, \delta \ddot{x}_i + (m_i \ddot{y}_i - Y_i)\, \delta \ddot{y}_i + (m_i \ddot{z}_i - Z_i)\, \delta \ddot{z}_i]\, dt$$

[cf., **J. Schenkel**, "Über eine dem Gaußschen Prinzip des kleinsten Zwanges entsprechende Integralform," Wien Sitzungsber. **122**, II[a1] (1913), pp. 721] and seek to convert it into an integral principle, whereby one will naturally not arrive at a variational problem, but an integrated differential formula. In that way, one must perform a variation that leaves the position and velocity along a piece of the space-time line unchanged, while the accelerations shall be varied everywhere along the entire time interval. Naturally, that can only lead to the most fruitless artifices in one's definition of the variation, but it is basically nothing but the logical continuation of the so-called "**Hölder** type of variation," i.e., the method that leads to the desired result by understanding the variation to mean something suitable in each case.

For that, cf., also **A. Wassmuth**, "Studien über Jourdain's Prinzip der Mech.," Wien Sitzungsber. **128**, II[a] (1919), pp. 365, who showed how an integral formula related to **Jourdain**'s principle [no. **5** ([40.b])].

(67.b) $$\delta \int_{t_1}^{t_2} (T-\Phi)\, dt \;=\; \left[\frac{\partial (T-\Phi)}{\partial \dot{q}_1}\, \delta q_1 + \cdots + \frac{\partial (T-\Phi)}{\partial \dot{q}_n}\, \delta q_n \right]_{t_1}^{t_2}.$$

The variation of the definite integral $\int_{t_1}^{t_2} (T-\Phi)\, dt$ (except for the limits) is represented in that form when the variation is performed with time held constant, which then says that the first variation that belongs to the variational problem:

(68) $$\int_{t_1}^{t_2} (T-\Phi)\, dt \;= \text{extrem.}$$

must vanish. However, that means nothing but the idea that the equations of motion of the mechanical system must be given as the **Euler** equations of the variational problem (68), i.e., from the demand that the integral of the so-called ***Lagrangian*** *function* $L = T - \Phi$ (for fixed limits) must be an extremum. That variational problem is referred to as ***Hamilton's*** *principle* (on the Continent), although that terminology (which has not been accepted in England, moreover) is not justified historically. [Cf. IV 1 (**A. Voss**), no. **42**, esp., footnote 243.]

Of course, in going from the variational problem to the Ansatz for the equations of motion, the only extremum conditions that will come into question are expressed by the vanishing of the *first* variation. One can restrict oneself to that as long as one has only the Ansatz for the equations of motion in mind. If one holds the limits of the integral (68) fixed under that variation then that will imply the formula:

(68.a) $$\int_{t_1}^{t_2} \delta (T-\Phi)\, dt \;= 0\,,$$

which is nothing but the formula (67.b) when one drops the right-hand side, which originates in the variation of the limits. Many times, in the literature, it is not actually the variational problem (68) that one appeals to as an expression of **Hamilton**'s principle for the Ansatz of the equations of motion, but the variational formula (68.a). That is closely related to referring to formula (67) (correspondingly distorted while fixing the limits) for the integrated central equation:

(67*) $$\int_{t_1}^{t_2} [\delta T + (Q_1\, \delta q_1 + \cdots + Q_n\, \delta q_n)\, dt \;= 0 \qquad \begin{cases} \delta q_\rho (t_1) = 0, \\ \delta q_\rho (t_2) = 0, \end{cases}$$

as the formula for **Hamilton**'s principle, although here one certainly cannot speak of a variational problem, at least in the sense of the classical calculus of variations ([76]). By contrast, the variational problem shall be understood in the narrow sense here.

That difference in viewpoint will take on a deeper meaning as soon as one no longer assumes, as before, that the q_ρ are independent coordinates, but assumes that *auxiliary conditions* exist between them. From the standpoint of the classical calculus of variations, one must then impose condition equations on the functions $q_\rho\,(t)$ that are allowable for concurrence with the extremum in the variational problem, i.e., perform the variation of the integral in such a way that not only the extremal itself, but also the varied curve will fulfill the prescribed condition equations. By contrast, in the integrated central equation (67*), the variations are chosen such that they will represent virtual displacements in the previously-defined sense (cf., no. **3**). Those two types of variations (which are performed on the space-time lines of motion in the R_{n+1} of $q_1, \ldots, q_n, t$ for fixed time) are probably compatible for holonomic constraints, *but not for non-holonomic constraints* (no. **2**). If one would then like to restrict oneself to the applications of the classical calculus of variations that would seem appropriate to the variational principles then one would have to say that *Hamilton's principle will imply the Ansatz for the equations of motions for holonomic auxiliary conditions, but not for non-holonomic ones* ([77]). In addition, one can properly speak of a variational

---

([76]) The integral formula can also serve to exhibit the equations of motion for *quasi-coordinates*. Cf., **G. Hamel**, "Über die virtuellen Verschiebungen in der Mechanik," Math. Ann. **59** (1904), pp. 416, esp., pp. 426 and 427, as well as **Th. Pöschl**, C. R. Acad. Sci. Paris **156** (1913), pp. 1829. In that way, it takes the form:

$$\int \left[\sum_\rho \left(\frac{\partial (T^* - \Phi)}{\partial \kappa_\rho}\, \delta \kappa_\rho + \frac{\partial T^*}{\partial \omega_\rho}\, \delta \frac{d\kappa_\rho}{dt} \right) \right] dt = 0\,,$$

in which one has correspondingly substituted the transitivity equations (17.a) from no. **2**:

$$\delta \frac{d\kappa_\rho}{dt} = \frac{d}{dt}(\delta \kappa_\rho) - \sum_{\sigma,\tau} \gamma^\rho_{\sigma\tau} \left(\frac{d\kappa_\sigma}{dt}\, \delta \kappa_\tau - \delta \kappa_\sigma \frac{d\kappa_\tau}{dt} \right).$$

One will then get:

$$\frac{d}{dt}\left(\frac{\partial T^*}{\partial \omega_\rho} \right) - \sum_{\sigma,\tau} \gamma^\tau_{\sigma\rho}\, \omega_\rho \frac{\partial T^*}{\partial \omega_\tau} - \frac{\partial (T^* - \Phi)}{\partial \kappa_\rho} = 0$$

as the form of the equations of motion. **Pöschl** correspondingly restricted himself to the scleronomic case, which is the usual custom in the literature.

([77]) For this way of looking at things, cf., **H. Hertz**, *Prinzipien der Mechanik, Ges. Werke* III (2nd ed.), Leipzig, 1910, pp. 23. That state of affairs is characterized by the following comparison: If a holonomic M_{n-k} is embedded in a **Riemannian** space R_n then the "geodetic" lines in that M_{n-k}, which one defines to be curves with parallel-translated tangent directions, will be the "shortest" lines in the M_{n-k}, which one characterizes as the extremals of the variational problem:

$$\int ds \;= \text{extrem.}$$

By contrast, if a non-holonomic M_n^{n-k} is embedded in the **Riemannian** space M_n then the "geodetic" lines in M_n^{n-k} that are defined by parallel translation of the tangent direction will be *different* from the "shortest" lines, which one defines to be the extremals of the variational problem:

principle only when the general components Q_ρ of the applied forces arise from a potential Φ (q_1, …, q_n, t). That standpoint is not adopted in the literature, as a rule, but rather one wishes to apply **Hamilton**'s principle to the widest-possible range of applications. In order to do that, one does not just speak of the variational problem (68) as **Hamilton**'s principle, but the variational formula (67*) of the integrated central equation and demands that the variations that appear in it must be performed at fixed times, and that they shall be subject to the same conditions as virtual displacements (cf., no. **3**), moreover ([78]).

Here one says that the term ***Hamilton**'s principle* is understood *in the narrow sense*, so it will only apply to mechanical systems for which the *forces derive from a potential* and *the constraints are holonomic*. Furthermore, one might eliminate the associated auxiliary conditions from the outset by the introduction of suitable general coordinates. With that assumption, **Hamilton**'s principle says that the equations of motion of the mechanical system will be given by the ***Euler** equations* of the variational problem:

(69) $$\int_{t_1}^{t_2} (T - \Phi)\, dt \;= \text{extrem.}$$

with no further auxiliary conditions. The function:

(69.a) $$L(\dot{q}_1, \ldots, \dot{q}_n, q_1, \ldots, q_n, t) \;= T - \Phi$$

$$\int ds \;= \text{extrem.}$$

with the auxiliary conditions defined to be non-holonomic conditions.

([78]) Cf., **O. Hölder**, "Über die Prinzipien von Hamilton und Maupertuis," Gött. Nachr. (1896), pp. 122. See also **Cl. Schaefer**, *Prinzipe*, **§ 8**. Should the varied space-time lines satisfy the prescribed auxiliary conditions, as is required of a variation in the classical calculus of variations, then the variation *could not be virtual displacements* in the case of non-holonomic constraints. On pp. 126, **Hölder** admitted that the term "the principle of least action" was basically no longer appropriate to his type of variation. **Hölder** exhibited a further variational formula from which one can arrive at the variational formula of **Hamilton**'s principle, as well as the **Euler-Maupertuis** principle that will be treated below (no. **10**). However, he had to appeal to gimmickry in that, on the one hand, he chose the variations of the position coordinates to be virtual displacements relative to auxiliary conditions, while on the other hand, in the variational formula for the integral, the time ordering between any two position points that are coupled by a virtual displacement was established arbitrarily, which would properly contradict the concept of a virtual displacement. **C. Schaefer** (*Prinzipe*, **§ 11**) took up **Hölder**'s formula, but without mentioning that contradiction. **A. Voss** adapted **Hölder**'s argument to general coordinates in "Über die Prinzipe von Hamilton und Maupertuis," Gött. Nachr. (1900), pp. 322. Moreover, the **L. Maurer** explained the different types of variation in "Über die Differentialgl. der Mech.," Gött. Nachr. (1905), pp. 91.

The difficulties in applying **Hamilton**'s principle in the case of non-holonomic constraints were also treated thoroughly by **G. Morera**, "Sulle equazioni dinamiche di Lagrange," Turin Atti **38** (1902), pp. 121. He appealed to the bilinear covariants of the **Pfaff**ian expressions (9) in order to examine when the consideration of the **Pfaff** equations (9) as the auxiliary conditions of the variational problem in the sense of the classical calculus of variations would lead to the correct form for the equations of motion and found that that would happen if and only if the **Pfaff** equations define a completely-integrable system.

More recently, **M. Kerner** has once more treated the problem in "Le principe de Hamilton et l'holonomisme," Prace mat.-fis. **38** (1931), pp. 1, without adding anything essentially new.

that appears under the integral, which is the difference between the kinetic energy and the force potential, was referred to by **E. J. Routh** as the ***Lagrangian*** *function* ([79]), and by **H. von Helmholtz** as the *kinetic potential* ([80]). Upon introducing that function L, the equations of motion for the mechanical system will take the form:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\rho}\right) - \frac{\partial L}{\partial q_\rho} = 0 \qquad (\rho = 1, \ldots, n). \tag{70}$$

The connection between those equations and the variational problem (69), which also requires its special form, leads to some results that are important for their *integration*, as will be made clear in what follows. To that extent, the variational principle overlaps with the arguments regarding the differential principle. The results concerning the integration of equations (70) can obviously be adapted to the general form of the **Lagrange** equations of motion:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\rho}\right) - \frac{\partial T}{\partial q_\rho} = Q_\rho \qquad (\rho = 1, \ldots, n)$$

to the extent that they arise from the formula for the first variation (68.a) since the general formula (67) indeed runs parallel to the integrated central equation. The meaning of the integrated central equation (67), as opposed to the original form of the central equation in no. **4**, consists of precisely the fact that it makes it possible to adapt part of that result, because, as will be shown, formula (67.b) is important for the theory of integration in particular, and it represents a special case of the so-called *boundary formula for the calculus of variations*. In order to get that *boundary formula* in general, one extends the integral of the variational problem (69) along an extremal, i.e., along a space-time line of motion that connects two given space-time points $P_1(q_1^{(1)}, \ldots, q_n^{(1)}, t_1)$ and $P_2(q_1^{(2)}, \ldots, q_n^{(2)}, t_2)$, and goes over to a second extremal that connects two points that are infinitely close to P_1 (P_2, resp.). The change in the value of the integral along the extremal will then be ([†]):

$$\begin{aligned}\delta \mathcal{E}\int_{t_1}^{t_2} L\,dt \;=\; & \left(\frac{\partial L}{\partial \dot{q}_1}\right)_2 \delta q_1^{(2)} + \cdots + \left(\frac{\partial L}{\partial \dot{q}_n}\right)_2 \delta q_n^{(2)} + \left(L - \dot{q}_1 \frac{\partial L}{\partial \dot{q}_1} - \cdots - \dot{q}_n \frac{\partial L}{\partial \dot{q}_n}\right)_2 \delta t_2 \\ & - \left(\frac{\partial L}{\partial \dot{q}_1}\right)_1 \delta q_1^{(1)} + \cdots + \left(\frac{\partial L}{\partial \dot{q}_n}\right)_1 \delta q_n^{(1)} + \left(L - \dot{q}_1 \frac{\partial L}{\partial \dot{q}_1} - \cdots - \dot{q}_n \frac{\partial L}{\partial \dot{q}_n}\right)_1 \delta t_1 \;,\end{aligned}$$

which is a relation that emerges from the formula (67.b) when one sets $\delta t_1 \equiv 0$, $\delta t_2 \equiv 0$.

([79]) **E. J. Routh**, *Dynamik* I, pp. 357.

([80]) **H. von Helmholtz**, "Die physikalische Bedeutung des Prinzips der kleinsten Wirkung," J. f. Math. **100** (1887), pp. 137. = *Ges. Abh.* III, pp. 203; cf., moreover, **H. von Helmholtz**, *Dynamik*, pp. 359.

([†]) Translator: I could not duplicate Prange's notation for an integral along an extremal, so I substituted the script E as a prefix to the integral.

The quantities that appear in that formula (71) have an immediate physical meaning, because since $L = T - \Phi$, and only the velocity components $\dot{q}_\rho$ appear in T, the derivatives:

(72) $$\frac{\partial L}{\partial \dot{q}_\rho} = \frac{\partial T}{\partial \dot{q}_\rho} = p_\rho$$

will be the *impulse components* of the mechanical system. Moreover, if one sets:

(73) $$L - \dot{q}_1 \frac{\partial L}{\partial \dot{q}_1} - \cdots - \dot{q}_n \frac{\partial L}{\partial \dot{q}_n} = - H$$

then H will also admit an intuitive mechanical interpretation. Namely, if T is a quadratic form in the velocity components $\dot{q}_\rho$:

$$T = \tfrac{1}{2} \sum g_{\lambda\mu}\, \dot{q}_\lambda\, \dot{q}_\mu$$

then one will get:

$$\dot{q}_1 \frac{\partial L}{\partial \dot{q}_1} + \cdots + \dot{q}_n \frac{\partial L}{\partial \dot{q}_n} = \dot{q}_1 \frac{\partial T}{\partial \dot{q}_1} + \cdots + \dot{q}_n \frac{\partial T}{\partial \dot{q}_n} = 2\, T\,.$$

One then has:

$$- H = T - \Phi - 2T = - (T + \Phi)\,,$$

or

(73.a) $$H = T + \Phi\,,$$

resp., i.e. (in the event that T and Φ do not include time t explicitly), *H is the sum of the kinetic and potential energy of the system.* On the other hand, if T is a general quadratic function of the $\dot{q}_\rho$:

$$T = T_0 + T_1 + T_2$$

then:

$$\dot{q}_1 \frac{\partial L}{\partial \dot{q}_1} + \cdots + \dot{q}_n \frac{\partial L}{\partial \dot{q}_n} = 2\, T_2 + T_1\,,$$

so:

$$L - \dot{q}_1 \frac{\partial L}{\partial \dot{q}_1} - \cdots - \dot{q}_n \frac{\partial L}{\partial \dot{q}_n} = (T_0 - \Phi) - T_2\,,$$

and that will imply the following expression for H:

(73.b) $$H = T_2 + (\Phi - T_0)\,,$$

which can likewise be referred to as the total energy of the system ([81]). With the introduction of the impulse coordinates p_ρ and the energy H, the boundary formula (71) will take the form:

$$(71.a) \qquad \delta_{\mathcal{E}} \int_{t_1}^{t_2} L\,dt = [\,p_1^{(2)}\,\delta q_1^{(2)} + \cdots + p_n^{(2)}\delta q_n^{(2)} - (H)_2\,\delta t_2\,] - [\,p_1^{(1)}\,\delta q_1^{(1)} + \cdots + p_n^{(1)}\delta q_n^{(1)} - (H)_1\,\delta t_1\,] \ .$$

In so doing, it is convenient to not regard the function H as a function of t, q_ρ, $\dot{q}_\rho$, but to replace the velocity components $\dot{q}_\rho$ with the impulse components p_ρ. In the case (73.a), since one has:

$$T = \tfrac{1}{2} \sum_{\rho,\sigma=1}^{n} g_{\rho\sigma}\,\dot{q}_\rho\,\dot{q}_\sigma = \tfrac{1}{2} \sum_{\rho=1}^{n} \dot{q}_\rho \frac{\partial T}{\partial \dot{q}_\rho} = \tfrac{1}{2} \sum_{\rho=1}^{n} \dot{q}_\rho\,p_\rho = \tfrac{1}{2} \sum_{\rho,\sigma=1}^{n} g^{\rho\sigma} p_\rho\,p_\sigma \ ,$$

that will then give:

$$(74.a) \qquad H\,(p_1, \ldots, p_n, q_1, \ldots, q_n, t) = \tfrac{1}{2} \sum_{\rho,\sigma=1}^{n} g^{\rho\sigma} p_\rho\,p_\sigma + \Phi\,(q_1, \ldots, q_n, t) \ .$$

In the general case (73.b), one has:

$$T_2 = \tfrac{1}{2} \sum_{\rho,\sigma=1}^{n} g_{\rho\sigma}\,\dot{q}_\rho\,\dot{q}_\sigma = \tfrac{1}{2} \sum_{\rho=1}^{n} \dot{q}_\rho \frac{\partial T_2}{\partial \dot{q}_\rho} = \tfrac{1}{2} \sum_{\rho=1}^{n} \dot{q}_\rho\,(p_\rho - g_{0\rho}) = \tfrac{1}{2} \sum_{\rho=1}^{n} g^{\rho\sigma}(p_\rho - g_{0\rho})(p_\sigma - g_{0\sigma}) \ ,$$

and therefore ([82]):

$$(74.b) \qquad H\,(p_1, \ldots, p_n, q_1, \ldots, q_n, t) = \tfrac{1}{2} \sum_{\rho=1}^{n} g^{\rho\sigma}(p_\rho\,p_\sigma - 2\,g_{0\rho}\,p_\sigma - g_{0\rho}\,g_{0\sigma}) + (\Phi - \tfrac{1}{2}\,g_{00}) \ .$$

([81]) Cf., say, **G. D. Birkhoff**, *Dynamical Systems*, Chap. 1, no. 6, pp. 14. Namely, in the energy equation (power equation):

$$\sum_{\rho=1}^{n} \dot{q}_\rho \left[\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\rho} \right) - \frac{\partial T}{\partial q_\rho} \right] = \sum_{\rho=1}^{n} Q_\rho\,\dot{q}_\rho \ ,$$

the terms that originate in T_1 drop out when T does not include the time t explicitly. On the other hand, one can imagine that the term T_0 in:

$$L = T - \Phi = T_2 + T_1 + T_0 - \Phi$$

that is free of $\dot{q}_\rho$ is combined with Φ, so one can write:

$$L = T_2 + T_1 - (\Phi - T_0) \ .$$

([82]) Cf., e.g., **L. P. Eisenhart**, "Dynamical trajectories and geodesics," Ann. Math. (2) **30** (1929), pp. 591.

9. Cyclic coordinates. The canonical form of the equations of motion. The Routh-Helmholtz transformation. – From the form (70) of the equations of motion, it will be immediately obvious that one can give a *first integral* of those equations when one of the coordinates, say q_n, *does not* appear in the **Lagrangian** function L, because one will then have:

$$(75) \qquad \frac{\partial L}{\partial q_n} = 0 , \qquad \text{so} \qquad \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_n} \right) = 0 \qquad \text{or} \qquad \frac{\partial L}{\partial \dot{q}_n} = \text{const.} = c_n .$$

Now, since a coordinate will not generally appear in the expression for the **Lagrangian** function when it belongs to a motion that returns to itself (i.e., a cyclic one) ([83]), **H. von Helmholtz** proposed the name of *cyclic coordinate* for that coordinate ([84]). The expression *hidden* coordinate has also been chosen ([85]), which is patterned on the term *ignored coordinate* that **W. Thomson (Lord Kelvin)** used ([86]). Now what is of fundamental significance for the integration of the equations of motion is the fact that, as **E. J. Routh** and **H. von Helmholtz** observed (and probably independently of each other)([87]), *by means of the first integral* (72), *the system of equations of motion* (70) *can be reduced to an analogous system with only* $(n - 1)$ *unknown functions that is once more the Euler system of equations that arises from a variational problem that is easy to specify.*

Namely, if one imagines a space-time line of the motion of the mechanical system, so one constructs an integral curve of the equations of motion (70) in the manifold of $(q_1, \ldots, q_n, t)$, then obviously the absence of q_n from the equations of motion (70) would have the consequence that every integral curve of the equations (70) will once more arise from an integral curve of those equations when one displaces it parallel to itself by an arbitrary segment in the q_n-direction ([88]). The set of all ∞^{2n} space-time curves is then arranged in such a way that ∞^1 curves will lie on a cylindrical M_2 whose "generators" are parallel to the q_n-direction, and indeed those ∞^1 curves will be *parallel curves* on the cylinder. When one then projects all integral curves parallel to the q_n-direction onto the n-dimensional manifold of $(q_1, \ldots, q_n, t)$, the ∞^1 curves that lie on the same cylinder will possess the same projection, such that one will get only ∞^{2n-1} curves in the M_n of $q_1, \ldots, q_{n-1}, t$ as the projections of the space-time lines of motion in the R_{n+1} of the $q_1, \ldots, q_n, t$. Now, obviously the integration constant c_n in (72) will have the same numerical value for all space-time lines in R_{n+1} that lie on the same cylinder, such that one can also assign a numerical value c_n to each projection onto M_n . Those projections that belong to the same numerical value of the constant

([83]) For example, the angle of rotation φ plays the role of such a coordinate for a rotating flywheel.

([84]) In particular, **Helmholtz** referred to a mechanical system with one such coordinate as *monocyclic*, while he called a mechanical system in which several coordinates are cyclic *polycyclic*. **H. von Helmholtz**, "Studien zur Statik monozyklischer Systeme," Berlin Sitzungsber. (1884), pp. 159; J. f. Math. **97** (1884), pp. 111 – *Ges. Abhandl. III* (1895), pp. 119, and also in **Helmholtz**, *Dynamik*, pp. 362.

([85]) The expression *hidden coordinate* will be explained later on, and likewise the term (that **J. J. Thomson** introduced) *kinosthenic coordinate*, which is still prevalent in England.

([86]) Cf., e.g., **Thomson-Tait**, *Natural Philosophy*, v. I, no. 319, pp. 320 [cf., also IV 1 (**A. Voss**), nos. **26-28**].

([87]) Cf., **E. J. Routh**, *Dynamik*, v. II, Chap. 10, **§ 450**, pp. 331; **H. von Helmholtz**, "Statik monozyklischer Systems," *Ges. Abhandl. III*, pp. 130 or also *Dynamik*, **§ 77**, pp. 361.

([88]) The set of all integral curves then admits the one-parameter group of parallel displacement in the q_n-direction. That group is characteristic of the existence of first integrals (75), cf. *infra*, no. 25.

c_n must then define a family of ∞^{2n-2} curves within the set of ∞^{2n-1} projections such that that set seems to be subdivided into ∞^1 families of ∞^{2n-1} curves. The ∞^{2n-2} curves in one such family of projections with a fixed numerical value of c_n once more define precisely the extremals of a variational problem with $(n-1)$ unknown functions. In order to see that, one puts the equations of motion (70) into the form of a system of $2n$ first-order differential equations, namely, the so-called *canonical system* ([89]), in which one juxtaposes the impulse components p_ρ with the position coordinates q_ρ. (For the historical development of that, cf., *infra*, no. **14**.)

If one writes the relation between H and L in the form:

$$H(p_1, \ldots, p_n, q_1, \ldots, q_n, t) = \dot{q}_1 p_1 + \cdots + \dot{q}_b p_n - L \ , \tag{75.a}$$

corresponding to (73), in which the $\dot{q}_\rho$ are now regarded as functions of p_σ, q_σ, t, then one can read off the relations:

$$\frac{\partial H}{\partial p_\rho} = \dot{q}_\rho \ , \qquad \frac{\partial H}{\partial q_\rho} = -\frac{\partial L}{\partial q_\rho} \ , \qquad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \ , \tag{75.b}$$

the first group of which is only another form of the relations (72) that couple impulse and velocity components. The equations of motion (70) then take the form:

$$\frac{dp_\rho}{dt} + \frac{\partial H}{\partial q_\rho} = 0 \ .$$

The first group of equations (75.b), which give the coupling between the $\dot{q}_\rho$ and p_ρ, are added to that. The system of equations of motion (70) can then be represented by the *canonical system* ([90]):

([89]) In regard to that terminology, which was introduced by **C. G. J. Jacobi** [cf., "Note sur l'intégration des équ. diff. de la dynamique," C. R. Acad. Sci. Paris **5** (1837), pp. 61 = *Werke IV*, pp. 129, esp. pp. 135], **Thomson-Tait** (*Natural Phil.*, I, np. 319, pp. 307) remarked "why it has been so called it would be hard to say."

([90]) The canonical form of the equations of motion was first achieved in a special problem of perturbation theory by **J. L. Lagrange**, cf., *Mécanique analytique* (2nd ed., 1811), 2. part., sect. 5, no. 14 = *Œuvres* XI, pp. 357. **S. D. Poisson** also met up with the canonical form of the equations of motion in his work on perturbation calculations, cf., **S. D. Poisson**, "Sur les inégalités séculaire des moyens mouvements des planètes," J. Éc. Polyt. **8** (1809), pp. 1. In general, the system was then exhibited by **A. Cauchy**, Bull. de la soc. philomath. (1819), pp. 10 and **W. R. Hamilton**, Trans. London Phil. Soc. (1835), pt. I, pp. 95.

The canonical equations have also been written down for more general mechanical problems. For instance, **Th. Pöschl**, "Sur les équations canoniques des systèmes non holonomes," C. R. Acad. Sci. Paris **156** (1913), pp. 1829, presented it for quasi-coordinates. He set:

$$J_\rho = \frac{\partial T^*}{\partial \omega_\rho}$$

and then defined:

$$T = \sum_\rho J_\rho \omega_\rho - T^* \ , \qquad \omega_\lambda = \frac{\partial T}{\partial J_\lambda} \ , \qquad -\left(\frac{\partial T^*}{\partial \kappa_i}\right) = \left(\frac{\partial T}{\partial \kappa_i}\right) \ ,$$

(76) $$\frac{dq_\rho}{dt} = \frac{\partial H}{\partial p_\rho}, \qquad \frac{dp_\rho}{dt} = -\frac{\partial H}{\partial q_\rho} \qquad (\rho = 1, \ldots, n),$$

which represents a system of $2n$ first-order differential equations for the $p_\rho\,(t)$, $q_\rho\,(t)$. Naturally, the special form of that system is required by the form that the equations of motion (70) have as the **Euler** equations of a variational problem. A special advantage in terms of integration also arises from that canonical form.

Now, if q_n is a cyclic coordinate, in particular, then from (75.b), since $\partial L / \partial q_n = 0$, one will also have $\partial H / \partial q_n = 0$, so the canonical system will have the first integral:

(77) $$p_n = \text{const.} = c_n\ .$$

If one substitutes that value of p_n in H then only the $2n - 1$ variables $p_1, \ldots, p_{n-1}$; $q_1, \ldots, q_{n-1}$; t will still appear in H, and one can solve the canonical system (76) with $2n$ unknown functions in such a way that one first solves the canonical system with $2\,(n - 1)$ unknown functions:

(77.a) $$\frac{dq_\rho}{dt} = \frac{\partial H}{\partial p_\rho}, \qquad \frac{dp_\rho}{dt} = -\frac{\partial H}{\partial q_\rho} \qquad (\rho = 1, \ldots, n-1),$$

in which the constant c_n is introduced in place of p_n in H, and then q_n is determined as a function of time by a quadrature from the equation:

(77.b) $$\frac{dq_n}{dt} = \frac{\partial H}{\partial p_n},$$

---

in which T must be a function of the J_ρ . If one then sets:

$$H = T + \Phi$$

then one will get the analogue of the canonical system:

$$\omega_\rho = \frac{d\kappa_\rho}{dt} = \frac{\partial H}{\partial J_\lambda}, \qquad \frac{dJ_\rho}{dt} = -\sum_{\sigma,\tau} \gamma^\tau_{\rho\sigma} \frac{\partial H}{\partial J_\sigma} \frac{\partial H}{\partial \omega_\tau} - \left(\frac{\partial H}{\partial \kappa_\sigma} \right).$$

When H is independent of t, and the defining equations of ω_ρ are independent of t, one can also derive the energy integral $H = k$ from that. Moreover, **G. Hamel** had already expressed the idea of converting the equations of motion for quasi-coordinates into canonical form in "Die *Lagrange-Euler* Gleichungen der Mechanik," Zeit. Math. u. Phys. **50** (1904), pp. 1 (esp., pp. 17, footnote).

A more formal Ansatz for the canonical equations for non-holonomic constraints is also found in **J. Quanjel**, Rend. Palermo **22** (1906), pp. 263. **R. Dautheville**, Bull. Soc. math. Fr. **37** (1909), pp. 120, already formally posed the **Appell** corrected form of the **Lagrange** equations in the canonical form under the assumption of the existence of a potential for the forces. Cf., moreover, **M. Bilimovitsch**, "Sur les transf. canon. des équ. du mouv. d'un system. non holonome," C. R. Acad. Sci. Paris **158** (1914), pp. 1064.

whose right-hand side is now a known function of time t.

However, the integration of the canonical system (77.a) with the 2 $(n-1)$ unknown functions $p_1(t), \ldots, p_{n-1}(t), q_1(t), \ldots, q_{n-1}(t)$ is nothing but the determination of the projections $q_1(t), \ldots, q_{n-1}(t)$ of the space-time lines of motion onto the manifold of $(q_1, \ldots, q_{n-1}, t)$ that were spoken of above, and indeed the ones that belong to the chosen numerical value of c_n. The quadrature (77.b) then yields the function $q_n(t)$ ([91]), and thus each projection of the space-time line itself.

Now, if one performs the transformation that converts the **Euler** equations (70) into the canonical system (76) backwards and applies it to the canonical system (77.a), so one calculates the $p_1, \ldots, p_{n-1}$ from:

$$\dot{q}_\rho = \frac{\partial H}{\partial p_\rho} \qquad (\rho = 1, \ldots, n-1)$$

and then defines:

$$L^*(\dot{q}_1, \ldots, \dot{q}_{n-1}, q_1, \ldots, q_{n-1}, t, c_n) = p_1 \frac{\partial H}{\partial p_1} + \cdots + p_{n-1} \frac{\partial H}{\partial p_{n-1}} - H \tag{78}$$

then one will get the following system of $(n-1)$ second-order differential equations from the canonical system (77.a):

$$\frac{d}{dt}\left(\frac{\partial L^*}{\partial \dot{q}_\rho}\right) - \frac{\partial L^*}{\partial q_\rho} = 0 \qquad (\rho = 1, \ldots, n-1), \tag{79}$$

which are the **Euler** equations of the variational problem:

$$\int_{t_1}^{t_2} L^*(\dot{q}_1, \ldots, \dot{q}_{n-1}, q_1, \ldots, q_{n-1}, t, c_n)\, dt = \text{extrem.} \tag{79.a}$$

A comparison of (78) and (73) will show that:

$$L^*(\dot{q}_1, \ldots, \dot{q}_{n-1}, q_1, \ldots, q_{n-1}, t, c_n) = L(\dot{q}_1, \ldots, \dot{q}_{n-1}, q_1, \ldots, q_{n-1}, t, c_n) - \dot{q}_n \frac{\partial L}{\partial \dot{q}_n}, \tag{79.b}$$

in which $\dot{q}_n$ is replaced with the function of $q_1, \ldots, q_{n-1}, \dot{q}_1, \ldots, \dot{q}_{n-1}, t$ that one calculates from ([92]):

([91]) In which the integration constants that appear additively say just that one will obtain a one-parameter family from a space-time line by parallel translation in the q_n-direction.

([92]) The fact that the projections of the extremals of the variational problem (69) with the same values of c_n prove to be the extremals of the variational problem (79.a) is also easy to see directly, perhaps in the following way: If one has a segment along an extremal of $\int L\, dt$ = extrem. then one can imagine that the extremal is given by the projection e and then determine $q_n(t)$ from equation (79.c) by a quadrature. Now, one has:

(79.c) $$\frac{\partial L}{\partial \dot{q}_n} = c_n \,.$$

That transformation (79.b), (79.c), which yields the integrand of the new variational problem, is the one that one refers to as the *Routh transformation* ([79]) or also the *Helmholtz transformation* ([80]). Moreover, it is easy to see how one might proceed analogously when *two or more cyclic coordinates* are present ([93]).

$$\mathcal{E}\int_{P_1}^{P_2} L\,dt = \mathcal{E}\int_{P_1}^{P_2} (L^* + c_n\,\dot{q}_n)\,dt = \int_{\Pi_1}^{\Pi_2} L^*\,dt + c_n\,(q_n^{(2)} - q_n^{(1)})\ ,$$

in which the last integral can be take over the projection, since q_n does not appear in it. If one now replaces the projection e with another curve ε that connects the points Π_1 and Π_2 then it will belong to a certain function $q_n\,(t)$ that one determines by means of (79.c), and thus a curve in the manifold of (q_1, …, q_n, t) [that goes through P_1] that must end on the line $\Pi_2\,P_2$. The endpoint might be P^* , so let the curve be E. Obviously, one will then have:

$$\mathsf{E}\int_{P_1}^{P^*} L\,dt = \mathsf{E}\int_{P_1}^{P^*} (L^* + c_n\,\dot{q}_n)\,dt = \varepsilon\int_{\Pi_1}^{\Pi_2} L^*\,dt + c_n\,(q_n^* - q_n^{(1)})\ .$$

On the other hand, only q_n varies along the line P_2P^* , such that the integral will become:

$$\int_{P_1}^{P^*} L\,dt = c_n\,(q_n^* - q_n^{(1)})$$

when it is extended along the line. If one the extends the integral $\int L\,dt$ over the broken path P_1P^* and P^*P_2 then that will give:

$$\varepsilon\int_{\Pi_1}^{\Pi_2} L^*\,dt + c_n\,(q_n^{(2)} - q_n^{(1)})\,.$$

However, that must be greater than the integral over the extremal $P_1\,P_2$, such that:

$$\varepsilon\int_{\Pi_1}^{\Pi_2} L^*\,dt + c_n\,(q_n^{(2)} - q_n^{(1)}) > \mathcal{E}\int_{P_1}^{P_2} L\,dt = e\int_{\Pi_1}^{\Pi_2} L^*\,dt + c_n\,(q_n^{(2)} - q_n^{(1)})\ ,$$

and therefore:

$$\varepsilon\int_{\Pi_1}^{\Pi_2} L^*\,dt > e\int_{\Pi_1}^{\Pi_2} L^*\,dt\,,$$

i.e., the projection e is an extremal of the variational problem (79.a).

([93]) If q_{k+1}, q_{k+2}, …, q_n are cyclic coordinates then one will have $L^* = T^* - \Phi$, in which one sets:

(a) $$T^* = T - c_{k+1}\,\dot{q}_{k+1} - \cdots - c_n\,\dot{q}_n$$

In the most important case in practice, namely, when *the kinetic energy T* in $L = T - \Phi$ *is a quadratic form in the* $\dot{q}_\rho$, the relation (79.c) will yield $\dot{q}_\rho$ as a linear (but *not* homogeneous) expression in the $\dot{q}_1, \ldots, \dot{q}_{n-1}$. If one substitutes the calculated value for $\dot{q}_n$ from (79.b) in:

$$L^* = T - \Phi - \dot{q}_n \dot{c}_n = T^* - \Phi$$

then the function:

$$T^* = T - \dot{q}_n \dot{c}_n$$

will *no longer* be a quadratic form in the $\dot{q}_1, \ldots, \dot{q}_{n-1}$, but a more general quadratic expression in the $\dot{q}_1, \ldots, \dot{q}_{n-1}$ in which both a term that depends upon the velocity components linearly and a term that is *completely free* of them will appear. Obviously, one can combine the term in *T* that is free of the velocity components $\dot{q}_1, \ldots, \dot{q}_{n-1}$ with the function Φ. It will then have the character of a potential. From here on, one can try to interpret the applied forces that act upon the mechanical system and arise from a potential as purely kinetic in origin [for that, cf., IV 1 (**A. Voss**), nos. **27** and **28**]. In order to do that, one must regard the potential of the applied forces as one term in the kinetic energy of the system that originates in an unknown cyclic motion and then try to construct a theory of mechanics in which the concept of force has been eliminated completely. In that sense, **H. Hertz** ([94]) had, in fact, presented the program for explaining potential energy in terms of the cyclic motion of hidden masses that exist along with the visible masses, which is admittedly a program whose implementation has not met with great success. That interpretation has also given rise to the English term *ignored coordinate* (*hidden coordinate*, resp.). What is especially important in that is the case in which the acyclic coordinates change so slowly in time *t* that the associated velocity components in the kinetic energy can be set equal to something close to zero. The name of *cyclic system* has been proposed for such a mechanical system ([95]) since the acyclic

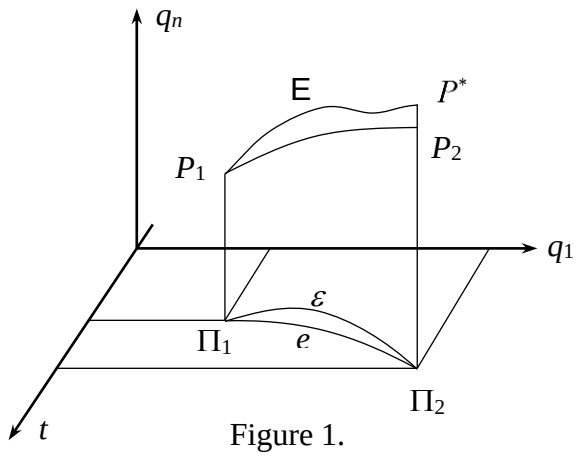

Figure 1.

---

as a generalization of (79.b). The recalculation must proceed here in such a way that one calculates $\dot{q}_{k+1}, \ldots, \dot{q}_n$ from the equations:

$$c_{k+1} - (g_{k+1,1}\dot{q}_1 + \cdots + g_{k+1,k}\ \dot{q}_k) = g_{k+1,k+1}\dot{q}_{k+1} + \cdots + g_{k+1,n}\ \dot{q}_n,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$c_n - (g_{n1}\ \dot{q}_1 + \cdots + g_{nk}\ \dot{q}_k) = g_{n,k+1}\ \dot{q}_{k+1} + \cdots + g_{nn}\ \dot{q}_n,$$

and substitute the values that one finds in the right-hand side of (a). On this, cf., also, **L. Koenigsberger**, "Das Prinzip der verborgenen Bewegung," Heidelberg Sitzungsber. **3** (1912), no. 10.

([94]) **H. Hertz**, *Prinzipien der Mechanik*, Book 2, Sect. 5, II = *Werke III*, pp. 252.

([95]) Cf., **H. Hertz**, *Prinzipien der Mechanik*, Book 2, Sect. 5, I = *Werke III*, pp. 235.

Furthermore, one can also speak of cyclic coordinates when no force function exists, so the equations of motion possess the form:

coordinates then will play the role of only parameters. The *linear terms in* T^* are referred to as *gyroscopic terms* since cyclic motions appear for certain motions of tops in an especially characteristic way, and one can then interpret the terms in the equations of motion that originate in the part of the kinetic energy that is linear in the velocity components in an especially intuitive way ([96]).

However, from the standpoint of this book, *understanding the meaning of the cyclic coordinates* does not lie in the direction that we have in mind. Rather, they are important here due to the fact that they open up the possibility of simplifying the system of equations of motion. *The systematic theory of integration can be oriented towards the model of eliminating the cyclic velocity components*. Knowing an arbitrary first integral can also, in fact, be utilized in a similar manner to knowing the constancy of a cyclic impulse. The simple relationship between a one-parameter group of transformations and the constancy of cyclic impulses can likewise be adapted to the more general cases for a cyclic coordinate. A first integral of the equations of motion and a one-parameter group of transformations that takes the set of space-time lines into itself are mutually implicit (cf., *infra*, no. **25**). As group theory teaches us, every one-parameter group of transformations possesses a normal form in which it appears to be a parallel displacement along a coordinate direction. The coordinate transformation that produces that normal form for the group must then, at the same time, give a form to the system of equations of motion in which the associated coordinate is a cyclic coordinate. In that way, the associated first integral must

---

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_\lambda}\right) - \frac{\partial T}{\partial q_\lambda} = Q_\lambda \qquad (\lambda = 1, \ldots, n),$$

and indeed, one refers to a coordinate q_n as cyclic in this case when q_n does not appear in the kinetic energy T itself, such that the associated equation of motion will read:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_n}\right) = Q_n .$$

If all cyclic impulses (so, e.g., $p_n = \partial T / \partial \dot{q}_n$) remain constant under the motion ($Q_n = 0$) then the motion is called *adiabatic*. By contrast, if the derivatives of the cyclic coordinates (viz., the cyclic velocity components) remain constant then the system will be referred to as *isocyclic*. That terminology was introduced by **Helmholtz**, who attempted to give a mechanical interpretation to the laws of thermodynamics by appealing to cyclic coordinates. [In his studies on the statics of monocyclic systems, cf., ([84]).]

([96]) One ascribes those terms to the system of forces that are in effect, and one correspondingly calls them *gyroscopic reaction forces*, which originate in the hidden top motions, cf., **W. Thomson** and **P. G. Tait**, *Natural Philosophy I*, no. 345[IV], pp. 392. This interpretation of the terms as forces is the basis for the term *kinosthenic coordinate*, moreover.

The following should be pointed out: If one were to replace the derivatives of the cyclic coordinates $\dot{q}_{k+1}, \ldots, \dot{q}_n$ with the constant cyclic impulses $c_{k+1}, \ldots, c_n$ in T itself, rather than T^*, then after recalculation, T would be the sum of two quadratic forms, the first of which is a quadratic form in the velocity components of the acyclic coordinates, and the second of which is a quadratic form in the cyclic impulses $c_{n-k+1}, \ldots, c_n$. Terms that are linear in the $\dot{q}_\rho$ would not appear. Cf., **Thomson-Tait**, *Natural Philosophy I*, no. 319, pp. 322.

Even when the $q_{k+1}, \ldots, q_n$ are not cyclic coordinates, such a replacement of the $\dot{q}_{k+1}, \ldots, \dot{q}_n$ in T with the associated impulses can be advantageous. Cf., **A. B. Bassett**, Quart. J. **38** (1907), pp. 367.

simultaneously assume a form in which it expresses the constancy of the associated cyclic impulse. One next treats the integrals that are linear in the velocity components ([96.a]) (cf., *infra*, no. **29**).

The cyclic coordinates are generally distinguished in a special way for this relationship between groups and integrals of the equations of motion. Namely, if several (say k) cyclic coordinates appear in a mechanical problem and one has, correspondingly, k first integrals then the set of k one-parameter groups of parallel displacements in the directions of the individual cyclic coordinates will define a *k-parameter group*. By contrast, if one has k arbitrary first integrals of the equations of motion then the k one-parameter groups that belong to the individual integrals will not, by any means, determine a k-parameter group, since the necessary conditions [cf., II A 6 (**L. Maurer** and **H. Burkhardt**), no. **5**] do not need to be fulfilled. That will happen only when the k integrals are in involution with each other (cf., *infra*, no. **26**). The first integrals of the equations of motion that are obtained by setting the k cyclic impulses equal to constants will then represent k integrals that lie *in involution*.

10. Jacobi's principle of least action. – If time t does not appear explicitly in the **Lagrangian** L then, from (75.b), it cannot appear in H either. As one sees immediately from the system (76) of canonical equations:

$$H(p_1, \ldots, p_n, q_1, \ldots, q_n) = \text{const.} = k \tag{80}$$

is a *first* integral of the canonical system ([97]), so:

$$L - \dot{q}_1 \frac{\partial L}{\partial \dot{q}_1} - \cdots - \dot{q}_n \frac{\partial L}{\partial \dot{q}_n} = -k \tag{80.a}$$

is a first integral of the **Lagrange** equations of motion (70), and indeed it is the *energy integral*, from the meaning H. The appearance of the energy integral as the first integral of the equations of motion is then connected with the fact that t does not appear explicitly in L (H, resp.), i.e., that t appears in the equations of motion only as *a differential dt*, and therefore the totality of all space-time lines of the motion must go to itself under a "parallel displacement" in the time direction ([98]).

([96.a]) **V. Voronets**, "Transformation of the dynamical equations by means of linear integrals of the equations of motion," Kyiv 1906 (Russian). In particular, for a linear integral, one has a one-parameter group of transformations of only the position coordinates.

([97]) In fact, one has:

$$\frac{dH}{dt} = \frac{\partial H}{\partial p_1}\frac{dp_1}{dt} + \cdots + \frac{\partial H}{\partial p_n}\frac{dp_n}{dt} + \frac{\partial H}{\partial q_1}\frac{dq_1}{dt} + \cdots + \frac{\partial H}{\partial q_n}\frac{dq_n}{dt} = 0\,,$$

as one will see immediately upon appealing to the canonical system.

([98]) That interpretation was first made known when the theory of relativity made it customary to regard the position coordinates and time (the impulse and energy, resp., which amounts to the same thing) as equivalent.

One sees immediately that the fact that the constancy of energy corresponds to the fact that time does not appear explicitly corresponds completely to the constancy of the impulse component that belongs to a cyclic coordinate when

As one sees, the independent variable t plays a role that corresponds to that of the coordinate q_n in the previous section. *The time t is then a type of cyclic coordinate here.* I lay ∞^1 space-time lines on the same cylinder M_2 with generators that are parallel to the t-direction, which cuts out the associated *trajectory* in the M_n of position coordinates $q_1, \ldots, q_n$. Mechanically, the appearance of the energy integral then means that every trajectory can be traversed in ∞^1 ways. Since the same numerical value of the energy constant k belongs to all space-time lines on the same cylinder, *one* numerical value of k will belong to each of the ∞^{2n-1} trajectories such that a fixed value of k will single out ∞^{2n-2} trajectories, and the totality of all trajectories will seem to be subdivided into ∞^1 families of ∞^{2n-2} trajectories with a fixed numerical value of the energy constant.

Now, the ∞^{2n-2} trajectories of one such family are also extremals of a variational problem here that emerges from the variational principle of **Hamilton**'s principle by *eliminating time*. In that way, **Hamilton**'s variational principle for space-time lines will then be juxtaposed with the *variational principle for the trajectories*, namely, the so-called *principle of least action*.

The transition to that new variational problem also takes place here most conveniently by starting from the canonical system (76) ([98.a]). Namely, if one gives the energy integral (80) the form:

$$H^*(p_1,\ldots,p_n,q_1,\ldots,q_n,k) = H(p_1,\ldots,p_n,q_1,\ldots,q_n) - k = 0 \tag{81}$$

then the canonical system (76), which one can now write as:

$$dq_1 : \ldots : dq_n : dp_1 : \ldots : dp_n = \frac{\partial H^*}{\partial p_1} : \cdots : \frac{\partial H^*}{\partial p_n} : -\frac{\partial H^*}{\partial q_1} : \cdots : -\frac{\partial H^*}{\partial q_n}, \tag{82}$$

---

one converts the variational problem of **Hamilton**'s principle, which is represented as a *function problem* in the calculus of variations, into a *parametric problem* by introducing a parameter u :

$$\int L(\dot q_1,\ldots,\dot q_n,q_1,\ldots,q_n)\,dt = \int L\left(\frac{q_1'}{t'},\ldots,\frac{q_n'}{t'},q_1,\ldots,q_n\right)t'\cdot du = \int L\,(q_1',\ldots,q_n',q_1,\ldots,q_n)\,du = \text{extrem.}$$

Since t is a cyclic coordinate in that parameter problem, one has the first integral:

$$\frac{\partial L}{\partial t'} = \text{const.} = -k,$$

or since:

$$\frac{\partial L}{\partial t'} = L(\dot q_1,\ldots,\dot q_n,q_1,\ldots,q_n) - \frac{\partial L}{\partial \dot q_1}\dot q_1 - \cdots - \frac{\partial L}{\partial \dot q_n}\dot q_n,$$

one will have:

$$H = k,$$

from (73).

([98.a]) This was pointed out to the author by **C. Carathéodory**. On this subject, cf., also **F. D. Murnaghan**, "The principle of Maupertuis," Washington Proc. of the Acad. **17** (1931), pp. 128.

can be regarded as the canonical system of a *parametric problem in the calculus of variations* ([99]):

([99]) Since the explanations from the calculus of variations that are required here are not found in the articles II A 8 (**A. Kneser**) and II A 8.a (**E. Zermelo-H. Hahn**), they might be briefly summarized here.

Let:

(1) $$\int f\ (q_1', \ldots, q_n', q_1, \ldots, q_n)\, du\ = \text{extrem.} \qquad \left(q_\rho' = \frac{dq_\rho}{du} \right),$$

in which f is homogeneous of degree one in the q_ρ', so:

(2) $$f = \frac{\partial f}{\partial q_1'} q_1' + \cdots + \frac{\partial f}{\partial q_n'} q_n' \ ,$$

be a parametric problem in the calculus of variations with the **Euler** equations:

(3) $$\frac{d}{du}\left(\frac{\partial f}{\partial q_1'} \right) - \frac{\partial f}{\partial q_1} = 0 \qquad (\rho = 1, \ldots, n)$$

(which are not mutually independent). If one sets:

(4) $$p_1 = \frac{\partial f}{\partial q_1'}, \ \ldots, \ p_n = \frac{\partial f}{\partial q_n'}$$

here, then since the right-hand sides are homogeneous of degree zero in the $q_1', \ldots, q_n'$, one can eliminate the $q_1', \ldots, q_n'$, and one will get a relation:

(5) $$H\,(p_1, \ldots, p_n, q_1, \ldots, q_n) = 0$$

(in which the form of the function H is not completely fixed).

Since:

$$\frac{\partial H}{\partial p_1}\delta p_1 + \cdots + \frac{\partial H}{\partial p_n}\delta p_n + \frac{\partial H}{\partial q_1}\delta q_1 + \cdots + \frac{\partial H}{\partial q_n}\delta q_n = 0$$

is then equivalent to these equations, which follow from (2) and (4):

$$q_1'\,\delta p_1 + \cdots + q_n'\,\delta p_n - \frac{\partial f}{\partial q_1}\delta q_1 + \cdots + \frac{\partial f}{\partial q_n}\delta q_n = 0\ .$$

Therefore, if λ is a proportionality factor then the following relations will be true:

(6) $$\begin{cases} q_1' = \lambda \dfrac{\partial H}{\partial p_1}, & \ldots, & q_n' = \lambda \dfrac{\partial H}{\partial p_n}, \\ -\dfrac{\partial f}{\partial q_1} = \lambda \dfrac{\partial H}{\partial q_1}, & \ldots, & -\dfrac{\partial f}{\partial q_n} = \lambda \dfrac{\partial H}{\partial q_n}, \end{cases}$$

which was first found (in essence) by **W. R. Hamilton** (cf., *infra*, no. **13**). With that, the **Euler** equations (3) can be converted into the *canonical system:*

(83) $$\int f(q_1',\ldots,q_n',q_1,\ldots,q_n)\,du = \text{extrem.} \qquad \left(q_\rho' = \frac{dq_\rho}{du}\right).$$

One then gets the integrand *f* when one next sets:

(84) $$\begin{cases} q_1' = \lambda \dfrac{\partial H^*}{\partial p_1} = \lambda \sum\limits_{\sigma=1}^{n} g^{1\sigma}(p_\sigma - g_{0\sigma}), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ q_n' = \lambda \dfrac{\partial H^*}{\partial p_n} = \lambda \sum\limits_{\sigma=1}^{n} g^{n\sigma}(p_\sigma - g_{0\sigma}), \end{cases}$$

and calculates:

(84.a) $$p_\sigma = g_{0\sigma} + \frac{1}{\lambda}\sum_{\tau=1}^{n} g_{\sigma\tau}\, q_\tau'$$

from that. If one substitutes those values in $H^* = 0$ then one will find that:

$$\frac{1}{2}\frac{1}{\lambda^2}\sum_{\rho,\sigma=1}^{n} g_{\rho\sigma}\, q_\rho'\, q_\sigma' + \left(\Phi - \tfrac{1}{2}\, g_{00}\right) - k = 0\,,$$

or

(84.b) $$\frac{1}{\lambda} = \sqrt{\frac{2\,(k - (\Phi - \frac{1}{2}\, g_{00}))}{\sum\limits_{\rho,\sigma} g_{\rho\sigma}\, q_\rho'\, q_\sigma'}}\,.$$

One can then get *f* from that by means of:

$$f = p_1\, q_1' + \cdots + p_n\, q_n',$$

on the basis of (84.a):

---

(7) $$\begin{cases} q_1' = \lambda \dfrac{\partial H}{\partial p_1}, & \ldots, \quad q_n' = \lambda \dfrac{\partial H}{\partial p_n}, \\ p_1' = -\lambda \dfrac{\partial H}{\partial q_1}, & \ldots, \quad p_n' = -\lambda \dfrac{\partial H}{\partial q_n}. \end{cases}$$

If one has, conversely, a canonical system in the form (7) with the relation (5) then it will belong to a parametric problem in the calculus of variations of the form (1), which one will prove in the following way: One determines the dependency of the p_ρ on the $q_1, \ldots, q_n, , q_1', \ldots, q_n'$ and λ from the first group of equations in (7), and then determines λ by substituting the values of p_ρ that one finds into the relation (5). One correspondingly puts *f* into the form:

(8) $$f(q_1',\ldots,q_n',q_1,\ldots,q_n) = p_1\, q_1' + \ldots + p_n\, q_n',$$

corresponding to (2).

$$f = \sum_{\sigma=1}^{n} g_{0\sigma}\, q'_\sigma + \frac{1}{\lambda} \sum_{\sigma,\tau=1}^{n} g_{\sigma\tau}\, q'_\sigma\, q'_\tau \,,$$

or furthermore, from (84.b):

(84.c) $$f = \sum_{\sigma=1}^{n} g_{0\sigma}\, q'_\sigma + \sqrt{2(k-(\Phi-\tfrac{1}{2}g_{00}))\sum_{\rho,\sigma} g_{\rho\sigma}\, q'_\rho\, q'_\sigma} \,.$$

One will then find that the *principle of least action* is the variational problem of:

(85) $$\int \left\{ \sum_{\sigma=1}^{n} g_{0\sigma}\, dq_\sigma + \sqrt{2(k-(\Phi-\tfrac{1}{2}g_{00}))\sum_{\rho,\sigma} g_{\rho\sigma}\, dq_\rho\, dq_\sigma} \right\} = \text{extrem.}$$

whose extremals are the trajectories. In it, the g_{00}, $g_{0\rho}$, and $g_{\rho\sigma}$ are functions of only the position coordinates $q_1, \ldots, q_n$, and are independent of time t.

If the kinetic energy T is a quadratic form in the $\dot{q}_\rho$, in particular, so the M_n of the coordinates $q_1, \ldots, q_n$ becomes a **Riemannian** space with the arc-length element:

$$ds^2 = \sum_{\rho,\sigma} g_{\rho\sigma}\, dq_\rho\, dq_\sigma \,,$$

then this variational problem will take the form:

(86) $$\int \sqrt{2(k-\Phi)}\, ds = \text{extrem.}$$

On the basis of that energy integral, the individual points of the trajectory that proves to be the extremal of the variational problem (85) [(86), resp.] can be associated with the time t by way of the differential relation:

(87) $$dt = \sqrt{\frac{\sum_{\rho,\sigma} g_{\rho\sigma}\, dq_\rho\, dq_\sigma}{2(k-(\Phi-\tfrac{1}{2}g_{00}))}} \,,$$

while for (86) it is given by ([99.a]):

(87.a) $$dt = \frac{ds}{\sqrt{2(k-\Phi)}} \,.$$

([99.a]) Here as well, an argument that is analogous to the one in (92) will show directly that the projections of the extremals of the **Hamilton**'s principle onto the spatial M_n of the $q_1, \ldots, q_n$ are the extremals of the variational problem (85) [(86), resp.].

The variational problem for the trajectories is the *principle of least action* in the form that **C. G. J. Jacobi** obtained it in ([100]) [cf., IV 1 (**A. Voss**), no. **44**]. Historically, the principle was first expressed by **M. de Maupertuis**, if only in a very unclear (if not outright false) form. [For the historical development, cf., IV 1 (**A. Voss**), nos. **43** and **44**]. **L. Euler** then first gave it a precise formulation by way of an example: The space-time curves of the motion are obtained as the extremals of the variational problem:

$$\int 2T\,dt = \text{extrem.}\,, \tag{88}$$

for which the constancy of the energy:

$$T + \Phi = k \tag{88.a}$$

is prescribed. In so doing, the lower limit of the integral is a given space-time point $P_1(q_1^{(1)}, \ldots, q_n^{(1)}, t_1)$, while only the position coordinates $q_1^{(2)}, \ldots, q_n^{(2)}$ are given for the upper limit P_2, and the associated time t_2 will first be determined by the extremum requirement itself ([101]). One can therefore refer to that formulation as *Euler's principle*. **Lagrange** ([102]) then generalized the **Euler** formulation for arbitrary holonomic and scleronomic mechanical systems ([103]). In the **Jacobi** formulation, it is almost self-explanatory that the principle can yield trajectories only when the *auxiliary conditions*, which perhaps come about as a result of the constraints on the system, are *holonomic and scleronomic*. Therefore, it would seem appropriate to also exclude non-holonomic constraints for the **Euler** form of the principle, just as one does with **Hamilton**'s principle, and allow only those auxiliary conditions that are holonomic, as well as scleronomic. That is because for a variation in the sense of classical calculus of variations, one would not get the correct equations of motion from the principle with auxiliary conditions that are indeed holonomic, but rheonomic. By contrast, in the literature, there are those who have also attempted to express that principle in the broadest-possible domain of application ([104]). In order to do that, the type of

([100]) Cf., **C. G. J. Jacobi**, *Vorlesungen*, 6. Vorlesung = *Werke*, Supp.-Bd., pp. 43. Using that *Jacobi form of the principle*, the problem of determining the trajectories in a mechanical problem will run parallel to that of determining the *geodetic lines* in a manifold with a general arc-length element. **Jacobi** has already pursued that parallel, which has made each advance in the one domain just as useful in the other, to worthwhile effect. It was systematically utilized later by **P. Stäckel** and others. Cf., also **J. L. Synge**, "On the geometry of dynamics," Trans. London Phil. Soc. (A) **226** (1927), pp. 31, esp., Chap. 4.

([101]) **L. Euler**, *Methodus inveniendi lineas curvas*…, Lausanne 1744, additamentum II.

H. von Helmholtz, "Zur Geschichte des Prinzips der kleinsten Aktion." Berlin Sitz. der Akad. (1887), pp. 225 = *Wissensch. Abhandl. III*, pp. 249 has proved that, in general, the **Euler** form of the principle will yield the equations of motion by the methods of the calculus of variations.

([102]) **J. L. Lagrange**, *Mécanique analytique*, 2. part. sect. III, **§ 6** = *Œuvres XI*, pp. 315.

([103]) In the literature, the principle is occasionally referred to as ***Lagrange's*** *principle*. The term *principle of Maupertuis* is also very widespread.

([104]) Cf., above all, **O. Hölder**, "Über die Prinzipien von Hamilton und Maupertuis," Gött. Nachr. (1896), pp. 122, as well as **Ph. E. B. Jourdain**, "On those principles of mechanics which depend upon processes of variation," Math. Ann. **65** (1908), pp. 513. **A. Voss**, "Über die Prinzipe von Hamilton und Maupertuis," Gött. Nachr. (1900), pp. 322, has also adapted the **Hölder** approach to general coordinates. A somewhat-different formulation was given by **H.**

variation that is appropriate to the variational problem must once more be replaced with another one that one establishes by a special convention. Basically, one naturally does nothing but replace the variational formula of the variational principle with the integrated central equation ([105]) and then employ the variations as the virtual displacements.

Brell, Wien Sitzungber. **122** (1913) II$^{a\,1}$, pp. 1031. One can also confer **H. Brell** and **E. Schenkel**, "Über die Prinzipien von Hamilton und Maupertuis," Verhandl. der Deutsch. phys. Ges. **15** (1913), pp. 1082.

A. Voss generalized those ideas by giving a general form to the variational problem:

$$\int (\alpha T + \beta U)\, dt \;= \text{extrem.}$$

and sought to derive the various forms that it would take, like **Hamilton**'s principle and **Euler**'s principle. Cf., **A. Voss**, "Bemerkungen über die Prinzipien der Mechanik," München Sitzungsber. **31** (1901), pp. 167. **R. Leitinger**, "Über Jourdains Prinzip der Mechanik und dessen Zusammenhang with der kleinsten Aktion." Wien Sitzungsber. **122** (1913) II$^{a\,1}$, pp. 635, would even like to arrive at **Euler** principle by integrating the differential formulas of **Jourdain**'s principle, whereby the definition of the variation must naturally be made heuristically. Similar arguments that started from the principle of least constraint were presented by **H. Brell**, Wien Sitzungber. **122** (1913) II$^{a\,1}$, pp. 933.

([105]) Naturally, in a form that takes the constancy of energy into account.

CHAPTER III

PRELIMINARY ANSÄTZE FOR THE GENERAL THEORY OF INTEGRATION.

11. Introductory remarks. – One finds the course of motion for a mechanical system from the differential equations of motion by integrating them, whereas examining the bases for obtaining them was the subject of the previous sections. Historically, the systematic theory of that integration has been constructed by treating examples, and indeed, there were basically only two problems that were treated repeatedly in the research, namely, the "celestial" mechanics of the *n-body problem* [cf., VI 2, 12 (**E. T. Whittaker**)] and the "terrestrial" mechanics of the *motion of rigid bodies* (i.e., tops). Of the two, the *n*-body problem has had a significantly greater influence on the general theory of integrating the equations of motion than the motion of rigid bodies. The beginnings of one such theory of integration are almost identical to the perturbation calculations of the astronomers, whose foundations were laid by **L. Euler**, **J. L. Lagrange**, **P. S. Laplace**, **S. D. Poisson** ([106]). **W. R. Hamilton** (from 1824 on) was the one we have to thank for showing us that defining the source of the equations of motion in the form of a variational problem would lead to some special simplifications in its analytical treatment. However, it was not his research in mechanics, but in *ray optics*, that inspired his new ideas. In the *characteristic function*, he discovered the suitable tool for tracing an entire ray system that was emitted from a luminous point through an optical instrument. He then soon recognized that the existence of the characteristic function would depend upon only the fact that the phenomena of ray optics could be summarized mathematically in *Fermat's principle of the shortest light-path*, and therefore (he concluded) for every realm of phenomena that was based upon a variational problem, the adaptation of the concept of the characteristic function to it must bring with it some simplifications that are analogous to the ones that were achieved in geometrical optics. In particular, he applied those ideas to the differential equations of mechanics ([107]), and in so doing developed a *new form for the theory of perturbations* in astronomy.

Hamilton's ideas were taken up in Germany by **C. G. J. Jacobi**, but of course, not in their original form. Rather, he was following up on his own preconceived notions that he had formed in his investigations into the *integration of first-order partial differential equations*. From there, the genesis of a systematic theory of integration for the differential equations of mechanics arose that was further built up by the **Jacobi** school. In so doing, the work of **Lagrange** and **Poisson** on the theory of perturbations had considerable influence. Some historical remarks on the work of **Lagrange** and **Poisson**, as well as **Hamilton**, are thus necessary for a deeper understanding of the

([106]) For a detailed report on those beginnings of a theory of integration, cf., **A. Cayley**, "Report on the recent progress of theoretical dynamics," Report on the British Assoc. for the Advancement of Science (1857), pp. 1 = *Coll. Papers III*, pp. 156. Cf., also **E. O. Lovett**, "The theory of perturbations and Lie's theory of contact transformations," Quart. J. of pure and appl. Math. **30** (1899), pp. 47.

([107]) The adaptation to mechanics was especially obvious when one stood on the base of **Newton**'s emissive optics, which researchers in England had long sought to establish in opposition to **Huyghen**'s wave theory.

conceptual structure of the general theory of integration of the equations of motion that the **Jacobi** school had built up.

12. The variation of constants due to Lagrange and Poisson. – The motion of the individual planets around the Sun (two-body problem) can be reduced to the motion of a mass-point about a fixed attracting point, and as such, they can be described by the equations of motion of a one-body problem:

$$(89)\qquad \begin{cases} \ddot{x} - \dfrac{1+m}{r^3}x = 0, \\ \ddot{y} - \dfrac{1+m}{r^3}y = 0, \qquad (r = \sqrt{x^2+y^2+z^2}\,) \\ \ddot{z} - \dfrac{1+m}{r^3}z = 0 \end{cases}$$

[cf., VI 2, 15 (**Karl F. Sundman**), no. **3**]. When one considers the influence of the other planets, a more general potential function $\Phi\,(x, y, z, t)$ will enter in place of the potential:

$$(89.a)\qquad \Phi^* = \frac{1+m}{r}\ .$$

However, the effect of the other planets is relatively minor in comparison to the effect of the Sun, such that the potential Φ can be split into the "unperturbed" motion and a *perturbing function* $\Omega\,(x, y, z, t)$:

$$(89.b)\qquad \Phi\,(x, y, z, t) = \frac{1+m}{r} - \Omega\,(x, y, z, t)\ .$$

The equations of motion for the actual ("perturbed") motion will then have the form:

$$(89.c)\qquad \ddot{x} - \frac{1+m}{r^3}x = \frac{\partial\Omega}{\partial x}, \qquad \ddot{y} - \frac{1+m}{r^3}y = \frac{\partial\Omega}{\partial y}, \qquad \ddot{z} - \frac{1+m}{r^3}z = \frac{\partial\Omega}{\partial z}$$

[cf., VI 2, 15 (**Karl F. Sundman**), no. **2**]. **Lagrange** was the next to address them ([108]). Meanwhile, he could also adapt his arguments to the more general problem in which the equations of motion in unperturbed form possessed the general form:

$$(90)\qquad \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\rho}\right) - \frac{\partial L}{\partial q_\rho} = 0 \qquad (\rho = 1, \ldots, n)$$

([108]) **J. L. Lagrange**, "Mémoires sur la théorie des variations des éléments des planètes et en particulier des variations des grands axes de leurs orbites," Paris Mém. de l'inst. (1808), pp. 1 = *Œuvres VI*, pp. 711.

so the perturbation equations will become ([109]):

(90.a) $$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\rho}\right) - \frac{\partial L}{\partial q_\rho} = \frac{\partial \Omega}{\partial q_\rho} \qquad (\rho = 1, \ldots, n).$$

Perturbation theory assumes that equations (90) are solved for the unperturbed problem, and indeed **Lagrange** assumed that the q_ρ were functions of t and determined from $2n$ integration constants ([110]):

(91) $$q_\rho = q_\rho\,(t, c_1, \ldots, c_{2n}) \qquad (\rho = 1, \ldots, n).$$

At the same time, the velocity components $\dot{q}_\rho$ (the associated impulse components p_ρ, resp.) will be functions of t and the $2n$ constants $c_1, \ldots, c_{2n}$:

(91.a) $$p_\rho = p_\rho\,(t, c_1, \ldots, c_{2n}) \qquad (\rho = 1, \ldots, n).$$

Now, in order to solve the equations (90.a) of the perturbed problems, **Lagrange**, learning from the method *variation of constants* that had already been employed by **L. Euler**, imagined that the $2n$ integration constants $c_1, \ldots, c_{2n}$ were (slowly-varying) functions of time and sought to ascertain them in such a way that equations (90.a) would be satisfied by the perturbed motion if one had substituted (91) ([111]). However, one will get only n equations for the $2n$ unknown functions $c_1, \ldots, c_{2n}$ in that way. For that reason, in order to fix the c_ρ completely, **Lagrange** added the condition that the velocity of the unperturbed motion and that of the perturbed motion should coincide at every moment t ([112]), which is a condition that one might also care to formulate in the law that the unperturbed motion shall be an *osculating motion* ([113]) to the perturbed motion. It will lead to the equations:

$$\frac{dq_\rho}{dt} = \frac{\partial q_\rho}{\partial t}$$

or

(92) $$\frac{\partial q_\rho}{\partial c_1}\frac{dc_1}{dt} + \cdots + \frac{\partial q_\rho}{\partial c_{2n}}\frac{dc_{2n}}{dt} = 0 \qquad (\rho = 1, \ldots, n)\,.$$

([109]) **J. L. Lagrange**, "Mémoires sur la théorie générale de la variation des constantes arbitrarires dans tous les problèmes de la mécanique," Paris Mém. de l'inst. (1809), pp. 257 = *Œuvres VI*, pp. 769.

([110]) For a planet, e.g., the six elliptic elements might be such a system of integration constants [cf., VI 2, 15 (**Karl F. Sundman**), no. **6**].

([111]) On this subject, cf., VI 2, 15 (**Karl F. Sundman**), no. **5**.

([112]) I. e., the $\dot{q}_\rho$ shall take the same numerical values at every moment t, regardless of whether one regards the $c_1, \ldots, c_{2n}$ in (91) as functions of time or constants in the differentiation with respect to t, to the extent that one only assigns those numerical values to the c_λ that belong to the value of t in question.

([113]) Cf., VI 2, 15 (**Karl F. Sundman**), no. **6**.

On the other hand, when one introduces the impulses $p_\rho = \partial L / \partial \dot{q}_\rho$ into equations (90) and (90.a), one will get:

$$\frac{dp_\rho}{dt} - \frac{\partial p_\rho}{\partial t} = \frac{\partial \Omega}{\partial q_\rho} \ ,$$

since, in fact, the $\partial L / \partial q_\rho$ will have the same values for the perturbed and unperturbed motions when the $c_1, \ldots, c_{2n}$ are osculating elements. Thus, the n equations:

$$\frac{\partial p_\rho}{\partial c_1}\frac{dc_1}{dt} + \cdots + \frac{\partial p_\rho}{\partial c_{2n}}\frac{dc_{2n}}{dt} = \frac{\partial \Omega}{\partial q_\rho} \qquad (\rho = 1, \ldots, n) \tag{93}$$

will be added to equations (92), and together with (92), they will represent a system of $2n$ first-order differential equations for the $2n$ unknown functions $c_1(t), \ldots, c_{2n}(t)$.

From (91) and (91.a), the factors that multiply the derivatives dc_λ / dt are functions of time t and the $c_1, \ldots, c_{2n}$. Likewise, the right-hand sides of (93), in which one imagines that the expressions (91) have been substituted for the q_ρ, are also functions of t and the $c_1, \ldots, c_{2n}$. One can imagine that the expressions (91) have been substituted for the q_ρ directly in the perturbation function $\Omega(q_1, \ldots, q_n, t)$, which will make it into a function of time t and the $c_1, \ldots, c_{2n}$:

$$\Omega = \Omega(t, c_1, \ldots, c_{2n}) \ . \tag{94}$$

However, it will then be convenient to introduce the derivatives of the perturbation functions (94) with respect to the $c_1, \ldots, c_{2n}$ into the "perturbation equations" (92) and (93). In order to do that, **Lagrange** multiplied the individual equations (93) by $\partial q_\rho / \partial c_\lambda$ and subtracted from them the corresponding terms in (92) multiplied by $\partial p_\rho / \partial c_\lambda$, by which, he then got:

$$\left(\frac{\partial p_\rho}{\partial c_1}\frac{\partial q_\rho}{\partial c_\lambda} - \frac{\partial q_\rho}{\partial c_1}\frac{\partial p_\rho}{\partial c_\lambda}\right)\frac{dc_1}{dt} + \cdots + \left(\frac{\partial p_\rho}{\partial c_{2n}}\frac{\partial q_\rho}{\partial c_\lambda} - \frac{\partial q_\rho}{\partial c_{2n}}\frac{\partial p_\rho}{\partial c_\lambda}\right)\frac{dc_{2n}}{dt} = \frac{\partial \Omega}{\partial q_\rho}\frac{\partial q_\rho}{\partial c_\lambda} \ .$$

Upon summing over ρ from 1 to n, one will then get the system of $2n$ so-called *Lagrangian perturbation equations:*

$$[c_1, c_\lambda]\frac{dc_1}{dt} + \cdots + [c_{2n}, c_\lambda]\frac{dc_{2n}}{dt} = \frac{\partial \Omega}{\partial c_\lambda} \qquad (\lambda = 1, \ldots, 2n) \ . \tag{95}$$

The $[c_\mu, c_\lambda]$ are the so-called *Lagrange brackets* ([114]), which are defined by:

([114]) Those constructions, which first appeared in perturbation calculations in the way that was given in the text, have since then taken on great significance in the theory of first-order partial differential equations and ancillary theories in analysis [cf., also II A 5 (**E. von Weber**)].

(96) $$[c_\mu, c_\lambda] = \sum_{\rho=1}^{n} \left(\frac{\partial p_\rho}{\partial c_\mu} \frac{\partial q_\rho}{\partial c_\lambda} - \frac{\partial q_\rho}{\partial c_\mu} \frac{\partial p_\rho}{\partial c_\lambda} \right),$$

and with that definition, they define a *skew-symmetric system:*

(96.a) $$[c_\mu, c_\lambda] = -[c_\lambda, c_\mu], \quad \text{so} \quad [c_\lambda, c_\lambda] \equiv 0,$$

such the terms in the principal diagonal will be missing from the perturbation equations (95). In general, one should expect that these **Lagrange** brackets will be functions of $c_1, \ldots, c_{2n}$, and time t. Meanwhile, as **Lagrange** could further show by calculation, *time t does not appear explicitly* in them ([115]); they are functions of only the $c_1, \ldots, c_{2n}$.

Almost simultaneously with **Lagrange**, **S. D. Poisson** presented his *perturbation formulas* in a different way ([116]), and they were, *to some extent, reciprocal to Lagrange's perturbation formulas*. Whereas **Lagrange** determined the *solutions* to equations (90) for the unperturbed problems, i.e., he thought of representing the q_ρ, p_ρ as functions of time t and the $2n$ integration constants $c_1, \ldots, c_{2n}$, **Poisson** started from a system of $2n$ *first integrals* of the unperturbed equations of motion (90), which he imagined to be posed in the form ([117]):

(97) $$\varphi_\rho(q_1, \ldots, q_n, p_1, \ldots, p_n, t) = c_\rho \qquad (\rho = 1, \ldots, 2n).$$

Now, in order to integrate the perturbation equations, **Poisson**, like **Lagrange**, regarded the constants $c_1, \ldots, c_{2n}$ of the unperturbed motion as slowly-varying functions of time. With that assumption, it will follow from (97) that:

(98) $$\frac{dc_\rho}{dt} = \frac{\partial \varphi_\rho}{\partial q_1}\frac{dq_1}{dt} + \cdots + \frac{\partial \varphi_\rho}{\partial q_n}\frac{dq_n}{dt} + \frac{\partial \varphi_\rho}{\partial p_1}\frac{dp_1}{dt} + \cdots + \frac{\partial \varphi_\rho}{\partial p_n}\frac{dp_n}{dt} + \frac{\partial \varphi_\rho}{\partial t},$$

while, on the other hand, it will emerge from (97) that one has:

(98.a) $$0 = \frac{\partial \varphi_\rho}{\partial q_1}\frac{\delta q_1}{\delta t} + \cdots + \frac{\partial \varphi_\rho}{\partial q_n}\frac{\delta q_n}{\delta t} + \frac{\partial \varphi_\rho}{\partial p_1}\frac{\delta p_1}{\delta t} + \cdots + \frac{\partial \varphi_\rho}{\partial p_n}\frac{\delta p_n}{\delta t} + \frac{\partial \varphi_\rho}{\partial t}$$

for the unperturbed motion ([118]). If one demands that the constants $c_1, \ldots, c_{2n}$ should be *osculating constants* then one would need to have:

([115]) The intrinsic reason for this will be given below. (Cf., no. **21**)

([116]) **S. D. Poisson**, "Mémoire sur la variation des constantes arbitraires dans les questions de mécanique," J. Éc. Polyt. **8** (1809), pp. 266.

([117]) One can obtain them as solutions to (91), (91.a) when one solves the $2n$ functions p_ρ, q_ρ for the $2n$ constants $c_1, \ldots, c_{2n}$.

([118]) Here, the derivatives with respect to time while the values of c_λ are held constant are distinguished from the derivatives with varying c_λ at one instant by δ.

$$\frac{\delta q_\rho}{\delta t} = \frac{dq_\rho}{dt},$$

while:

$$\frac{dp_\rho}{dt} - \frac{\delta p_\rho}{\delta t} = \frac{\partial \Omega}{\partial q_\rho}.$$

Thus, subtracting (98) and (98.a) will yield:

(98.b)
$$\begin{cases} \dfrac{dc_\rho}{dt} = \dfrac{\partial \varphi_\rho}{\partial p_1}\dfrac{\partial \Omega}{\partial q_1} + \cdots + \dfrac{\partial \varphi_\rho}{\partial p_n}\dfrac{\partial \Omega}{\partial q_n} = \displaystyle\sum_{\lambda=1}^{n} \frac{\partial c_\rho}{\partial p_\lambda}\frac{\partial \Omega}{\partial q_\lambda} = \sum_{\lambda=1}^{n}\left(\frac{\partial c_\rho}{\partial p_\lambda}\cdot\sum_{\mu=1}^{2n}\frac{\partial \Omega}{\partial c_\mu}\frac{\partial c_\rho}{\partial q_\lambda}\right) \\ \qquad = \displaystyle\sum_{\mu=1}^{2n}\left[\frac{\partial \Omega}{\partial c_\mu}\left(\sum_{\lambda=1}^{n}\frac{\partial c_\rho}{\partial p_\lambda}\frac{\partial c_\mu}{\partial q_\lambda}\right)\right]. \end{cases}$$

On the other hand, since the perturbing function is independent of the impulse components, one will have:

$$0 = \frac{\partial \Omega}{\partial c_1}\frac{\partial c_1}{\partial p_\lambda} + \cdots + \frac{\partial \Omega}{\partial c_{2n}}\frac{\partial c_{2n}}{\partial p_\lambda},$$

or when one multiplies by $\partial c_\rho / \partial q_\lambda$ and sums over λ :

$$0 = \sum_{\mu=1}^{2n}\frac{\partial \Omega}{\partial c_\mu}\left(\sum_{\lambda=1}^{n}\frac{\partial c_\rho}{\partial q_\lambda}\frac{\partial c_\mu}{\partial p_\lambda}\right).$$

If one subtracts that from (98.b) then that will yield the system of so-called *Poisson perturbation equations:*

(99)
$$\frac{dc_\rho}{dt} = (c_\rho, c_1)\frac{\partial \Omega}{\partial c_1} + \cdots + (c_\rho, c_{2n})\frac{\partial \Omega}{\partial c_{2n}} \qquad (\rho = 1, \ldots, 2n).$$

The *Poisson brackets* (c_ρ, c_μ) in that are defined by:

(100)
$$(c_\rho, c_\mu) = \sum_{\lambda=1}^{n}\left(\frac{\partial c_\rho}{\partial p_\lambda}\frac{\partial c_\mu}{\partial q_\lambda} - \frac{\partial c_\rho}{\partial q_\lambda}\frac{\partial c_\mu}{\partial p_\lambda}\right) = \left(\frac{\partial c_\rho}{\partial p_1}\frac{\partial c_\mu}{\partial q_1} - \frac{\partial c_\rho}{\partial q_1}\frac{\partial c_\mu}{\partial p_1}\right) + \cdots + \left(\frac{\partial c_\rho}{\partial p_n}\frac{\partial c_\mu}{\partial q_n} - \frac{\partial c_\rho}{\partial q_n}\frac{\partial c_\mu}{\partial p_n}\right).$$

Like the **Lagrange** brackets, the **Poisson** brackets also define a *skew-symmetric system* with that definition:

(101.a)
$$(c_\rho, c_\mu) = -(c_\mu, c_\rho), \quad \text{so} \qquad (c_\rho, c_\rho) \equiv 0,$$

such that the terms in the principal diagonal will also be missing from the **Poisson** perturbation equations. Similarly, as **Poisson** could show by calculation, the **Poisson** brackets have the property that they are functions of time t and the $c_1, \ldots, c_{2n}$, as one might expect, but time t *does not* appear in them explicitly ([119]).

Just as formulas (91), (91.a), on the one hand, and (97), on the other, which together determine the unperturbed motion, are solutions to each other, the *Poisson perturbation equations* (100) *are also the solutions to the Lagrangian perturbation equations* (95), *and conversely.* It then follows that *the relations:*

$$\sum_{\sigma=1}^{2n} [c_\sigma, c_\lambda](c_\sigma, c_\mu) = \begin{cases} 0 & (\lambda \neq \mu), \\ 1 & (\lambda = \mu) \end{cases} \tag{102}$$

must exist between the Lagrange and Poisson brackets.

Lagrange soon remarked ([120]) that for a special choice of the constants $c_1, \ldots, c_{2n}$ (e.g., when he chose them to be the initial values of the position coordinates $q_1, \ldots, q_n$, and that of the impulse components $p_1, \ldots, p_n$), all of the bracket expressions in each of the perturbation equations can vanish, except for one of them that will assume the value + 1 or – 1, and the perturbation equations will take the form:

$$\frac{d\alpha_\lambda}{dt} = \frac{\partial \Omega}{\partial \beta_\lambda}, \qquad \frac{d\beta_\lambda}{dt} = -\frac{\partial \Omega}{\partial \alpha_\lambda}, \tag{103}$$

when one introduces such constants $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$, which is then the *canonical form* ([121]). Nevertheless, even though that form for the perturbation equations seemed remarkable to him, he did not succeed in putting the general equations of motion into canonical form. Thus, he had overlooked the fact that he had basically posed the "problem of the canonical transformation of the equations of motion" here (cf., *infra*, no. **31**, *et seq.*).

13. W. R. Hamilton's investigations into geometrical optics ([122]). – The arguments by which **W. R. Hamilton** blazed new trails into the analytical treatment of the differential equations of motion arose from *geometrical optics* (cf., no. **11**), and indeed he was already basically toying with

([119]) That is ***Poisson**'s theorem*, whose significance for the theory of integration was first emphasized by **Jacobi** (cf., *infra*, no. **26**).

([120]) Cf., **J. L. Lagrange**, "Second mémoire sur la théorie de la variation des constantes arbitraires dans les problèmes de mécanique, dans lequel on simplifie l'application des formules générales à ces problèmes," Paris Mémoires de l'inst. (1809), pp. 343 = *Œuvres VI*, pp. 807.

([121]) Cf., ([90]).

([122]) **W. R. Hamilton**, "Essay on the theory of systems of rays," Dublin Trans. R. Irish Acad. **15** (1828), pp. 69 = *Papers I*, pp. 1, as well as "Supplement to an essay…," Dublin Trans. R. Irish Acad. **16**[1] (1830), pp. 1 = *Papers I*, pp. 107, "Second supplement to an essay…," Dublin Trans. R. Irish Acad. **16**[2] (1831), pp. 93 = *Papers I*, pp. 145, "Third supplement to an essay…," Dublin Trans. R. Irish Acad. **17** (1837), pp. 1 = *Papers I*, pp. 164. The first thing to precipitate from his ideas was the treatise "On caustics," which was submitted to the Royal Society of Ireland in its nineteenth year 1824. It was not printed at the time but was first published more recently as a supplement to the first volume of his works (**W. R. Hamilton**, *Papers I*, pp. 345), which also included all of **Hamilton**'s other papers on ray optics.

the idea of fusing the emissive and undulatory optics [cf., V 21 (**A. Wangerin**), no. **1**] into a greater whole ([123]), which has now (a hundred years later) been realized by wave mechanics. In order to do that, in the bundles of light rays that are emitted from a luminous point and pass through an optical system, *along with the light rays*, he always considered *the wave surfaces of the bundle* and observed the relationships between rays and wave surfaces. Now, since one prefers to summarize the laws of light propagation in *Fermat's principle of the shortest light path* ([124]), he was then led to regard the light rays as the extremals of the variational problem:

$$\int n\,ds = \text{extrem.} \tag{104}$$

The n in that is the index of refraction, and in an anisotropic, inhomogeneous medium, it will be a function of the direction and position:

$$n = n\,(\alpha, \beta, \gamma, x, y, z) \qquad \left(\alpha = \frac{dx}{ds}, \beta = \frac{dy}{ds}, \gamma = \frac{dz}{ds}\right), \tag{104.a}$$

such that the *light rays are the integral curves of the Euler equations for the variational problem* (104):

$$\frac{d}{ds}\left(\frac{\partial n}{\partial \alpha}\right) - \frac{\partial n}{\partial x} = 0\,, \quad \frac{d}{ds}\left(\frac{\partial n}{\partial \beta}\right) - \frac{\partial n}{\partial y} = 0\,, \quad \frac{d}{ds}\left(\frac{\partial n}{\partial \gamma}\right) - \frac{\partial n}{\partial z} = 0\,. \tag{104.b}$$

That function n simplifies considerably in the individual cases. In particular, the index of refraction n will be *a constant* ([125]) in a homogeneous isotropic medium, such that the extremals will be straight lines ([126]). In a single bundle of light rays that is emitted from a luminous point, and is therefore initially homocentric, but will no longer remain homocentric when it passes through the optical instrument (as a result of reflections or refractions, resp.), **Hamilton** could now find the wave surfaces from the variational problem (104) without appealing to the undulatory theory of light. Namely, one will get the *individual wave surfaces* when one determines the point on each light ray that delimits a prescribed *light-path length* with the luminous point and can be

---

([123]) Cf., **E. Schrödinger**, "Quantisierung als Eigenwertproblem," Ann. Phys. (4) **79** (1926), pp. 489 = **E. Schrödinger**, *Abhandlungen zur Wellenmechanik*, Leipzig, 1927, pp. 17. **F. Klein** had already referred to the coupling of the emission theory and the wave theory of light in **Hamilton**'s work on ray optics. Cf., **F. Klein**, "Über neuere englische Arbeiten zur Mechanik," Jahresber. d. Deutsch. Math.-Ver. **1** (1891/92), pp. 35 = *Ges. math. Abhandl. II*, pp. 601.

([124]) Cf., e.g., **M. Herzberger**, *Strahlenoptik*, Berlin 1931, pp. 5.

([125]) The numerical value of that constant is required by the color of the light. From **Hamilton**'s practical viewpoint (e.g., construction of optical instruments that are free of chromatic aberration), it is characterized by the fact that he imagined that n was not only an independent variable, but also a "color parameter."

([126]) The extremum requirement for the variational problem, which simplifies to an extremum problem for a function of finitely-many variables here:

$$\sum_{\sigma} n_\sigma\, l_\sigma = \text{extrem.}, \tag{104.c}$$

is correctly established by the kinking of the light ray under reflection and refraction.

introduced into the ray bundle as surfaces of constant light-path in the ray-bundle ([127]). If one thinks of the numerical value of the constant as variable then one will get an entire family of such surfaces in the ray bundle:

(105) $$V(x, y, z) = \text{const.}, \qquad \text{in which} \quad V(x, y, z) = \sum_{\sigma} n_{\sigma} l_{\sigma} ,$$

and the sum is taken over the extent of the path along the light ray that connects the luminous point with the point $P(x, y, z)$. Hamilton called the function $V(x, y, z)$ the *characteristic function* of the bundle of light rays. Now, the surfaces of constant light-path have the property that the light rays intersect them perpendicularly, such that conversely, when one knows a surface $V(x, y, z) = \text{const.}$, one can get the light rays as the system of normals to that surface. One has ([128]):

(106) $$\frac{\partial V}{\partial x} = n\,\alpha , \qquad \frac{\partial V}{\partial y} = n\,\beta , \qquad \frac{\partial V}{\partial z} = n\,\gamma .$$

For the generalization in which the *index of refraction n* is not constant, but a *function of position* (e.g., light ray in the atmosphere):

(107) $$n = n(x, y, x) ,$$

the argument will remain valid with no changes, except that now the light rays will no longer be straight lines, but curves, and the characteristic function is defined by:

(107) [*sic*] $$V(x, y, z) = \mathcal{E} \int_{P_0}^{P_1} n(x, y, z)\, ds$$

accordingly, and the integral is taken along the light ray. What is more significant is the generalization that relates to the study of the propagation of light in *crystals*, in which the *index of refraction is a function of the direction of the ray:*

(108) $$n = n(\alpha, \beta, \gamma) .$$

The surfaces of constant light path ([129]):

$$V(x, y, z) = \text{const.}$$

([127]) Since, according to the undulatory theory, the *time* that light requires to traverse a segment of the light ray is proportional to the length of the light path, the surfaces of constant light-path will be reached by the light at equal times, and will then be the wave surfaces, in the sense of the undulatory theory.

([128]) In order to obtain those formulas directly, one must replace the index of refraction n in the variation problem (104) with the equal quantity:

(106.a) $$n \cdot \sqrt{\alpha^2+\beta^2+\gamma^2} .$$

([129]) In which V is defined as in (105).

will no longer intersect the rays at right angles, in the sense of Euclidian metric geometry, i.e., the light rays and wave normals will no longer coincide, but rather the direction of the ray will be coupled with the direction of the surface normal by the law:

(109) $$\frac{\partial V}{\partial x} = \frac{\partial n}{\partial \alpha}, \qquad \frac{\partial V}{\partial y} = \frac{\partial n}{\partial \beta}, \qquad \frac{\partial V}{\partial z} = \frac{\partial n}{\partial \gamma},$$

which says, in the terminology of the calculus of variations [cf., II A 8.a (**E. Zermelo-H. Hahn**), no. **1**], that the surfaces are *transverse* to the light rays ([130]). From there, it is only one more step to the general case of inhomogeneous and anisotropic optical media, in which the surfaces of constant light path in a bundle of (curvilinear) rays will be obtained setting the function:

(110) $$V(x, y, z) = \mathcal{E} \int_{P_0}^{P_1} n(\alpha, \beta, \gamma, x, y, z)\, ds$$

equal to a constant, in which P_0 is understood to mean the common starting point of the rays in the bundle. The curvilinear light rays are also coupled with the normals to the surfaces $V(x, y, z) =$ const. by formulas of the same form as (109):

(1111) $$\frac{\partial V}{\partial x} = \frac{\partial n}{\partial \alpha}, \qquad \frac{\partial V}{\partial y} = \frac{\partial n}{\partial \beta}, \qquad \frac{\partial V}{\partial z} = \frac{\partial n}{\partial \gamma},$$

([130]) That relationship between the light rays and the surfaces of constant light path can also be referred to as *orthogonality*, in the generalized sense of the word. In order to see that, one must no longer regard $ds = \sqrt{dx^2+dy^2+dz^2}$ as the arc-length element of the metric that prevails in space, but the metric in the sense of optics, which is defined with the help of the index of refraction (108) (which is a homogeneous function of degree one in its variables) by:

(108.a) $$d\sigma = n(\alpha, \beta, \gamma)\, ds = n(dx, dy, dz)\,.$$

In place of the unit sphere in Euclidian space:

$$\sqrt{\xi^2+\eta^2+\zeta^2} = 1$$

that serves as a "gauge surface," one will then find the surface:

$$n(\alpha, \beta, \gamma) = 1$$

as the "gauge surface in space" that **Hamilton** referred to as a *spheroid*. In isotropic media, the gauge surface has the corresponding form ([128]):

$$n \cdot \sqrt{\xi^2+\eta^2+\zeta^2} = 1\,,$$

so it is a sphere with radius $1 / n$. From the standpoint of wave optics, those gauge surfaces are the *unit waves* that the light attains from a luminous point in a unit time. Physically, one then represents the introduction of the new metric in space in such a way that all lengths are measured in the time that light needs in order to traverse it.

but in which *n* now depends upon the *position coordinates*, along with the direction cosines. Those formulas, which give the relationship between the (tangent to the) light ray and the wave normals can also be expressed in the terminology of the calculus of variation by saying that the light rays intersect the surfaces of constant light path transversally ([131]).

Now, **Hamilton** eliminated the direction cosines α, β, γ for the rays from equations (111), and in that way obtained the *first-order partial differential equation:*

$$\Omega\left(\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial z}, x, y, z\right) = 0 \tag{113}$$

for the characteristic function $V(x, y, z)$, in which V itself did not appear ([132]). The function Ω that one introduces in that way takes on a very intuitive interpretation in terms of the unit waves (112.a):

$$n(\alpha, \beta, \gamma, x, y, z) = 1 ,$$

because the derivatives:

([131]) Here, transversality is orthogonality in the sense of the metric that is defined by the arc-length:

$$ds = n(\alpha, \beta, \gamma, x, y, z)\, ds = n(dx, dy, dz, x, y, z) . \tag{112}$$

That metric is defined in only a *differential-geometric* sense, so in the sense of ***Riemannian*** *geometry* [cf., III D 11 (**L. Berwald**), no. **17**]. If one then imagines that one assigns a space to each individual point that is filled with a homogeneous medium of the same index of refraction that prevails at that point then one will have the gauge surface:

$$n(\alpha, \beta, \gamma, x, y, z)\, ds = 1 , \tag{112.a}$$

for one such space, in the sense of the metric (112), which arises from the "infinitesimal" gauge surface (112) by a dilatation with a ratio of $d\sigma : 1$.

The light rays are geodetic lines for the metric (112). In terms of the history of the mathematical ideas, it might not be uninteresting to note that perhaps simultaneous to **Hamilton**, **C. F. Gauss** studied the pencil of geodetic lines through a point on a surface in his investigations into surface theory and introduced the associated family of geodetic circles, which are completely analogous to the surfaces of constant light path in the bundle of light rays. **Gauss** published his investigations in regard to that in 1827 in his *Disquisitiones generales circa superficies curvas*, *Werke IV*, pp. 217 [cf., III D 3 (**R. von Lilienthal**), no. **15**].

([132]) For a *homogeneous* medium, in particular, it will be free of x, y, z:

$$\Omega\left(\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial z}\right) = 0 .$$

If the medium is also *isotropic*, in addition, then it will read simply:

$$\frac{1}{n}\sqrt{\left(\frac{\partial V}{\partial x}\right)^2 + \left(\frac{\partial V}{\partial y}\right)^2 + \left(\frac{\partial V}{\partial z}\right)^2} = 1 ,$$

or

$$\left(\frac{\partial V}{\partial x}\right)^2 + \left(\frac{\partial V}{\partial y}\right)^2 + \left(\frac{\partial V}{\partial z}\right)^2 = n^2 ,$$

resp.

(114) $$\pi = \frac{\partial n}{\partial \xi}, \qquad \kappa = \frac{\partial n}{\partial \eta}, \qquad \rho = \frac{\partial n}{\partial \zeta}$$

are the components of the altitude that one can drop from the tangent plane to the point ξ, η, ζ in question on the unit wave to the center, since:

$$\pi\,\xi + \kappa\,\eta + \rho\,\zeta = 1\,,$$

i.e., π, κ, ρ are the coordinates of the tangent plane at the point in question. Hence, if ξ, η, ζ are eliminated from (114) then that will give the equation of the *unit wave in planar coordinates* in the form of ([133]):

(115) $$\Omega\,(\pi, \kappa, \rho, x, y, z) = 0\,.$$

Geometrically, the fact that the characteristic function $V\,(x, y, z)$ of a bundle of rays satisfies the partial differential equation (113), means that the tangent plane to the surface of constant light path (i.e., the wave surface) $V\,(x, y, z)$ = const. that goes through a point $P\,(x, y, z)$ is parallel to the tangent plane to the unit wave that belongs to the point that contacts the direction of the ray at the point where the ray pierces the unit wave. If one imagines that π, κ, ρ are not the coordinates of a tangent plane to the unit wave, but the components of a vector that might be referred to as the *wave vector*, then the endpoints of all wave vectors of the unit wave of a point (x, y, z) will likewise sweep out a surface ([134]) whose radii will give the direction and reciprocal magnitudes of the speed

---

([133]) In so doing, **Hamilton** chose the normalization that would make:

(115.a) $$\pi\frac{\partial\Omega}{\partial\pi} + \kappa\frac{\partial\Omega}{\partial\kappa} + \rho\frac{\partial\Omega}{\partial\rho} - 1 = 0\,.$$

Cf., the form of Ω in ([132]). One correspondingly obtains the coordinates of the contact point of the individual tangent plane (π, κ, ρ) from the formulas:

(115.b) $$\frac{\partial\Omega}{\partial\pi} = \frac{\xi}{n}, \qquad \frac{\partial\Omega}{\partial\kappa} = \frac{\eta}{n}, \qquad \frac{\partial\Omega}{\partial\rho} = \frac{\zeta}{n},$$

and comes back to equation (112.a) for the unit wave in point coordinates by eliminating π, κ, ρ from (115.a), say in the form:

$$\Psi\left(\frac{\alpha}{n}, \frac{\beta}{n}, \frac{\gamma}{n}, x, y, z\right) = 0\,.$$

(Cf., **Hamilton**, *Papers I*, pp. 171.)

([134]) In **Fresnel**'s crystal optics, one juxtaposes that surface as the *normal surface* to the *ray surface* that the unit wave defined. [cf., e.g., **P. Drude**, *Lehrbuch der Optik*, 3rd ed., Leipzig, 1912, pp. 311 and pp. 303, as well as V 21 (**A. Wangerin**), no. **7**]. Due to formulas (115.b), in conjunction with:

$$\pi\,\xi + \kappa\,\eta + \rho\,\zeta = 1\,,$$

of propagation of the wave. When viewed from that standpoint, the *partial differential equation* (113) means that the *wave vector of the surface* $V(x, y, z) = \text{const.}$ *at the point* (x, y, z) *is simultaneously a wave vector to the unit wave* that belongs to that point, and indeed it is the one that is determined by the direction of the ray.

Now, along the individual light ray, the direction of the ray, or better yet, the *ray vector* ([135]) with the components $\frac{\alpha}{n}, \frac{\beta}{n}, \frac{\gamma}{n}$ $[n = n(\alpha, \beta, \gamma, x, y, z)]$ and the *wave vector* with the components π, κ, ρ are coupled to each other, and indeed correspond to the formulas (114), which now read:

$$\pi = \frac{\partial n}{\partial \alpha}, \qquad \kappa = \frac{\partial n}{\partial \beta}, \qquad \rho = \frac{\partial n}{\partial \gamma}, \tag{116}$$

or the formulas (115.b), which are equivalent to them and might now be written in the form:

$$\frac{\alpha}{n} = \frac{\partial \Omega}{\partial \pi}, \qquad \frac{\beta}{n} = \frac{\partial \Omega}{\partial \kappa}, \qquad \frac{\gamma}{n} = \frac{\partial \Omega}{\partial \rho}. \tag{117}$$

Since, on the one hand, $\alpha = dx / ds$, $\beta = dy / ds$, $\gamma = dz / ds$, and on the other [cf., ([131])]:

$$n\, ds = ds = dV,$$

they will possess the form ([136])

$$\frac{dx}{dV} = \frac{\partial \Omega}{\partial \pi}, \quad \frac{dy}{dV} = \frac{\partial \Omega}{\partial \kappa}, \quad \frac{dz}{dV} = \frac{\partial \Omega}{\partial \rho}. \tag{118}$$

Therefore, when one knows how the wave vector changes as one advances along the light ray, one will also have the light ray. In order to determine that change, **Hamilton** took the difference of the values of the characteristic function at two infinitely-close points:

$$V(x + dx, y + dy, z + dz) - V(x, y, z) = \bar{V}(dx, dy, dz, x, y, z) = n(dx, dy, dz, x, y, z)$$
$$= n(\alpha, \beta, \gamma, x, y, z)\, ds,$$

and one will then find that ([137]):

$$d\pi = \pi_2 - \pi_1 = \frac{\partial \bar{V}}{\partial x} = \frac{\partial n}{\partial x}\, ds,$$

---

that will lead back to the same construction that one arrives at when one applies the ray surface to the normal surface as when one applies the normal surface to the ray surface. Moreover, the simultaneous consideration of the ray surface and the normal surface was what led **Hamilton** to the discovery of *conical refraction*, cf., "Third Supplem.," no. **28**, **29** = *Papers I*, pp. 283, *et seq.*

([135]) which gives the direction and reciprocal magnitude of the ray velocity.

([136]) Cf., **Hamilton**, *Papers I*, pp. 208.

([137]) Cf., **Hamilton**, *Papers I*, pp. 173, cf., also the presentation in **Thomson-Tait**, *Nat. phil. I*, no. **330**, pp. 346, *et seq.*

$$d\kappa = \kappa_2 - \kappa_1 = \frac{\partial \bar{V}}{\partial y} = \frac{\partial n}{\partial y}\, ds\,,$$
$$d\rho = \rho_2 - \rho_1 = \frac{\partial \bar{V}}{\partial z} = \frac{\partial n}{\partial z}\, ds\,,$$

in which one understands π_2, κ_2, ρ_2 [π_1, κ_1, ρ_1, resp.] to mean the wave vectors at neighboring points, or since $n\, ds = ds = dV$:

$$\frac{d\pi}{dV} = \frac{1}{n}\frac{\partial n}{\partial x}\,, \qquad \frac{d\kappa}{dV} = \frac{1}{n}\frac{\partial n}{\partial y}\,, \qquad \frac{d\rho}{dV} = \frac{1}{n}\frac{\partial n}{\partial z}\,,$$

or since ([138]):

$$\frac{\partial \Omega}{\partial x} = -\frac{1}{n}\frac{\partial n}{\partial x}, \qquad \frac{\partial \Omega}{\partial y} = -\frac{1}{n}\frac{\partial n}{\partial y}, \qquad \frac{\partial \Omega}{\partial z} = -\frac{1}{n}\frac{\partial n}{\partial z},$$

one will have:

(119)
$$\frac{d\pi}{dV} = -\frac{\partial \Omega}{\partial x}, \qquad \frac{d\kappa}{dV} = -\frac{\partial \Omega}{\partial y}\,, \qquad \frac{d\rho}{dV} = -\frac{\partial \Omega}{\partial y}\,,$$

resp. Together, *equations* (118) *and* (119) *represent the conversion of Euler's equations* (104.b) *into a canonical system*. At the same time, one has obtained an *intuitive interpretation* of that conversion. The system (118) and (119) determines the propagation of light in the sense that the ray vector and wave vector are known at each point, and their integration basically ensues from the following construction: One first constructs the unit wave $\Omega = 0$ at the starting point (x, y, z). The initial values π, κ, ρ will then give a well-defined tangent plane to the unit wave, and the connecting line of that point (x, y, z) with the contact point of the tangent plane will give the direction of the ray. The differential of the independent variable dV determines a neighboring point $x + dx, y + dy, z + dz$, and one will get the new value $\pi + d\pi, \kappa + d\kappa, \rho + d\rho$ of the wave vector from (119). One again constructs the unit wave at the neighboring point, determines the tangent plane that belongs to the new wave vector, etc.

Now, it was not actually that determination of the light rays that **Hamilton** had generally attempted to do in his approach to the problems of practical optics. Rather, he dealt with the problem of mastering the map from object space to image space that an optical instrument produced. Thus, he immediately focused on the set of ∞^4 light rays, which he imagined to be divided into ray bundles with ∞^2 rays in each of them that were generated by the luminous point in object space. The characteristic functions that belonged to the individual bundles, which represented the wave surfaces of each bundle, must be distinguished from each other by giving the associated luminous point is object space:

$$V = V\,(x, y, z\,;\, x_0, y_0, z_0)\,.$$

([138]) Cf., **Hamilton**, *Papers I*, pp. 171.

If one now focuses upon only the light ray that connects the point $P_0\,(x_0, y_0, z_0)$ in object space and the point $P\,(x, y, z)$ in image space then the two points P_0 and P will seem to be interchangeable for that ray, and one can regard the characteristic function:

$$V(x, y, z\,;\, x_0, y_0, z_0) = \mathcal{E}\int_{P_0}^{P_1} n(\alpha, \beta, \gamma, x, y, z)\, ds \tag{120}$$

as a *function of the point-pair* $\{P_0, P\}$. If one varies both limits of the integral then the boundary formula of the calculus of variations will yield the variational formula:

$$\delta V = \left(\frac{\partial n}{\partial \alpha}\,\delta x + \frac{\partial n}{\partial \beta}\,\delta y + \frac{\partial n}{\partial \gamma}\,\delta z\right) - \left(\frac{\partial n_0}{\partial \alpha_0}\,\delta x_0 + \frac{\partial n_0}{\partial \beta_0}\,\delta y_0 + \frac{\partial n_0}{\partial \gamma_0}\,\delta z_0\right), \tag{121}$$

in which one understands $n_0\,(\alpha_0, \beta_0, \gamma_0, x_0, y_0, z_0)$ to mean the index of refraction in object space. It can be split into the six relations:

$$\left\{\begin{array}{lll} \dfrac{\partial V}{\partial x} = \dfrac{\partial n}{\partial \alpha} = \pi, & \dfrac{\partial V}{\partial y} = \dfrac{\partial n}{\partial \beta} = \kappa, & \dfrac{\partial V}{\partial z} = \dfrac{\partial n}{\partial \gamma} = \rho, \\ \dfrac{\partial V}{\partial x_0} = -\dfrac{\partial n_0}{\partial \alpha_0} = -\pi_0, & \dfrac{\partial V}{\partial y_0} = -\dfrac{\partial n_0}{\partial \beta_0} = -\kappa_0, & \dfrac{\partial V}{\partial z_0} = -\dfrac{\partial n_0}{\partial \gamma_0} = -\rho_0. \end{array}\right. \tag{121.a}$$

Just as eliminating α, β, γ from the first three of those equations was what led to the partial differential equation (113) above, one can also eliminate $\alpha_0, \beta_0, \gamma_0$ from the last three. When:

$$\Omega\,(\pi_0, \kappa_0, \rho_0, x_0, y_0, z_0) = 0 \tag{122}$$

refers to the unit wave in object space, the two rows in equations (121.a) will then yield the two first-order partial differential equations:

$$\left\{\begin{array}{l} \Omega\left(\dfrac{\partial V}{\partial x}, \dfrac{\partial V}{\partial y}, \dfrac{\partial V}{\partial z}, x, y, z\right) = 0, \\ \Omega_0\left(-\dfrac{\partial V}{\partial x_0}, -\dfrac{\partial V}{\partial y_0}, -\dfrac{\partial V}{\partial z_0}, x_0, y_0, z_0\right) = 0 \end{array}\right. \tag{123}$$

that the *characteristic function V, which is now regarded as a function of the point-pair* P_0, P, must satisfy. Equations (121.a) give the wave vectors, and therefore also the dependency of the ray vectors at the two boundary points P and P_0 of the light ray on the coordinates of those two boundary points. They then determine – and this is the *fundamental problem in ray optics* – the changes that the directions of the light ray at the two boundary points will experience under a change in position of one of the two boundary points. In particular, the direction of the light ray at

P_0, and therefore the wave vector π_0, κ_0, ρ_0, must now remain unchanged when the point P varies along the light ray itself. Hence, the equations of the second group in (121.a) [the second of them, resp.] ([139]) must be the *finite equations of the light ray* in image space ([140]).

In order to get an *intuitive geometric interpretation of formulas* (121.a), one must observe that a point (x_0, y_0, z_0) in object space is associated with a surface (which is just the wave surface) in image space by the relation ([141]):

(124) $$V(x, y, z, x_0, y_0, z_0) = \text{const.}$$

If one takes a wave element in object space that goes through a point P_0 (x_0, y_0, z_0) instead of P_0, which might be established by two further points P_0', P_0'' that are infinitely close to P_0, then each of the points P_0, P_0', and P_0'' in image space will belong to a wave surface (124), and the three surfaces will intersect at a point P. At the same time, that point of intersection P belongs to the enveloping surface of that two-parameter family of surfaces that one will get from (124) when the starting point (x_0, y_0, z_0) in object space varies on the element $P_0 P_0' P_0''$ such that the point of intersection of the three wave surface will be associated with the tangent plane to the enveloping surface of a wave element. The intersection of the three surfaces is then given by the equations:

$$\frac{\partial V}{\partial x_0} + \frac{\partial V}{\partial z_0}\frac{\partial z_0}{\partial x_0} = 0\,, \qquad \frac{\partial V}{\partial y_0} + \frac{\partial V}{\partial z_0}\frac{\partial z_0}{\partial y_0} = 0\,,$$

so when (π_0, κ_0, ρ_0) is the wave vector in object space, it will be given by:

$$\frac{\partial V}{\partial x_0} : \frac{\partial V}{\partial y_0} : \frac{\partial V}{\partial z_0} = \pi_0 : \kappa_0 : \rho_0\,,$$

and the proportionality factor is determined to be – 1, since:

$$(\pi_0, \kappa_0, \rho_0, x_0, y_0, z_0) = 0\,,$$

when one recalls (121.a).

If one varies the constant in (124) with which one has constructed the three wave surfaces (so one considers the entire wave-train, instead of the individual wave surface) then one will traverse the whole light ray. The second group of equations (121.a) will then say that *the individual light rays are generated by the intersection of three infinitely-close wave-trains*. The first group will

([139]) The third of them is determined from two of them, respectively, by (123).

([140]) Correspondingly, the equations of the first group in (121.a) would be the finite equations of the light ray in object space when one regards $x, y, z, \pi, \kappa, \rho$ as constants. Cf., **Hamilton**, *Papers I*, pp. 173.

([141]) Naturally, since a point in image space also conversely corresponds to a surface in object space, one is dealing with a *contact transformation* for which (124) represents the "directrix equation" [cf., III D 7 (**H. Liebmann**), esp. no. **3**]. For this and the following developments in the text, cf., also **E. Vessiot**, "Sur l'interprétation mécanique des transformations de contact infinitésimales," Bull. de la soc. math. de France **34** (1906), pp. 230.

then determine wave vectors that belong to the light rays that are generated at the individual points in that way ([142]).

Equations (121.a) then give a representation of the individual light ray in image space, along with the wave element that propagates along it, and indeed its dependency upon the associated initial wave element in object space. The interpretation of the formulas (121.a) is based upon that association of corresponding wave elements in object space and image space. The fact that such as association is possible with the help of the *one* characteristic function V is the source of some important *reciprocity theorems* that were known previously in ray optics. Namely, when one partially-differentiates the function V with respect to two of its variables, the result will be independent of the sequence of differentiations. That leads to *two different types of reciprocity relations*.

For *the first type*, one has relations of the form:

$$\frac{\partial \pi}{\partial y} = \frac{\partial \kappa}{\partial x} = \frac{\partial^2 V}{\partial x \, \partial y}, \qquad \text{etc.,} \tag{125}$$

or

$$\frac{\partial \pi_0}{\partial y_0} = \frac{\partial \kappa_0}{\partial x_0} = -\frac{\partial^2 V}{\partial x_0 \, \partial y_0}, \qquad \text{etc.,} \tag{125.a}$$

resp., while the *second type* has the form:

$$\frac{\partial \pi}{\partial x_0} = -\frac{\partial \pi_0}{\partial x} = \frac{\partial^2 V}{\partial x_0 \, \partial y}, \qquad \text{etc.,} \tag{126}$$

as well as ([143]):

$$\frac{\partial \pi}{\partial y_0} = -\frac{\partial \kappa_0}{\partial x_0} = \frac{\partial^2 V}{\partial x \, \partial y_0}, \qquad \text{etc.} \tag{126.a}$$

Equations (125) once more give a strange relationship between the changes in the components of the wave vectors (the ray direction that it implies, resp.) that they experience under a change in position of the associated point. One can infer from it the way by which every ray in the image space is associated with the ray that is infinitely close to it. On the other hand, equations (126) and (126.a) show that the changes that the components of the wave vector at a point in object space experience are connected with the changes that a displacement of the point in image space will produce in the components of the wave vector at the luminous point in object space. Briefly stated, the latter reciprocity emerges from the fact that the map that an optical instrument mediates will remain unchanged when one switches the locations of the eye and the subject being observed while the position of the instrument remains unchanged ([144]).

([142]) The ray vector that is coupled with it, resp., i.e., the direction of the light ray.
([143]) Cf., **W. R. Hamilton**, *Papers I*, pp. 256.
([144]) Cf., also **Thomson-Tait**, *Mat. Phil.*, pp. 358.

14. Introduction of the characteristic function into mechanics and its application to the calculation of perturbations. – In order to adapt those results that he had obtained in geometrical optics to mechanics ([145]), **Hamilton** ([146]) replaced the variational problem:

$$\int n\,ds = \text{extrem.}$$

with the *variational problem for the Euler principle:*

(127) $$\int 2T\,ds = \text{extrem.}$$

with the energy integral:

(128) $$T + \Phi = \text{const.} = k$$

as an auxiliary condition ([147]). He then juxtaposed the *principle of least* or (better yet, since one is only dealing with the vanishing of the first variation) *stationary action*, which yields the Ansatz for the equations in the form of the *principle of varying action*, for which he introduced the extremal integral:

(129) $$V = \mathcal{E}\int 2T\,ds$$

which extends over a trajectory (space-time line, resp.) of the system, and which he regarded as a function of the coordinates of the initial position $\Pi_0(q_1^{(0)},\ldots,q_n^{(0)})$ and the coordinates of the final position Π (q_1, …, q_n), as well as the energy constant k ([148]): $V = V(q_1, \ldots, q_n, q_1^{(0)},\ldots,q_n^{(0)},k)$.

One will get the derivatives of that function from the boundary formula in the calculus of variation ([149]):

Any adaptation of the concept of characteristic function to a different variational problem must exhibit reciprocity theorems that are analogous to these (cf., no. **16**). Such theorems have often been pointed out in theoretical physics without ever mentioning the intrinsic reason for that, which is just the existence of a variational problem and the associated extremal integral.

([145]) Cf., footnote ([107]).

([146]) **W. R. Hamilton**, "On a general method in dynamics by which the study of all free systems of attracting or repelling points is reduced to the search and differentiation of one central relation, or characteristic function," Trans. London Phil. Soc. (1834), Part 2, pp. 247. "Second essay on a general method in dynamics," Trans. London Phil. Soc. (1835), Part 1, pp. 95.

([147]) In particular, he had the *n*-body problem in mind with:

$$\Phi = \sum m_\rho m_\sigma f(r_{\rho\sigma}) \ .$$

([148]) In which the time t_0 that belongs to the initial location can be prescribed arbitrarily.

([149]) Here, since a variation of the energy constant for fixed limits Π_0 and Π_1 will have only a change in the duration $(t - t_0)$ of the motion as a consequence, it will read:

$$(130)\qquad \begin{cases} \dfrac{\partial V}{\partial q_1} = \dfrac{\partial T}{\partial \dot q_1}, & \dots, & \dfrac{\partial V}{\partial q_n} = \dfrac{\partial T}{\partial \dot q_n}, \\ \dfrac{\partial V}{\partial q_1^{(0)}} = -\left(\dfrac{\partial T}{\partial \dot q_1}\right)_0, & \dots, & \dfrac{\partial V}{\partial q_n^{(0)}} = -\left(\dfrac{\partial T}{\partial \dot q_n}\right)_0, \\ \dfrac{\partial V}{\partial q_1} = t - t_0 . \end{cases}$$

It would be convenient to introduce the impulse components into equations (130):

$$(131)\qquad p_\rho = \frac{\partial T}{\partial \dot q_\rho},$$

which will make them take the form:

$$(132)\qquad \begin{cases} \dfrac{\partial V}{\partial q_1} = p_1, & \dots, & \dfrac{\partial V}{\partial q_n} = p_n, \\ \dfrac{\partial V}{\partial q_1^{(0)}} = -p_1^{(0)}, & \dots, & \dfrac{\partial V}{\partial q_n^{(0)}} = -p_n^{(0)}, \\ \dfrac{\partial V}{\partial k} = t - t_0 . \end{cases}$$

If one also introduces the energy $H = T + \Phi$ in impulse components:

$$(133)\qquad H = H\,(p_1, \dots, p_n, q_1, \dots, q_n)$$

then one will see immediately that the characteristic function V satisfies the two first-order partial differential equations:

$$(134)\qquad \begin{cases} H\left(\dfrac{\partial V}{\partial q_1}, \dots, \dfrac{\partial V}{\partial q_n}, q_1, \dots, q_n\right) = k, \\ H\left(-\dfrac{\partial V}{\partial q_1^{(0)}}, \dots, -\dfrac{\partial V}{\partial q_n^{(0)}}, q_1^{(0)}, \dots, q_n^{(0)}\right) = k. \end{cases}$$

If one knows the characteristic function V then the n equations of the second row of (130) [(132), resp.] will yield a representation of the finite equations of the trajectory, while the equations of the first row will yield the impulse components (the velocity components, resp.) at the individual points of the trajectory. Finally, the last of the equations describes the time evolution of the motion

---

(129.a) $$\mathcal{E}\int T\,dt = \sum \frac{\partial T}{\partial \dot q_\rho}\delta q_\rho - \sum \left(\frac{\partial T}{\partial \dot q_\rho}\right)_0 \delta q_\rho^{(0)} + (t - t_0)\delta k .$$

along the trajectory ([150]). Everything is exactly as it is in ray optics, except that the trajectories in an *n*-dimensional **Riemannian** space whose arc-length is determined from $ds^2 = 2T\,dt^2$ enter in place of the light rays in three-dimensional space, and in place of the wave vector, one finds the impulse vector, which is perpendicular to the surfaces V = const. in the sense of the metric. Just as in optics, **Hamilton**'s goal in this case was to develop a systematic method for the integration of equations of motion, and all it took for him to achieve that goal was to represent the motion with the help of *one* function V ([151]).

Now, to **Hamilton**, formulas (130) was closely related to the idea of introducing the time t as a variable in place of the energy constant k by means of:

$$\frac{\partial V}{\partial k} = t - t_0 \,.$$

In that way, he had replaced the characteristic function V with the *principal function:*

([150]) In the case of the *n*-body problem, that representation of the motion by the formulas (130) [(132), resp.] led **Hamilton** (if only casually) to the connection between the "first center of mass integral" and the "area integral" with the "invariance of the equations of motion under parallel displacements and rotations."

That is because if one imagines, say, replacing all coordinates x_ρ with $(x_\rho + a)$ and correspondingly replacing $x_\rho^{(0)}$ with $\left(x_\rho^{(0)} + a\right)$ then the derivative of V with respect to a must vanish, so:

$$\left(\frac{\partial V}{\partial x_\rho} + \frac{\partial V}{\partial x_\rho^{(0)}}\right) = 0$$

or

$$\Sigma \frac{\partial V}{\partial x_\rho} = \text{const.}$$

However, from (130), that is nothing but the first center of mass integral for the x-direction:

$$\frac{\partial T}{\partial \dot{x}_1} + \cdots + \frac{\partial T}{\partial \dot{x}_n} = \text{const.}$$

One will get analogous statements for the first center of mass integral for the other two coordinate directions, and when one appeals to rotations about the coordinate axes instead of parallel displacements, one will get the three area integrals (for that, cf. *infra*, no. **23**).

C. G. J. Jacobi has also emphasized that connection: **C. G. J. Jacobi**, *Vorlesungen, Werke Supplement. Bd.,* Vorles. 3, pp. 15 and Vorles. 5, pp. 31.

([151]) He himself made the remark that one must feel a certain "intellectual pleasure" in all cases where the motion can be represented in the form (130) with the help of *one* characteristic function, even when "no practical facility" is achieved in the problem of integrating the equations of motion.

For the motion of the *n*-body problem, he showed that the representation (130) subsumed known properties of the motion when he correspondingly represented the decomposition of the motion of the system into the motion of the center of mass and the motion relative to the center of mass that is customary in the celestial mechanics by decomposing the characteristic function V as a sum of two corresponding summands.

(135) $$S(t,q_1,\ldots,q_n,t_0,q_1^{(0)},\ldots,q_n^{(0)}) = V(q_1,\ldots,q_n,q_1^{(0)},\ldots,q_n^{(0)},k) - k\frac{\partial V}{\partial k},$$

in the sense of the so-called **Legendre** transformation ([152]). From (129) and (130), one has:

$$S = -k(t-t_0) + \mathcal{E}\int 2T\,dt = \mathcal{E}\int (2T-k)\,dt = \mathcal{E}\int (2T-(T+\Phi))\,dt$$

or

(136) $$S = \mathcal{E}\int_{P_1}^{P_2} (T-\Phi)\,dt,$$

i.e.: *the principal function S is the extremal integral of Hamilton's principle* ([153]), *when regarded as a function of the limits.* The derivatives of the principal function are:

([152]) The application of the **Legendre** transformation to the characteristic function was common in **Hamilton**'s investigations into optics. Namely, since a bundle of rectilinear rays that is emitted by a luminous point in object space will become a bundle of lines in image space (which is no longer homocentric) under the optical map (in the case of homogeneous media), it would seem convenient to characterize an individual ray in one such general bundle of rays in image space by its direction cosines. The function:

$$V(\alpha, \beta, \gamma, x_0, y_0, z_0) = x\frac{\partial n}{\partial \alpha} + y\frac{\partial n}{\partial \beta} + z\frac{\partial n}{\partial \gamma} - V$$

will then enter in place of the characteristic function $V(x, y, z, x_0, y_0, z_0)$ (cf., **Hamilton**, *Papers I*, pp. 111), in which the x, y, z on the right-hand side are eliminated with the help of equations (111). That function is best suited to the applications in optical practice. By contrast, by a systematic process, one has not introduced (and that is what **Hamilton** did in the third Supplement) the direction cosines of the ray as the variables, in the sense of the **Legendre** transformation, but the components of the wave vector in image space. Moreover, **Hamilton** went even further in the third Supplement. He then replaced the characteristic function V with a function T that depended on the components of the wave vector in object space and image space by performing the Legendre transformation in object space, as well as image space. That is the appropriate instrument for the study of the general line map from the object space to image space.

For an attempt to duplicate that in mechanics, one might cf., **E. J. Routh**, *Dynamik II*, Chap. 10, **§ 487**, pp. 362.

([153]) On the Continent, that recalculation has given rise to the introduction of the term "**Hamilton**'s principle" for the variational principle $\int (T-\Phi)\,dt$ = extrem., despite the fact that **Lagrange** has already expressed the principle in *Mécanique analytique*. The fact that one was dealing with a *new* principle in the formulation that he had achieved whose range of applications was more extensive that that of **Euler**'s principle did not occur to **Hamilton**, since he also assumed that the energy integral was valid after recalculation. If the energy integral is fulfilled then time will appear in the principal function S only in the combination $(t - t_0)$ such that for **Hamilton**:

$$\frac{\partial S}{\partial t} = -\frac{\partial S}{\partial t_0} = \frac{\partial S}{\partial (t-t_0)}.$$

$$(137)\qquad \begin{cases} \dfrac{\partial S}{\partial q_1} = \dfrac{\partial T}{\partial \dot q_1} = p_1, & \ldots, \quad \dfrac{\partial S}{\partial q_n} = \dfrac{\partial T}{\partial \dot q_n} = p_n, \\ \dfrac{\partial S}{\partial q_1^{(0)}} = -\left(\dfrac{\partial T}{\partial \dot q_1}\right)_0 = -\,p_1^{(0)}, & \ldots, \quad \dfrac{\partial S}{\partial q_n^{(0)}} = -\left(\dfrac{\partial T}{\partial \dot q_n}\right)_0 = -\,p_n^{(0)}, \end{cases}$$

to which the two derivatives:

$$(137.a)\qquad \frac{\partial S}{\partial t} = -\,H\,, \qquad \frac{\partial S}{\partial t_0} = +\,H_0$$

are added ([154]). The elimination of impulse components from equations (137) and (137.a) yields the two first-order partial differential equations for the principal function *S* ([155]):

$$(138)\qquad \begin{cases} \dfrac{\partial S}{\partial t} + H\left(\dfrac{\partial S}{\partial q_1}, \ldots, \dfrac{\partial S}{\partial q_n},\ q_1, \ldots, q_n, t \right) = 0, \\ -\dfrac{\partial S}{\partial t_0} + H\left(-\dfrac{\partial S}{\partial q_1^{(0)}}, \ldots, -\dfrac{\partial S}{\partial q_n^{(0)}}, q_1^{(0)}, \ldots, q_n^{(0)}, t \right) = 0, \end{cases}$$

in which time *t* is included explicitly in *H*, which goes beyond what **Hamilton** did.

The appearance of the impulse components p_ρ as the derivatives of the principal function *S* (as is already true of the derivatives of the characteristic function *V*, resp.) led **Hamilton** to pose *the impulse components as independent variables, along with the position coordinates* q_ρ and thus (no. **9**) to arrive at the canonical form:

([154]) For **Hamilton**, the two equations (137.a) coincided in the one equation:

$$\frac{\partial S}{\partial (t - t_0)} = \frac{\partial S}{\partial t} = -\,\frac{\partial S}{\partial t_0}$$

since the energy integral has the form $H = \text{const.} = k$.

If one juxtaposes the formulas (137) and (137.a):

$$\frac{\partial S}{\partial q_1} = p_1\,, \ \ldots,\ \frac{\partial S}{\partial q_n} = p_n\,, \qquad \frac{\partial S}{\partial t} = -\,H$$

then one will see that the assumption that the position coordinates are on a par with time must necessarily imply that impulse and energy must also be building blocks in a higher unity, as the theory of relativity has accomplished by introducing the impulse-energy tensor [cf., V 19 (**W. Pauli, Jr**.), no. **37**].

([155]) Since time *t* did not appear explicitly in *H* for **Hamilton**, he wrote the second of equations (138) in the form [cf., ([154])]:

$$\frac{\partial S}{\partial t} + H\left(-\frac{\partial S}{\partial q_1^{(0)}}, \ldots, -\frac{\partial S}{\partial q_n^{(0)}}, q_1^{(0)}, \ldots, q_n^{(0)} \right) = 0\,.$$

(139) $$\frac{dq_\rho}{dt} = \frac{\partial H}{\partial p_\rho}, \qquad \frac{dp_\rho}{dt} = -\frac{\partial H}{\partial q_\rho} \qquad (\rho = 1, \ldots, n)$$

for the equations of motion.

As an application of his theory, **Hamilton** gave a *new representation of the equations of perturbation*. In order to do that, he represented the energy H as the sum of the energy H^* of the unperturbed motion and the perturbing function Ω :

(140) $$H = H^* (p_1, \ldots, p_n, q_1, \ldots, q_n, t) + \Omega (p_1, \ldots, p_n, q_1, \ldots, q_n, t) ,$$

and started by integrating the equations of motion of the unperturbed motion:

(141) $$\frac{dq_\rho}{dt} = \frac{\partial H^*}{\partial p_\rho}, \quad \frac{dp_\rho}{dt} = -\frac{\partial H^*}{\partial q_\rho},$$

in general, such that one defines the associated principal function as the function S^* of the coordinates of the limits $q_1, \ldots, q_n, t, , \ldots, t_0, q_1^{(0)}, \ldots, q_1^{(0)}, t_0$, and one can then represent the unperturbed motion in the form (137) with their help.

In order to then find the *principal function S for the perturbed problem*, he then imagined that it was represented as the sum of S^* and a correction term U :

(142) $$S = S^* + U .$$

In so doing, one assumes that the correction term U is known as a function of $q_1, \ldots, q_n, t, , \ldots, t_0, q_1^{(0)}, \ldots, q_1^{(0)}, t_0$ at the initial moment and then calculates the change that U will experience when one advances along a space-time line of the unperturbed problem that starts from the space-time point $(q_1^{(0)}, \ldots, q_1^{(0)}, t_0)$. For that change, one gets in the desired degree of approximation ([156]):

([156]) Namely, one has:

(a) $$\frac{dU}{dt} = \frac{\partial U}{\partial t} + \sum_\rho \frac{\partial U}{\partial q_\rho} \dot{q}_\rho = \frac{\partial U}{\partial t} + \sum_\rho \frac{\partial H^*}{\partial p_\rho} \frac{\partial U}{\partial q_\rho} ,$$

in which p_λ naturally replaces with the derivatives $\frac{\partial S^*}{\partial q_\lambda}$ in $\frac{\partial H^*}{\partial p_\rho}$.

On the other hand, from (138):

$$\frac{\partial U}{\partial t} = -\frac{\partial S^*}{\partial t} - H = -\frac{\partial S^*}{\partial t} - H^* \left(\frac{\partial S^*}{\partial q_\rho} + \frac{\partial U}{\partial q_\rho}, q_\rho, t \right) - \Omega \left(\frac{\partial S^*}{\partial q_\rho} + \frac{\partial U}{\partial q_\rho}, q_\rho, t \right) ,$$

or, when one develops H^* and Ω and considers the fact that S^* , as the principal function of the unperturbed problem, satisfies the equation:

(143) $$\frac{dU}{dt} = -\,\Omega\,(t, q_1^{(0)}, \ldots, q_n^{(0)}, p_1^{(0)}, \ldots, p_n^{(0)}, t_0),$$

such that one can find U by a quadrature:

(143.a) $$U = \bar{U}\,(t, q_1^{(0)}, \ldots, q_n^{(0)}, p_1^{(0)}, \ldots, p_n^{(0)}, t_0)\ .$$

If the $p_\rho^{(0)}$ in that were replaced with $-\dfrac{\partial S^*}{\partial q_\rho^{(0)}}$ then the correction term U would be represented in the desired form as a function of the coordinates of the two space-time points P_0 and P :

(144) $$U\,(q_1^{(0)}, \ldots, q_n^{(0)}, t_0, q_1, \ldots, q_n, t) = \bar{U}\,(t, q_1^{(0)}, \ldots, q_n^{(0)}, -\frac{\partial S^*}{\partial q_1^{(0)}}, \ldots, -\frac{\partial S^*}{\partial q_n^{(0)}}, t_0)\ .$$

However, if one has determined the correction term in the principal function then one will get the integral of the perturbed motion by applying the formulas (137) in the form:

$$\frac{\partial S^*}{\partial t} + H^*\left(\frac{\partial S^*}{\partial q_\rho}, q_\rho, t\right) = 0,$$

one will have:

$$\frac{\partial U}{\partial t} = -\sum_\rho \frac{\partial U}{\partial q_\rho}\frac{\partial H^*}{\partial p_\rho} - \cdots - \Omega\left(\frac{\partial S^*}{\partial q_\rho}, q_\rho, t\right) - \sum_\rho \frac{\partial U}{\partial q_\rho}\frac{\partial \Omega}{\partial p_\rho} + \cdots .$$

Here, in the sense of the approximate calculations, the terms of order two and higher in $\partial U / \partial q_\rho$, as well as the first-order terms, that arise from the intrinsically-small perturbing function can be neglected, such that one will have:

$$\frac{\partial U}{\partial t} = -\sum_\rho \frac{\partial U}{\partial q_\rho}\frac{\partial H^*}{\partial p_\rho} - \Omega\ .$$

If then follows from (a) that, in the same approximation, one will have:

$$\frac{dU}{dt} = -\,\Omega\,(p_1, \ldots, p_n, q_1, \ldots, q_n, t)$$

along a space-time curve of the unperturbed problem. Since p_ρ and q_ρ are known as functions of time and the initial values $p_\rho^{(0)}$, $q_\rho^{(0)}$, t_0 along such a space-time line, one will have:

$$\Omega\,(p_1, \ldots, p_n, q_1, \ldots, q_n, t) = \bar{\Omega}\,(t, q_1^{(0)}, \ldots, q_n^{(0)}, p_1^{(0)}, \ldots, p_n^{(0)}, t_0)$$

here.

$$(145)\qquad \begin{cases} \dfrac{\partial S^*}{\partial q_1}+\dfrac{\partial U}{\partial q_1}= p_1, & \ldots, & \dfrac{\partial S^*}{\partial q_n}+\dfrac{\partial U}{\partial q_n}= p_n, \\ \dfrac{\partial S^*}{\partial q_1^{(0)}}+\dfrac{\partial U}{\partial q_1^{(0)}}=-p_1^{(0)}, & \ldots, & \dfrac{\partial S^*}{\partial q_n^{(0)}}+\dfrac{\partial U}{\partial q_n^{(0)}}=-p_n^{(0)}, \end{cases}$$

and indeed, those would be exact formulas if $U = S - S^*$ were determined exactly. If U is determined only approximately, in the sense of the calculations that were just described, then those formulas will also give the perturbed motion to the same degree of approximation ([157]).

The connection between that treatment of the perturbed motion and its treatment by the variation of constants is implied immediately by the fact that the formulas of the unperturbed motion ([158]):

$$(146)\qquad \begin{cases} q_\rho = \varphi_\rho(t, t_0, q_1^{(0)}, \ldots, q_n^{(0)}, p_1^{(0)}, \ldots, p_n^{(0)}), \\ p_\rho = \psi_\rho(t, t_0, q_1^{(0)}, \ldots, q_n^{(0)}, p_1^{(0)}, \ldots, p_n^{(0)}) \end{cases}$$

will yield the perturbed motion when one replaces the initial impulse $p_\rho^{(0)}$ for the unperturbed motion with $p_\rho^{(0)} + \dfrac{\partial U}{\partial q_\rho^{(0)}}$ and replaces the impulse p_ρ of the unperturbed motion at time t with $p_\rho - \dfrac{\partial U}{\partial q_\rho}$, such that one has to use the new formulas:

$$(147)\qquad \begin{cases} q_\rho = \varphi_\rho\left(t, t_0, q_1^{(0)}, \ldots, q_n^{(0)}, p_1^{(0)} + \dfrac{\partial U}{\partial q_1^{(0)}}, \ldots, p_n^{(0)} + \dfrac{\partial U}{\partial q_n^{(0)}}\right), \\ p_\rho = \dfrac{\partial U}{\partial q_\rho} + \psi_\rho\left(t, t_0, q_1^{(0)}, \ldots, q_n^{(0)}, p_1^{(0)} + \dfrac{\partial U}{\partial q_1^{(0)}}, \ldots, p_n^{(0)} + \dfrac{\partial U}{\partial q_n^{(0)}}\right), \end{cases}$$

([157]) The second group in those formulas says how one must change the initial impulse of the unperturbed motion at the space-time point P_0 in order to reach the same space-time point P under the perturbed motion as one does with the unperturbed motion. One then infers from the first group how the impulse components at the space-time point P that one reaches under the perturbed motion will change in comparison to the unperturbed motion.

([158]) They will become identities when one substitutes:

$$p_\rho = \frac{\partial S^*}{\partial q_\rho}, \qquad p_\rho^{(0)} = -\frac{\partial S^*}{\partial q_\rho^{(0)}}.$$

in place of formulas (156), in order to be calculate the position and impulse of the perturbed motion at time t for given initial values t_0, $q_1^{(0)}, \ldots, q_n^{(0)}, p_1^{(0)}, \ldots, p_n^{(0)}$ ([159]).

15. The intervention of Jacobi. – **Hamilton**'s work on ray optics was hardly noticed outside of England ([160]). Indeed, his treatises on mechanics were greatly noticed, but since their roots in his work on ray optics was not known, there was a danger that **Hamilton**'s basic concepts that arose from those works would be misunderstood. That danger was avoided by the first great researcher to take up those work in an autonomous spirit, as opposed to other more subjective works, namely, **C. G. J. Jacobi**. In fact, **Jacobi** gave a new twist to **Hamilton**'s development, and its further construction by the **Jacobi** school would then create a ***Hamilton-Jacobi*** *theory* that would no longer coincide with **Hamilton**'s original ideas but was believed to grow out of **Hamilton**'s basic ideas on the Continent ([161]). That theory will be presented systematically in the following sections. Here, we shall only go briefly into *the basis for Jacobi's reshaping of Hamilton's ideas*, because the following development will become more understandable in that way.

Ever since **Jacobi** did some work in his youth that addressed the integration of first-order partial differential equations ([162]), he became fascinated with **Hamilton**'s work on mechanics, and especially with the connection between the two partial differential equations and the canonical system of the equations of motion that is mediated by the principal function ([163]). In contrast to

---

([159]) In so doing, in the correction terms $\frac{\partial U}{\partial q_\rho^{(0)}}$ or $\frac{\partial U}{\partial q_\rho}$, resp., the $q_1, \ldots, q_n$ are replaced with the values that would enter into the unperturbed motion at time t if they had possessed the same initial values $q_1^{(0)}, \ldots, q_n^{(0)}, p_\rho^{(0)}, \ldots, p_\rho^{(0)}$ that were prescribed for the perturbed motion.

Moreover, with a simple conversion of (147), one can obtain the changes Δq_ρ, Δp_ρ that would be required of the values q_ρ, p_ρ of the unperturbed problem in order for one to arrive at the position coordinates and the impulse components of the perturbed motion (with the same initial values) at time t. That would give, e.g.:

$$\Delta q_\rho = \sum_\sigma \left(\frac{\partial \varphi_\rho}{\partial p_\sigma^{(0)}} \Delta^{(0)} p_\sigma + \frac{\partial \varphi_\rho}{\partial q_\sigma^{(0)}} \Delta^{(0)} q_\sigma \right),$$

in which one has:

$$\Delta^{(0)} p_\sigma = -\int_{t_0}^{t_1} \frac{\partial \Omega}{\partial q_\sigma^{(0)}}\, dt, \quad \Delta^{(0)} q_\sigma = -\int_{t_0}^{t_1} \frac{\partial \Omega}{\partial p_\sigma^{(0)}}\, dt.$$

([160]) The repercussions of that work were found only in the English literature (e.g., **Thomson-Tait**, *Natural Phil.*).

([161]) For the contrast between **Jacobi**'s and **Hamilton**'s approaches, cf., also **A. F. Conway** and **A. J. M'Connell**, "On the determination of Hamilton's principal function," Proc. R. Irish Acad. Dublin (A) **41** (1932), pp. 18.

([162]) **C. G. J. Jacobi**, "Über die Integration der partiellen Differentialgleichungen erster Ordnung," J. f. Math. **2** (1827), pp. 317 = *Werke IV*, pp. 1 and **C. G. J. Jacobi**, "Über die Pfaffsche Methode, eine gewöhnliche lineare Differentialgleichung zwischen 2*n* Variabeln durch ein System von *n* Gleichungen zu integrieren," J. f. Math. **2** (1827), pp. 347 = *Werke IV*, pp. 17.

([163]) **Jacobi**'s first reference to **Hamilton**'s ideas was found in a letter to **Encke**, cf., **C. G. J. Jacobi**, "Zur Theorie der Variationsrechung und der Differentialgleichungen," J. f. Math. **17** (1837), pp. 68 = *Werke IV*, pp. 39.

Hamilton's earlier conception of things, in the spirit of his era, he saw the solution to the problem of integrating a partial differential equation in the reduction to the integration of a system of ordinary differential equations, like what **J. F. Pfaff** had first achieved [cf., II A 5 (**E. von Weber**), no. **28**]. From that standpoint, *the canonical system represents the first Pfaffian system for each of the two Hamiltonian partial differential equations*. **Hamilton**'s argument then shows that one can already complete the integration of a first-order partial differential equation with the help of that first **Pfaffian** system and not necessarily have to appeal to the "higher **Pfaffian** systems" ([164]). As a result of that *line of reasoning*, **Jacobi** had to take offence at the fact that **Hamilton** had introduced *two* partial differential equations for his principal function, and he wished to eliminate the second one as superfluous, since he did not regard the *principal function* as a function of the two equally-justified space-time points P_0 and P, but *as a function of the space-time point P alone*, into which the coordinates of the space-time point P_0 would enter only as parameters. Therefore, he believed that a new idea had to be introduced, while he had only resumed **Hamilton**'s view of things that the latter had originally started from in his optical investigations ([165]).

On the other hand, **Jacobi** interpreted the **Hamiltonian** representation of the motion with the help of the principal function that is given by formulas (137) (the corresponding representation in terms of the characteristic function, resp.) by saying that *it should precede the first-order differential equation*, and indeed in such a way that the 2*n* integrals of the canonical system in the form (137) (a suitably-generalized form, resp.) would arise from a "complete" solution to that partial differential equation by mere differentiation and eliminations. Naturally, his systematic view of such a direct procedure seemed impossible from the standpoint of partial differential equations since the integration of a system of *ordinary* differential equations could not be based upon the integration of a *partial* differential equation. Nonetheless (so he concluded), it offered the possibility of constructing a systematic method for integrating the canonical system that would make it possible to simplify the integration essentially. The 2*n* integrals (137) of the canonical system decompose into two classes, one of which is written in the first row of (137), and the other of which is written in the second row. Now, if one has found the *n* integrals of the first row:

$$\frac{\partial S}{\partial q_\rho} = p_\rho ,$$

into which the *n* constants $q_1^{(0)}$, …, $q_n^{(0)}$ enter, in any way then one can obtain the principal function *S* itself by a quadrature, from which one will get the other *n* integrals [viz., the second row of (137)] of the canonical system

$$\frac{\partial S}{\partial q_\rho^{(0)}} = - p_\rho^{(0)} \qquad (\rho = 1, \ldots, n)$$

([164]) **C. G. J. Jacobi**, "Über die Reduktion der Integration der partiellen Differentialgleichungen erster Ordnung zwischen irgendeiner Zahl von Variablen auf die Integration eines einzigen Systems gewöhnlicher Differentialgleichungen," J. f. Math. **17** (1837), pp. 97 = *Werke IV*, pp. 57.

([165]) As long as **Hamilton** considered only an isolated bundle of rays with his wave surfaces (e.g., in the "Theory of Rays" itself), naturally he also had *only one* partial differential equation.

by mere differentiation. The integration will then be essentially complete when one knows only one-half of the set of all $2n$ integrals that are present, which are generally selected suitably. That argument agrees with a result to which **Jacobi** arrived in 1836. Namely, in his study of the motion of an individual mass-point in the plane ([166]), he had recognized that in addition to the energy integral:

$$H(p_1, p_2, q_1, q_2) = k ,$$

one needs to know *only one* further integral:

$$F(p_1, p_2, q_1, q_2) = c .$$

Namely, if one calculates the impulse components p_1, p_2, as functions of the q_1, q_2, and the two constants k and c from those two integrals then:

(148) $$p_1\, dq_1 + p_2\, dq_2$$

will be a complete differential, and one can then obtains the two missing integrals by the quadratures:

$$\int \left(\frac{\partial p_1}{\partial c}\, dq_1 + \frac{\partial p_2}{\partial c}\, dq_2 \right) = b , \quad \int \left(\frac{\partial p_1}{\partial k}\, dq_1 + \frac{\partial p_2}{\partial k}\, dq_2 \right) = t - \tau .$$

Naturally, the expression (148) is nothing but the differential of the characteristic function, such that **Jacobi** had already arrived at the other half of the integrals from one half of them along the path of the characteristic function. The generalization of that argument, namely, how it facilitated the relationship between the principal function and the canonical system, led to the realization that when one knows one first integral, the degree of the canonical system can be lowered by *two* units. The $(2n - 2)$ first integrals of the reduced system are all integrals of the original system, as well, such that *only one* integral, namely, the *conjugate* to the starting integral, will not appear among the integrals of the reduced system. That way of looking at things systematically reduces the integration of the given canonical system with $2n$ unknown functions to the search for a first integral of the integration of system of that sort with $(2n - 2)$ unknown functions, and then the search for a first integral of that reduced system reduces it to such an integration with $(2n - 4)$ unknown functions, etc. ([167]) (Cf., no. **27**)

Jacobi already had a wealth of further arguments for organizing those integration methods, which also essentially simplified the application to the perturbation calculations for astronomy. The **Jacobi** tradition has created a systematic theory from them ([168]).

([166]) **C. G. J. Jacobi**, "Sur le movement d'un point et sur un particulier cas du problème des trois corps," C. R. Acad. Sci. Paris **3** (1836), pp. 59 = *Werke IV*, pp. 35. Cf., also ([282]).

([167]) In the literature of the theory of first-order partial differential equations, that type of integration of the canonical system (the associated partial differential equation, resp.) is referred to as ***Jacobi**'s second method* [cf., II A 5 (**E. von Weber**), no. **36**].

([168]) The foundations of that research are defined by two presentations that go back to **Jacobi**: **C. G. J. Jacobi**, "Über diejenigen Probleme der Mechanik, in welchen eine Kräftefunktion existiert und über die Theorie der Störungen," *Werke V*, pp. 217 and **C. G. J. Jacobi**, *Vorlesungen über Dynamik*, *Werke Supplementband*. Both of them were published in the edition of **A. Clebsch**.

CHAPTER IV

THE VARIED ACTION.

16. Hamilton's principal function.

16.a The derivatives of the principal function. – The general analytical treatment of the equations of motion has been developed on the model of the *n*-body problem. Therefore, the preferred problems are the ones for which a *Lagrangian function exists:*

(148) $$L(\dot{q}_1,\ldots,\dot{q}_n,q_1,\ldots,q_n,t) = T - \Phi ,$$

in which T can be a general quadratic function of the velocity components $\dot{q}_1$, …, $\dot{q}_n$:

$$T = T_0 + T_1 + T_2$$

(cf., no. **4**). The *equations of motion*, as the *Euler equations of the variational problem* of **Hamilton**'s principle:

(149.a) $$\int L(\dot{q}_1,\ldots,\dot{q}_n,q_1,\ldots,q_n,t)\,dt = \text{extrem.}$$

will then possess the from:

(150) $$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\rho}\right) - \frac{\partial L}{\partial q_\rho} = 0 \qquad (\rho = 1, \ldots, n).$$

Starting from the proof of the existence of solution to a system of differential equations, which says that one can establish a solution in the neighborhood of a location by $2n$ givens – say, the initial position and velocity – one can, without always having to clearly address the difficulties associated with adapting a statement "in the small" to a corresponding one "in the large" ([169]), represent the solutions to the equations of motion (150) as a family of ∞^{2n} curves:

(150.a) $$q_\lambda = q_\lambda\,(t, c_1, \ldots, c_{2n}) \qquad (\lambda = 1, \ldots, n)$$

in the space-time manifold of the $(q_1, \ldots, q_n, t)$. From that standpoint, the individual extremal space-time lines are generally determined uniquely in such a way that one prescribes two of its space-time points:

([169]) Cf., on this, **T. Levi-Civita**, "Drei Vorlesungen über adiabatische Invarianten," Hamburg Abhandl. aus. d. math. Sem. **6** (1928), pp. 323.

(151) $$P_1(q_1^{(1)},\ldots,q_n^{(1)},t_1) \qquad \text{and} \qquad P_2(q_1^{(2)},\ldots,q_n^{(2)},t_1)\ .$$

If one now constructs the integral (149.a) between the two space-time points P_1 and P_2 along that extremal as an integration path:

(152) $$S(P_1, P_2) = \mathcal{E}\int_{P_1}^{P_2} L\,dt$$

then one will have *Hamilton's principal function* for the variational problem (149.a), which is regarded as a function of the point-pair $\{P_1, P_2\}$ [as a function of its $(2n + 2)$ coordinates:

(152.a) $$S(q_1^{(2)},\ldots,q_n^{(2)},t_2\,;q_1^{(1)},\ldots,q_n^{(1)},t_1) = \mathcal{E}\int_{P_1}^{P_2} L\,dt\ ,$$

resp.], in the spirit of no. **8**. When one goes to a "neighboring" point-pair:

$$P_1^*(q_1^{(1)}+\delta q_1^{(1)},\ldots,q_n^{(1)}+\delta q_n^{(1)},t_1+\delta t_1)$$

and

$$P_2^*(q_1^{(2)}+\delta q_1^{(2)},\ldots,q_n^{(2)}+\delta q_n^{(2)},t_2+\delta t_2),$$

the change in S according to (71.a) will be:

(153) $$\delta S = (q_1^{(2)}\delta q_1^{(2)}+\ldots+q_n^{(2)}\delta q_n^{(2)}-(H)_{P_2}\delta t_2)-(q_1^{(1)}\delta q_1^{(1)}+\ldots+q_n^{(1)}\delta q_n^{(1)}-(H)_{P_1}\delta t_1)\,,$$

from which one can infer the *derivatives of the principal function* corresponding to (137):

(154) $$\begin{cases} \dfrac{\partial S}{\partial q_1^{(2)}} = p_1^{(2)}, \quad \ldots, \quad \dfrac{\partial S}{\partial q_n^{(2)}} = p_n^{(2)}, \quad \dfrac{\partial S}{\partial t_2} = -(H)_{P_2}, \\ \dfrac{\partial S}{\partial q_1^{(1)}} = p_1^{(1)}, \quad \ldots, \quad \dfrac{\partial S}{\partial q_n^{(1)}} = p_n^{(1)}, \quad \dfrac{\partial S}{\partial t_1} = +(H)_{P_1}\,. \end{cases}$$

The derivatives of the principal function with respect to the position coordinates are the impulse components, and its derivative with respect to time is energy ([169.a]).

([169.a]) The change in sign in the formulas in the second of (154) is explained by the fact that one leaves the interval P_1P_2 from the point P_2 as t increases, while one enters it from the point P_1 .

16.b. The reciprocity theorems. – If one regards the impulse components in (154) as functions of the coordinates of the two boundary points (151) ([169.b]) then that will give *two classes of reciprocity relations* ([169.c]) that correspond to the two classes that **W. R. Hamilton** had developed in ray optics (cf., no. **13**). One gets the relations of the *first* class when one focuses on one of the two boundary points and considers the influence of the changes in position coordinates (time, resp.) on the impulse components (energy. resp.). For the boundary point P_2, they will read:

$$\frac{\partial p_\rho^{(2)}}{\partial q_\sigma^{(2)}} = \frac{\partial p_\sigma^{(2)}}{\partial q_\rho^{(2)}} = \frac{\partial^2 S}{\partial q_\rho^{(2)}\,\partial q_\sigma^{(2)}}\,, \tag{155}$$

and

$$\frac{\partial p_\rho^{(2)}}{\partial t_2} = -\frac{\partial (H)_{P_2}}{\partial q_\rho^{(2)}} = \frac{\partial^2 S}{\partial q_\rho^{(2)}\,\partial t_2}\,, \tag{156}$$

resp., in which:

$$\frac{\partial (H)_{P_2}}{\partial q_\rho^{(2)}} = \left(\frac{\partial H}{\partial p_1}\right)_{P_2}\frac{\partial p_1^{(2)}}{\partial q_\rho^{(2)}} + \cdots + \left(\frac{\partial H}{\partial p_n}\right)_{P_2}\frac{\partial p_n^{(2)}}{\partial q_\rho^{(2)}} + \left(\frac{\partial H}{\partial q_\rho}\right)_{P_2}. \tag{156.a}$$

Corresponding relations are true for the boundary point P_1 .

One will arrive at the *second* class of reciprocity theorems when one asks how the change in a position coordinate (time, resp.) at one boundary point influences any of the impulse components or the energy at the other boundary point. That will lead to the relations:

$$\frac{\partial p_\rho^{(2)}}{\partial q_\sigma^{(1)}} = -\frac{\partial p_\sigma^{(1)}}{\partial q_\rho^{(2)}} = \frac{\partial^2 S}{\partial q_\rho^{(2)}\,\partial q_\sigma^{(1)}}\,, \tag{157}$$

and

$$\frac{\partial p_\rho^{(2)}}{\partial t_1} = \frac{\partial (H)_{P_1}}{\partial q_\rho^{(2)}} = \frac{\partial^2 S}{\partial q_\rho^{(2)}\,\partial t_1}\,, \qquad \frac{\partial p_\rho^{(1)}}{\partial t_2} = \frac{\partial (H)_{P_2}}{\partial q_\rho^{(1)}} = -\frac{\partial^2 S}{\partial q_\rho^{(1)}\,\partial t_2}\,, \tag{158}$$

resp., in which:

$$\frac{\partial (H)_{P_1}}{\partial q_\rho^{(2)}} = \left(\frac{\partial H}{\partial p_1}\right)_{P_1}\frac{\partial p_1^{(1)}}{\partial q_\rho^{(2)}} + \cdots + \left(\frac{\partial H}{\partial p_n}\right)_{P_1}\frac{\partial p_n^{(1)}}{\partial q_\rho^{(2)}}\,, \tag{158.a}$$

and $\dfrac{\partial (H)_{P_2}}{\partial q_\rho^{(1)}}$ has the corresponding meaning.

In so doing, the position coordinates (time, resp.) are thought of as arbitrary quantities at the boundary points, corresponding to the definition of the principal function *S*. From the standpoint of mechanics, it is often more reasonable to prescribe the changes in the impulse components

---

([169.b]) By means of it, $(H)_{P_1}$, as well as $(H)_{P_2}$, will become functions of the $(2n + 2)$ coordinates of the boundary points.

([169.c]) **H. von Helmholtz**, "Über die physikalische Bedeutung des Prinzips der kleinsten Wirkung," J. f. Math. **100** (1887), pp. 137 and pp. 213, esp., pp. 213, *et seq.* = *Ges. Abh. III*, pp. 203, esp., pp. 238, *et seq.*

arbitrarily, which one can think of as being created by impacts. The given reciprocity relations will also remain preserved in that conception of things, except that one needs to interpret the partial derivatives as quotients of infinitesimal quantities, and in that sense, construct their “reciprocal

values" ([170]). With that, the reciprocity relations will take on a form that can be expressed in an especially intuitive way from the standpoint of mechanics ([171]).

([170]) In order to derive that, one applies the so-called **Legendre** transformation to the function S in the same way that **W. R. Hamilton** did in ray optics [cf., ([152])], i.e., one calculates the position coordinates $q_\rho^{(1)}$, $q_\rho^{(2)}$ of the two boundary functions as functions of the impulse components $p_\rho^{(1)}$, $p_\rho^{(2)}$, and the time coordinates t_1, t_2 [cf., **H. von Helmholtz**, *loc. cit.* ([169.c]), *Ges. Abh.*, pp. 246] from the $2n$ equations:

$$\frac{\partial S}{\partial q_\rho^{(2)}} = p_\rho^{(1)}, \qquad \frac{\partial S}{\partial q_\rho^{(1)}} = -\,p_\rho^{(1)},$$

and in that way define the function:

$$U(p_1^{(1)},\ldots,p_n^{(1)},t_1\,;p_1^{(2)},\ldots,p_n^{(2)},t_2) = \sum_{\rho=1}^{n}(p_\rho^{(2)}\,q_\rho^{(2)} - p_\rho^{(1)}\,q_\rho^{(1)}) - S\,. \tag{159}$$

One gets the associated position coordinates as derivatives of that function with respect to the impulse coordinates:

$$\frac{\partial U}{\partial p_\rho^{(2)}} = q_\rho^{(1)}, \qquad \frac{\partial U}{\partial p_\rho^{(1)}} = -\,q_\rho^{(1)}, \tag{160}$$

and for the derivatives with respect to t_1, t_2, one will get:

$$\frac{\partial U}{\partial t_2} = -\,\frac{\partial S}{\partial t_2} = (H)_{P_2}, \qquad \frac{\partial U}{\partial t_1} = -\,\frac{\partial S}{\partial t_1} = -\,(H)_{P_1}. \tag{161}$$

If one imagines that the position coordinates of the two boundary points $q_\rho^{(1)}$, $q_\rho^{(2)}$ are expressed as functions of the impulse components $p_\rho^{(1)}$, $p_\rho^{(2)}$ (and the time coordinates t_1, t_2) and also introduces those functions $(H)_{P_1}$ [$(H)_{P_2}$, resp.] then one will get the following equations as the *first* class of reciprocity relations:

$$\begin{cases} \dfrac{\partial q_\rho^{(2)}}{\partial p_\sigma^{(2)}} = \dfrac{\partial q_\sigma^{(2)}}{\partial p_\rho^{(2)}} = \dfrac{\partial^2 U}{\partial p_\rho^{(2)}\,\partial p_\sigma^{(2)}}, \\ \dfrac{\partial q_\rho^{(2)}}{\partial t_2} = \dfrac{\partial (H)_{P_2}}{\partial p_\rho^{(2)}} = \dfrac{\partial^2 U}{\partial p_\rho^{(2)}\,\partial t_2}, \end{cases} \tag{162}$$

while the following equations will be the reciprocity relations of the *second* class:

$$\begin{cases} \dfrac{\partial q_\rho^{(2)}}{\partial p_\sigma^{(1)}} = -\dfrac{\partial q_\sigma^{(1)}}{\partial p_\rho^{(2)}} = \dfrac{\partial^2 U}{\partial p_\rho^{(2)}\,\partial p_\sigma^{(1)}}, \\ \dfrac{\partial q_\rho^{(2)}}{\partial t_1} = -\dfrac{\partial (H)_{P_1}}{\partial p_\rho^{(2)}}, \quad \dfrac{\partial q_\rho^{(1)}}{\partial t_2} = -\dfrac{\partial (H)_{P_2}}{\partial p_\rho^{(1)}}. \end{cases} \tag{163}$$

Moreover, one can replace the time coordinates t_1 and t_2 of the boundary points in the function U with the energy values $(H)_{P_1}$ [$(H)_{P_2}$, resp.] by a further extension of the **Legendre** transformation. One would then also get the direct inversion of the relations (156) [(158), resp.].

16.c. The field of extremal space-time lines. – A *bundle of* ∞^n extremal *space-time lines* emanate from *every isolated space-time point* $P_0(q_1^{(0)},\ldots,q_n^{(0)},t_0)$, and indeed, on the basis of the existence theorems, one and only one of those curves will go through each space-time point P (q_1, …, q_n, t) that belongs to a certain neighborhood of P_0 and does not leave that neighborhood between P_0 and P. In such a bundle, one has the simplest example of a so-called *field of extremals*

It should be pointed out that **C. G. J. Jacobi** had already developed reciprocity relations that were basically equivalent to the ones that were given (cf., **C. G. J. Jacobi**, *Probleme der Mech.*, § 21, *Werke V*, pp. 315).

([171]) Those reciprocity theorems should not be confused with other reciprocity theorems that **H. von Helmholtz** had presented [cf., **H. von Helmholtz**, *loc. cit.* ([169.c]), *Ges. Abh. III*, pp. 231, as well as *Dynamik diskreter Massenpunkte*, pp. 373]. In them, **Helmholtz** considered a mechanical system with the **Lagrangian** function L. However, it did not move according to the **Euler** equations (150), but further external influences would be in effect, such that the equations:

(164) $$R_\rho = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot q_\rho}\right) - \frac{\partial L}{\partial q_\rho}$$

must enter in place of (150) in order to describe the motion. In that way, **Helmholtz** did not think of the quantities R_ρ as being prescribed, but likewise defined by equations (164), in which he made the further assumption that time t did not appear explicitly in L. The R_ρ will then seem to depend upon the position coordinates q_ρ themselves, their first derivatives $\dot q_\rho$, and their second derivatives $\ddot q_\rho$. If one considers the dependencies of those three types of quantities in succession then that will give three classes of reciprocity relations, namely:

(165) $$\frac{\partial R_\rho}{\partial \ddot q_\sigma} = \frac{\partial R_\sigma}{\partial \ddot q_\rho} = \frac{\partial^2 L}{\partial \dot q_\rho\,\partial \dot q_\sigma},$$

(165.a) $$\begin{cases} \dfrac{\partial R_\rho}{\partial \dot q_\sigma} - \dfrac{\partial R_\rho}{\partial \dot q_\sigma} = 2\left(\dfrac{\partial^2 L}{\partial \dot q_\rho\,\partial q_\sigma} - \dfrac{\partial^2 L}{\partial \dot q_\sigma\,\partial \dot q_\rho}\right), \\ \dfrac{\partial R_\rho}{\partial \dot q_\sigma} + \dfrac{\partial R_\rho}{\partial \dot q_\sigma} = 2\dfrac{d}{dt}\left(\dfrac{\partial^2 L}{\partial \dot q_\rho\,\partial \dot q_\sigma}\right) = 2\dfrac{d}{dt}\left(\dfrac{\partial R_\rho}{\partial \ddot q_\sigma}\right) = 2\dfrac{d}{dt}\left(\dfrac{\partial R_\sigma}{\partial \ddot q_\rho}\right), \end{cases}$$

(165.b) $$\frac{\partial R_\rho}{\partial q_\sigma} - \frac{\partial R_\rho}{\partial q_\sigma} = \frac{d}{dt}\left(\frac{\partial^2 L}{\partial \dot q_\rho\,\partial q_\sigma} - \frac{\partial^2 L}{\partial \dot q_\sigma\,\partial \dot q_\rho}\right) = \frac{1}{2}\frac{d}{dt}\left(\frac{\partial R_\rho}{\partial \dot q_\sigma} - \frac{\partial R_\rho}{\partial \dot q_\sigma}\right).$$

Naturally, the R_ρ are linear in the second derivatives $\ddot q_\lambda$, corresponding to their definition (164), such that one must have:

(166) $$\frac{\partial^2 R_\rho}{\partial \ddot q_\sigma\,\partial \ddot q_\tau} = 0\,.$$

Conversely, **Helmholtz** could also show that n functions $R_\rho(\ddot q_\sigma, \dot q_\sigma, q_\sigma)$ that satisfy the conditions (165, etc.) can be expressed in the form (164) in terms of a function $L(\dot q_1,\ldots,\dot q_n,q_1,\ldots,q_n)$. Cf., the treatise that was released by the estate to **L. Koenigsberger**: **H. von Helmholtz**, "Über die physikalische Bedeutung des Prinzips der kleinsten Wirkung," Berlin Sitzungsber. (1905), pp. 863, as well as **G. D. Birkhoff**, *Dynamical Systems*, pp. 26.

Physical examples of that reciprocity were given in **J. J. Thomson**, *Applications of dynamics to physics and chemistry*, Cambridge, 1888, German, Leipzig, 1890.

for the variational problem ([172]), in which the point P_0 itself might still be removed from that region in the space-time manifold. The space-time line (field extremal) that runs through an arbitrary space-time point P (q_1, …, q_n, t) of the field defines the velocity vector $(\dot{q}_1, \ldots, \dot{q}_n)$ there [the impulse vector (p_1, …, p_n) that is coupled with it, resp.] as functions of the "position," i.e., the velocity components $\dot{q}_\rho$, like the impulse components p_ρ, as functions of q_1, …, q_n, t :

$$\text{(167)} \qquad \begin{cases} \dot{q}_\rho = \dot{q}_\rho(q_1, \ldots, q_n, t), \\ p_\rho = p_\rho(q_1, \ldots, q_n, t). \end{cases}$$

If one constructs the *extremal integral of the variational problem* (149.a) along the individual extremal space-time lines in the field and chooses the lower limit to be the midpoint P_0 of the extremal bundle then the value of the extremal integral will become a function of only the upper limit P of the integral, i.e., a function of q_1, …, q_n, t :

$$\text{(168)} \qquad S(q_1, \ldots, q_n, t) = \mathcal{E} \int_{P_0}^{P} L\, dt\,,$$

whose differential is, from (153):

$$\text{(168.a)} \qquad \begin{aligned} \delta S &= \frac{\partial L}{\partial \dot{q}_1} \delta q_1 + \cdots + \frac{\partial L}{\partial \dot{q}_n} \delta q_n + \left(L - \dot{q}_1 \frac{\partial L}{\partial \dot{q}_1} - \cdots - \dot{q}_n \frac{\partial L}{\partial \dot{q}_n} \right) \delta t \\ &= p_1\, \delta q_1 + \ldots + p_n\, \delta q_n - H\, \delta t\,. \end{aligned}$$

In place of the bundle of space-time lines that emanate from a point, one can now use a *family of ∞^n extremal space-time lines, in general,* which might likewise have the property that one and only one space-time line of the family goes through every point in a certain region of the space-time manifold, such that the point (as well every space-time point) is assigned a *velocity vector* (*impulse vector*, resp.) as it was given in the formulas (167). However, such an n-parameter family of extremals is not generally a field in the same sense as the extremal family through a point.

If one selects any extremal space-time line from that family and chooses two space-time points P_1 and P_2 along it then one can define the principal function (152) for the pair of space-time points P_1, P_2. Now, if one imagines that a simple ([173]) closed curve C_1 has been drawn in the region of the space-time manifold in question then the extremal space-time lines of the n-parameter family that runs through the individual points of that curve will define the generators of a *tube* ([174]). If one draws a second closed curve C_2 on such a tube that twines around the tube in such a way that it likewise meets each of the extremal space-time lines of the tube only once then the two curves are in one-to-one correspondence with each other by way of the space-time lines that serve as

([172]) Cf., e.g., **O. Bolza**, *Vorlesungen über Variationsrechnung*, Leipzig, and Berlin, 1909, § 77, pp. 635.

([173]) That is intended to mean that an extremal space-time line through a point of the curve can have no other point in common with the curve.

([174]) For this subject, cf., **H. Poincaré**, *Mèthodes nouv. III*, pp. 4, *et seq.*

generators. Since any two associated points bound a segment of an extremal space-time, one can think of defining the principal function (152.a) for one such point-pair. The formula (153) then gives the change in the principal function when one goes from one generator to the neighboring generator, and in that way, the differentials $\delta q_1^{(1)}, \ldots, \delta q_n^{(1)}, \delta t_1$ for an advance along one of them and the ones $\delta q_1^{(2)}, \ldots, \delta q_n^{(2)}, \delta t_2$ for an advance along the other will both determine closed curves.

One can now go completely around the tube with the points P_1 and P_2 on the two closed curves until one again returns to the starting location. Since the associated total change in the function S is equal to zero, (153) will give the relation:

$$(169) \qquad \int_{C_2} [\, p_1^{(2)}\, \delta q_1^{(2)} + \cdots + p_n^{(2)}\, \delta q_n^{(2)} - (H)_{P_2}\, \delta t_2\,] - \int_{C_1} [\, p_1^{(1)}\, \delta q_1^{(1)} + \cdots + p_n^{(1)}\, \delta q_n^{(1)} - (H)_{P_1}\, \delta t_1\,] = 0\,,$$

in which the first integral is taken over the curve C_2, while the second is taken over the curve C_1. Since the curves C_1 and C_2 are twined around the tube in a completely arbitrary way, it will then follow that:

$$(170) \qquad \int_C (p_1\, \delta q_1 + \cdots + p_n\, \delta q_n - H\, \delta t) = \text{const.}$$

for the tube, which is a relation that is true in general for every tube through a closed curve in the region of the space-time manifold in question, since one can easily free it from the restrictions that were introduced later.

The *numerical value of the constant* in (170) can be interpreted very simply. In order to do that, one advances from a point on the surface of the tube in such a direction that the direction of advance will also be coupled with the direction of the extremal space-time lines by the relation:

$$(171) \qquad p_1\, \delta q_1 + \ldots + p_n\, \delta q_n - H\, dt = 0\,,$$

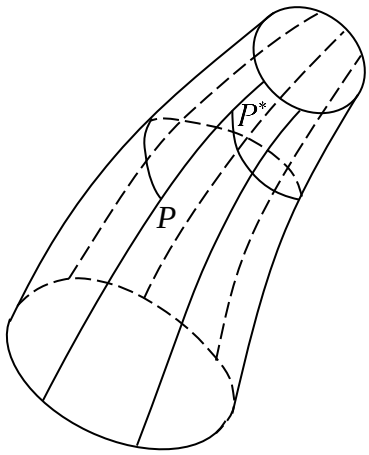

Figure 2.

so in a direction of advance that one cares to call *transversal* to the direction of the extremal in the calculus of variations [cf., II A 8.a (**E. Zermelo-H. Hahn**), no. **1**]. Under such a circuit around the tube, one will once more meet up with the extremal space-time line from which one started at the point P for the first time at the point P^*. If one extends the curve segment PP^*, which encircles the tube like a type of helix, with the segment PP^* of the extremal space-time line itself to a *closed curve* then equation (170) will be true for that closed curve. However, since the integrand in (170) will vanish on the thread PP^*, due to transversality, that thread will make no contribution to the integral (170). On the other hand, the integrand of (170) will reduce to the function L itself along the extremal space-time line

PP^*, such that the numerical value of the constant in formula (170) will be:

(170.a) $$\text{const.} = \mathcal{E}\int_{P^*}^{P} L\,dt\,.$$

That quantity is nothing but the *pitch of the screw* PP^* (as measured in the sense of the variational problem). In general, one then has that the curve that one will obtain when one advances along the tube transversally to the extremal space-time line (which serves as a generator) is a helix that possesses a *constant extremal pitch*. That should mean: If one selects two successive points on an arbitrary extremal space-time line on the surface of the tube where it intersects the transverse helix then **Hamilton**'s principal function will have the same constant value for such a point-pair everywhere on the tube, which one can refer to as the "tube constant."

If the tube constant is equal to zero, in particular, then when one encircles the tube transversally to the extremal generators, that will yield a *closed* curve instead of a helix. What is especially important is the case in which for *every* tube with extremal generators in the *n*-parameter family of space-time lines a transverse circuit will give a closed curve, so the constant in (170) will always be equal to zero. In the region that is filled up by the extremal space-time lines of the *n*-parameter family, the integral:

$$\int (p_1\,\delta q_1 + \cdots + p_n\,\delta q_n - H\,\delta t)$$

will obviously be independent of the integration path, and:

(172) $$p_1\,\delta q_1 + \ldots + p_n\,\delta q_n - H\,\delta t$$

will be a complete differential. Thus, the differential equation:

(173) $$p_1\,\delta q_1 + \ldots + p_n\,\delta q_n - H\,\delta t = 0$$

will determine a one-parameter family of *n*-dimensional manifolds in the *n*-parameter family of extremal space-time lines, and indeed according to the differential equation (173), those M_n will intersect the extremal space-time lines of the *n*-parameter family transversally. Any two of the M_n will cut out segments on all extremal space-time lines for which $\mathcal{E}\int L\,dt$ has the same value.

Moreover, the expression (172) for the $(n + 1)$ variables $q_1, \ldots, q_n, t$ will already be a complete differential when:

(174) $$p_1\,\delta q_1 + \ldots + p_n\,\delta q_n$$

is a complete differential for a manifold t = const. That is because since the integral (170) will have the same value for every closed curve that winds around a tube that is constructed from extremal space-time lines once, one needs to require the vanishing of the tube constant for only all closed curves in the manifold t = const. In particular, the expression (172) will then be a complete

differential for the *n*-parameter bundle of extremal space-time lines that emanate from a space-time point, because at the center of the bundle itself, the corresponding differential equation:

$$p_1\,\delta q_1 + \ldots + p_n\,\delta q_n = 0 \tag{174.a}$$

will, however, be certainly fulfilled, since $\delta q_1 = \ldots = \delta q_n = 0$.

Just like the family of extremals, one also refers to this special case of a general *n*-parameter family of extremal space-time lines as a *field* of extremals ([175]) in the event that one can construct a one-parameter family of transversals M_n in them. From (173), the transverse manifolds of the field are given by:

$$S\,(q_1, \ldots, q_n, t) = \int_{\Pi}^{P} (p_1\,\delta q_1 + \cdots + p_1\,\delta q_1 - H\,\delta t) = \text{const.}, \tag{175}$$

in which the integral is independent of the path of integration in the field, which extends from a fixed, but arbitrary, point $\Pi(q_1^{(0)},\ldots,q_n^{(0)},t_0)$ in the field to a varying point P (q_1, …, q_n, t). **D. Hilbert** ([176]) has referred to the significance of that integral that is independent of the path of integration, which one can also write in the form:

$$\int_{\Pi}^{P} \left[L + \frac{\partial L}{\partial \dot{q}_1}\left(\frac{\delta q_1}{\delta t} - \dot{q}_1 \right) + \cdots + \frac{\partial L}{\partial \dot{q}_n}\left(\frac{\delta q_n}{\delta t} - \dot{q}_1 \right) \right] \delta t\,. \tag{176}$$

It is then referred to as *Hilbert's independent integral*.

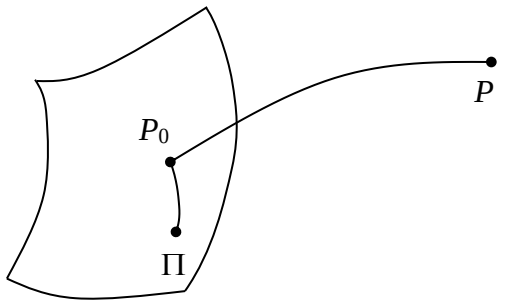

Figure 3.

The fact that the value of the path-independent integral in (175) is denoted by *S*, which is also the function symbol for the *extremal integral*, is actually justified. Namely, if one lays the transverse manifold through Π in the field the one can arrive at the arbitrarily-chosen point *P* in the field in the following way: The extremal space-time line of the field that runs through *P* meets the transverse manifold through Π at a point P_0. One then goes from Π to P_0 along the transverse manifold and then from P_0 to *P* along the space-time line of the field. Since the integral of (175) [(176), resp.] is equal to zero on the transverse manifold, while it reduces to *L* along the extremal space-time line of the field, for which:

([175]) and indeed a *field* in the narrow sense (independence field).

([176]) Cf., **D. Hilbert**, "Mathematische Probleme," talk given at the International Congress of Mathematicians, Paris, 1900. French translation in the C. R. du congrès, pp. 58. Original version in Gött. Nachr. (1900), pp. 253. The general theorem was first proved by **A. Mayer**, "Über den Hilbertschen Unabhängingkeitsatz in d. Theorie d. Max. u. Min. d. einf. Int.," Math. Ann. **58** (1904), pp. 235. Then **D. Hilbert**, "Zur Variationsrechnung," Gött. Nachr. (1905), pp. 159 = Math. Ann. **62** (1906), pp. 351. Moreover, the theorem was previously remarked in passing by **E. Beltrami**, "Sulla teoria delle line geodetiche," Mailand Rend. del istit. Lombardo (2) **1** (1868), pp. 708 = *Opere I*, pp. 366.

$$\frac{\delta q_1}{\delta t} = \dot{q}_1 , \qquad \ldots , \qquad \frac{\delta q_n}{\delta t} = \dot{q}_n ,$$

the function (175) will be precisely the *extremal integral* that is taken along the field extremal through P from P_0 to the point P :

$$\mathcal{E} \int_{P_0}^{P} L\, dt = S\,(q_1, \ldots, q_n, t) . \tag{177}$$

The coordinates $(q_1^{(0)}, \ldots, q_n^{(0)}, t_0)$ of the space-time point P_0 are regarded as functions of the coordinates $q_1, \ldots, q_n, t$ of the point P in the field in that.

The differential of that extremal integral is, in fact, equal to the expression (172). Namely, if one varies the point P in the field then the point P_0 will indeed vary with it, but since it remains on the transverse manifold, one will have:

$$p_1^{(0)}\, \delta q_1^{(0)} + \cdots + p_n^{(0)}\, \delta q_n^{(0)} - (H)_{P_0}\, \delta t_0 = 0 \tag{178}$$

for it, no matter what values the $\delta q_1^{(0)}, \ldots, \delta q_n^{(0)}, \delta t_0$ might take on as a result of the variations $\delta q_1, \ldots, \delta q_n, \delta t$. From (153), the differential of the function (177), like that of the function (168) of the field that emanates from a point, is:

$$\delta S = p_1\, \delta q_1 + \ldots + p_n\, \delta q_n - H\, \delta t . \tag{179}$$

Upon decomposing that relation, one will then get:

$$\frac{\partial S}{\partial q_1} = p_1 , \ \ldots , \qquad \frac{\partial S}{\partial q_n} = p_n , \qquad \frac{\partial S}{\partial t} = - H , \tag{180}$$

i.e., in fields of extremal space-time lines, the value of the extremal integral that is defined by a transversal M_n of the field will define a function $S\,(q_1, \ldots, q_n, t)$ whose level surfaces are the transversals M_n of the field and whose gradient will determine the impulse components p_ρ and the energy H in the field.

17. The Hamilton-Jacobi partial differential equation. – The elimination of the impulse components from equations (180) will yield the *first-order partial differential equation* for the function $S\,(q_1, \ldots, q_n, t)$ ([177]):

([177]) Since H is normally quadratic in the impulse components in mechanics, the partial differential equation would be quadratic in the $\partial S / \partial q_\rho$.

(181) $$\frac{\partial S}{\partial t} + H\left(\frac{\partial S}{\partial q_1}, \ldots, \frac{\partial S}{\partial q_n}, q_1, \ldots, q_n, t\right) = 0 ,$$

which coincides with one of the two differential equations that were given in the previous section [cf., no. **14**, eq. (138)] for **Hamilton**'s principal function. Only one of the two partial differential equations can appear here, since indeed one of the points here, namely P_0, seems to be established by the other one, namely the varying point P, while from the approach that **Hamilton** took in his mechanical treatises, both points of the point-pair that serves to delimit the extremal integral are thought of as independently varying, which was suggested by the notation in no. **16**: $S(q_1^{(2)}, \ldots, q_n^{(2)}, t_2, q_1^{(1)}, \ldots, q_n^{(1)}, t_1)$. Thus, in his treatises on *mechanics*, to a certain extent, **Hamilton** considered the set of ∞^{2n} extremal space-time lines simultaneously, while here, corresponding to the approach that **Hamilton** took in his treatises on *ray optics*, a family of ∞^n curves is selected from it that defines a *field*. In the *simplest case* of such a field, namely the extremals that go through a fixed point (so when one regards one of the two space-time points, namely the *lower limit of the integral*, as fixed:

(182) $$q_1^{(1)} = q_1^{(0)}, \quad \ldots, \quad q_n^{(1)} = q_n^{(0)}, \quad t_1 = t_0 ,$$

while the other one is regarded as varying in the field:

(182.a) $$q_1^{(2)} = q_1 , \quad \ldots, \quad q_n^{(2)} = q_n , \quad t_2 = t)$$

the transverse manifolds:

$$S(q_1, \ldots, q_n, t) = \text{const.}$$

will be the immediately generalizations of the light waves that belong to the bundle that emanates from a luminous point in ray optics. However, even in a general field of the type that was introduced in the previous section, the transverse manifolds can be spoken of as generalizations of the light waves, since indeed a bundle of light rays that is originally centered will become one such general field as long as it experiences reflections or refractions. Since **Jacobi** (cf., no. **15**) refused to regard **Hamilton**'s principal function as a function of two equally-justified points of a point-pair, in order to do that, he introduced the concept of a field with its extremal integral whose level-M_n would be the transversal manifolds to the field, and he arrived ([178]) at the function $S(q_1, \ldots, q_n, t)$, which is a solution to the *one* partial differential equation (181). That one equation is then referred to in the literature as the *Hamilton-Jacobi partial differential equation*.

([178]) **C. G. J. Jacobi**, "Über die Reduktion der Integr. d. partiellen Differentialgl. 1. Ordn. zwischen irgendeiner Zahl Variabl. auf die Integr. eines einzigen Systems gewöhnlicher Differentialgleichungen," J. f. Math. **17** (1837), pp. 97 = *Werke IV*, pp. 57, as well as **C. G. J. Jacobi**, *Vorlesungen über Dynamik*, *Werke Supplementband*, pp. 148, *et seq.*, in which **Jacobi** derived the partial differential equation (181) by initially considering "centered" fields whose extremal space-time lines all ran through a fixed space-time point and then asking what an arbitrary solution to that equation would mean. Cf., **C. G. J. Jacobi**, *Probleme der Mech.*, *Werke V*, pp. 240.

As was done in no. **15**, *Jacobi's approach* basically arose from the *converse of that result*, namely, the recognition that every solution of the *Hamilton-Jacobi* equation (181) must be obtained from a field of space-time lines in the given way ([178.a]). That is because, from the general theorem on the existence of solutions [cf., II A 5 (**E. von Weber**), no. **1**], a solution of the partial differential equation (181) is determined uniquely when it assumes a prescribed value on an n-dimensional manifold. Now, if any solution S (q_1, …, q_n, t) of (181) is known then one can determine the manifold on which $S = 0$ ([179]) and construct an n-parameter family of extremal space-time lines of the variational problem $\int L\,dt$ = extrem. that is *transverse to it* ([180]), which will obviously represent a *field*, since the condition (171) is indeed fulfilled on $S = 0$, and therefore the tube constant will be equal to zero for all tubes that are defined by extremal space-time lines. The function:

$$S^*(q_1,\ldots,q_n,t) = \mathcal{E}\int_{P_0}^{P} L\,dt$$

that is defined in the field is a solution to the partial differential equation (181) now that likewise assumes the value $S^* = 0$ on $S = 0$. Since S and S^* are identical on that n-dimensional manifold, from the existence theorem, they must be identical everywhere. As a result, the given solution $S(q_1, \ldots, q_n, t)$ will arise from the field of extremal space-time lines that was constructed.

([178.a]) That is the basis for the so-called *Jacobi first method* [II A 5 (**E. von Weber**), no. **30**] for integrating first-order partial differential equations. **C. G. J. Jacobi**, "Über die Reduktion…," *loc. cit.* ([178]), esp., *Werke IV*, pp. 100, *et seq.*

([179]) Or a suitable manifold S = const., resp., when $S = 0$ is not supposed to be favorable. Such a change will come about when one changes the function S by an additive constant, which is obviously inessential.

([180]) I.e., one constructs the extremal space-time line at each point of the manifold $S = 0$ whose direction is determined by the equations:

$$\frac{\partial L}{\partial \dot{q}_\rho} = \frac{\partial S}{\partial q_\rho} \qquad (\rho = 1, \ldots, n).$$

The construction of the field that belongs to a particular solution S (q_1, …, q_n, t) of the *Hamilton-Jacobi* equation requires the solution to a system of n first-order ordinary differential equations. Namely, if one replaces the **Euler** equations (150) with the associated canonical system:

$$\frac{dq_\rho}{dt} = \frac{\partial H}{\partial p_\rho}, \qquad \frac{dp_\rho}{dt} = -\frac{\partial H}{\partial q_\rho}$$

then since $p_\rho = \partial S / \partial q_\rho$, one will need only to integrate the n equations:

$$\frac{dq_\rho}{dt} = \frac{\partial H}{\partial p_\rho},$$

whose right-hand sides are now functions of q_1, …, q_n, t. For that, cf., **R. Lehmann-Filhés**, "Über die Verwendung unvollständing Integrale der Hamilton-Jacobischen part. Differentialgl.," Astron. Nachr. **165** (1904), pp. 209.

If one now starts from a centered field, so from a bundle of extremal space-time lines that go through a space-time point $P_0\,(q_1^{(0)},\ldots,q_n^{(0)},t_0)$, then one can assume that the center P_0 in the associated solution:

$$S\,(q_1,\ldots,q_n,t) = \mathcal{E}\int_{P_0}^{P} L\,dt$$

of the **Hamilton-Jacobi** equation (181) is not fixed, but variable. The coordinates $(q_1^{(0)},\ldots,q_n^{(0)},t_0)$ of the center enter into the function S as constants then, and one will get a solution to the **Hamilton-Jacobi** equation that depends upon $(n+1)$ constants in the form of $S(q_1,\ldots,q_n,t;q_1^{(0)},\ldots,q_n^{(0)},t_0)$. That is a *complete* solution of that equation [cf., II A 5 (**E. von Weber**), no. **7**], in which one constant is superfluous ([181]). One will then get a complete solution with the necessary number of n arbitrary constants when one gives a *fixed* value to – say – t_0 and then regards $q_1^{(0)}$, …, $q_n^{(0)}$ as chosen arbitrarily from the n-dimensional manifold $t = t_0$. The set of ∞^{2n} extremal space-time lines then seems to be divided into ∞^n bundles of ∞^n space-time lines whose centers are the ∞^n points of a manifold t = const.

Now, if arbitrary parameters c_1, …, c_n enter in place of the coordinates $q_1^{(0)}$, …, $q_n^{(0)}$ of the center of the bundle then:

$$S = S\,(q_1,\ldots,q_n,t,c_1,\ldots,c_n) \tag{182}$$

will represent the most general Ansatz for a *complete solution to the Hamilton-Jacobi equation*. However, since the function (182) can likewise be interpreted as the value of an extremal integral for a field whose extremals are transverse to an initial M_n when the parameters c_1, …, c_n possess fixed values, the complete solution of the most general kind (182), just like the special one, will find its interpretation in a family of ∞^n fields, except that now a family of ∞^n n-dimensional manifolds will enter in place of the centers of the bundle, each of which can serve as the initial M_n for a field of extremals that are transverse to it ([182]). The set of ∞^{2n} extremal space-timelines also seems to be once more divided into ∞^n *fields* here, in the same way that they were divided into ∞^n bundles with centers in a manifold t = const. above. In one such chosen field that belongs to fixed values of the parameters c_1, …, c_n, a certain extremal space-time line will now go through a space-time point $P\,(q_1,\ldots,q_n,t)$ and determine the point P_0 as the point of intersection with the transverse manifold $S = 0$, such that for every choice of the c_1, …, c_n, the coordinates of the point P_0 will be functions of the coordinates q_1, …, q_n, t of the point P. One will then have:

$$\left\{\begin{array}{c} q_1^{(0)} = q_1^{(0)}(q_1,\ldots,q_n,t,c_1,\ldots,c_n),\quad \ldots,\quad q_n^{(0)} = q_n^{(0)}(q_1,\ldots,q_n,t,c_1,\ldots,c_n), \\ t_0 = t_0(q_1,\ldots,q_n,t,c_1,\ldots,c_n). \end{array}\right. \tag{183}$$

([181]) Cf., **C. G. J. Jacobi**, *Probleme der Mechanik*, *Werke V*, esp., pp. 309.
([182]) Cf., **D. Hilbert**, "Zur Variationsrechnung," Gött. Nachr. (1905), esp., pp. 170.

By definition, the point P_0 *will remain unchanged when one replaces the point P with another point of the same extremal space-time line of the field*.

If one now interprets the function $S\,(q_1, \ldots, q_n, t, c_1, \ldots, c_n)$ as the value of the extremal integral of the field:

$$S\,(q_1, \ldots, q_n, t, c_1, \ldots, c_n) = \mathcal{E}\int_{P_0}^{P} L\,dt \tag{184}$$

then one can easily infer the partial derivatives of S with respect to the $c_1, \ldots, c_n$ from the boundary formula for the calculus of variations. That is because if one fixes the upper limit of the integral then the boundary formula will imply the relation:

$$\delta S = -\,p_1^{(0)}\,\delta q_1^{(0)} - \cdots - p_n^{(0)}\,\delta q_n^{(0)} + H\,(p_1^{(0)}, \ldots, p_n^{(0)}, q_1^{(0)}, \ldots, q_n^{(0)}, t_0)\,\delta t_0, \tag{184.a}$$

in which $p_1^{(0)}, \ldots, p_n^{(0)}$ are the impulse components at the point P_0 that belongs to the extremal space-time line that runs through it. They are also functions of $q_1, \ldots, q_n, t, c_1, \ldots, c_n$:

$$p_1^{(0)} = p_1^{(0)}\,(q_1, \ldots, q_n, t, c_1, \ldots, c_n)\,, \ldots,\; p_n^{(0)} = p_n^{(0)}\,(q_1, \ldots, q_n, t, c_1, \ldots, c_n)$$

that have the property that they will not change when the point $P\,(q_1, \ldots, q_n, t)$ on the extremal of the field in question that goes through P_0 changes its position.

The n relations:

$$\frac{\partial S}{\partial c_\rho} = -\,p_1^{(0)}\,\frac{\partial q_1^{(0)}}{\partial c_\rho} - \cdots - p_n^{(0)}\,\frac{\partial q_n^{(0)}}{\partial c_\rho} + H\,(p_1^{(0)}, \ldots, p_n^{(0)}, q_1^{(0)}, \ldots, q_n^{(0)}, t_0)\,\frac{\partial t_0}{\partial c_\rho} \qquad (\rho = 1, \ldots, n) \tag{185}$$

follow immediately from (184.a). However, from what was said, the right-hand sides will now remain unchanged when the space-time point $P\,(q_1, \ldots, q_n, t)$ varies along one and the same extremal of a field, such that one will have:

$$\frac{\partial S}{\partial c_1} = \gamma_1\,, \qquad \ldots, \qquad \frac{\partial S}{\partial c_n} = \gamma_n \tag{186}$$

for the individual extremal space-time lines of the fields, in which $\gamma_1, \ldots, \gamma_n$ are understood to mean constants. Conversely, *those n equations* (186) *can be spoken of as the finite equations of the extremal space-time lines*. Geometrically, that means that the individual points P of an extremal space-time line, as the intersections of the $(n + 1)$ manifolds S = const. that belong to the parameter values $(c_1, \ldots, c_n)$, $(c_1 + \Delta c_1, c_2, \ldots, c_n)$, …, $(c_1, \ldots, c_n + \Delta c_n)$, resp., and that one will give the entire space-time line when one appeals, not to the $(n + 1)$ manifolds S = const. = λ (with a fixed value of λ), but to the $(n + 1)$ paths in M_n with varying λ . With that, the determination of the light ray

as the intersection of three wave-trains (cf., no. **13**) is generalized to the general problem of determining the space-time lines of the motion.

Along with the finite equations (186) of the extremal space-time lines, which **Jacobi** referred to as *second integrals* of the equations of motion, one can also infer the impulse components p_ρ (velocity components $\dot{q}_\rho$, resp.) from the complete solution S (q_1, …, q_n, t, c_1, …, c_n). That is because, along with (185), the boundary formula will further imply the relations:

$$\frac{\partial S}{\partial q_1} = p_1 = \frac{\partial L}{\partial \dot{q}_1}, \ldots, \frac{\partial S}{\partial q_n} = p_n = \frac{\partial L}{\partial \dot{q}_n}, \tag{187}$$

which couple the components of the impulse vector (velocity vector, resp.) with the position coordinates and time, and represent *n first integrals* of the equations of motion, with **Jacobi**'s terminology. The n first integrals (187) and the n second integrals (186), with their $2n$ arbitrary constants c_1, …, c_n, γ_1, …, γ_n, will then produce the general solution to the equations of motion. With that, one has arrived at a connection between the solutions of the equations of motion (150) and a complete solution to the **Hamilton-Jacobi** partial equation (181) that **Jacobi** expressed by saying that: *One will get the solution to the system of equations of motion from a complete solution of the Hamilton-Jacobi partial differential equation by mere differentiations and eliminations using equations* (186) *and* (187) ([183]).

18. Simplifying the Hamilton-Jacobi equation when an integral of the equations of motion is known. –

([183]) **C. G. J. Jacobi**, *Vorlesungen*, *Werke Supplement-Bd.*, pp. 157, cf., also *Probleme der Mech.*, *Werke V*, pp. 240. The theorem was first expressed in the treatise by **C. G. J. Jacobi**, "Über die Reduktion der Integration der partiellen Differentialgleichungen erster Ordnung zwischen irgendeinen Zahl Variablen auf die Integration eines einzigen Systems gewöhnlicher Differentialgleichungen," J. f. Math. **17** (1837), pp. 97 = *Werke IV*, pp. 57, cf., esp., pp. 71.

If one has an incomplete solution to the **Hamilton-Jacobi** equations that includes less than n constants:

(a) $$S = S\,(q_1, \ldots, q_n, t, c_1, \ldots, c_r) \qquad (r < n)$$

then the derivatives of S with respect to the c_1, …, c_r along the individual extremals will be:

(b) $$\frac{\partial S}{\partial c_1} = \gamma_1, \ldots, \frac{\partial S}{\partial c_r} = \gamma_r .$$

For every choice of the c_1, …, c_r, from ([180]), the construction of the field whose extremal integral is the function (a) will require the integration of the system of n ordinary differential equations:

$$\frac{dq_r}{dt} = \frac{\partial H}{\partial p_\rho}.$$

However, one already known the r integrals (b) of that system. Cf., **R. Lehmann-Filhés**, *loc. cit.* ([180]).

18.a. Cyclic coordinates and the energy integral. – The **Hamilton-Jacobi** partial differential equation simplifies when one addresses a *problem with one cyclic coordinate*, say, the coordinate q_n . Namely, since the coordinate q_n will not appear in the function H then, as was shown before in no. **9**, then obviously:

$$S = c_n\, q_n + S^*(q_1,\ldots,q_{n-1},t,c_1,\ldots,c_{n-1}) \tag{188}$$

will represent a complete solution to the **Hamilton-Jacobi** partial differential equation (181), assuming that S^* is determined to be a complete solution of the simplified partial differential equation:

$$\frac{\partial S^*}{\partial t} + H\left(\frac{\partial S^*}{\partial q_1},\ldots,\frac{\partial S^*}{\partial q_{n-1}},c_n,q_1,\ldots,q_{n-1},t\right) = 0\,. \tag{189}$$

However, *this partial differential equation* (189) can once more be regarded as the ***Hamilton-Jacobi*** *equation for a new variational problem that it determines*, and indeed it will follow precisely from the variational problem:

$$\int L^*(\dot q_1,\ldots,\dot q_{n-1},c_n,q_1,\ldots,q_{n-1},t)\,dt = \text{extrem.}, \tag{190}$$

in which

$$\begin{aligned} L^* &= \sum_{\rho=1}^{n-1} p_\rho\,\dot q_\rho - H(p_1,\ldots,p_{n-1},c_n,q_1,\ldots,q_{n-1},t) \\ &= L - p_n\,\dot q_n = L - c_n\,\dot q_n \end{aligned} \tag{190.a}$$

is the function that arises from L by the *Routh-Helmholtz transformation*, as in no. **9** ([184]). If one interprets the function S that was written out in (188) (for fixed values of the constants c_1, …, c_n)

([184]) One gets the equations of the space-time lines from (188) in the form:

$$\frac{\partial S^*}{\partial c_1} = \gamma_1\,,\ \ldots,\quad \frac{\partial S^*}{\partial c_{n-1}} = \gamma_{n-1}\,,\quad q_n + \frac{\partial S^*}{\partial c_n} = \gamma_n\,, \tag{188.a}$$

to which the relations

$$p_1 = \frac{\partial S^*}{\partial q_1}\,,\ \ldots,\quad p_{n-1} = \frac{\partial S^*}{\partial q_{n-1}}\,,\quad p_n = c_n \tag{188.b}$$

will then be added. The last of them is the first integral that belongs to the cyclic coordinate q_n . The first $(n-1)$ of equations (188.a) represent the extremals of the variational problem (190) that are given by projecting the extremals of the variational problem $\int L\,dt$ = extrem. in the q_n-direction. The last of equations (188.a) then implies such a projection onto the associated value of q_n at each point, which will show that the additive constant belongs to a projection onto ∞^1 space-time lines.

as the value of the extremal integral of a field of extremals of the variational problem $\int L\, dt$ = extrem. then obviously all transversal manifolds S = const. of the field will emerge from one of them by a parallel displacement in the q_n-direction, since two manifolds $S = C_1$ and $S = C_2$ will possess the constant distance $\Delta p_n = (C_2 - C_1) / c_n$. Just as in no. **9**, when q_n is a cyclic coordinate, the set of all space-time lines will go to itself under parallel translation in the q_n-direction, so *every individual field*, along with its transverse manifolds, will also go to itself under a parallel translation in the q_n-direction.

From no. **9**, the ∞^{n-1} projections of the ∞^n extremal space-times of the field in question in the q_n-direction will produce a family of ∞^{n-1} extremals of the variational problem (190) $\int L^*\, dt$ = extrem. Those ∞^{n-1} projections are a *field* of the simplified variational problem in their own right. That is because the intersection of the transverse manifold S (q_1, …, q_n, t) = const. of the original field with $q_n = 0$ ([185]):

$$S^*(q_1, \ldots, q_{n-1}, t) = \text{const.}, \tag{191}$$

and since that function S^* is a solution of the **Hamilton-Jacobi** partial differential equation (189) that belongs to simplified variational problem (190), the manifolds (191) will be a one-parameter family of M_{n-1} that can define the transverse manifolds of a field of the simplified variational problem (190). However, since one has:

$$\frac{\partial L^*}{\partial \dot{q}_\rho} = \frac{\partial L}{\partial \dot{q}_\rho} = p_\rho \qquad (\rho = 1, \ldots, n-1),$$

from (190.a), on the basis of (188.b) in footnote ([184]), the projections of the extremals space-time liness of the initial field in the sense of new variational problem will be transverse to the M_{n-1} (191), so they will then define precisely the field that belongs to them. That connection, which is easy to predict from the standpoint of the **Hamilton-Jacobi** partial differential equation, make the convenience of the *Routh-Helmholtz* transformation when cyclic coordinates are present immediately understandable.

An entirely-analogous simplification of the **Hamilton-Jacobi** partial differential equation will come about when the function L (H, resp.) does not include *time t explicitly*. One will obviously have that:

$$S = -k\, t + V(q_1, \ldots, q_n) \tag{192}$$

is a solution to (181) then, as long as V is a solution of the partial differential equation:

([185]) One will also get those intersection manifolds when one determines the intersections M_{n-1} of an individual n-dimensional transverse manifold S = const. with all manifolds q_n = const. and projects them onto the M_n of the q_1, …, q_{n-1}, t in the q_n-direction.

(193) $$H\left(\frac{\partial V}{\partial q_1},\ldots,\frac{\partial V}{\partial q_n},q_1,\ldots,q_n\right)=k\,,$$

and indeed (192) will be a complete solution when the function $V(k, q_1, \ldots, q_n, c_1, \ldots, c_{n-1})$ includes precisely $(n-1)$ arbitrary constants $c_1, \ldots, c_{n-1}$, so it is a complete solution of (193) in its own right.

The partial differential equation now represents the **Hamilton-Jacobi** equation of a *variational problem in parametric form* ([186]), and indeed, according to (85), that variational problem has the form ([187]):

(194) $$\int\left\{\sum_{\sigma=1}^{n}g_{0\sigma}\,q'_\sigma+\sqrt{\left[2k-\left(\Phi-\tfrac{1}{2}\,g_{00}\right)\right]}\sum_{\rho,\sigma=1}^{n}g_{\rho\sigma}\,q'_\rho\,q'_\sigma\right\}du\;=\text{extrem.},$$

$$q'_\rho\;=\;\frac{dq_\rho}{du}\,,$$

which will simplify to ([187.a]):

(194.a) $$\int\sqrt{2\,(k-\Phi)\sum_{\lambda,\mu=1}^{n}g_{\lambda\mu}\,q'_\lambda\,q'_\mu}\;du\;=\text{extrem.}$$

in the special case. Conversely, the partial differential equation (193) is the **Hamilton-Jacobi** differential equation of that variation problem. The **Hamilton-Jacobi** equation then immediately implies that *when the energy integral is valid, Hamilton's principle can be replaced with the variational principle for the trajectories*, namely, ***Jacobi**'s principle* ([188]). The conceptual arguments in no. **10** obviously have their roots in that connection between the two equations (181) and (193).

([186]) Thus, it is a variational problem:

$$\int f\left(\frac{dq_1}{du},\ldots,\frac{dq_n}{du},q_1,\ldots,q_n\right)du\;=\text{extrem.},$$

in which f is homogeneous of degree one in the $dq_1 / du, \ldots, dq_n / du$.

In order to calculate the integrand f, one must put equation (193) into the form:

$$H^*\left(\frac{\partial V}{\partial q_1},\ldots,\frac{\partial V}{\partial q_n},q_1,\ldots,q_n,k\right)=H-k=0$$

and then proceed as in ([99]).

([187]) In that way, H will have the form (74.b), but time t cannot appear explicitly in $g^{\rho\sigma}$, $g_{0\rho}$, g_{00}, and Φ.

([187.a]) Where H possesses the simple form (74.a) in which t cannot appear in $g^{\rho\sigma}$ and Φ, as before.

([188]) As was shown in no. **10**, ***Euler**'s principle*:

$$\int 2T\,dt\;=\text{extrem.}$$

is then equivalent to the auxiliary condition:

$$T+\Phi=\text{const.}$$

If the energy integral is valid for the mechanical problem in question then one correspondingly cares to write out the **Hamilton-Jacobi** equation in the simplified form (193) from the outset and start with a "complete" integral:

$$V = V(q_1, \ldots, q_n, k, c_1, \ldots, c_{n-1}) \tag{195}$$

of that simplified equation (193) directly. The integral of the equations of motion can then be written in the form ([189]):

$$\left\{\begin{array}{l} \dfrac{\partial V}{\partial q_1} = p_1 = \dfrac{\partial L}{\partial \dot{q}_1}, \ldots, \dfrac{\partial V}{\partial q_n} = p_n = \dfrac{\partial L}{\partial \dot{q}_n}, \\ \dfrac{\partial V}{\partial c_1} = \gamma_1, \ldots\ldots\ldots\ldots, \dfrac{\partial V}{\partial c_{n-1}} = \gamma_{n-1}, \dfrac{\partial V}{\partial k} = t - \tau, \end{array}\right. \tag{195.a}$$

instead of the form (186) and (187), with the help of *this characteristic function V*, in which τ is an arbitrary constant ([190]).

The simplification of the **Hamilton-Jacobi** partial differential equation that comes about in the two cases considered is obviously coupled with the fact in each case, one can find a first integral of the equations of motion (the associated canonical system, resp.). If q_n is, in fact, a cyclic coordinate then equations of motion will possess the first integral:

$$p_n = \text{const.} = c_n .$$

If one replaces the impulse component with the derivative of the function S that it equals here then one will see that, in addition to the **Hamilton-Jacobi** equation (181), the function S must satisfy the further first-order partial differential equation:

$$\frac{\partial S}{\partial q_n} = c_n .$$

Those two partial differential equations will be fulfilled immediately by the Ansatz (188).

It follows analogously from the energy integral:

$$H(p_1, \ldots, p_n, q_1, \ldots, q_n) = k$$

that in addition to the **Hamilton-Jacobi** equation (181), the function S must satisfy the further partial differential equation:

([189]) Cf., **C. G. J. Jacobi**, *Vorlesungen*, *Werke Supplementband*, pp. 167.

([190]) It will enter in place of γ_n in the same way that k enters in place of c_n .

$$H\left(\frac{\partial S}{\partial q_1},\ldots,\frac{\partial S}{\partial q_n},q_1,\ldots,q_n\right)=k\,.$$

The Ansatz (192) for S obviously employs the fact that the further partial differential equation for S:

$$\frac{\partial S}{\partial t}+k\;=0$$

emerges immediately from those two partial differential equations, and therefore the system of the two partial differential equations:

$$\text{(196)}\qquad \begin{cases} \dfrac{\partial S}{\partial t}+H\left(\dfrac{\partial S}{\partial q_1},\ldots,\dfrac{\partial S}{\partial q_n},q_1,\ldots,q_n\right)=0, \\ H\left(\dfrac{\partial S}{\partial q_1},\ldots,\dfrac{\partial S}{\partial q_n},q_1,\ldots,q_n\right)=k \end{cases}$$

can be replaced with the system:

$$\text{(196.a)}\qquad \begin{cases} \dfrac{\partial S}{\partial t}+H\left(\dfrac{\partial S}{\partial q_1},\ldots,\dfrac{\partial S}{\partial q_n},q_1,\ldots,q_n\right)=0, \\ \dfrac{\partial S}{\partial t}+k=0, \end{cases}$$

in which time t appears as an analogue of a cyclic coordinate and makes it clear that any field of extremal space-time lines, along with its transverse M_n, will go to itself under the one-parameter group of parallel translations in the t-direction.

Meanwhile, in order for the argument that was posed for the energy integral to be adaptable to an arbitrary first integral of the equations of motion (150), it is necessary to appeal to the system (196) itself directly and regard the second equation in it as the **Hamilton-Jacobi** equation of a parametric problem of the calculus of variations. From no. **10**, that is just the variational problem of the **Jacobi** principle of least action, whose extremals are trajectories. Now, it is important that the cylindrical M_2 with generators parallel to the t-axis that projects ∞^1 space-time lines onto a trajectory can also be generated by this variational problem. That is because each cylindrical M_2 indeed intersects the manifolds t = const. in congruent curves that one can regard as extremals of the variational problem for **Jacobi**'s principle. One will also get the cylindrical M_2 then when one draws curves through the individual points of an extremal space-time line in the manifold t = const. that are congruent to the associated trajectory. If, as before [in the spirit of (196.a)], the individual M_2 are thought of as being covered by a net of ∞^1 space-time lines of the field and ∞^1 parallels to the t-axis then one can think of them, in the spirit of the new viewpoint [corresponding to (196)],

as being covered by ∞^1 space-time lines of the field and ∞^1 curves in the manifolds t = const., that are all congruent to the associated trajectory (and might also be called trajectories, for brevity).

Now, if the extremal integral of the parameter problem is denoted by V then the field of extremal space-time lines with its transverse manifolds will go to itself when one goes forward along all "trajectories" by the same segment V, which would follow immediately from (192). If one thinks of V as variable then one will obviously have a one-parameter group of transformations that transform the field into itself just like the one-parameter group of parallel translations in the t-direction.

18.b. Existence of an arbitrary first integral. – If one knows *any first integral* of the equations of motion (150):

$$F(\dot{q}_1,\ldots,\dot{q}_n,q_1,\ldots,q_n,t) = c \tag{197}$$

then when one replaces velocity components with the impulse components, one can give it the form:

$$G(p_1,\ldots,p_n,q_1,\ldots,q_n,t) = c\ . \tag{198}$$

Now, if one has a field of extremal space-time lines, each of which has the same numerical value for the constant in (198), then the associated solution $S(q_1,\ldots,q_n,t)$ to the **Hamilton-Jacobi** equation (181) must likewise satisfy the first-order partial differential equation:

$$G\left(\frac{\partial S}{\partial q_1},\ldots,\frac{\partial S}{\partial q_n},q_1,\ldots,q_n,t\right) = c\ . \tag{198.a}$$

In a way that is analogous to how the equation arose from the energy integral in no. **18**.a, *that equation can be employed to simplify the Hamilton-Jacobi equation and therefore the variational problem.* Namely, if one regards t as a constant in (198.a) ([191]) then the partial differential equation (198.a) can be regarded as the **Hamilton-Jacobi** differential equation of a parametric problem in the calculus of variations in each manifold t = const., which might have the form ([192]):

$$\int g\left(\frac{dq_1}{du},\ldots,\frac{dq_n}{du},q_1,\ldots,q_n,t\right)du = \text{extrem.} \tag{199}$$

([191]) In the literature, as a rule, one treats the case in which t does not appear explicitly in H, such that one can consider the trajectories directly in place of the extremal space-time lines of the motion. One will correspondingly have only integrals (197) [(198), resp.] in which time t does not appear explicitly.

([192]) The integrand g in that is calculated from G in the way that was given in ([99]). t appears in the integrand as a variable parameter, which corresponds to the fact that one has such a variational problem (199) for every manifold t = const.

The **Euler** equations for the problem ([193]):

(199.a) $$\frac{d}{du}\left(\frac{\partial g}{\partial q'_\rho}\right) - \frac{\partial g}{\partial q_\rho} = 0 \qquad \left(q'_\rho = \frac{dq_\rho}{du}\right)$$

$$(\rho = 1, \ldots, n)$$

can, by introducing the associated "impulse" ([194]):

(200) $$\overline{p}_\rho = \frac{\partial g}{\partial q'_\rho},$$

be put into the canonical form ([195]):

(201) $$\frac{dq_\rho}{du} = \lambda \frac{\partial G}{\partial \overline{p}_\rho}, \qquad \frac{d\overline{p}_\rho}{du} = -\lambda \frac{\partial G}{\partial q_\rho}.$$

Now, the function $S(q_1, \ldots, q_n, t)$ is simultaneously a solution to the **Hamilton-Jacobi** equation (181) of the problem of motion itself and the **Hamilton-Jacobi** equation (198.a) of the variational problem (199). The manifolds S = const. are not just the transverse M_n to the field of extremal space-time lines in question, but when one intersects them with an M_n :

$$t = \text{const.} = t^*,$$

they will give, at the same time, a family of ∞^1 manifolds $S(q_1, \ldots, q_n, t^*)$ = const. in that M_n that represent the transverse M_{n-1} of a field of extremals of the variational problem (199). If one thinks of constructing the extremals of that field for the variational problem (199) then two extremals will go through each space-time point $(q_1, \ldots, q_n, t)$, one of which is an extremal space-time line, while the other is an extremal of the variational problem (199). Since the same function $S(q_1, \ldots, q_n, t)$ will belong to both fields as the value of the extremal integral, one will have:

$$p_\rho = \frac{\partial S}{\partial q_\rho}$$

([193]) Which are not mutually independent.

([194]) The impulses are independent of the choice of parameters, since the derivatives $\partial g / \partial q'_\rho$ are indeed homogeneous of degree zero in $q'_\rho = dq_\rho / du$.

([195]) In the parametric problem (in contrast to the function problems of the calculus of variations), the function of the canonical system (and therefore the **Hamilton-Jacobi** partial differential equation, as well) is not determined uniquely. It is obtained from the relations (200) by eliminating the q'_ρ and can be changed in many different ways. That is the reason for the appearance of the parameter λ.

for the impulse vector of the motion $p_1, \ldots, p_n$, as well as having:

$$\bar{p}_\rho = \frac{\partial S}{\partial q_\rho}$$

for the "impulse vector" (200) $\bar{p}_1, \ldots, \bar{p}_n$ of the field of the variational problem (199), such that one will have ([196]):

$$p_\rho = \bar{p}_\rho$$

at every point $(q_1, \ldots, q_n, t)$, and one briefly says that *both fields have the same field elements* (i.e., space-time point + associated impulse vector).

Now, the variational problem that arises mediates a transformation of the field elements, which would follow directly from the foregoing. That is because if one starts from an arbitrary space-time point $P^*(q_1^*, \ldots, q_n^*, t^*)$ whose impulse vector $p_1^*, \ldots, p_n^*$ is associated by means of the field of the extremal space-time lines then a well-defined extremal of the field of the variational problem that arose will go through that point, and every point P of the extremal will determine a field element $q_1, \ldots, q_n, t = t^*, p_1, \ldots, p_n$ that one can associate with the field element of the starting point P^*. One can then use the value of the extremal integral $W_{P^*}^{P}$ of the variational problem (199) that is bounded by the two points P^* and P of the extremals as a parameter. One then sees immediately that the family of transformations that one obtains when one thinks of W as variable possesses the group property. Since an extremal of the variational problem (199) goes through every point of the field of space-time lines of the motion, one will then have a one-parameter group of transformations that take every field element $(q_1, \ldots, q_n, t, p_1, \ldots, p_n)$ to another field element. The canonical system (201) represents the *infinitesimal transformation* of that one-parameter group of transformations [cf., II A 6 (**L. Maurer** and **H. Burkhardt**), no. **4**] ([197]), and in order to suggest that fact, it might take the form:

(201.a) $$\delta q_\rho = \lambda \frac{\partial G}{\partial p_\rho} \delta u\,, \qquad \delta p_\rho = -\lambda \frac{\partial G}{\partial q_\rho} \delta u\,, \qquad \delta t = 0\,.$$

([196]) In the case where the energy integral:

$$H(p_1, \ldots, p_n, q_1, \ldots, q_n) = k$$

is the known integral (198), that will mean: *The same impulse components p_ρ will belong to the fields of the two variational problems, namely, that of Hamilton's principle and that of Jacobi's principle*.

([197]) One should only note that in that way a different parameter is employed in place of the value W of the extremal integral, and the factor λ is introduced. That is based upon the fact that G is not homogeneous of degree one in the impulse components. If G were replaced with a suitable function that is homogeneous of order one in the impulse components then the value W of the extremal integral would appear as a parameter, and one would set $\lambda = 1$ (cf., no. **13**).

The *extremals of the field* of the variational problem (199) are correspondingly the *orbits* (*Überführungskurven*) *of that one-parameter group of transformations* ([198]).

The fact that the transverse manifold S = const. of the field of space-time lines of the motion will go to another such transversal M_n under a transformation of that group follows immediately from the fact that the value W of the extremal integral that is bounded by two associated points of an orbit is directly equal to the difference ΔS between the two values of S that belong to the two points. On the other hand, an extremal space-time line of the field will also go to another extremal space-time line. That is because the time derivatives of the position coordinates dq_ρ / dt along the transformed curve are equal to the velocity components $\dot{q}_\rho$, which one calculates from the transformed impulse components at the transformed space-time point ([199]):

$$\dot{q}_\rho = \frac{\partial H}{\partial p_\rho}.$$

([198]) The expression *orbits* (*Überführungskurven*) might be chosen in place of the usual term *as trajectories* of the group (cf., e.g., **S. Lie**, *Theorie der Transformationsgruppen I*, Leipzig, 1888, pp. 60) in order to prevent any confusion with the trajectories of the mechanical system itself.

([199]) In fact, from (201.a):

$$\frac{d\,\delta q_\rho}{du} = \lambda\,\delta u \left\{ \sum_\sigma \left(-\frac{\partial^2 G}{\partial p_\rho\,\partial p_\sigma}\frac{\partial H}{\partial q_\sigma} + \frac{\partial^2 G}{\partial p_\rho\,\partial q_\sigma}\frac{\partial H}{\partial p_\sigma} \right) + \frac{\partial^2 G}{\partial p_\rho\,\partial t} \right\},$$

while on the other hand, $\dot{q}_\rho = \partial H / \partial p_\rho$ will yield the relation:

$$\delta\dot{q}_\rho = \lambda\,\delta u \sum_\sigma \left(-\frac{\partial^2 H}{\partial p_\rho\,\partial p_\sigma}\frac{\partial G}{\partial q_\sigma} + \frac{\partial^2 H}{\partial p_\rho\,\partial q_\sigma}\frac{\partial G}{\partial p_\sigma} \right).$$

On the other hand, (198) will imply the identity:

$$\sum_\sigma \left(-\frac{\partial G}{\partial p_\sigma}\frac{\partial H}{\partial q_\sigma} + \frac{\partial G}{\partial q_\sigma}\frac{\partial H}{\partial p_\sigma} \right) + \frac{\partial G}{\partial t} = 0,$$

and when one differentiates that with respect to p_ρ, that will give:

$$\sum_\sigma \left(-\frac{\partial^2 G}{\partial p_\rho\,\partial p_\sigma}\frac{\partial H}{\partial q_\sigma} + \frac{\partial^2 G}{\partial p_\rho\,\partial q_\sigma}\frac{\partial H}{\partial p_\sigma} \right) + \frac{\partial^2 G}{\partial p_\rho\,\partial t} = \sum_\sigma \left(\frac{\partial G}{\partial p_\sigma}\frac{\partial^2 H}{\partial p_\rho\,\partial q_\sigma} - \frac{\partial G}{\partial q_\sigma}\frac{\partial^2 H}{\partial p_\rho\,\partial p_\sigma} \right).$$

It will then follow that:

$$\frac{d\,\delta q_\rho}{dt} = \delta\dot{q}_\rho,$$

i.e., the field elements that are produced from the field elements of an extremal space-time line by a transformation of the group will all belong to one and the same extremal space-time line again.

Every space-time line of the motion, along with the orbits that emanate from its individual space-time points, will then determine an M_2 to which an entire family of space-time lines is assigned, and together with the orbits, they will define a net on the M_2 in such a way that an extremal space-time line and an orbit will run through each point of the M_2 ([200]). In the field of the space-time lines of the motion, one will then have a family of ∞^{n-1} such M_2, along each of which ∞^1 of the space-time lines of the field will lie. One can imagine introducing t and q_n as independent variables along them:

$$q_1 = q_1\,(t, q_n)\,, \quad \ldots, \quad q_{n-1} = q_{n-1}\,(t, q_n)\,, \tag{202}$$

which will also make the impulse components p_ρ that are associated with the individual points of the M_2 by the field become functions of t and q_n :

$$p_1 = p_1\,(t, q_n)\,, \quad \ldots, \quad p_{n-1} = p_{n-1}\,(t, q_n)\,. \tag{202.a}$$

If one now introduces an arbitrary M_n into the R_{n+1} of the $(q_1, \ldots, q_n, t)$ by setting:

$$t = t\,(\lambda)\,, \qquad q_n = q_n\,(\lambda) \tag{203}$$

then each of the ∞^{n-1} M_2 will intersect that M_n in a curve, and one will get a family of ∞^{n-1} intersection curves in the M_n . One can say that the orbits project the extremal space-time lines of the field into the M_n (203) here, in the adapted sense, as in the case of cyclic coordinates above, which will make the ∞^1 space-time lines that belong to one and the same M_2 possess the same projection.

If one now makes the transverse manifolds S = const. of the field of the space-time lines likewise intersect the M_n (203) then one will get a one-parameter family of M_{n-1} :

$$S^*(q_1,\ldots,q_{n-1},\lambda) = S\,(q_1, \ldots, q_{n-1}, q_n\,(\lambda), t\,(\lambda)) = \text{const.} \tag{204}$$

Once more, a variational problem can be given for which the projected curves represent a field of extremals, while the intersection manifolds (204) define the associated transverse manifolds. In order to get that variational problem, one solves equation (198.a) for $\partial S / \partial q_n$:

$$\frac{\partial S}{\partial q_n} + h\left(\frac{\partial S}{\partial q_1},\ldots,\frac{\partial S}{\partial q_{n-1}}, q_1,\ldots,q_{n-1},t,q_n,c\right) = 0\,, \tag{205}$$

and then substitutes the value for $\partial S / \partial q_n$ that is obtained in the **Hamilton-Jacobi** equation:

([200]) That M_2 was basically introduced before by **S. Lie** in his investigation into the integration of partial differential equations. The individual extremal, with the impulse vectors that are associated with its points, is the analogue of the "characteristic strip" that **Lie** introduced [cf., II A 5 (**E. von Weber**), no. **34**], while M_2 are the characteristic M_2 for the two-parameter system in involution (cf., say, the remarks of **F. Engel** in **S. Lie**, *Ges. Abh. III*, pp. 607).

$$\frac{\partial S}{\partial t} + H\left(\frac{\partial S}{\partial q_1}, \ldots, \frac{\partial S}{\partial q_n}, q_1, \ldots, q_n, t\right) = 0 ,$$

which will make it go to:

(205.a) $$\frac{\partial S}{\partial t} + H^*\left(\frac{\partial S}{\partial q_1}, \ldots, \frac{\partial S}{\partial q_{n-1}}, q_1, \ldots, q_{n-1}, t, q_n, c\right) = 0 .$$

If one introduces the functions (203) for q_n and t then the two equations:

(206) $$\begin{cases} \dfrac{\partial S}{\partial t} + H^*\left(\dfrac{\partial S}{\partial q_1}, \ldots, \dfrac{\partial S}{\partial q_{n-1}}, q_1, \ldots, q_{n-1}, \lambda, c\right) = 0 , \\ \dfrac{\partial S}{\partial q_n} + h\left(\dfrac{\partial S}{\partial q_1}, \ldots, \dfrac{\partial S}{\partial q_{n-1}}, q_1, \ldots, q_{n-1}, \lambda, c\right) = 0 \end{cases}$$

will be fulfilled at each point of M_n .

If one now goes over to the function S^* then from (204), one will have:

(204.a) $$\begin{cases} \dfrac{\partial S^*}{\partial \lambda} = \dfrac{\partial S}{\partial t} t'(\lambda) + \dfrac{\partial S}{\partial q_n} q_n'(\lambda), \\ \dfrac{\partial S^*}{\partial q_\rho} = \dfrac{\partial S}{\partial q_\rho} \qquad (\rho = 1, \ldots, n-1), \end{cases}$$

and when one introduces the function:

(207) $$K(p_1, \ldots, p_{n-1}, q_1, \ldots, q_{n-1}, \lambda, c)$$
$$= t'(\lambda) H^*(p_1, \ldots, p_{n-1}, q_1, \ldots, q_{n-1}, \lambda, c) + q_n'(\lambda) h(p_1, \ldots, p_{n-1}, q_1, \ldots, q_{n-1}, \lambda, c) ,$$

one will get the partial differential equation for S^*:

(208) $$\frac{\partial S^*}{\partial \lambda} + K\left(\frac{\partial S^*}{\partial q_1}, \ldots, \frac{\partial S^*}{\partial q_{n-1}}, q_1, \ldots, q_{n-1}, \lambda, c\right) = 0 .$$

The intersection manifolds (204) are then transverse manifolds of a variational problem in the M_n (203) for which this partial differential equation (208) is the **Hamilton-Jacobi** equation. If one puts the **Euler** equations for that variational problem into canonical form then they will read:

(208.a) $$\frac{dq_\rho}{d\lambda} = \frac{\partial K}{\partial p_\rho}, \qquad \frac{dp_\rho}{d\lambda} = -\frac{\partial K}{\partial q_\rho} \qquad (\rho = 1, \ldots, n-1),$$

and one will easily see that the projected curves are a family of solutions to that canonical system (208.a), so they are extremals to the variational problem that belongs (208) ([201]), as well as that, due to the fact that:

$$p_\rho = \frac{\partial S^*}{\partial q_\rho} \qquad (\rho = 1, \ldots, n-1),$$

([201]) In fact, the differentials of the position coordinates and the impulse components for the projections must be linear combinations of the differentials of the space-time lines and their orbits on M_2, since the projections belong to M_2, such that if d^* means the direction of advance along the projection then d will mean the direction of advance along the space-time line of the motion, and δ means the direction of advance along the orbit then one will have the relations:

$$d^* q_\rho = \mu_1\, dq_\rho + \mu_2\, \delta q_\rho\,, \qquad d^* p_\rho = \mu_1\, dp_\rho + \mu_2\, \delta p_\rho \qquad (\rho = 1, \ldots, n-1),$$

in which μ_1 and μ_2 are determined from the conditions:

$$d^* t = \mu_1\, dt = t'(\lambda)\, d\lambda\,, \qquad d^* q_n = \mu_1\, dp_n + \mu_2\, \delta p_n = q_n'(\lambda)\, d\lambda$$

to be:

$$\mu_1 = \frac{t'(\lambda)}{dt} d\lambda\,,$$
$$\mu_2 = \frac{q_n'(\lambda) - t'(\lambda)\dot{q}_n}{\delta t}\, d\lambda\,.$$

Thus:

$$\frac{d^* q_\rho}{d\lambda} = t'(\lambda)\left(\dot{q}_\rho - \frac{\delta q_\rho}{\delta q_n}\dot{q}_n\right) + q_n'(\lambda)\frac{\delta q_\rho}{\delta q_n}\,,$$
$$\frac{d^* p_\rho}{d\lambda} = t'(\lambda)\left(\dot{p}_\rho - \frac{\delta p_\rho}{\delta q_n}\dot{q}_n\right) + q_n'(\lambda)\frac{\delta p_\rho}{\delta q_n}\,,$$

or, since one has:

$$\frac{\delta q_\rho}{\delta q_n} = \frac{\partial h}{\partial p_\rho}, \qquad \frac{\delta p_\rho}{\delta q_n} = -\frac{\partial h}{\partial q_\rho},$$

from (205), and further when one recalls (205.a):

$$\dot{q}_\rho - \frac{\delta q_\rho}{\delta q_n}\dot{q}_n = \frac{\partial H}{\partial p_\rho} - \frac{\partial h}{\partial p_\rho}\frac{\partial H}{\partial p_n} = \frac{\partial H^*}{\partial p_\rho}\,,$$
$$\dot{p}_\rho - \frac{\delta p_\rho}{\delta q_n}\dot{q}_n = -\frac{\partial H}{\partial q_\rho} + \frac{\partial h}{\partial q_\rho}\frac{\partial H}{\partial p_n} = -\frac{\partial H^*}{\partial q_\rho}\,,$$

one will have:

$$\frac{d^* q_\rho}{d\lambda} = t'(\lambda)\frac{\partial H^*}{\partial p_\rho} + q_n'(\lambda)\frac{\partial h}{\partial p_\rho}\,,$$
$$\frac{d^* p_\rho}{d\lambda} = -\left(t'(\lambda)\frac{\partial H^*}{\partial q_\rho} + q_n'(\lambda)\frac{\partial h}{\partial q_\rho}\right).$$

However, from (207), that is just the system (208.a).

they will define just the field that belongs to the intersection manifolds (204). One has *once more reduced the variational problem in one step by means of the known integral* then.

Since the M_n (203) can be chosen arbitrarily, the new variational problem will determine the M_2 immediately. If one has found them then all that will remain is to calculate the family of ∞^1 extremal space-time lines of the original variational problem on the individual M_2. In order to do that, one substitutes the expression (202.a) for $p_1, \ldots, p_n$ and the expressions (202) for $q_1, \ldots, q_{n-1}$ in the equation:

$$\frac{dq_n}{dt} = \frac{\partial H}{\partial p_n},$$

which will give one a differential equation:

$$\frac{dq_n}{dt} = f(q_n, t).$$

Whereas in the case of cyclic coordinates (and the energy integral), the determination of the extremal space-time line on the M_2 will require only a quadrature, here, it seems necessary to solve a first-order differential equation. Meanwhile, one can immediately give an **Euler** multiplier for that differential equation, such that one also comes down to a quadrature here (cf., also no. **27**)([202]).

([202]) The simplification of the variational problem when a first integral is present was achieved in a somewhat-different way by **P. Woronetz**, "Sur l'intégration des équations aux dérivées partielles," Bull. des sciences mathémat. (2) **47** (1923), pp. 113, cf., also the (Russian) treatise by **P. Woronetz**, "On the question of integrating the Lagrange differential equations," Proc. math. Labor. Crimean Univ. **3** (1921), pp. 39. Referenced in Fortschritte der Mathematik **48**, pp. 901.

He immediately addressed the partial differential equation (205.a) and imagined that a solution $\bar{S}(q_1, \ldots, q_{n-1}, t, q_n)$ was determined, in which q_n appeared as a parameter. In a manifold q_n = const., that solution can differ from the desired function $S(q_1, \ldots, q_n, t)$ that simultaneously satisfies the **Hamilton-Jacobi** equation (181) and the partial differential equation (198.a) that arises from the first integral by only a constant, i.e., one will have:

(209)
$$\frac{\partial S}{\partial q_1} = \frac{\partial \bar{S}}{\partial q_1}, \quad \ldots, \quad \frac{\partial S}{\partial q_{n-1}} = \frac{\partial \bar{S}}{\partial q_{n-1}}.$$

Therefore, the difference $S - \bar{S}$ must be a function of only q_n. If one then defines:

$$dS - d\bar{S} = \left(\frac{\partial S}{\partial q_n} - \frac{\partial \bar{S}}{\partial q_n} \right) dq_n,$$

and replaces the $\partial S / \partial q_n$ in it with h, corresponding to equation (205), then when one introduces h into the derivatives of $\bar{S}$, the variables $q_1, \ldots, q_{n-1}, t$ must be given by the expression:

$$\left[h\left(\frac{\partial S}{\partial q_1}, \ldots, \frac{\partial S}{\partial q_{n-1}}, q_1, \ldots, q_{n-1}, t, q_n \right) + \frac{\partial S}{\partial q_n} \right],$$

and one will get the function S by the quadrature:

If one has yet another first integral to the equations of motion, in addition to the integral (197) [(198), resp.], then it can be employed to further simplify the **Hamilton-Jacobi** differential equation (208) if and only if the function that arises from it by the substitution:

$$(210) \qquad p_n = -h\,(p_1, \ldots, p_{n-1}, q_1, \ldots, q_{n-1}, t, q_n, c)\,, \quad q_n = q_n\,(\lambda)\,, \quad t = t\,(\lambda)$$

represents an integral of the reduced canonical system (207). The necessary and sufficient condition for that is that the *Poisson bracket* that is defined by the two integrals must vanish ([203]), which will be explained in more detail below (cf., no. **25**). The energy integral $H = k$ (when it exists) has the property that each of the integrals (that are free of time t) that are added to it will fulfill that condition. One can then use the existence of such a further integral, along with the energy integral, to simplify the partial differential equation (191), which enters in place of the original **Hamilton-Jacobi** equation (181) when an energy integral exists, in a way that is analogous to what one does with just the **Hamilton-Jacobi** equation itself. If one knows two (or more) integrals of the canonical system that are free of t then if they are to be usable to further simplify the partial differential equation (191), it will once more be necessary and sufficient that the **Poisson** bracket must vanish.

19. Integrating the Hamilton-Jacobi equation by separation of variables. –

19.a. General statement of the problem. – The problem of motion is solved when one knows a complete solution to the **Hamilton-Jacobi** equation (cf., nos. **17** and **18**). Therefore, **Jacobi** had turned his attention to the cases in which such a solution could be given in a simple way, and in particular, treated mechanical problems for which a *complete solution* to the **Hamilton-Jacobi** partial differential equation can be achieved by *a quadrature alone* ([204]). With the application of

---

$$(209.a) \qquad S\,(q_1, \ldots, q_n, t) = \bar{S} - \int \left(h + \frac{\partial \bar{S}}{\partial q_n} \right) dq_n\,.$$

([203]) The problem statement is then identical to the question of whether one can give the one n-parameter family of space-time lines of motion that define a *field* when one demands that the impulse components $p_\rho\,(q_1, \ldots, q_n, t)$ of the field should make both integrals into identities (for prescribed numerical values of the constants).

([204]) The simplest example is the so-called *one-body problem* in celestial mechanics, i.e., the motion of a mass-point that is attracted to a fixed center according to **Newton**'s law of gravitation (cf., **C. G. J. Jacobi**, *Vorlesungen über Dynamik*, Vorlesungen 24 and 25 = *Werke Suppl.-Band*, pp. 183, *et seq.*). In rectangular coordinates, since the energy integral for the equations of motion exists, the **Hamilton-Jacobi** equation will have the form:

$$(211) \qquad \frac{1}{2}\left\{ \left(\frac{\partial W}{\partial x} \right)^2 + \left(\frac{\partial W}{\partial y} \right)^2 + \left(\frac{\partial W}{\partial z} \right)^2 \right\} - \frac{\gamma^2}{r} = k \qquad \left(r = \sqrt{x^2 + y^2 + z^2} \right)$$

here. When one introduces *spatial polar coordinates*, that will go to:

$$(211.a) \qquad \frac{1}{2}\left\{ \left(\frac{\partial W}{\partial r} \right)^2 + \frac{1}{r^2}\left(\frac{\partial W}{\partial \varphi} \right)^2 + \frac{1}{r^2 \sin^2 \varphi}\left(\frac{\partial W}{\partial \psi} \right)^2 \right\} - \frac{\gamma^2}{r} = k\,,$$

quantum theory to atomic mechanics ([205]), for some time now, the question of whether and when a solution to the **Hamilton-Jacobi** equation by quadratures would be possible has taken on special importance because when such a solution is present, the Ansatz for the quantization conditions will become especial simple (cf., also no. **23**).

In the literature, when one investigates that question, one restricts oneself to mechanical problems for which the energy integral exists, and one further assumes that the kinetic energy is a quadratic form in the velocity components such that the **Hamilton-Jacobi** differential equation will possess the form:

$$\sum_{\lambda,\mu=1}^{n} \tfrac{1}{2} g^{\lambda\mu} \frac{\partial W}{\partial q_\lambda} \frac{\partial W}{\partial q_\mu} + \Phi(q_1,\ldots,q_n) = k \,. \tag{212}$$

---

and **Jacobi** could then split that into the two equations:

$$\frac{1}{2}\left(\frac{\partial W_1}{\partial r}\right)^2 = \frac{\gamma^2}{r} + k - \frac{c_1}{r^2}$$

and

$$\frac{1}{2}\left\{\left(\frac{\partial W}{\partial \varphi}\right)^2 + \frac{1}{\sin^2\varphi}\left(\frac{\partial W}{\partial \psi}\right)^2\right\} = c_1 \,,$$

the latter of which is again replaced with the system of two equations:

$$\frac{1}{2}\left(\frac{\partial W_2}{\partial \varphi}\right)^2 = c_1 - \frac{c_2}{\sin^2\varphi}$$

and

$$\frac{1}{2}\left(\frac{\partial W_3}{\partial \psi}\right)^2 = c_2 \,.$$

If $W_1\,(r)$ is then a solution to the first equation, $W_2\,(\varphi)$ is a solution to the second one, and $W_3\,(\psi)$ is a solution to the third then one will obviously have a solution to equation (211.a) in:

$$W = W_1\,(r) + W_2\,(\varphi) + W_3\,(\psi) \,.$$

Now, since the functions W_1, W_2, W_3 are determined immediately by quadratures, one will have found a solution to the **Hamilton-Jacobi** equation by quadratures alone in:

$$W = \int \sqrt{\frac{2\gamma^2}{r} + 2k - \frac{2c_1}{r^2}}\, dr + \int \sqrt{2c_1 - \frac{2c_2}{\sin^2\varphi}}\, d\varphi + \sqrt{2}\, c_2 \psi \,, \tag{211.b}$$

and indeed, it is a solution that is complete solution, due to the appearance of the arbitrary constants c_1, c_2, and k. **Jacobi** gave an extension of the process to n variables (*loc. cit.*, pp. 185, *et seq.*). At the same time, he recognized (Lecture 25, pp. 190) that there was a *second way* by which one could bring the partial differential equation (211) into a form from which one could get a complete integral by three quadratures.

One should observe that *the reduction of the solution to quadratures is possible only when one employs suitable coordinates*. One cannot arrive at a complete solution by quadratures alone with the form (211) of the **Hamilton-Jacobi** equation for the one-body problem, i.e., with the use of ordinary rectangular coordinates.

([205]) Cf., e.g., **M. Born**, *Vorlesungen über Atommechanik I*, Berlin, 1925.

The goal is to give a complete solution to that equation that has the form:

(213) $$W = W_1 (q_1, k, c_1, \ldots, c_{n-1}) + W_2 (q_1, k, c_1, \ldots, c_{n-1}) + \ldots + W_n (q_1, k, c_1, \ldots, c_{n-1}) ,$$

so a solution that is represented as a sum of n summands in which each summand depends upon only one of the n variables, while constants can appear in all summands. In the event that the process with the *Ansatz* (213) is possible, one refers to it as *the integration of the Hamilton-Jacobi differential equation by separation of variables*.

19.b. Levi-Civita's theorem. The "essentially geodetic" case and the case of Stäckel's theorem. – In order to find the conditions that the $g^{\lambda\mu}$ [the coefficients $g_{\lambda\mu}$, resp.] in the kinetic energy:

(214) $$T = \sum_{\lambda,\mu} g_{\lambda\mu}\, \dot{q}_\lambda\, \dot{q}_\mu$$

and the potential function $\Phi (q_1, \ldots, q_n)$ must satisfy in order for a solution of the form (213) to be able to exist in the coordinates $q_1, \ldots, q_n$ ([206]), one interprets the solution W as the value of an extremal integral for a field of trajectories of the system, i.e., extremals of the variational problem of the principle of least action in the **Jacobi** form. From (213), the *impulse components* in the field:

(215) $$p_\rho = \frac{\partial W}{\partial q_\rho} = \frac{dW}{dq_\rho} = p_\rho (q_\rho, k, c_1, \ldots, c_{n-1})$$

will then depend upon *each* (and indeed, the one with the same index) position coordinate, such that one will then have:

(215.a) $$\frac{\partial p_\rho}{\partial q_\lambda} = 0 \qquad (\lambda \neq \rho).$$

The equation:

$$H (p_1, \ldots, p_n, q_1, \ldots, q_n) = k ,$$

which will become an identity when one replaces the impulse components in it with the functions (215) of the field, will then imply that:

$$\frac{\partial H}{\partial p_\lambda}\frac{dp_\lambda}{dq_\lambda} + \frac{\partial H}{\partial q_\lambda} = 0 ,$$

([206]) Cf., **T. Levi-Civita**, "Sulla integrazione della equazione di Hasmilton-Jacobi per separazione di variabili," Math. Ann. **59** (1904), pp. 383.

i.e. ([207]):

$$\frac{dp_\lambda}{dq_\lambda} = -\frac{\dfrac{\partial H}{\partial q_\lambda}}{\dfrac{\partial H}{\partial p_\lambda}} \,. \tag{215.b}$$

If one differentiates that relation with respect to a position coordinate q_ρ that is different from q_λ then it will further follow that:

$$\frac{\partial H}{\partial q_\lambda}\frac{\partial}{\partial q_\rho}\left(\frac{\partial H}{\partial p_\lambda}\right) - \frac{\partial H}{\partial p_\lambda}\frac{\partial}{\partial q_\rho}\left(\frac{\partial H}{\partial q_\lambda}\right) = 0 \qquad (\rho \neq \lambda),$$

and therefore, one will finally have:

$$\frac{\partial^2 H}{\partial q_\lambda \partial q_\rho}\frac{\partial H}{\partial p_\lambda}\frac{\partial H}{\partial p_\rho} - \frac{\partial^2 H}{\partial q_\lambda \partial p_\rho}\frac{\partial H}{\partial p_\lambda}\frac{\partial H}{\partial q_\rho} - \frac{\partial^2 H}{\partial p_\lambda \partial q_\rho}\frac{\partial H}{\partial q_\lambda}\frac{\partial H}{\partial p_\rho} + \frac{\partial^2 H}{\partial p_\lambda \partial p_\rho}\frac{\partial H}{\partial q_\lambda}\frac{\partial H}{\partial q_\rho} = 0 \quad (\rho \neq \lambda). \tag{216}$$

Those are the $n\,(n-1)/2$ necessary and sufficient conditions for the partial differential equation (212) to possess a complete solution of the form (213) ([208]).

Since the function H is a sum of two terms, according to (212), one of which has degree two in the impulse components p_ρ and the other of which has degree zero:

$$\begin{aligned} H(p_1, \ldots, p_n, q_1, \ldots, q_n) &= Q(p_1, \ldots, p_n, q_1, \ldots, q_n) + \Phi(q_1, \ldots, q_n) \\ &= \sum_{\lambda,\mu} \tfrac{1}{2}\, g^{\lambda\mu}\, p_\lambda\, p_\mu + \Phi(q_1, \ldots, q_n)\,, \end{aligned} \tag{217}$$

each of the condition equations (216) will include terms of degree four, terms of degree two, and terms of degree zero in the impulse components $p_1, \ldots, p_n$. However, each of equations (216) must be true identically in the impulse components, so each of them must split into three equations such that the necessary and sufficient condition for the possibility of integrating the **Hamilton-Jacobi** partial differential equation (212) by separation of variables will yield the following three classes of $n\,(n-1)/2$ condition equations ([209]):

$$\frac{\partial^2 Q}{\partial q_\lambda \partial q_\rho}\frac{\partial Q}{\partial p_\lambda}\frac{\partial Q}{\partial p_\rho} - \frac{\partial^2 Q}{\partial q_\lambda \partial p_\rho}\frac{\partial Q}{\partial p_\lambda}\frac{\partial Q}{\partial q_\rho} - \frac{\partial^2 Q}{\partial p_\lambda \partial q_\rho}\frac{\partial Q}{\partial q_\lambda}\frac{\partial Q}{\partial p_\rho} + \frac{\partial^2 Q}{\partial p_\lambda \partial p_\rho}\frac{\partial Q}{\partial q_\lambda}\frac{\partial Q}{\partial q_\rho} = 0\,, \tag{218.a}$$

([207]) $\partial H / \partial p_\lambda = 0$ is excluded, since the partial differential equation can be simplified directly in the spirit of the previous section when H is free of p_λ .

([208]) **T. Levi-Civita**, *loc. cit.* ([206]), pp. 385.

([209]) These condition equations were converted into condition equations for the coefficients g_{ik} in **F. A. Dall'Acqua**, "Sulla integrazione delle equazioni di Hamilton-Jacobi per separazione di variabili," Math. Ann. **66** (1909), pp. 398.

(218.b)
$$\frac{\partial^2\Phi}{\partial q_\lambda\,\partial q_\rho}\frac{\partial Q}{\partial p_\lambda}\frac{\partial Q}{\partial p_\rho}-\frac{\partial^2 Q}{\partial q_\lambda\,\partial p_\rho}\frac{\partial Q}{\partial p_\lambda}\frac{\partial\Phi}{\partial q_\rho}-\frac{\partial^2 Q}{\partial p_\lambda\,\partial q_\rho}\frac{\partial\Phi}{\partial q_\lambda}\frac{\partial Q}{\partial p_\rho}+\frac{\partial^2 Q}{\partial p_\lambda\,\partial p_\rho}\left(\frac{\partial Q}{\partial q_\lambda}\frac{\partial\Phi}{\partial q_\rho}+\frac{\partial Q}{\partial q_\rho}\frac{\partial\Phi}{\partial q_\lambda}\right)=0\,,$$

(218.c)
$$\frac{\partial^2 Q}{\partial p_\lambda\,\partial p_\rho}\frac{\partial\Phi}{\partial q_\lambda}\frac{\partial\Phi}{\partial q_\rho}=0\,.$$

As a comparison with (216) will show, the first group (218.a) of those conditions says that the partial differential equation:

(219)
$$2Q\left(\frac{\partial W}{\partial q_1},\ldots,\frac{\partial W}{\partial q_n},q_1,\ldots,q_n\right)=\sum_{\lambda,\mu}g^{\lambda\mu}\frac{\partial W}{\partial q_\lambda}\frac{\partial W}{\partial q_\mu}=1$$

can be integrated by separation of variables. However, that equation is just the **Hamilton-Jacobi** partial differential equation for the variational problem:

(219.a)
$$\int\sqrt{\sum g_{\lambda\mu}\,dq_\lambda\,dq_\mu}\;=\text{extrem.}$$

of the geodetic lines in **Riemannian** space ([210]) with the arc-length element:

(219.b)
$$ds^2=\sum g_{\lambda\mu}\,dq_\lambda\,dq_\mu$$

that is defined by the kinetic energy (214) of the mechanical problem. That will lead to the *theorem of* ***T. Levi-Civita*** ([211]):

If a mechanical problem is to be soluble by separation of variables for a certain choice of coordinates then it will be necessary that the geodetic lines of the arc-length element that arises from the kinetic energy can also be determined by separation of variables in those coordinates.

For a mechanical problem whose geodetic problem is soluble by separation of variables, the two groups (218.b) and (218.c) will then give conditions that the potential function must satisfy if the general mechanical problem with applied forces is to also be soluble by separation of variables.

An examination of the condition equations (218.a), which one can write in the form:

([210]) One can refer to that variational problem for geodetic lines briefly as the geodetic problem that is associated with the mechanical problem.

([211]) **T. Levi-Civita**, *loc. cit.* ([206]), pp. 386.

$$\frac{\partial Q}{\partial q_\lambda}\left(\frac{\partial Q}{\partial p_\rho}\frac{\partial^2 Q}{\partial q_\lambda\,\partial q_\rho}-\frac{\partial Q}{\partial q_\rho}\frac{\partial^2 Q}{\partial q_\lambda\,\partial p_\rho}\right)=\frac{\partial Q}{\partial q_\lambda}\left(\frac{\partial Q}{\partial p_\rho}\frac{\partial^2 Q}{\partial p_\lambda\,\partial q_\rho}-\frac{\partial Q}{\partial q_\rho}\frac{\partial^2 Q}{\partial p_\lambda\,\partial p_\rho}\right),$$

will lead to the question of whether or not:

$$\frac{\partial Q}{\partial q_\lambda}=\frac{\partial T}{\partial q_\lambda}=\sum_{\sigma,\rho}\tfrac{1}{2}\frac{\partial g_{\sigma\tau}}{\partial q_\lambda}\,\dot q_\sigma\,\dot q_\tau$$

is divisible by:

$$\frac{\partial Q}{\partial p_\lambda}=\dot q_\lambda\,.$$

If the divisibility exists for the first ν of the λ indices ([212]), but not for the remaining ($n-\nu$), then **F. A. Dall'Acqua** ([213]) referred to the first ν indices as "indices of the first kind" and the remaining ($n-\nu$) ones as "indices of the second kind" ([214]). Since one can have $\nu = 0, \ldots, n$, there will then be ($n+1$) cases to distinguish.

If λ is an index of the first kind then one will initially have:

(220) $$\frac{\partial g_{\rho\sigma}}{\partial q_\lambda}=0 \qquad (\rho\neq\sigma;\ \sigma=1,\ldots,n),$$

and one will infer immediately from (218.b) that:

(221) $$\frac{\partial\Phi}{\partial q_\lambda}=0 \qquad (\lambda = \text{index of the first kind}).$$

Only coordinates with indices of the second kind can then appear in the potential function, such that in the case of $\nu = n$, i.e., when all indices are of the first kind, the potential function must

([212]) Which means that the second factor:

$$\left(\frac{\partial Q}{\partial p_\rho}\frac{\partial^2 Q}{\partial p_\lambda\,\partial q_\rho}-\frac{\partial Q}{\partial q_\rho}\frac{\partial^2 Q}{\partial p_\lambda\,\partial q_\rho}\right)$$

should *not* be divisible by $\dot q_\lambda$.

([213]) **F. A. Dall'Acqua**, "Le equazioni di Hamilton-Jacobi, che si integrano per separazione di variabili," Rend. circ. mat. Palermo **33** (1912), pp. 341.

([214]) For an index of the second kind:

$$\left(\frac{\partial Q}{\partial p_\rho}\frac{\partial^2 Q}{\partial p_\lambda\,\partial q_\rho}-\frac{\partial Q}{\partial q_\rho}\frac{\partial^2 Q}{\partial p_\lambda\,\partial q_\rho}\right) \qquad (\rho \text{ arbitrary, but} \neq \lambda)$$

must be divisible by $\dot q_\lambda$.

reduce to a constant, and only the *geodetic problem* will be soluble by separation of variables ([215]). One then refers to that case as *essentially geodetic*.

The other limiting case $\nu = 0$, in which *all n indices are ones of the second kind*, is also easy to grasp. That is because it would follow from (218.c) that one has:

$$g^{\rho\sigma} = 0 \tag{222}$$

when the two coordinates of the second kind q_ρ and q_σ both appear in the potential function. If all coordinates are present in the potential function then the **Hamilton-Jacobi** differential equation (212) must possess the form:

$$\frac{1}{2}\left\{ g^{11}\left(\frac{\partial W}{\partial q_1}\right)^2 + \cdots + g^{nn}\left(\frac{\partial W}{\partial q_n}\right)^2 \right\} + \Phi(q_1,\ldots,q_n) = k\,, \tag{223}$$

i.e., it must relate to the arc-length element that arises from kinetic energy in orthogonal coordinates ([216]):

$$ds^2 = g_{11}\,dq_1^2 + \cdots + g_{nn}\,dq_n^2\,. \tag{224}$$

([215]) Its analytical treatment is in **T. Levi-Civita** ([206]), pp. 388. The **Christoffel** three-index symbols in this case are:

$$\left\{ \begin{matrix} \rho\,\sigma \\ \lambda \end{matrix} \right\} = 0 \qquad\qquad (\rho \neq \sigma),$$

and in the relations (215.b), the right-hand side will be a linear function of the p_τ, namely:

$$\frac{dp_\lambda}{dq_\lambda} = \sum_\tau \left\{ \begin{matrix} \lambda\,\lambda \\ \tau \end{matrix} \right\} p_\tau\,.$$

Furthermore, that will give the conditions:

$$\frac{\partial}{\partial q_\sigma}\left\{ \begin{matrix} \lambda\,\lambda \\ \tau \end{matrix} \right\} + \left\{ \begin{matrix} \lambda\,\lambda \\ \sigma \end{matrix} \right\}\left\{ \begin{matrix} \sigma\,\sigma \\ \tau \end{matrix} \right\} = 0\,,$$

from which, it will follow that the **Riemannnian** curvature tensor of the arc-length will vanish, i.e., that the *arc-length must be one of a Euclidian manifold*.

([216]) In that way, one has:

$$g^{\lambda\lambda} = \frac{1}{g_{\lambda\lambda}}\,. \tag{224.a}$$

That is the case that was treated systematically, with some preliminary remarks, by **C. G. J. Jacobi** ([217]) and **J. Liouville** ([218]), as well as by **P. Stäckel** ([219]).

When **Stäckel** applied the Ansatz (213) to the partial differential equation (223) and set:

$$\left(\frac{\partial W}{\partial q_\rho}\right)^2 = \left(\frac{dW}{dq_\rho}\right)^2 = w_\rho\,(q_\rho, k, c_1, \ldots, c_{n-1})\,, \tag{225}$$

as well as ([220]):

$$\frac{\partial w_\rho}{\partial k} = \psi_\rho\,(q_\rho)\,, \qquad \frac{\partial w_\rho}{\partial c_\lambda} = \varphi_\rho^{(\lambda)}(q_\rho)\,, \tag{226}$$

he obtained a system of linear equations for the $g^{\lambda\lambda}$:

$$\left\{\begin{array}{ll} g^{11}\psi_1 + g^{22}\psi_2 + \cdots + g^{nn}\psi_n = 2, \\ g^{11}\psi_1^{(1)} + g^{22}\psi_2^{(1)} + \cdots + g^{nn}\psi_n^{(1)} = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ g^{11}\psi_1^{(n-1)} + g^{22}\psi_2^{(n-1)} + \cdots + g^{nn}\psi_n^{(n-1)} = 0\,. \end{array}\right. \tag{227}$$

If one denotes the determinant that one obtains, from the matrix:

$$\begin{pmatrix} \varphi_1^{(1)} & \varphi_2^{(1)} & \cdots & \varphi_n^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1^{(n-1)} & \varphi_2^{(n-1)} & \cdots & \varphi_n^{(n-1)} \end{pmatrix}$$

by dropping the row by D_ρ then the determinant of the system of equations (227) will become:

([217]) **C. G. J. Jacobi**, *Vorlesungen, Werke Suppl.-Bd.*, pp. 185, *et seq.* In connection with **Jacobi**, **E. Rosochatius** treated the motion of a point in the plane and on a second-order surface in "Über Bewegungen eines Punktes," Diss. Göttingen, 1877.

([218]) **J. Liouville**, "Sur quelques cas particuliers où les équations du mouvement d'un point matériel peuvent s'intégrer," J. de math. **11** (1846), pp. 345, as well as J. de math. **12** (1847), pp. 410, in which he restricted himself to two and three degrees of freedom. For the case of *n* degrees of freedom, cf., **J. Liouville**, "L'intégration des équations différentielles du movement d'un nombre quelconque des points matériels," J. de math. **14** (1849), pp. 257.

([219]) **P. Stäckel**, "Über die Integration der *Hamilton-Jacobischen* Differentialgleichung mittels Separation der Varänderlichen," Habilitat.-Schrift Halle 1891.

([220]) The upper index in $\varphi_\rho^{(\lambda)}(q_\rho)$ is placed in parentheses in order to avoid confusion with the fact that one is dealing with contravariant and covariant indices.

On this, cf., also the presentation in **C. L. Charlier**, *Die Mechanik des Himmels I*, pp. 77, *et seq.*

(228) $$D = \begin{vmatrix} \psi_1 & \psi_2 & \cdots & \psi_n \\ \varphi_1^{(1)} & \varphi_2^{(1)} & \cdots & \varphi_n^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1^{(n-1)} & \varphi_2^{(n-1)} & \cdots & \varphi_n^{(n-1)} \end{vmatrix} = \psi_1 D_1 + \ldots + \psi_n D_n ,$$

and the solution of (227) will give the $g^{\lambda\lambda}$:

(229) $$g^{\lambda\lambda} = \frac{2D_\lambda}{D} .$$

Thus, the coefficients of the arc-length element will become:

(229.a) $$g_{\lambda\lambda} = \frac{D}{2D_\lambda} ,$$

and therefore, the arc-length element itself will be:

(230) $$ds^2 = \frac{D}{2}\left(\frac{dq_1^2}{D_1} + \cdots + \frac{dq_n^2}{D_n} \right) .$$

For the potential function Φ, it will then follow from (223), (225), and (229) that:

$$\Phi\,(q_1, \ldots, q_n) = k - \frac{1}{D}(D_1 w_1 + \ldots + D_n w_n) ,$$

or due to the identity:

$$1 = \frac{\psi_1 D_1 + \cdots + \psi_n D_n}{D} ,$$

it will also follow that:

$$\Phi\,(q_1, \ldots, q_n) = \frac{1}{D}[D_1 (k\ \psi_1 - w_1) + \ldots + D_n (k\ \psi_n - w_n)]$$

when one sets:

$$k\ \psi_\rho - w_\rho = \chi_\rho\,(q_\rho) ,$$

resp., so:

(231) $$\Phi\,(q_1, \ldots, q_n) = \frac{1}{D}(D_1 \chi_1 + \ldots + D_n \chi_n) = \frac{D_1 \chi_1 + \cdots + D_n \chi_n}{D_1 \psi_1 + \cdots + D_n \psi_n} .$$

Conversely, one also has: If one chooses the $n\ (n + 1)$ functions $\varphi_\rho^{(\lambda)}(q_\rho)$, $\psi_\rho(q_\rho)$, $\chi_\rho(q_\rho)$ arbitrarily and makes the Ansatz (230) for the arc-length element and the Ansatz (231) for the

potential function then the associated **Hamilton-Jacobi** equation (223) will be integrable by separation of variables, i.e., one will get a complete solution by k quadratures. Namely, with the Ansatz (213), the function W_ρ will be given by quadrature:

$$W_\rho = \int \sqrt{k\psi_\rho(q_\rho) + \{c_1\varphi_\rho^{(1)}(q_\rho) + \cdots + c_{n-1}\varphi_\rho^{(n-1)}(q_\rho)\} - \chi_\rho(q_\rho)}\, dq_\rho \,. \tag{232}$$

That is ***Stackel**'s theorem* ([221]).

19.c. The discussion of the (n + 1) individual cases. – In the general case where ν of the coordinates are of the first kind and the other ($n - \nu$) coordinates are of the second kind ([222]), the potential function will depend upon only the coordinates of the second kind, such that in the general case, one will have:

$$\Phi = \Phi(q_{\nu+1}, q_{\nu+2}, \ldots, q_n) \,. \tag{235}$$

([221]) For $n = 2$, one has $D = \psi_1\varphi_2 - \psi_2\varphi_1$, such that the arc-length element will have the form:

$$ds^2 = \frac{1}{2}\left(\frac{\psi_1}{\varphi_1} - \frac{\psi_2}{\varphi_2}\right)(\varphi_1 dq_1^{\,2} - \varphi_2 dq_2^{\,2}) \,,$$

and for a somewhat-different choice of parameters, it will have the form:

$$ds^2 = \{F_1(q_1) - F_2(q_2)\}\{f_1(q_1)dq_1^{\,2} + f_2(q_2)dq_2^{\,2}\} \,, \tag{233}$$

resp., while the potential function will assume the form:

$$\Phi = \frac{\varphi_2\chi_1 - \varphi_1\chi_2}{\varphi_2\psi_1 - \varphi_1\psi_2} = \frac{\dfrac{\chi_1}{\varphi_1} - \dfrac{\chi_2}{\varphi_2}}{\dfrac{\psi_1}{\varphi_1} - \dfrac{\psi_2}{\varphi_2}} \,,$$

so

$$\Phi = \frac{G_1(q_1) - G_2(q_2)}{F_1(q_1) - F_2(q_2)} \,. \tag{234}$$

J. Liouville [cf., J. de math. **11** (1846), pp. 345] has already given that form. Should the arc-length element in the Euclidian form take the form (233), then it would be necessary to introduce *elliptic coordinates* (their degenerate forms: parabolic, ordinary polar, rectangular coordinates, resp.)

Liouville's result generalizes to three-dimensional and n-dimensional spaces. The special form of **Stäckel**'s arc-length element that arises from (233) by generalizing to more dimensions:

$$ds^2 = \{F_1(q_1) + \cdots + F_n(q_n)\}\{dq_1^{\,2} + \cdots + dq_n^{\,2}\}$$

will then be referred to as the ***Liouville** arc-length element*.

([222]) **F. A. Dall'Acqua**, "Le equazioni di Hamilton-Jacobi, che si integrano per separazione di variabili," Rend. circ. mat. Palermo **33** (1922), pp. 341.

In the same way as what was done in the special cases that were treated ([223]), it will now follow that one has:

$$g^{\lambda\mu} = 0 \qquad (\lambda = \mu), \tag{235}$$

when λ and μ are both indices of the second kind, while the relation:

$$\frac{\partial g^{\lambda\lambda}}{\partial q_\rho} = 0 \tag{236.a}$$

will be true when λ is an index of the second kind and ρ is one of the first kind ([224]).

([223]) Namely, $\nu = n$ (all indices are of the first kind, which is the essentially geodetic *case of* ***T. Levi-Civita***) and $\nu = 0$ (all indices of are the second kind, which is the ***Stäckel*** *case*).

([224]) From (220), the $g_{\rho\sigma}$ $(\rho \neq \sigma)$ include only coordinates with indices of the second kind. If μ is again an index of the second kind and τ is one of the first kind then, if one understands λ to mean an index of the second kind then:

$$\frac{\partial}{\partial q}\left(\frac{g^{\mu\tau}}{g^{\mu\mu}}\right) = 0\,,$$

i.e., the quotients $g^{\mu\tau} / g^{\mu\mu}$ (τ is an index of the first kind and μ is an index of the second kind) depend upon only coordinates of the first kind.

$$g^{\mu\tau} = g^{\mu\mu} l_\tau^{(\mu)}(q_1, \ldots, q_\nu)\,. \tag{236.b}$$

In that case, the **Hamilton-Jacobi** equation will have the form:

$$\left\{\begin{array}{l} \frac{1}{2}\sum\limits_{\sigma,\tau=1}^{\nu} g^{\sigma\tau}(q_1,\ldots,q_n)\,p_\sigma p_\tau \\ \qquad + \sum\limits_{\lambda=\nu+1}^{n}\left\{ g^{\lambda\lambda}(q_1,\ldots,q_n)\,p_\lambda \cdot \sum\limits_{\tau=1}^{\nu} l_\tau^{(\lambda)}(q_1,\ldots,q_\nu)\,p_\tau \right\} \\ \qquad + \frac{1}{2}\sum\limits_{\lambda=\nu+1}^{n}\left\{ g^{\lambda\lambda}(q_{\nu+1},\ldots,q_n)\,p_\lambda^2 + \Phi(q_{\nu+1},\ldots,q_n)\right\} = k\,. \end{array}\right. \tag{237}$$

When λ and μ are indices of the second kind, the following conditions will further emerge from equations (218):

$$\frac{\partial^2\Phi}{\partial q_\lambda \partial q_\mu} - \frac{\partial\Phi}{\partial q_\lambda}\frac{\partial \ln g^{\lambda\lambda}}{\partial q_\mu} - \frac{\partial\Phi}{\partial q_\mu}\frac{\partial \ln g^{\mu\mu}}{\partial q_\lambda} = 0 \qquad (\lambda \neq \mu)$$

and

$$\frac{\partial^2 Q}{\partial q_\lambda \partial q_\mu} - \frac{\partial Q}{\partial q_\lambda}\frac{\partial \ln g^{\lambda\lambda}}{\partial q_\mu} - \frac{\partial Q}{\partial q_\mu}\frac{\partial \ln g^{\mu\mu}}{\partial q_\lambda} = 0 \qquad (\lambda \neq \mu)\,.$$

Now, should the form of the arc-length element and the potential function Φ be characterized for arbitrary ν in the **Stäckel** case $\nu = 0$, then one would have to start from the Ansatz:

$$W = W_1(q_1) + \ldots + W_n(q_n) , \tag{238}$$

as one does in that case. In that way, it is important that for an index of the first kind, one must have:

$$\frac{\partial Q}{\partial q_\rho} = \dot{q}_\rho L_\rho , \tag{239}$$

in which L_ρ means a linear function of the impulse components of the first kind, and that one will then have:

$$\frac{dp_\rho}{dq_\rho} = -\frac{\dfrac{\partial Q}{\partial q_\rho}}{\dfrac{\partial Q}{\partial p_\rho}} = -\frac{\dfrac{\partial Q}{\partial q_\rho}}{\dot{q}_\rho} = -L_\rho \tag{240}$$

for an index of the first kind ([225]).

In the special case that was first treated by **T. Levi-Civita**, namely, the one in which all indices were of the first kind, one can evaluate those relations in the following way: If the Ansatz (238) is a complete solution then it will represent a field of extremals for a certain choice of the constants in the value of the extremal integral ([226]). For example, one can choose a certain field when one prescribes the impulse components of the field:

$$p_1^{(0)} = \alpha_1 , \ldots, \; p_n^{(0)} = \alpha_n \tag{241}$$

arbitrarily at any point, say, the initial point ([227]). Now, since the individual impulse components:

$$p_\rho = \frac{\partial W}{\partial q_\rho} = \frac{dW_\rho}{dq_\rho}$$

When κ, λ, μ are all indices of the second kind, it will then follow from the latter that:

$$\frac{\partial^2 g^{\kappa\kappa}}{\partial q_\lambda \partial q_\mu} - \frac{\partial g^{\kappa\kappa}}{\partial q_\lambda}\frac{\partial \ln g^{\lambda\lambda}}{\partial q_\mu} - \frac{\partial g^{\kappa\kappa}}{\partial q_\mu}\frac{\partial \ln g^{\mu\mu}}{\partial q_\lambda} = 0 .$$

([225]) **P. Burgatti**, "Determinazione dell'equazione di Ham.-Jac. integrabili mediante la seperazione delle variabili," Rend. Accad. Lincei (5) **20**[1] (1911), pp. 108; **F. A. Dall'Acqua**, Rend. circ. mat. Palermo **33** (1912), pp. 341.

([226]) Since that case is "essentially geodetic," the extremals are the geodetic lines of the arc-length element.

([227]) That is because the constants $k, c_1, \ldots, c_{n-1}$ in the complete solution can be expressed in terms of the initial values and conversely.

depend upon only q_ρ, so they will have the same value in the entire manifold q_ρ = const., it will suffice that p_ρ is well-defined along the M_1 ($q_1 = 0$, …, $q_{\rho-1} = 0$, $q_{\rho+1} = 0$, …, $q_n = 0$), i.e., along the q_ρ-axis. However, upon advancing along that M_1, all impulse components, with the exception of p_ρ, will keep the same value that they had at the initial point, such that one can introduce:

$$q_1 = 0\ ,\ \ldots,\ q_{\rho-1} = 0\ ,\ q_{\rho+1} = 0\ ,\ \ldots,\ q_n = 0\ ,$$

$$p_1 = \alpha_1\ ,\ \ldots,\ p_{\rho-1} = \alpha_{\rho-1}\ ,\ p_{\rho+1} = \alpha_{\rho+1}\ ,\ \ldots,\ p_n = \alpha_n$$

into (240). However, that will then give the equation for determining $p_\rho\,(q_\rho)$:

$$\frac{dp_\rho}{dq_\rho} + b_\rho\,(q_\rho)\,p_\rho = \lambda_\rho\,(q_\rho)\ , \tag{242}$$

in which λ_ρ is a linear form in the constants α_1 , …, $\alpha_{\rho-1}$, $\alpha_{\rho+1}$, …, α_n whose coefficients depend upon q_ρ. The integration of that linear differential equation ([228]) will produce p_ρ as a linear form in the α_1 , …, α_n whose coefficients are functions of q_ρ . If one were to replace one of the initial values of the n impulse components – say α_n – with the constant k that appeared before in the **Hamilton-Jacobi** equation ([229]) then one will get:

$$\frac{\partial W}{\partial q_\rho} = \frac{dW_\rho}{dq_\rho} = p_\rho = \alpha_1\,\varphi_\rho^{(1)} + \cdots + \alpha_{n-1}\,\varphi_\rho^{(n-1)} + \sqrt{2k - A\cdot\psi_\rho}\ , \tag{243}$$

in which A is a quadratic form in the α_1 , …, α_{n-1} , and $\varphi_\rho^{(\lambda)}$ and ψ_ρ are functions that depend upon only the coordinate q_ρ .

The general case (ν indices of the first kind, $n - \nu$ indices of the second kind) can be easily resolved in an analogous way. The relations (240) are valid for the indices of the first kind, in which the L_ρ are linear forms in just the impulse components with indices of the first kind. Corresponding to (243), one will get the impulse components from that as linear forms in the initial values of the impulse components α_1 , …, α_ν:

([228]) Upon considering the initial value of p_ρ :

$$p_\rho\,(0) = \alpha_\rho\ .$$

([229]) Due to the **Hamilton-Jacobi** equation, α_n can be determined from the quadratic equation:

$$g^{nn}\,\alpha_n^2 + 2\,(g^{n1}\alpha_1 + \cdots + g^{n,n-1}\alpha_{n-1})\,\alpha_n = 2k - \sum_{\lambda,\mu=1}^{n-1} g^{\lambda\mu}\alpha_\lambda\,\alpha_\mu\ ,$$

in which one naturally introduces the $q_1^{(0)}$, …, $q_n^{(0)}$ into $g^{\lambda\mu}$.

(244) $$\frac{\partial W}{\partial q_\rho} = \frac{dW_\rho}{dq_\rho} = p_\rho = \alpha_1 \varphi_\rho^{(1)} + \cdots + \alpha_\nu \varphi_\rho^{(\nu)} \quad (\rho = \text{index of the first kind}).$$

Now, if λ is an index of the second kind and one advances in the q_λ-direction such that only q_λ and p_λ are variable then the relation will follow from the **Hamilton-Jacobi** equation (236) ([230]):

(245) $$g^{\lambda\lambda} p_\lambda^2 + 2p_\lambda (g^{\lambda 1} \alpha_1 + \cdots + g^{\lambda n} \alpha_n) = 2\,(k - \Phi) - \sum_{\sigma,\tau \neq \lambda} g^{\sigma\tau} \alpha_\sigma \alpha_\tau ,$$

from which one will get p_λ in the form:

(245.a) $$p_\lambda = \alpha_1 \varphi_\lambda^{(1)} + \cdots + \alpha_\nu \varphi_\lambda^{(\nu)} \pm \sqrt{\chi_\lambda + F_\lambda(\alpha_1, \ldots, \alpha_\nu) + \eta_\lambda(\beta_1, \ldots, \beta_{n-\nu})} \ .$$

The χ_λ in that is a function of only q_λ, while F_λ represents a quadratic form in the α_1, …, α_ν with coefficients that are functions of q_λ. Moreover, the relation (236.b) is used in that, and the quantities:

(245.b) $$\alpha_{\nu+\sigma}^2 + 2\alpha_{\nu+\sigma} \sum_{\tau=1}^{\nu} l_\tau \alpha_\tau = \beta_\sigma$$

are introduced into it as new constants ([231]) that shall enter in place of $\alpha_{\nu+1}$, …, α_n. The η_λ are then linear forms in the β_1, …, $\beta_{n-\nu}$ with coefficients that are functions of q_λ. Along with the relations (244) that hold for the indices of the first kind, one also has the equations:

(246) $$\frac{\partial W}{\partial q_\lambda} = \frac{dW_\lambda}{dq_\rho} = \alpha_1 \varphi_\lambda^{(1)} + \cdots + \alpha_\nu \varphi_\lambda^{(\nu)} \pm \sqrt{\chi_\lambda + F_\lambda(\alpha_1, \ldots, \alpha_\nu) + g_\lambda(\beta_1, \ldots, \beta_\nu)}$$

for the indices of the second kind. If no indices of the first kind are present then the equations (246) will obviously go to the **Stäckel** expressions (232). That will come to mind all the more when the energy constant k is introduced in place of, say $\beta_{n-\nu}$ ([232]), which will give:

(246.a) $$\frac{\partial W}{\partial q_\lambda} = \frac{dW_\lambda}{dq_\lambda}$$

([230]) In which, k is naturally expressed in terms of α_1, …, α_n. The coefficients $g^{\rho\sigma}$ and Φ in (245) are thought of as functions that depend upon only q_λ.

([231]) Naturally, the l_τ are constants here, since from (236.b), they depend upon only coordinates with indices of the first kind.

([232]) One easily achieves that with the help of the relation:

$$\sum_{\sigma,\tau=1}^{\nu} \bar{g}^{\sigma\tau} \alpha_\sigma \alpha_\tau + \sum_{\rho=1}^{n-\nu} \bar{g}^{\nu+\rho,\nu+\rho} \beta_\rho = 2(k - \bar{\Phi}) ,$$

in which the overbar means that one has set $q_1 = 0$, …, $q_n = 0$.

$$= \alpha_1 \varphi_\lambda^{(1)} + \cdots + \alpha_\nu \varphi_\lambda^{(\nu)} \pm \sqrt{\chi_\lambda^* + F^*(\alpha_1, \ldots, \alpha_\nu) + (\beta_1 \varphi_\lambda^{(1)} + \ldots + \beta_{n-\nu-1} \varphi_\lambda^{(n-\nu-1)}) + k \psi_\lambda^*} \ .$$

Conversely, one can arrive at the form of the arc-length element and potential function for which the mechanical system is soluble by separation of variables by eliminating the constants from those formulas ([233]).

If one has established the general form of the arc-length element and potential function for which the separation of variables then the question would arise of whether a given mechanics problem (with given arc-length element and potential) can be converted into the desired form by a coordinate transformation. The investigation of that question has been taken up in only some individual cases ([234]).

([233]) For $n = 3$, all possible cases were worked out by **F. A. Dall'Acqua**, *loc. cit.* ([209]), pp. 398.

([234]) Cf., **O. Haupt** and **E. Hilb**, "Über die Transformation Liouvillescher Mannigfaltigkeiten," Gött. Nachr. (1924), pp. 77. One should also confer the work of **J. Weinacht**, "Über die bedingt periodische Bewegung eines Massenpunktes," Math. Ann. **91** (1924), pp. 279, in which the cases of $n = 2$ and $n = 3$ were treated.