From: Ed., H. Geiger and K. Scheel, *Handbuch der Physik*, Bd. V, *Grundlagen der Mechanik*, *Mechanik der Punkte und starren Körper*, Chap. 3, Springer, Berlin, 1927.

The Hamilton-Jacobi Theory of Dynamics

BY

L. NORDHEIM, Göttingen and **E. FUES**, Stuttgart

Translated by D. H. Delphenich

Contents

---

1. General problem statement. – In the previous chapter, the principles of mechanics were presented and discussed in their most-general form, as well as the equations of motion that arise from them. Therefore, the most natural question to ask next would be how one might actually carry out the integration of those equations, and in particular whether one cannot infer some essential conclusions from their character as differential equations of mechanics. In fact, that is the case, to a large extent, at least for problems for which a kinetic potential exists (cf., Chap. 2, no. **10**).

In order to do that, the main ideas of the theory of integration of **Jacobi** ([1]) and **Hamilton** ([2]) will be developed systematically. It has great significance, on the one hand, for celestial mechanics, and on the other, for that of the atom since neither constraints nor non-conservative forces exist for either of them, as least as long as one ignores tidal forces or radiation reactions, resp.

The structure of that theory is described in three steps: First of all, one tries to obtain the simplest-possible form for the differential equations. That leads to the canonical equations of mechanics. Secondly, one can ask about the general laws of the transformations of those differential equations that preserve their form. That leads to canonical transformations and the theory of their most-important invariants. Thirdly, the actual theory of integration of the system of canonical equations is presented, which consists of exhibiting and integrating the Hamiltonian partial differential equation.

The restriction to systems with a kinetic potential that was introduced before is the same one that makes Hamilton's principle into a true variational principle. Therefore, an application of the methods of the calculus of variations will lead to a great simplification, and the deeper meaning of the peculiar Hamilton-Jacobi integration procedure will also be revealed by it, which we shall return to in the conclusion ([3]).

Above all, the book by **Whittaker** ([4]) should be cited as a modern reference. Jacobi ([5]) gave the first systematic development that also had fundamental significance in his famous lecture on dynamics. Many important connections with it, especially in regard to the theory of canonical transformations, are also included in **Lie**'s ([6]) investigations.

Our starting point is Hamilton's principle. We shall then assume that a kinetic potential exists (cf., Chap. 2, no. **10**), which is a function of the coordinates and velocities $L(q_k, \dot{q}_k, t)$ that is supposed to satisfy the equations of the system of Hamilton's principle (see Chap. 2, no. **22**):

$$\int_{t_1}^{t_2} L(q_k, \dot{q}_k, t)\, dt = \text{extremum}. \tag{1}$$

According to the rules of the calculus of variations, they read:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} = 0 \qquad (k = 1, 2, \ldots, f). \tag{2}$$

L can have the most-general form in it, so it can also include time t, and it likewise allows forces that depend upon velocities in the sense of Chap. 2, no. **10**. For example, for an isolated electron

([1]) **G. C. Jacobi**, *Vorlesungen über Dynamik*, *Werke Supplementband*, 2nd ed., Berlin, 1888.

([2]) **W. A. Hamilton**, Brit. Assn. Rep., 1834, pp. 513; Phil. Trans. (1835), pp. 95.

([3]) The following presentation is connected, in many respects, and in particular, the employment of the calculus of variations, with the one that one of us (**Nordheim**) heard about in **Hilbert**'s lectures. At this point, we would also like to warmly thank Herrn Geh.-Rat **Hilbert** for his kind permission to use them.

([4]) **E. T. A. Whittaker**, *Analytical Mechanics*, 2nd ed., Cambridge, 1917. German translation by **F.** and **K. Mittelstein-Scheid**, Berlin, Springer, 1924.

([5]) See rem. 1 on pp. 1.

([6]) **S. Lie**, *Theorie de Transformationsgruppen*, Bd. I-III, Leipzig, 1888-1890, in particular, Bd. II.

in the most-general case, i.e., when one considers the theory of relativity and the influence of arbitrary electric and magnetic fields that arise from potentials φ and $\mathfrak{A}$, the Lagrangian will be:

$$L = m_0 c^2 \left(1 - \sqrt{1 - \frac{\mathfrak{v}^2}{c^2}}\right) + \frac{e}{c} \mathfrak{A} \mathfrak{v} - e\varphi . \tag{3}$$

One calls the expression on the left in (2) the variational derivative of L with respect to q_k. We would like to denote it with the abbreviation $[L]_{q_k}$:

$$[L]_{q_k} \equiv \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} . \tag{4}$$

2. Reduction of the problem to canonical form. – We shall now take the first step and look for the new and simpler form for the variational problem. In formula (1) of no. **1**, L was a function of the q_k, $\dot{q}_k$, and possibly t. Obviously, one will get a problem that is simpler, in a certain sense, if one can eliminate the derivatives $\dot{q}_k$. To that end, we simply introduce the $\dot{q}_k$ as new variables that are to be varied independently by setting:

$$\dot{q}_k - k_k = 0 . \tag{1}$$

The variational problem will then read:

$$\int_{t_1}^{t_2} L(q_k, k_k, t)\, dt = \text{extremum}, \tag{2}$$

which will generally mean that equations (1) must now be added as an auxiliary condition. We will now have a variational problem with $2f$ unknowns and f auxiliary conditions.

The latter can be treated in the known way by the method of Lagrange factors ([1]). One multiplies them by the still-undetermined factors λ_k and treats the absolute variational problem that now has $3f$ unknowns:

$$\int_{t_1}^{t_2} \{L + \sum_k \lambda_k (\dot{q}_k - k_k)\}\, dt = \text{extremum}. \tag{3}$$

([1]) Naturally, in the present problem, the neighboring curves must also satisfy the auxiliary conditions (1). One must them make use of that fact *before* the variation, as opposed to the situation with ordinary non-holonomic auxiliary conditions, for which the neighboring curves do not satisfy the auxiliary conditions, as was shown in Chap. 2, no. **20** and **27**.

Here, one can determine the λ_k from the demand that the variational derivatives with respect to the new variables k_l must vanish:

$$\left[L + \sum_k \lambda_k (\dot{q}_k - k_k)\right]_{k_l} = 0 .$$

Since the $\dot{k}_k$ do not, in fact, appear in the bracketed expression, those equations will reduce to:

$$\frac{\partial L}{\partial k_l} - \lambda_l = 0 , \qquad \lambda_l = \frac{\partial L}{\partial k_l} .$$

The λ_l are thus determined by that. One can substitute their values and then obtain a free variational problem with $2f$ undetermined functions:

$$\int_{t_1}^{t_2} \{L(q_k, k_k, t) + \sum_k \frac{\partial L}{\partial k_l} (\dot{q}_k - k_k)\}\, dt = \text{extremum}. \tag{4}$$

In that way, the extremum is to be chosen from all functions $q_k(t)$ and $k_k(t)$, but the k_k need not be subject to any boundary conditions since their derivatives do not enter into the integral, nor does (1) of no. **1** include any conditions on the $\dot{q}_k$. One can see that the requirement (4) is actually completely equivalent to (1) in no. **1** as follows: The conditions for the desired functions read:

$$\left[L + \sum_k \frac{\partial L}{\partial k_k} (\dot{q}_k - k_k)\right]_{q_k} = 0 ,$$

$$\left[L + \sum_k \frac{\partial L}{\partial k_k} (\dot{q}_k - k_k)\right]_{k_k} = - \frac{\partial}{\partial k_k}\left\{L + \sum_k \frac{\partial L}{\partial k_k} (\dot{q}_k - k_k)\right\} = - \frac{\partial^2 L}{\partial k_k^2} (\dot{q}_k - k_k) = 0 .$$

The second of them says nothing besides the fact that one must have $\dot{q}_k = k_k$, except for the singular cases where $\partial^2 L / \partial k_k^2 = 0$, which were excluded here. If we substitute that into the first equation then we will return to the original form (1) of no. **1**.

That proof of equivalence is necessary since (3) [(4), resp.] by itself does not by any means coincide with (1) of no. **1**. That is because in (1) of no. **1**, one seeks the extremum from among all quantities that arise when one substitutes all arbitrary functions $q_k(t)$ in L. In that way, the $\dot{q}_k$ are naturally determined, as well. By contrast, the k_k are still taken to be arbitrary functions. The domain from which the extremum must sought will be much larger correspondingly. In fact, it can also be shown that in the event that the arbitrary trajectory makes the integral (1) of no. **1** into a true minimum, that cannot at all be the case for (4), but only that the integral assumes a saddle value then in such a way that it is initially made a maximum relative to the $k_k(t)$ for a fixed, but arbitrarily-chosen q_k (t), and it is only after that determination that the q_k (t) are chosen in such a way that the integral will become a minimum relative to all of its variations. That was shown by **Hilbert** in his lectures.

However, for the purposes of mechanics, the character of an extremum, i.e., whether it is a maximum, minimum, or (as it is here) a saddle value, is entirely irrelevant. The only thing that matters is whether the variational derivatives are identical for the various forms that the variational problem can take, and therefore the curves that make the integral an extremum, which are just the desired trajectories. That is why we shall not go into the details of that here, but only remark that for a sufficiently-small neighborhood of the true motion, Hamilton's integral (2) will become a true minimum ([1]).

We will address the form (4) later on. Here, we shall first take yet another step further by introducing the *generalized impulses* (see Chap. 2, no. **2**):

$$p_k = \frac{\partial L(q_l, k_l)}{\partial k_k} = \frac{\partial L(q_l, \dot{q}_l)}{\partial \dot{q}_k} \tag{5}$$

as new unknowns in place of the k_k. The k_k will become functions of p_k, q_k, and possibly t by means of (5), and (4) will take the form:

$$\int_{t_1}^{t_2} \left\{ \sum_k p_k \, \dot{q}_k - L(p_k, q_k, t) \right\} dt = \text{extremum}, \tag{6}$$

in which:

$$H = -L + \sum_k k_k \frac{\partial L}{\partial k_k} \equiv -L + \sum_k \dot{q}_k \frac{\partial L}{\partial \dot{q}_k} \tag{7}$$

means the so-called *Hamiltonian function*. The k_k in H are thought of as being expressed in terms of p_k, q_k, and t. Now, equation (6) has the simplest form that an absolute variational problem can assume since only the derivatives of a series of variables appear, and they appear only linearly and while multiplied by the other variables themselves. That is why they are also called *canonical*. One also calls the q_k and p_k *canonical variables* accordingly, and in particular, the p_k are the *canonically-conjugate impulses* to the q_k. Naturally, the proof that (4) is equivalent to (6) is no longer necessary here since (6) emerges from (4) by a direct transformation.

Moreover, one can easily return to the variables k_k ($\dot{q}_k$, resp.), p_k from the variables p_k, q_k. In order to do that, one partially differentiates H with respect to p_k:

$$\frac{\partial H}{\partial p_k} = \frac{\partial}{\partial p_k}\left(-L + \sum_l k_l \, p_l \right) = -\sum_l \frac{\partial L}{\partial k_k}\frac{\partial k_l}{\partial p_k} + \sum_l \frac{\partial k_l}{\partial p_k} p_l + k_k = k_k . \tag{8.a}$$

It further follows from this that:

([1]) See, e.g., the book by **Whittaker**, *Analytische Dynamik*, pp. 265, that cited in rem. ([1]) on page 2.

$$\left.\begin{aligned} H &= -L + \sum_k k_k\, p_k = -L + \sum_k \frac{\partial H}{\partial p_k}\, p_k\,, \\ L &= -H + \sum_k \frac{\partial H}{\partial p_k}\, p_k\,. \end{aligned}\right\} \tag{8.b}$$

The transition from L to H will then have the same form as the inverse transition from H to L. One refers to it as the *Legendre transformation*, and it plays a role in many other branches of mathematics and physics. For example, in thermodynamics, it mediates the transition between the various thermodynamic potentials.

The differential equations of the variational problem, i.e., the equations of motion of the system, will take an especially simple form in the new variables. They will then read:

$$\left[\sum_l p_l\, \dot{q}_l - H\right]_{p_k} = 0\,,$$

$$\left[\sum_l p_l\, \dot{q}_l - H\right]_{q_k} = 0\,,$$

and as one sees immediately, they will reduce to:

$$\left.\begin{aligned} \frac{dq_k}{dt} &= \frac{\partial H}{\partial p_k}\,, \\ \frac{dp_k}{dt} &= -\frac{\partial H}{\partial q_k}\,. \end{aligned}\right\} \tag{9}$$

Those are the so-called *canonical equations of mechanics*, which define the starting point for most investigations of higher dynamics. Instead of the second-order system of f Lagrange differential equations (2) for the q_k that was discussed in no. **1**, they define a first-order system of $2f$ differential equations for the q_k and p_k . However, their derivation is completely equivalent to the former.

One can also perform the transformation of the differential equations of a mechanical system into canonical form when not all of the auxiliary conditions have been eliminated, but some of them are carried separately. If those auxiliary conditions are:

$$\varphi_r\,(q_k, t) = 0$$

then the corresponding Hamiltonian equations will read:

$$\left.\begin{aligned} \dot{q}_k &= \frac{\partial H}{\partial p_k}, \\ \dot{p}_k &= -\frac{\partial H}{\partial q_k} + \sum_r \lambda_r \frac{\partial \varphi_r}{\partial q_k} . \end{aligned}\right\} \tag{10}$$

If the constraints have the non-holonomic form:

$$\sum_r a_{rk}\, \delta q_k = 0$$

then the second row will be replaced with:

$$\dot{p}_k = -\frac{\partial H}{\partial q_k} + \sum_r \lambda_r\, a_{rk} . \tag{10.a}$$

However, the use of those equations would hardly be advantageous since they have lost their symmetry ([1]).

We now ask what the mechanical meaning of the quantity H might be. If the kinetic energy T is a homogeneous quadratic function of the $\dot{q}_k$, as in the rule, then from Euler's theorem for homogeneous functions:

$$T = \tfrac{1}{2} \sum_k \frac{\partial T}{\partial \dot{q}_k} \dot{q}_k . \tag{11}$$

Since one must have $L = T - U$, by assumption, one will then have:

$$\sum p_k\, \dot{q}_k = \sum_k \frac{\partial L}{\partial \dot{q}_k} \dot{q}_k = \sum_k \frac{\partial T}{\partial \dot{q}_k} \dot{q}_k = 2\, T ,$$

in the event that the potential energy V does not depend upon velocities. Therefore, under the stated assumptions:

$$H = -L + \sum_k p_k\, \dot{q}_k = -T + U + 2T = T + U \tag{12}$$

will be the total energy in the system.

The recipe for exhibiting the canonical equations is then conceptually simple. One needs only to know the energy as a function of the coordinates and impulse in order to be able to write them down directly. From (12), one must generally observe that this simple mechanical meaning of H is true only under the assumption that (11) is true. For other cases, e.g., when things are referred

([1]) On this, see **T. Pöschl**, C. R. Acad. Sci. Paris **156** (1913), pp. 1829; **S. Dautheville**, Bull. Soc. Math. France **37** (1909), pp. 120.

to a rotating coordinate system, H is no longer the energy by any means, and one must return to equation (7) in order to determine the Hamiltonian function ([1]).

One will get a first integral of the equations of motion directly when the Hamiltonian function does not include time explicitly. If one multiplies the canonical equations (9) by $\dot{q}_k$ ($\dot{p}_k$, resp.) then it will follow that:

$$\frac{dH}{dt} = \sum_k \frac{\partial H}{\partial p_k} \dot{p}_k + \sum_k \frac{\partial H}{\partial \dot{q}_k} \dot{q}_k = \sum_k \dot{q}_k \, \dot{p}_k - \sum_k \dot{p}_k \, \dot{q}_k = 0 . \tag{13}$$

$$H = \text{const.} = W$$

is then an integral of the canonical equations. In the simplest case that was mentioned above, that would be nothing but the *law of energy*.

Furthermore, if the Hamiltonian function does not include a coordinate (e.g., q_1) explicitly then it will follow immediately that:

$$\dot{p}_1 = - \frac{\partial H}{\partial q_1} = 0 , \qquad p_1 = \text{const.} \tag{14}$$

We will once more have an integral of the canonical equations then. For example, the law of areas p_φ = const. for *Keplerian motion* will follow in that way, and its Hamiltonian function is written:

$$H = \frac{1}{2m}\left(p_r^2 + \frac{1}{r^2} p_\varphi^2 \right) - \frac{\varphi}{r} \tag{11}$$

in plane polar coordinates r, φ. It is probably based upon that example, in which φ has the meaning of the azimuth in the orbital plane, that one calls coordinates that the Hamiltonian function does not depend upon *cyclic variables*. That case will always occur when the energy does not depend upon the incidental value of one coordinate, such as, e.g., when it does not change under a translation or rotation of the entire system. In that way, one would get, e.g., the law of the center of gravity and the law of areas for free systems with no further analysis. We shall come back to that idea from a more general standpoint in nos. **9** and **11**. (Cf., also no. **11** of the foregoing Chap. 2.)

3. Canonical transformations. – We now move on to our second question and examine what sort of transformations can be performed on the variables that will preserve the canonical form of the equations of motion.

We then seek the substitutions:

([1]) For the Hamiltonian function and the theory of integration in relativistic mechanics, see Chap. 10 of this volume of the *Handbuch*. Furthermore, see **J. Frenkel**, *Lehrbuch der Elektrodynamik*, Chap. 10, pp. 330, *et seq.*, Berlin, 1926.

$$\left.\begin{aligned} q_i &= q_i(Q_k, P_k, t), \\ p_i &= p_i(Q_k, P_k, t) \end{aligned}\right\} \tag{1}$$

that will take the variational problem [(6), no. **2**] into an equivalent one with a new Hamiltonian function K :

$$\int_{t_1}^{t_2} \left\{ \sum_k P_k \dot{Q}_k - K(P_k, Q_k, t) \right\} dt = \text{extremum}. \tag{2}$$

We will not actually succeed in making the two integrals identical to each other in that way, but only in making them assume their extrema simultaneously, i.e., if the integral (6) in no. **6** assumes its extremal value for the functions $q_k(t)$, $p_k(t)$ then the integral (2) shall do the same thing for the functions $Q_k(t)$, $P_k(t)$ that emerge from q_k and p_k under the inverse substitution to (1).

That will be guaranteed if and only if the two integrands differ by merely the complete derivative with respect to t of an otherwise-arbitrary function $\Phi(Q_k, P_k, t)$. The integral will be independent of the path for such a thing, and it will give a constant value that will in no way affect the occurrence of an extremum for all cases in which the integration limits are fixed. The condition that the Q_k and P_k must fulfill:

$$\sum_k p_k \dot{q}_k - H = \sum_k P_k \dot{Q}_k - K + \frac{d\Phi}{dt}(P, Q, t). \tag{3}$$

Naturally, that condition must also be true for all non-mechanical, varied integration paths in p, q, t-space. Now, since no kinematical constraints are supposed to exist between the q_k, one can also write (3) more clearly in the form:

$$\sum_k p_k \Delta q_k - H \Delta t = \sum_k P_k \Delta Q_k - K \Delta t + \Delta\Phi, \tag{4}$$

which is a condition that must be fulfilled for a completely-arbitrary choice of the differentials Δq_k, ΔQ_k, Δt. In that way, $\Delta\Phi$ will be explained by:

$$\Delta\Phi = \sum_k \frac{\partial\Phi}{\partial Q_k} \Delta Q_k + \sum_k \frac{\partial\Phi}{\partial P_k} \Delta P_k + \frac{\partial\Phi}{\partial t} \Delta t,$$

but the ΔP_k in that are always fixed by the Δq_k, ΔQ_k, Δt since (for a well-defined Δt) obviously the $2f$ relations:

$$\Delta q_k = \sum_k \frac{\partial q_k}{\partial Q_k} \Delta Q_k + \sum_k \frac{\partial q_k}{\partial P_k} \Delta P_k + \frac{\partial q_k}{\partial t} \Delta t,$$

$$\Delta p_k = \sum_k \frac{\partial p_k}{\partial Q_k} \Delta Q_k + \sum_k \frac{\partial p_k}{\partial P_k} \Delta P_k + \frac{\partial p_k}{\partial t} \Delta t$$

must always exist for the $4f$ differentials Δq_k, Δp_k, ΔQ_k, ΔP_k . Naturally, we must assume that the functional determinant of the transformation (1) is not equal to zero here.

In order to arrive at actual conditions for the transformation equations (1) from (4), we introduce q_k in place of the P_k in Φ by imagining that we have solved the relations:

$$q_k = q_k\,(P_i, Q_i, t)$$

for the P_k :

$$P_k = P_k\,(q_i, p_i, t)\ .$$

We assume that this solution is possible. In that way, Φ will go to a function $V\,(q_k, Q_k, t)$. From (4), one will then have:

$$\sum_k p_k\,\Delta q_k - H\,(p_k, q_k, t)\,\Delta t = \sum_k P_k\,\Delta Q_k - K\,(P_k, Q_k, t)\,\Delta t + \Delta V(q_k, Q_k, t)\,, \tag{4.a}$$

with:

$$\Delta V = \sum_k \frac{\partial V}{\partial q_k}\Delta q_k + \sum_k \frac{\partial V}{\partial Q_k}\Delta Q_k + \frac{\partial V}{\partial t}\Delta t\,.$$

In order for equation (4.a) to be fulfilled identically, we must set the factors of Δq_k, ΔQ_k, At on both sides equal to each other:

$$\left.\begin{aligned} p_k &= \frac{\partial V}{\partial q_k}, \\ P_k &= -\frac{\partial V}{\partial Q_k}, \\ K &= H + \frac{\partial V}{\partial t}. \end{aligned}\right\} \tag{5}$$

Since one can generally calculate the q_k as functions of the P_k, Q_k from the equations in the second row, and then calculate the p_k in the first row from them, *equations* (5) *will always give a canonical transformation for an arbitrary choice of the function* $V\,(q_k, Q_k, t)$, and in that way, the third row will yield the new Hamiltonian K. The function V is called the *generator of the transformation*. The new canonical equations read:

$$\frac{dP_k}{dt} = -\frac{\partial K}{\partial Q_k}, \qquad \frac{dQ_k}{dt} = \frac{\partial K}{\partial P_k}, \qquad K = H + \frac{\partial V}{\partial t}.$$

In particular, if V does not include time explicitly then one will have simply:

$$K = H\,.$$

It is very remarkable that the canonical transformations are independent of the special nature of mechanical problems. The property of a transformation that it is canonical then does not depend at all upon the nature of the problem in question but is peculiar to the transformation itself.

We have preferred to use the variables q_k, Q_k in the generator V. We would just as well have taken any f of the variables q_k, p_k and any f of the Q_k, P_k . The general result can then be expressed as follows ([1]): Let $V(x_k, X_k, t)$ be an arbitrary function of the $2f + 1$ variables x_k, X_k, t, such that the x_k ($k = 1, \ldots, f$) are any functions of the variables q_k, p_k, and the X_k are any functions of the Q_k, P_k:

$$\left.\begin{aligned} y_k &= \pm\frac{\partial V}{\partial x_k}, \\ Y_k &= \mp\frac{\partial V}{\partial X_k}, \\ K &= H + \frac{\partial V}{\partial t} \end{aligned}\right\} \tag{6}$$

will then be a canonical transformation. Therefore, y_k will be conjugate to x_k and Y_k will be conjugate to X_k, and the upper sign will be valid when one differentiates with respect to a coordinate, while the lower sign will be valid when one differentiates with respect to an impulse. One often needs the canonical transformation in the form, e.g.:

$$\left.\begin{aligned} V &= V(q_k, P_k, t), \\ p_k &= +\frac{\partial V}{\partial q_k}, \\ Q_k &= +\frac{\partial V}{\partial X_k}. \end{aligned}\right\} \tag{5.a}$$

Each transformation of the configuration coordinates alone:

$$q_k = q_k(Q_l, t),$$

which is referred to as a *point transformation* since it will take each point in configuration space of the q_k to another such thing, is also canonical. One only needs to take the transformation function to be:

$$V = -\sum q_k(Q_l)\, p_k, \tag{7}$$

and from (6), one will have:

([1]) See **M. Born**, *Vorlesungen über Atommechanik*, Berlin, 1925, pp. 35. Cf., in addition, the detailed exposition in the following Chap. 4, no. 3, of this volume of the *Handbuch*.

$$q_k = -\frac{\partial V}{\partial p_k} = q_k\,(Q_l)\;.$$

The identity transformation is obtained from:

$$V = -\sum_k Q_k\,p_k\;. \tag{8}$$

Above and beyond that, the theory of canonical transformations allows one to introduce more general dynamical coordinates in such an exceptionally free manner that their choice can mostly be suited to each problem precisely. Naturally, the character of the variables Q_k, P_k as configuration and impulse coordinates will be lost for the general transformations (6). It is only when taken all together that they will give a picture of the configuration and state of motion of the system in question. Due to their mathematical relationship to the contact transformations of geometry, those transformations will frequently be given the name of *contact transformations*.

One can also perform canonical transformation that fulfill certain auxiliary conditions when the latter can be brought into the form of a relationship between the old and new coordinates:

$$\Omega_r\,(q_k, Q_k, t) = 0\;. \tag{9}$$

That can be added to the identity (4) by simply introducing Lagrange multipliers λ_r , and one will then get the equations that determine the corresponding canonical transformations:

$$\left.\begin{aligned} P_k &= \frac{\partial V}{\partial q_k} + \sum_r \lambda_r \frac{\partial \Omega_r}{\partial Q_k}\,, \\ p_k &= -\frac{\partial V}{\partial q_k} - \sum_r \lambda_r \frac{\partial \Omega_r}{\partial q_k}\,, \\ K &= H + \frac{\partial V}{\partial t} + \sum_r \lambda_r \frac{\partial \Omega_r}{\partial t}\,, \end{aligned}\right\} \tag{10}$$

and together with the relations (9), they suffice precisely to determine the quantities q_k, p_k, λ_r as functions of the Q_k, P_k. One specialization of that is, e.g., the existence of an auxiliary condition:

$$\varphi\,(q_k, t) = 0$$

for the original coordinates.

Ultimately, one must also be able to multiply the left-hand side of (3) by a constant factor λ without affecting the property of the transformation that it is canonical. That will lead to, e.g., transformations of the type:

$$P_k = p_k\,, \qquad Q_k = \lambda\,q_k\,, \qquad K = \lambda\,H \tag{11}$$

that will be used many times. By contrast, the general form of the contact transformation that is customary in geometry, where λ is an arbitrary function of the variables, is inapplicable here.

As was said before, the canonical transformations are independent of the choice of special Hamiltonian function. Therefore, if one wishes to have only the conditions for the transformation of the p_k, q_k into the P_k, Q_k themselves then one can restrict oneself to the variations with $\Delta t = 0$ in (4), i.e., treat t like a constant parameter. If we characterize those variations by a δ to distinguish them, then we can write the conditions for canonical transformations in the form:

$$\sum_k p_k\,\delta q_k = \sum_k P_k\,\delta Q_k + \delta\,\Phi(P_k, Q_k, t)\,, \tag{12}$$

in which no mention of the special nature of the mechanical problem is found at all. The variations Δ and δ are then described by ([1]):

$$\left.\begin{aligned} \Delta F(p_k, q_k, t) &= \sum_k \frac{\partial F}{\partial q_k}\Delta q_k + \sum_k \frac{\partial F}{\partial p_k}\Delta p_k + \frac{\partial F}{\partial t}\Delta t, \\ \delta F(p_k, q_k, t) &= \sum_k \frac{\partial F}{\partial q_k}\delta q_k + \sum_k \frac{\partial F}{\partial p_k}\delta p_k, \end{aligned}\right\} \tag{13}$$

respectively. Equation (12) will then have the same degree of generality as (4) in terms of characterizing the transformation, and one needs only the latter form in order to determine the new Hamiltonian function. Naturally, one can also introduce the q into Φ in place of the P beforehand and obtain the explicit equations of transformation (5) with the help of the function $V(q_k, Q_k, t)$.

With the introduction of canonical transformations, the most important step has already been taken in regard to the theory of integration of the mechanical equations that will be presented in nos. **12**, *et seq.* Knowledge of what is contained in nos. **4** to **11**, namely, the further exposition of the properties of canonical transformations, is not necessarily required for an understanding of it. They can therefore be skipped over in an initial study of the topic.

4. Introducing time as a canonical variable. – One can arrive at a symmetric form for the general variational principle of mechanics by starting from the canonical variational problem when one strips time of its special role. Formally, one can initially eliminate the Hamiltonian function $H(p, q, t)$ that still remains in the integral in equation (6) or no. **2** by adding an auxiliary condition and requiring that:

$$\int \left(\sum_k p_k\,\dot{q}_k - W \right) dt = \text{extremum},$$

([1]) The symbols Δ and δ are chosen by analogy with the general and virtual displacements in Chap. 2, no. **23**. The difference between them is merely that now the p_k will also be varied since they also appear as variables in the variational problem.

under the auxiliary condition that:

$$W = H(p, q, t).$$

If we now introduce a new parameter τ in place of t, so $t = t(\tau)$, e.g., the arc-length along the trajectory or the proper time in the theory of relativity, then we will get the form:

$$\int \left(\sum_k p_k \frac{dq_k}{d\tau} - W \frac{dt}{d\tau} \right) d\tau = \text{extremum}, \tag{2}$$

with

$$W = H(p, q, t)$$

as the auxiliary condition. That form is closely related to introducing t itself as a new canonical variable $\mathfrak{q}$ that is conjugate to the impulse $\mathfrak{p} = -W$, and in that way, we will get the completely-symmetric form:

$$\int \left(\sum p_k q_k' + \mathfrak{p}\mathfrak{q}' \right) d\tau = \text{extremum}, \tag{3}$$

while one will also have:

$$F(p_k, q_k, \mathfrak{p}, \mathfrak{q}) = H + \mathfrak{p} = H = W = 0 . \tag{4}$$

The prime in that characterizes the derivative with respect to τ. The mechanical system is no longer characterized by a function then, namely, the Hamiltonian function, but by an equation, namely:

$$F(p_k, q_k, \mathfrak{p}, \mathfrak{q}) = H = W = 0 \tag{4}$$

between the $2l + 2$ canonical variables and impulses. That form of the variational problem can also be adapted to, e.g., the theory of relativity. In general, an arbitrary function $F(p, q, W, t) = 0$ can enter in place of $F \equiv H - W$, but it can always be forced to take the canonical form (4) by solving it for W.

With the multiplier prescription of no. **2**, the general equations of motion will become:

$$\left.\begin{aligned} \frac{dq_k}{d\tau} &= +\lambda \frac{\partial F}{\partial p_k}, & \frac{dt}{d\tau} &= +\lambda \frac{\partial F}{\partial \mathfrak{p}} = -\lambda \frac{\partial F}{\partial W}, \\ \frac{dp_k}{d\tau} &= -\lambda \frac{\partial F}{\partial q_k}, & \frac{d\mathfrak{p}}{d\tau} &= -\frac{dW}{d\tau} = -\lambda \frac{\partial F}{\partial t}, \end{aligned}\right\} \tag{5}$$

and for the canonical form $F = H - W$, since:

$$\frac{dt}{d\tau} = -\lambda \frac{\partial F}{\partial W} = -\lambda \frac{\partial (H - W)}{\partial W} = \lambda ,$$

they will reduce to the ordinary canonical equations:

$$\left.\begin{aligned} \frac{dq_k}{d\tau}\frac{d\tau}{dt} &= \frac{\partial(H-W)}{\partial p_k} = \frac{\partial H}{\partial p_k}, \quad \frac{dW}{d\tau}\frac{d\tau}{dt} = \frac{\partial(H-W)}{\partial t} = \frac{\partial H}{\partial t}, \\ \frac{dp_k}{d\tau}\frac{d\tau}{dt} &= -\frac{\partial(H-W)}{\partial q_k} = -\frac{\partial H}{\partial q_k}, \qquad \frac{dt}{d\tau} = \lambda. \end{aligned}\right\} \tag{6}$$

One can also generalize the canonical transformations in such a way that they subsume time. In order to do that, the necessary and sufficient condition is obviously that the differential form:

$$\sum_k p_k \, \Delta p_k + \mathfrak{p}\,\Delta t \, ,$$

in which the variables p_k, q_k, $\mathfrak{p}$, t are coupled by the auxiliary condition:

$$H + \mathfrak{p} = 0 \, , \tag{7}$$

should go to the differential form:

$$\sum P_k \, \Delta Q_k + \mathfrak{P}\,\Delta T + \Delta\Phi ,$$

whose variables are coupled by the corresponding auxiliary condition:

$$K + \mathfrak{P} = 0 \, .$$

That will imply any arbitrary canonical transformation of the $2f + 2$ variables q_k, p_k, t, $\mathfrak{p}$ into Q_k, P_k, T, $\mathfrak{P}$, which is therefore generated by an arbitrary function $V^*(q_k, Q_k, t, T)$. In so doing, the function K must be determined in such a way that one performs the transformation in equation (7) and solves the relation thus-obtained for $\mathfrak{P}$, and thus finds:

$$\mathfrak{P} = -\,K\,(Q_k, P_k, T) \, .$$

Should t in particular not be transformed, i.e., t goes to T, then V^* will have the form:

$$V^* = \mathfrak{P}\,t + V\,(q_k, Q_k, t) \, ,$$

since from equation (6) in no. **3**, one will have:

$$T = \frac{\partial V^*}{\partial \mathfrak{P}} = t \, , \quad \mathfrak{p} = -\,W = \frac{\partial V^*}{\partial t} = \mathfrak{P} + \frac{\partial V}{\partial t} ,$$

i.e.:

$$-\mathfrak{P} = K(Q_k, P_k, t) = W + \frac{\partial V}{\partial t} = H + \frac{\partial V}{\partial t}.$$

Naturally, one comes back to the formula in no. **3**.

5. Integral invariants. – Just like with every transformation, the question of invariants also has great importance for canonical transformations, i.e., the question of what functions will not change their values under the transformation. One can give a whole series of such invariants for all canonical transformations. We shall first discuss the *integral invariants* that **Poincaré** ([1]) was the first to consider.

When the integral:

$$J_1 = \iint \sum_k dp_k \, dq_k \tag{1}$$

is extended over an arbitrary two-dimensional region of the $2f$-dimensional phase space of p_k and q_k, it will be is an invariant of a canonical transformation. In order to prove that, we represent that two-dimensional region in such a way that we shall give p_k and q_k as functions of two parameters u and v. In it, we will have:

$$J_1 = \iint \sum_k \begin{vmatrix} \dfrac{\partial p_k}{\partial u} & \dfrac{\partial q_k}{\partial u} \\ \dfrac{\partial p_k}{\partial v} & \dfrac{\partial q_k}{\partial v} \end{vmatrix} du\, dv\,. \tag{2}$$

We assume that the canonical transformation has the form:

$$\left.\begin{aligned} p_k &= \frac{\partial V(q_k, P_k, t)}{\partial q_k}, \\ Q_k &= \frac{\partial V(q_k, P_k, t)}{\partial P_k}, \end{aligned}\right\} \tag{3}$$

and introduce the p_k as functions q_k, P_k, in J_1 by means of equations in the first row, whereby the value of t is fixed in (3), so t can be treated as a constant parameter. One will then have:

$$\sum_k \begin{vmatrix} \dfrac{\partial p_k}{\partial u} & \dfrac{\partial q_k}{\partial u} \\ \dfrac{\partial p_k}{\partial v} & \dfrac{\partial q_k}{\partial v} \end{vmatrix} = \sum_k \begin{vmatrix} \displaystyle\sum_l \frac{\partial^2 V}{\partial q_k \, \partial P_l} \frac{\partial P_l}{\partial u} & \dfrac{\partial q_k}{\partial u} \\ \displaystyle\sum_l \frac{\partial^2 V}{\partial q_k \, \partial P_l} \frac{\partial P_l}{\partial v} & \dfrac{\partial q_k}{\partial v} \end{vmatrix} = \sum_{k,l} \frac{\partial^2 V}{\partial q_k \, \partial P_l} \begin{vmatrix} \dfrac{\partial P_l}{\partial u} & \dfrac{\partial q_k}{\partial u} \\ \dfrac{\partial P_l}{\partial v} & \dfrac{\partial q_k}{\partial v} \end{vmatrix}.$$

([1]) **H. Poincaré**, *Les méthodes nouvelles de la mécanique céleste*, t. III, Chap. 22/23, Paris, 1899. The proof is in **E. Brody**, Zeit. Phys. **6** (1921), pp. 224.

By switching the indices, that will give:

$$\sum_{l,k} \frac{\partial^2 V}{\partial q_l \, \partial P_k} \begin{vmatrix} \frac{\partial P_k}{\partial u} & \frac{\partial q_l}{\partial u} \\ \frac{\partial P_k}{\partial v} & \frac{\partial q_l}{\partial v} \end{vmatrix} .$$

If we now take the q_k, P_k to the Q_k, P_k with the help of the second row of equations (3) then that will give:

$$\sum_k \begin{vmatrix} \frac{\partial P_k}{\partial u} & \sum_l \frac{\partial^2 V}{\partial P_k \, \partial q_l} \frac{\partial q_l}{\partial u} \\ \frac{\partial Q_k}{\partial v} & \sum_l \frac{\partial^2 V}{\partial P_k \, \partial q_l} \frac{\partial q_l}{\partial v} \end{vmatrix} = \sum_{l,k} \begin{vmatrix} \frac{\partial P_k}{\partial u} & \frac{\partial Q_l}{\partial u} \\ \frac{\partial P_k}{\partial v} & \frac{\partial Q_l}{\partial v} \end{vmatrix} ,$$

which also proves the invariance of the integral (1).

One can prove the invariance of:

$$J_2 = \iiint\!\!\int \sum_{k,l} dp_k \, dp_l \, dq_k \, dq_l \tag{5}$$

and in general, that of:

$$J_n = \int \overset{2n}{\cdots} \int \sum_{k_1,\ldots,k_n} dp_{k_1} \cdots dp_{k_n} \, dq_{k_1} \cdots dq_{k_n} \tag{6}$$

analogously. The last integral in this sequence is the volume of the phase space of p_k and q_k :

$$J_f = \int \overset{2f}{\cdots} \int dp_1 \cdots dp_f \, dq_1 \cdots dq_f \ , \tag{7}$$

so that is also an invariant under canonical transformations. In that way, it is also shown, at the same time, that the functional determinant of a canonical transformation is equal to 1.

As will be shown later (no. **9**), the time variation of the coordinates and impulses of a mechanical system can also be regarded as a canonical transformation of it. Therefore, all invariants of canonical transformations are also invariants of motion. That is understood to mean that the points of the corresponding $2n$-dimensional region in phase space are to be thought of as the image points of a corresponding manifold of the same mechanical system with a somewhat-different initial configuration. Under the motion of that system, the original domain of values p, q over which one must integrate will be taken to a different one that will have the same value, according to our theorems. Therefore, the worldlines of that system in p, q, t-space will define a tube of constant cross-section. For J_f, that is **Liouville**'s theorem, which is fundamental to statistical mechanics.

The integral invariants (1) and (6) to (7) are called *absolute* since no sort of assumptions about the domain of integration is made in them. They can be converted into *relative* ones with the help of Stokes's theorem, i.e., integral invariants that are extended over closed domains of integration whose order (i.e., whose number of integrations) is lower. For example, the invariance of the integral that is performed over a closed curve in p, q-space (that must lie on a plane t = const. in p, q, t-space):

$$J_1 = \oint \sum_k p_k \, dq_k \tag{8}$$

will enter in place of (1).

Moreover, as will be shown in no. **6**, it will follow conversely from the existence of the integral invariant (8) [(2), resp.] for a system of transformation equations:

$$\left. \begin{aligned} q_l &= q_l\,(Q_k, P_k, t), \\ p_l &= p_l\,(Q_k, P_k, t) \end{aligned} \right\} \tag{9}$$

that they can be brought into the form of equation (6) in no. **6**, so the transformation that one employs will be canonical.

If one chooses the domain of integration in (1) to be the parallelogram that is spanned by two infinitesimal vectors in pq-space whose components are dq_k, dp_k (δq_k, δp_k, resp.) then one will have the invariance of the *bilinear covariant:*

$$\sum_k (\delta p_k \, dq_k - dp_k \, \delta q_k) \tag{10}$$

that belongs to the differential form $\sum_k p_k \, dq_k$. From what was said before, its invariance is also sufficient for the transformation to be canonical in nature. Moreover, from what we remarked in regard to equation (3), the invariance of (10) will be true only when either V is independent of time or the two small vectors, along with their images in P, Q, t-space, lie on the planes t = const., i.e., when they are δ-variations, in the sense of no. **3**. On the other hand, it is not (10) that is invariant, but the covariant:

$$\sum_k (\Delta p_k \, dq_k - dp_k \, \Delta q_k) - (\Delta H \, dt - dH \, \Delta t), \tag{11}$$

which belongs to the differential form.

6. The conditions for canonical transformations, when expressed in terms of the Lagrange and the Poisson-Jacobi bracket symbols. – One refers to the expressions that appeared in (4) of no. **5**:

$$\left.\begin{aligned}[][u,v] &= \sum_k \left(\frac{\partial q_k}{\partial u}\frac{\partial p_k}{\partial v} - \frac{\partial p_k}{\partial u}\frac{\partial q_k}{\partial v}\right) \\ &= -\sum_k \begin{vmatrix} \dfrac{\partial p_k}{\partial u} & \dfrac{\partial q_k}{\partial u} \\ \dfrac{\partial p_k}{\partial v} & \dfrac{\partial q_k}{\partial v} \end{vmatrix}\end{aligned}\right\} \tag{1}$$

as *Lagrange brackets*. As we showed at the time, they are invariant under canonical transformations. In no. **5**, u and v were understood to mean the values of the coordinates that belong to the parameters of a two-dimensional section of pq-space. Naturally, the coordinate values themselves can also serve as such things. That will lead to the equations:

$$\left.\begin{aligned} &[p_i, p_k] = [q_i, q_k] = 0, \\ &[q_i, p_k] = \delta_{ik} = \begin{cases} 0 & \text{for} \quad i \neq k \\ 1 & \text{for} \quad i = k. \end{cases}\end{aligned}\right\} \tag{2}$$

Invariance means that the equations:

$$\left.\begin{aligned} &[P_i, P_k] = [Q_i, Q_k] = 0, \\ &[Q_i, P_k] = \delta_{ik}\end{aligned}\right\} \tag{3}$$

will also be valid as long are the transformation $(p, q) \rightarrow (P, Q)$ is canonical. Conversely, equations (3) will, in turn, suffice to ensure the canonical character of the transformation, as we will soon show. They are then the characteristic differential equations that the p, q must satisfy as functions of P, Q in order for the transformation to be canonical. The proof is obtained as follows:

When equations (3) are written out in detail, they will read:

$$[Q_k, P_j] = \sum_l \left(\frac{\partial q_l}{\partial Q_k}\frac{\partial p_l}{\partial P_j} - \frac{\partial p_l}{\partial Q_k}\frac{\partial q_l}{\partial P_j}\right) = \delta_{jk} ,$$

$$[Q_k, Q_j] = \sum_l \left(\frac{\partial q_l}{\partial Q_k}\frac{\partial p_l}{\partial Q_j} - \frac{\partial p_l}{\partial Q_k}\frac{\partial q_l}{\partial Q_j}\right) = 0 ,$$

$$[P_k, P_j] = \sum_l \left(\frac{\partial q_l}{\partial P_k}\frac{\partial p_l}{\partial P_j} - \frac{\partial p_l}{\partial P_k}\frac{\partial q_l}{\partial P_j}\right) = 0 .$$

They can be rewritten as follows:

$$\frac{\partial}{\partial P_j}\left(\sum_l p_l \frac{\partial q_l}{\partial Q_k} - P_k\right) - \frac{\partial}{\partial Q_k}\left(\sum_l p_l \frac{\partial q_l}{\partial P_j}\right) \qquad = 0 ,$$

$$\frac{\partial}{\partial Q_j}\left(\sum_l p_l \frac{\partial q_l}{\partial Q_k} - P_k\right) - \frac{\partial}{\partial Q_k}\left(\sum_l p_l \frac{\partial q_l}{\partial Q_j} - P_j\right) = 0 ,$$
$$\frac{\partial}{\partial P_j}\left(\sum_l p_l \frac{\partial q_l}{\partial P_k}\right) - \frac{\partial}{\partial P_k}\left(\sum_l p_l \frac{\partial q_l}{\partial P_j}\right) = 0 .$$

However, those equations mean that a function Φ (Q_k, P_k, t) exists for which:

$$\sum_l p_l \frac{\partial q_l}{\partial Q_k} - P_k = \frac{\partial \Phi}{\partial Q_k}$$

and

$$\sum_l p_l \frac{\partial q_l}{\partial P_k} = \frac{\partial \Phi}{\partial P_k} .$$

If one now defines the δ-variation of Φ:

$$\delta\Phi = \sum_k \frac{\partial \Phi}{\partial Q_k}\delta Q_k + \sum_k \frac{\partial \Phi}{\partial P_k}\delta P_k$$
$$= \sum_{k,l} p_l \frac{\partial q_l}{\partial Q_k}\delta Q_k + \sum_{k,l} p_l \frac{\partial \Phi}{\partial P_k}\delta P_k - \sum_k P_k\,\delta Q_k ,$$

and considers the fact that:

$$\delta q_l = \sum_k \frac{\partial q_l}{\partial Q_k}\delta Q_k + \sum_k \frac{\partial q_l}{\partial P_k}\delta P_k$$

then one will get:

$$\delta\Phi = \sum_l p_l\,\delta q_l - \sum_k P_k\,\delta Q_k .$$

Therefore, the transformation formulas:

$$q_k = q_k(Q_l, P_l, t) , \qquad p_k = p_k(Q_l, P_l, t) \tag{4}$$

will obey the relation (12) in no. **3**:

$$\sum_k p_k\,\delta q_k = \sum_k P_k\,\delta Q_k + \delta\Phi(P, Q, t) .$$

In other words, the transformation (4) is canonical.

With that, the statement that was made before in the context of [(8), no. **5**] that the existence of the invariant [(8), no. **5**] or [(2), no. **2**] is sufficient to ensure the canonical character of the transformation (4) can now be justified since that invariant has equations (3) as a consequence.

The Lagrange brackets are closely related to what are called ***Poisson*** or ***Jacobi*** *symbols:*

$$(u, v) = \sum_k \left(\frac{\partial u}{\partial q_k} \frac{\partial v}{\partial p_k} - \frac{\partial u}{\partial p_k} \frac{\partial v}{\partial q_k} \right) . \tag{5}$$

The connection between the two of them consists of the fact that for any $2f$ independent functions $u_1, \ldots, u_{2f}$, the following equations will be true for p_k, q_k :

$$\sum_{l=1}^{2f} (u_l, u_r)[u_l, u_r] = \delta_{rs} . \tag{6}$$

One confirms immediately by direct calculation that when one considers the fact that the sums:

$$\sum_{t=1}^{2f} \frac{\partial u_t}{\partial x} \frac{\partial y}{\partial u_t}$$

will be non-zero and equal to unity only when x and y mean *the same* quantity from the p_k, q_k.

Equations (3) and (6) will imply the further necessary and sufficient condition for characterizing a canonical transformation in the form of the system:

$$(P_i, P_k) = (Q_i, Q_k) = 0 , \qquad (Q_i, P_k) = \delta_{ik} \tag{7}$$

when one takes the u_i to be the P_k and Q_k themselves. They represent the differential equations that the new variables P, Q must fulfill as functions of the original p, q (so the inversion formula for the transformation) in order for them to be canonical. Equations (7) mean the same thing as the invariance of the special bracket symbols in question. However, the invariance of the Poisson bracket (u, v) for any two functions u and v of the q_k, p_k can also be proved from the invariance of $[u, v]$ with the help of (6)

7. Further properties of the bracket symbols. The theorems of Poisson and Lagrange. – In recent times, the **Poisson** brackets have taken on a special significance as a result of their adaptation to quantum mechanics ([1]). Some further rules of calculation and theorems that relate to them might find a place here then.

From the definition [(5), no. **6**], one initially has the identities:

([1]) Cf., especially the work of **P. A. M. Dirac** in the Proc. Roy. Soc. London (A) **109** (1925), pp. 642; *ibid.* **110** (1926), pp. 561; *ibid.* **111** (1926), pps. 281, 405.

$$\left.\begin{array}{ll} (u,u)=0, & (u,v)=-(v,u), \\ \dfrac{\partial u}{\partial q_j}=(u,p_j)=-(p_j,u), & \dfrac{\partial u}{\partial p_j}=(q_j,u)=-(u,q_j). \end{array}\right\} \tag{1}$$

Moreover, one has:

$$(u,(v,w))+(v,(w,u))+(w,(u,v))=0 \tag{2}$$

identically. Namely, the left-hand side is obviously linear and homogeneous in the second derivatives of u, v, w. We shall now combine only the terms that include the second derivatives of u. The first term in (2) certainly includes only first derivatives. From (1), the second and third ones can be written in the form:

$$(v,(w,u))+(w,(u,v))=(v,(w,u))-(w,(v,u)) .$$

If we introduce the differential operators:

$$D_1(f)=(v,f) , \qquad D_2(f)=(w,f)$$

then the terms that might contain the second derivatives can be combined into the form:

$$(D_1 D_2-D_2 D_1)\, u .$$

However, such a combination of two linear differential operators will never include second derivatives. Namely, if one has, say:

$$D_1=\sum_k \xi_k \frac{\partial}{\partial x_k}, \qquad D_2=\sum_k \eta_k \frac{\partial}{\partial x_k},$$

then one will have:

$$D_1 D_2=\sum_{k,l} \xi_k \eta_l \frac{\partial^2}{\partial x_k\, \partial x_l}+\sum_{k,l} \xi_k \frac{\partial \eta_l}{\partial x_k}\frac{\partial}{\partial x_l},$$

$$D_2 D_1=\sum_{k,l} \eta_k \xi_l \frac{\partial^2}{\partial x_k\, \partial x_l}+\sum_{k,l} \eta_k \frac{\partial \xi_l}{\partial x_k}\frac{\partial}{\partial x_l}.$$

Therefore:

$$D_1 D_2-D_2 D_1=\sum_l\left[\sum_k\left(\xi_k \frac{\partial \eta_l}{\partial x_k}-\eta_k \frac{\partial \xi_l}{\partial x_k}\right)\right]\frac{\partial}{\partial x_l}$$

is also an operator that contains only first derivatives. As a result, no terms at all with second derivatives of u can enter into (2), and since the same thing must be true for v and w, the entire expression must vanish identically. Eq. (2) is the so-called ***Jacobi*** *identity*.

As a result of (1), it is possible to give the canonical equations of motion [cf., (9), no. **2**]:

$$\dot{p}_k = -\frac{\partial H}{\partial q_k}, \quad \dot{q}_k = \frac{\partial H}{\partial p_k}, \tag{3}$$

the form:

$$\dot{p}_k = (p_k, H), \quad \dot{q}_k = (q_k, H), \tag{4}$$

which is a reasonable adaptation of the one that is employed in quantum mechanics.

If one considers (3) then one will further see that for every integral $F(q, p) = a$ of the motion that does not include time t explicitly, one will have:

$$(F, H) = 0 . \tag{5}$$

That theorem means only that the gradient of the hypersurface $F(q, p) = a$ in $2f$-dimensional pq-space will be perpendicular to the phase trajectory element:

$$dq_k = \dot{q}_k \, dt = \frac{\partial H}{\partial p_k} dt ,$$
$$dp_k = \dot{p}_k \, dt = -\frac{\partial H}{\partial q_k} dt ,$$

so the element lies completely in the surface.

Finally, we shall derive a noteworthy and important theorem of **Poisson** whose full significance was first known to **Jacobi**, in general. It allows one to find new integrals of the mechanical equations in some cases. It says: If F = const. and G = const. are two time-independent integrals of the canonical equations (3) then their Poisson bracket will be:

$$(F, G) = \sum_k \left(\frac{\partial F}{\partial q_k}\frac{\partial G}{\partial p_k} - \frac{\partial F}{\partial p_k}\frac{\partial G}{\partial q_k} \right) = \text{const.} \tag{6}$$

Equation (6) will then be an integral again.

The proof follows immediately from (2) when one considers the ffact that from (5), one will have:

$$(H, F) = 0 \qquad \text{and} \qquad (H, G) = 0 .$$

Namely, that will give:

$$(H, (F, G)) = 0 , \tag{7}$$

i.e., (F, G) = const. is an integral of the canonical equations.

Naturally, one will not always get new integrals by that process since there is only a restricted number of them at all, but rather one will often get only a trivial one or one that is a function of the first two F, G.

There is also an analogue of theorem (6) for the Lagrange brackets. If we employ the theorem that was mentioned before and will be established later that the change in coordinates of a mechanical system in the course of its motion can be regarded as the evolution of a canonical transformation then one get **Lagrange**'s theorem from the invariance of the brackets. It says that for any two-dimensional family of solutions of the canonical equations:

$$q_j = q_j\,(a, b, t)\,, \qquad p_j = p_j\,(a, b, t)\,,$$

in which a and b are arbitrary integral constants, the corresponding Lagrange bracket will satisfy:

$$[a, b] = \text{const.}$$

for all time, i.e., along the whole mechanical trajectory.

All of the theorems above can be easily generalized to systems (integrals, resp.) that include time explicitly when one regards time as a canonical variable, as in no. **4**. As a definition of the Poisson brackets (which will now be written with curly brackets, in order to distinguish them), one will then have:

$$\left.\begin{aligned} \{u, v\} &= (u, v) - \frac{\partial u}{\partial t}\frac{\partial v}{\partial W} + \frac{\partial u}{\partial W}\frac{\partial v}{\partial t} \\ &= (u, v) + \frac{\partial u}{\partial t}\frac{\partial v}{\partial \mathfrak{p}} - \frac{\partial u}{\partial \mathfrak{p}}\frac{\partial v}{\partial t}. \end{aligned}\right\} \tag{9}$$

One can also extend the Lagrange brackets correspondingly. The considerations of this section and no. **6** can then be adapted, word-for-word, except that one must replace H with $H - W$ $(H + \mathfrak{p})$ everywhere.

Thus, the form (4) of the canonical equations now corresponds to:

$$\left.\begin{aligned} \dot{p}_k &= \{p_k, (H - W)\} &&= -\frac{\partial (H - W)}{\partial q_k} &&= -\frac{\partial H}{\partial q_k}, \\ \dot{q}_k &= \{q_k, (H - W)\} &&= \frac{\partial (H - W)}{\partial p_k} &&= \frac{\partial H}{\partial p_k}, \\ \dot{t} &= \{t, (H - W)\} &&= 1 \\ \dot{W} &= \{W, (H - W)\} &&= \frac{\partial (H - W)}{\partial t} &&= \frac{\partial H}{\partial t}. \end{aligned}\right\} \tag{10}$$

It follows from them that for arbitrary functions $F\,(p_k, q_k, W, t)$:

$$\dot{F} = \sum_k \left(\frac{\partial F}{\partial q_k}\dot{q}_k + \frac{\partial F}{\partial p_k}\dot{p}_k\right) + \frac{\partial F}{\partial t} + \frac{\partial F}{\partial W}\dot{W} = \{F, (H - W)\}\,. \tag{11}$$

Every integral of the equations of motion will then fulfill the condition that is analogous to (5):

$$\{F, (H - W)\} = 0 , \tag{12}$$

which will reduce to:

$$\{F, H\} + \frac{\partial F}{\partial t} = 0 \tag{13}$$

for integrals that are independent of W. Now, Poisson's theorem says that when F = const. and G = const., one will also have that:

$$\{F, H\} = \text{const.} \tag{14}$$

is an integral of the canonical equations (10). The simple form (6) will follow from equation (14) only when F and G are both independent of W. Therefore, the restriction to time-independent integrals is not essential for (6) to be true.

8. Continuous transformation groups. – The question of what sort of meaning that the integrals of the canonical equations would have for the variational problem can be treated in a very elegant way with the help of the theory of transformation groups. In order to do that, we must first preface a few theorems about them.

We subject the mechanical system to a transformation of the form ([1]):

$$\left.\begin{aligned} P_k &= P_k(p_l, q_l, \alpha) = p_k + \sum_{n=1}^{\infty} \alpha^n p_k^{(n)}(p_l, q_l), \\ Q_k &= Q_k(p_l, q_l, \alpha) = q_k + \sum_{n=1}^{\infty} \alpha^n q_k^{(n)}(p_l, q_l). \end{aligned}\right\} \tag{1}$$

Thus, this transformation includes another parameter that will allow one to develop it in a power series, and it will go to the identity transformation when $\alpha = 0$. If α is very small then we will have a transformation in the neighborhood of the identity. One then calls it an *infinitesimal transformation*. For every value of α, we will have a certain transformation. A whole family of transformations is then determined by (1).

We would now like to demand that those transformations should define a group, i.e., that when two of the transformations with any values α_1, α_2 are performed in succession, that will again yield a transformation of the family. **Lie** ([2]) has shown that on the basis of that requirement, the linear terms in the development (1), which we would like to denote by $\mathfrak{p}_k$, $\mathfrak{q}_k$, will also determine all of the following terms, and therefore they will be characteristic of the transformation by themselves. Only a group would belong to a theorem about such terms. Proving that would take us too far afield. We restrict ourselves to merely specifying the transformations, so to showing how we will get the higher-order terms from the first-order ones.

([1]) It is entirely irrelevant in this whether one regards the p_k, q_k or the P_k, Q_k are the original variables. For the sake of the more-convenient application in no. **9**, we shall write it in the form above that corresponds to the solution of a transformation $p_k = p_k(P, Q)$, $q_k = q_k(P, Q)$.

([2]) **S. Lie**, *Theorie der Transformationsgruppen*, Bd. I, Leipzig, 1888, pp. 51, *et seq.*

One constructs the following differential operator with the help of the $\mathfrak{p}_k$, $\mathfrak{q}_k$:

$$D = \sum_k \mathfrak{p}_k \frac{\partial}{\partial p_k} + \sum_k \mathfrak{q}_k \frac{\partial}{\partial q_k}, \tag{2}$$

which one refers to as the *generating symbol of the group*. Therefore, D is also given by $\mathfrak{p}_k$, $\mathfrak{q}_k$. Now, there are three different ways of defining the transformations that form the group, which will naturally lead to identical results.

a) One defines the series:

$$\left.\begin{aligned} P_k &= [p_k] \equiv p_k + \alpha\, D\, p_k + \tfrac{1}{2}\alpha^2 D^2 p_k + \cdots = \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} D^n p_k, \\ Q_k &= [q_k] \equiv q_k + \alpha\, D\, q_k + \tfrac{1}{2}\alpha^2 D^2 q_k + \cdots = \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} D^n q_k. \end{aligned}\right\} \tag{3}$$

The D^n in them are operators that arise by an n-fold application of D. We introduce the symbol:

$$[F] = \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} D^n F, \tag{4}$$

here, to abbreviate. The series (3) can therefore be determined by only differentiations and multiplications with the help of the $\mathfrak{p}_k$, $\mathfrak{q}_k$ then, and will also become convergent for sufficiently-small α, as is easy to show. Moreover, we obviously have that for an arbitrary function $F(p_k, q_k)$:

$$F(P_k, Q_k) = F([p_k], [q_k]) = [F(p_k, q_k)]. \tag{5}$$

We also see from the representation (3) that the general transformation (1) can be constructed by continually repeating the linear (infinitesimal) transformation:

$$P_k = p_k + \alpha\, \mathfrak{p}_k, \qquad Q_k = q_k + \alpha\, \mathfrak{q}_k.$$

b) One defines the partial differential equation for the function F of the $2f + 1$ variables p_k, q_k, α:

$$\frac{\partial F}{\partial \alpha} = D\, F = \sum_k \mathfrak{p}_k \frac{\partial F}{\partial p_k} + \sum_k \mathfrak{q}_k \frac{\partial F}{\partial q_k}, \tag{6}$$

and seeks those integrals $F(p_k, q_k, \alpha)$ that go to the variables p_k, q_k themselves for $\alpha = 0$. The $2f$ integrals $P_k(\alpha, p_l, q_l)$, $Q_k(\alpha, p_l, q_l)$, thus-determined will again be precisely the desired transformation functions. One sees that this definition coincides with the first one from the definition (4), from which, it follows that:

$$D\,[F] = \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} D^{n+1} F\,,$$
$$\frac{\partial}{\partial \alpha}[F] = \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} D^{n+1} F$$

for any function $[F]$. Any function $[F]$ will then satisfy the differential equation (6) in its own right. Therefore, the functions $P_k(p_l, q_l)$, $Q_k(p_l, q_l)$ that are defined in both ways must also coincide for $\alpha = 0$, which will establish them uniquely, together with the differential equation (6).

c) The functions that represent the transformation are also the solutions of the system of $2f$ ordinary differential equations:

$$\left.\begin{aligned} \frac{dP_k}{d\alpha} &= \mathfrak{p}_k(P_l, Q_l)\,, \\ \frac{dQ_k}{d\alpha} &= \mathfrak{q}_k(P_l, Q_l)\,, \end{aligned}\right\} \tag{7}$$

which assume the values p_k, q_k for $\alpha = 0$. In that way, the new variables are thought of as being introduced into the right-hand sides by means of (3), while the old variables appear as integration constants of the system (7). One also sees that this definition agrees with the first one, and therefore the second one, as well, with the help of the series developments in (3) and the definitions (2), (4), and (5) because one will have, in succession, e.g.:

$$\frac{dP_k}{d\alpha} = \frac{d[p_k]}{d\alpha} = [D\,p_k] = [\mathfrak{p}_k] = \mathfrak{p}_k([p_l], [q_l]) = \mathfrak{p}_k(p_l, q_l)\,.$$

A connection between the different transformation of the group is likewise much simpler to show with the help of the representation (2). Namely, if $f_1, f_2, \ldots, f_f$ are solutions of a linear, homogeneous, partial differential equation like (6) then it is known that an arbitrary function $F(f_1, \ldots, f_f)$ will also be so. Now, since, e.g., if $[p_k]_{\alpha=\alpha_1}$ is a solution of (6) then the same thing will be true for $[[p_k]_{\alpha_1}]_{\alpha=\alpha_2}$, and since $[p_k]_0$ is the identity transformation, one will have $[[p_k]_0]_{\alpha_2} = [p_k]_{\alpha_1}$. However, the solution $[p_k]_{\alpha_1+\alpha_2}$ also has the same property that it will be equal to $[p_k]_{\alpha_2}$ for $\alpha_1 = 0$, but since there is only one solution of the partial differential equation that is equal to for $[p_k]_{\alpha_2}$ for $\alpha_1 = 0$, one must have:

$$[[p_k]_{\alpha_1}]_{\alpha_2} = [p_k]_{\alpha_1+\alpha_2}\,,$$

i.e., when the transformations with the parameters α_1 and α_2 are performed in succession, that will give the transformation with the parameter $\alpha_1 + \alpha_2$. With that, it is proved that our transformations actually define a group.

If one now considers a function $f(p_k, q_k)$ and applies the transformation (3) to it then it will go to:

$$f(P_k, Q_k) = [f(p_k, q_k)] = \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} D^n f(p_k, q_k) .$$

If f goes to itself in that way then one would call such a function an *invariant* of the group. In order for that to be true, it is obviously necessary and sufficient that one must have:

$$D f(p_k, q_k) = 0$$

identically in the p_k, q_k since all higher powers in the power series will vanish then, and only the zeroth-order term, i.e., the identity operator, will remain. The invariants of the group will then satisfy the partial differential equation:

$$D f = \sum_k \mathfrak{p}_k \frac{\partial f}{\partial p_k} + \sum_k \mathfrak{q}_k \frac{\partial f}{\partial q_k} = 0 . \tag{8}$$

9. The meaning of the integrals for canonical equations. – Following that preliminary excursion, we shall now return to mechanics and ask *when such a transformation group is canonical*, so it includes only canonical transformations. For the sake of simplicity, we restrict ourselves to the case in which the independent variable t does not appear in the Hamiltonian function. Otherwise, as in no. **4**, t must be likewise treated as a canonical variable and transform it, as well.

The condition for canonical transformations was equation [(12), no. **3**]:

$$\sum_k p_k \, \delta q_k = \sum P_k \, \delta Q_k + \delta \Phi , \tag{1}$$

in which the operation δ was defined by:

$$\delta f(p_k \, q_k) = \sum_k \frac{\partial f}{\partial p_k} \delta p_k + \sum_k \frac{\partial f}{\partial q_k} \delta q_k .$$

If we introduce the developments [(3), no. **8**] into this then when we recall equation [(2), no. **8**], that will give:

$$\sum_k p_k\,\delta q_k = \sum_k (p_k + \alpha\,\mathfrak{p}_k + \frac{\alpha^n}{2!} D\,\mathfrak{p}_k + \cdots)(\delta q_k + \alpha\,\delta\mathfrak{q}_k + \frac{\alpha^n}{2!} D\,\delta\,\mathfrak{q}_k + \cdots) + \sum_{n=0}^{\infty} \alpha^n\,\delta\Phi_n\,, \qquad (2)$$

in which Φ is also expanded into a power series in α:

$$\Phi = \sum_{n=0}^{\infty} \alpha^n\,\Phi_n\,.$$

In order for the relation (2) to be fulfilled identically, all powers of α must have the same coefficients on both sides of it. One must then have $\Phi_0 = 0$ to begin with. The linear terms yield:

$$\sum_k \mathfrak{p}_k\,\delta q_k + \sum_k p_k\,\delta\mathfrak{q}_k = \delta\Phi_1 \qquad (3)$$

identically in p_k, q_k. If one has chosen $\mathfrak{p}_k$, $\mathfrak{q}_k$ in such a way that this relation is fulfilled then a corresponding repeated application of the operator D to the first relation will give the higher powers, and one will easily see that equation (2) will be fulfilled in general when one sets:

$$\Phi = \alpha\,\Phi_1 + \frac{\alpha^2}{2!} D\,\Phi_1 + \frac{\alpha^3}{3!} D^2\,\Phi_2 + \cdots$$

If we now introduce the function:

$$-\Psi\,(p_k, q_k) = \Phi_1 - \sum_k p_k\,\mathfrak{q}_k\,,$$

so

$$-\delta\Psi = \delta\Phi_1 - \sum_k p_k\,\delta\mathfrak{q}_k - \sum_k \mathfrak{q}_k\,\delta p_k\,,$$

in place of Φ_1 then (3) will go to the condition:

$$\sum_k \mathfrak{p}_k\,\delta q_k - \sum_k \mathfrak{q}_k\,\delta p_k = -\,\delta\Psi. \qquad (4)$$

It is fulfilled identically in the p_k, q_k if and only if:

$$\mathfrak{p}_k = -\,\frac{\partial\Psi}{\partial q_k}\,, \qquad \mathfrak{q}_k = +\,\frac{\partial\Psi}{\partial p_k}\,.$$

$\Psi\,(p_k, q_k)$ can be chosen quite arbitrarily, and one will then get the most-general group of canonical transformations by means of the operator:

$$D = \sum_k \frac{\partial\Psi}{\partial p_k}\frac{\partial}{\partial q_k} - \sum_k \frac{\partial\Psi}{\partial q_k}\frac{\partial}{\partial p_k}, \tag{5}$$

so by equations (2) and (3) of no. **8**, the transformation formulas themselves will be given by:

$$\left.\begin{aligned} P_k &= p_k - \alpha\frac{\partial\Psi}{\partial q_k} + \frac{\alpha^2}{2!}D\frac{\partial\Psi}{\partial q_k} - \cdots, \\ Q_k &= q_k + \alpha\frac{\partial\Psi}{\partial p_k} + \frac{\alpha^2}{2!}D\frac{\partial\Psi}{\partial p_k} - \cdots \end{aligned}\right\} \tag{6}$$

From the results of no. **9**, those transformation functions are simultaneously the solutions of the partial differential equation:

$$\frac{\partial F}{\partial\alpha} = D\,F \tag{7}$$

that go to p_k, q_k, respectively, for $\alpha = 0$. Moreover, they will be the solutions of the system of differential equations:

$$\frac{dP_k}{d\alpha} = -\frac{\partial\Psi}{\partial Q_k}, \qquad \frac{dQ_k}{d\alpha} = \frac{\partial\Psi}{\partial P_k} \tag{8}$$

that assume the values p_k, q_k for $\alpha = 0$. In agreement with no. **3**, the canonical group depends upon a single arbitrary function, namely Ψ, which will be referred to as the *generating function of the group*.

Naturally, the Hamiltonian function for a mechanical problem will generally go to another function by means of the canonical group. We now ask (and this is the essential gist of the following investigation) whether there are also groups that take the problem to itself, i.e., ones for which H is invariant under them. From equation [(8), no. **8**], in order to do that, it is necessary for H to satisfy the partial differential equation:

$$D\,H \equiv \sum_k \left(\frac{\partial\Psi}{\partial q_k}\frac{\partial H}{\partial p_k} - \frac{\partial\Psi}{\partial p_k}\frac{\partial H}{\partial q_k}\right) \equiv (\Psi, H) = 0, \tag{9}$$

in which (Ψ, H) means the **Poisson** bracket (cf., no. **6**).

Therefore, if we would like to determine the transformation groups for a given Hamiltonian function H under which it will be invariant then we must only look for the corresponding functions Ψ that satisfy the partial differential equation (9). They will then be the generating functions of the group. *There will then be just as many canonical transformations of the problem to itself as there are integrals of that differential equation*.

From [(5), no. **7**], equation (9) means that Ψ is an integral of the equations of motion. We have thus arrived at the fundamental theorem that *the generating functions of those canonical* *transformation groups that leave H invariant will be integrals of the canonical equations*. Conversely, every such integral will obviously generate a group of canonical transformations that take the problem to itself. *Knowing the transformation groups of the system is then equivalent to knowing the integrals.*

As one sees from (8), the formulas that determine a transformation group have precisely the form of canonical equations. *Conversely, they can also be interpreted as a canonical transformation then for which t plays the role of the parameter α and H itself defines the generating function.* That transformation associates every system of values $p_k^{(0)}$, $q_k^{(0)}$ at a certain time t_0 with the system of values $p_k^{(t)}$, $q_k^{(t)}$ in which the mechanical system will be found after evolving with the motion from the initial state $p_k^{(0)}$, $q_k^{(0)}$, t_0 over a length of time $t - t_0$. One can regard the course of the motion of the mechanical system as development of a canonical transformation, We have already employed that theorem in nos. **5** and **7**.

The simplest special case is that of cyclic coordinates (cf., Chap. 2, no. **11**). If, say, q_1 is cyclic, so it does not appear into the Hamiltonian function, then:

$$q_1 = Q_1 + \alpha , \qquad q_l = Q_l , \qquad p_k = P_k \qquad (l = 2, \ldots, f), (k = 1, \ldots, f)$$

will be a transformation of the system into itself, while:

$$p_1 = \text{const.}$$

will be the corresponding integral of the canonical equations.

With the help of the general theory of transformation groups, one will also see the meaning of the ten general integrals of the system of free mass-points ([1]) with no further analysis since for that system, they are just the displacements, Galilean transformations, and rotations of the system into itself that do not change the energy. They correspond precisely to the laws of center of mass, impulse, and area. The law of energy itself corresponds to the transformation $T = t$ + const., which also takes the system to itself, but also includes time.

For example, let x_n, y_n, z_n be the x, y, z-coordinates of the n^{th} mass-point, so the first group of transformations will read:

$$\begin{aligned} x_n &= X_n + \alpha_n , & p_{x_n} &= P_{x_n} , \\ y_n &= Y_n + \alpha_n , & p_{y_n} &= P_{y_n} , \\ z_n &= Z_n , & p_{z_n} &= P_{z_n} . \end{aligned}$$

It means a simple displacement of the system in the x-direction. From (5) and (6), the corresponding symbol of the group:

([1]) See Chap. 7, no. **24** of this volume of the *Handbuch*. One might also cf., **F. Engel**, "Über die zehn allgemeinen Integrale der klassischen Mechanik," Göttinger Nachr. (1916) and (1917).

$$\Psi = \sum_n p_{x_n}\,, \quad D = \sum_n \frac{\partial}{\partial x_n}\,.$$

The corresponding integral then reads:

$$\sum_n p_{x_n} = \text{const.}$$

However, that is the first center of mass integral. One likewise finds the other two:

$$\sum_n p_{y_n} = \text{const.}, \qquad \sum_n p_{z_n} = \text{const.}$$

The second group of center of mass integrals:

$$\sum_n p_n\, x_n = t \sum_n p_{x_n} + \text{const.}$$

includes time explicitly. For that reason, in order to treat them, the previous considerations must be extended to transformations that include time.

The law of areas belongs to the group of rotations:

$$\begin{aligned} X_n &= \quad x_n \cos\alpha + y_n \sin\alpha\,, \\ Y_n &= -x_n \sin\alpha + y_n \cos\alpha\,, \\ P_{x_n} &= \quad p_{x_n} \cos\alpha + p_{y_n} \sin\alpha\,, \\ P_{y_n} &= -p_{x_n} \sin\alpha + p_{y_n} \cos\alpha\,. \end{aligned}$$

As one easily verifies by developing in α, the corresponding symbol is:

$$D = \sum_n \left(y_n \frac{\partial}{\partial x_n} - x_n \frac{\partial}{\partial y_n} + p_{x_n} \frac{\partial}{\partial p_{y_n}} - p_{y_n} \frac{\partial}{\partial p_{x_n}} \right).$$

It belongs to the integral:

$$\Psi = \sum_n (p_{y_n} x_n - p_{x_n} y_n) = \text{const.},$$

and that is the area integral for the z-axis. Corresponding results are true for the x and y axes.

10. Reducing the order with the help of known integrals. – The canonical transformations also put us into a position to take advantage of any possible prior knowledge of integrals of the canonical equations in order to bring down the order of the system of differential equations. For example, in very many cases, the energy integral will exist, as well as the center of mass and area integrals. In the three-body problem, one reduces the order of the system from 18 to 6 with the help of those integrals ([1]). In general, one can eliminate a canonical pair with the help of a known integral, so one will reduce the number of variables by two each time.

Therefore, let an integral be known:

$$G(p_k, q_k) = \text{const.} = g\,.$$

The problem is then to succeed in making a pair – e.g., P_1, Q_1 – drop out of the Hamiltonian integral:

$$\int_{t_1}^{t_2} \sum_k (P_k \dot{Q}_k - K)\,dt = \text{extremum}$$

by transforming to suitable new variables. That will obviously not be achieved when one demands that the new variable must be:

$$P_1 = G(p_k, q_k) = g\,. \tag{1}$$

That is because P_1 would be constant then, so $\dot{P}_1 = 0$ would be precisely an integral of the transformed problem, and since:

$$\dot{P}_1 = -\frac{\partial K}{\partial Q_1} = 0\,, \qquad \dot{Q}_1 = \frac{\partial K}{\partial P_1}$$

Q_1 would then have to drop out of K, while P_1 would no longer play the role of a constant parameter. The variables Q_l, P_l ($l = 2, \ldots, f$) would then define a canonical system with the Hamiltonian function K.

Now, from equation [(5), no. **3**], in order for (1) to be true, the transformation function V that is supposed to generate the desired canonical transformation shall satisfy the condition:

$$P_1 = -\frac{\partial V}{\partial Q_1} = G\left(\frac{\partial V}{\partial q_1}, q_k\right). \tag{2}$$

That is a partial differential equation that possesses corresponding integrals, and that will show the possibility of making the reduction. One can even proceed with that reduction without actually needing to look for a solution to the partial differential equation. If one has, in fact, first determined V according to (2) then Q_1 will drop out of K automatically under the corresponding canonical transformation. For the purpose of that transformation, one can then assign any arbitrary value to Q_1, and in particular, the value zero, and one must arrive at the correct function. Therefore, one does not at all need to know how the function V depends upon Q_1 beforehand. Rather, it is sufficient to possess its values $V(q_k, 0, Q_2, \ldots, Q_f)$ for $Q_1 = 0$. However, that is entirely arbitrary since from

([1]) Cf., Chap. 7, nos. **24**, **27**, and **28** of this volume of the *Handbuch*.

the existence theory for partial differential equations, one can always give an integral of (2) that will go to an arbitrarily-given function $V(q_k, Q_2, \ldots, Q_f)$ when $Q_1 = 0$.

We can then proceed as follows: We take a function $V(q_k, Q_2, \ldots, Q_f)$ of $2f - 1$ variables $q_1, \ldots, q_f, Q_2, \ldots, Q_f$, and possibly t that is arbitrary, except for a restriction that shall be given shortly, and then express the p_k as functions of the q_k and Q_k by means of the equations:

$$p_k = \frac{\partial V}{\partial q_k} = p_k(q_1, \ldots, q_f, Q_2, \ldots, Q_f) \,. \tag{3}$$

We substitute that value in the auxiliary condition (1) such that we will get:

$$G\left(q_1, \ldots, q_f, \frac{\partial V}{\partial q_1}, \ldots, \frac{\partial V}{\partial q_f}\right) = G(q_1, \ldots, q_f, Q_2, \ldots, Q_f) = g = P_1 \,. \tag{4}$$

We shall take that equation in place of $P_1 = \partial V / \partial Q_1$, which is permissible, from the consideration of limits above. If we now set:

$$P_l = -\frac{\partial V}{\partial Q_l} = P_l(q_1, \ldots, q_f, Q_2, \ldots, Q_f) \qquad (l = 2, \ldots, f) \tag{5}$$

then (3), (4), and (5) collectively define the formulas for the transformation of the p, q into P, Q. In that way, V is subject to only the restriction that equations (3) and (4) must be soluble for the q_k. The new Hamiltonian function will then be:

$$K = H + \frac{\partial V}{\partial t},$$

as usual, and will not include the variable Q_1, but $P_1 = g$, which is considered to be a constant parameter.

The simplest special case is again that of cyclic coordinates. If, say, q_1 is cyclic then it will not appear in L, and therefore not in H, either, while $\dot{q}_1$ (p_1, resp.) probably will. The integral will then be:

$$\frac{\partial L}{\partial \dot{q}_1} = p_1 = \text{const.} = c \,,$$

and the canonical problem will already have the form that we seek. We can then simply suppress p_1 and q_1, such that we will arrive at the variational problem:

$$\int_A^B \left\{ \sum_l p_l\, \dot{q} - K(p_l, q_l, c) \right\} dt = \text{extremum} \qquad (l = 2, \ldots, f),$$

in which we have set $H(p_1, p_l, q_l) = K(c, p_l, q_l)$. The entire process in this section then means just that one can make a variable cyclic with the help of an integral.

11. The connection between the various integral principles. – The argument that was just presented also allows one to clarify the connection between the different integral principles in a very instructive way when one regards the energy equation as an auxiliary condition. Those considerations, which borrow very heavily from the previous chapter, shall be justified here, since this is the first point at which the necessary mathematical tools have been available.

First of all, we must get back to the Hamiltonian variational problem from the canonical one. We therefore assume that we have eliminated the auxiliary conditions in the latter by introducing cyclic variables, as in the previous section, and we now apply the Legendre transformation in equation (8.b) of no. **2**. In that way, the new Lagrangian function (let q_1 be cyclic) will be:

$$L^* = \sum_l p_l \frac{\partial K}{\partial p_l} - K \qquad (l = 2, \ldots, l).$$

On the other hand, one had:

$$L = p_1 \frac{\partial H}{\partial p_1} + \sum_l p_l \frac{\partial H}{\partial p_l} - H \equiv \sum_l p_l \frac{\partial K}{\partial p_l} + c\,\dot{q}_1 - K .$$

One will then have:

$$L^* = L - c\,\dot{q}_1 ,$$

and the variational problem will next take the form:

$$\int_A^B \{L(q_l, \dot{q}_l) - c\,\dot{q}_l\}\,dt = \text{extremum}. \tag{1}$$

The quantity q_1 in this, which does not itself appear at all, in contrast to the other coordinates, is no longer subject to any boundary conditions, and therefore $\dot{q}_1$ will be a completely-arbitrary function. For that reason, the problem can be regarded as one that no longer includes one unknown $\dot{q}_1$ whose derivative will not appear and whose corresponding Lagrange equation will then read:

$$\frac{\partial L}{\partial \dot{q}_1} - c = 0 , \tag{2}$$

while the remaining Lagrange equations will not change, so they will give the same extremals. Since (2) must always be fulfilled, one can also require that relation as an auxiliary condition and then treat things just as one did in no. **2**. Obviously, that will imply that (1) is equivalent to the relative minimum principle:

$$\int_A^B \left\{ L - \frac{\partial L}{\partial \dot{q}_1} \dot{q}_1 \right\} dt = \text{extremum} \tag{3}$$

with equation (2) as the auxiliary condition.

Finally, one can eliminate $\dot{q}_1$ completely by solving (2) for $\dot{q}_1$ and substituting that in (1). One will then, in fact, get a simple minimal principle again:

$$\int_A^B F(c, p_l, q_l)\, dt = \text{extremum} \qquad (l = 2, \ldots, f), \tag{3.a}$$

but with one less desired function.

As was said before, we employ that argument in order to *go from the Hamiltonian principle to the remaining integral principles by applying the law of energy to it*. That process is generally valid for only conservative systems. In that case, t is itself cyclic since it does not appear in the kinetic potential. In order to be able to apply the method above, as before (no. **4**), we must introduce a parametric representation that gives t the same status as the remaining variables. If we assume that all quantities are functions of an auxiliary parameter τ:

$$t = t(\tau), \qquad q_k = q_k(\tau),$$

in such a way that $l(\tau_1) = l_1$, $l(\tau_2) = l_2$, and if we denote the derivative with respect to t by a prime then we will have:

$$\dot{q}_k = \frac{q_k'}{t'},$$

and therefore the kinetic energy T, which we assume to be a homogeneous quadratic function of the $\dot{q}_k$, will be:

$$T(\dot{q}_k) = \frac{1}{t'^2} T(q_k').$$

Hamilton's principle will then go to:

$$\int_{\tau_1}^{\tau_2} \left\{ \frac{1}{t'} T(q_k') - U(q_k)\, t' \right\} d\tau = \text{extremum},$$

in which the boundary condition that one requires is that the q_k and t go to well-defined values $q_k^{(1)}$ and $t^{(1)}$ ($q_k^{(2)}$, $t^{(2)}$, resp.) for $\tau = \tau_1$ ($\tau = \tau_2$, resp.). t is no longer distinguished now, and we can then apply the previous arguments. t enters in place of q_1, and τ, in place of t, while:

$$L = \frac{1}{t'} T - U\, t' .$$

One integral of that variational problem will be:

$$\frac{\partial L}{\partial t'} = -\frac{1}{t'^2} T(q_k') - U = -E , \tag{4}$$

so it is naturally the energy integral. With its help, one will get the form (1) for the equivalent to Hamilton's principle, which reads:

$$\int_{\tau_1}^{\tau_2} \left\{ \frac{1}{t'} T(q_k') - U\, t' + E\, t' \right\} d\tau = \text{extremum} \tag{5}$$

here, in which the boundary values of t are no longer prescribed then. If we again reintroduce t as a variable then that will make:

$$\int_A^B (T - U + E)\, dt = \text{extremum.} \tag{6}$$

That is a new principle of mechanics that is equivalent to Hamilton's that is probably still not known in the literature and shall be called ***Hilbert****'s principle*. It says:

A point system moves in such a way that of all motions that proceed with any temporal evolution from the initial location A with the coordinates $q_k = q_k^{(1)}$ *to the endpoint B with the coordinates* $q_k = q_k^{(2)}$, *the motion that actually occurs will make the integral* (6) *an extremum, where E is the value of the total energy that is given at the starting point*.

Naturally, the law of energy will follow from that principle since t does not appear explicitly in the integrand. However, it is not required as an auxiliary condition and accordingly stands in the middle between Hamilton's principle and the principle of least action.

Since E is constant, one can also write (6) as:

$$\int_A^B (T - U)\, dt + E\,(t_2 - t_1) = \text{extremum,}$$

in which $t_2 - t_1$ is the still-unknown time that the system needs along its path. One will then get back to Hamilton's principle when one gives the time $t_2 - t_1$ to the motion.

One will arrive at the principle of least action when one adds the law of energy $T + U = E$, which already follows from (6), as an auxiliary condition. One will then come to the form (3), which assumes the form:

$$2\int_A^B T\,dt = \text{extremum}, \qquad \text{while} \quad T + U = E,$$

due to (4), which is precisely the principle of least action (see Chap. 2, no. **25**). The extremum is sought from among all functions that go from the starting point to the endpoint in any length of time and satisfy the law of energy in so doing.

Finally, one can still eliminate t completely, so one will arrive at the form (3.a). In order to do that, one again employs the parametric representation appropriately. However, that is precisely the process that led to Jacobi's principle in Chapter 2, no. **26**, which can also be classified by this argument then.

12. The Hamilton-Jacobi partial differential equation. – We shall not turn to the theory of integration for the canonical equations of motion:

$$H = H(q_k, p_k, t), \qquad \dot{q}_k = \frac{\partial H}{\partial p_k}, \qquad \dot{p}_k = -\frac{\partial H}{\partial q_k}. \tag{1}$$

We have encountered fragments of such a thing many times already (in nos. **2**, **7**, **9**, and **10**), but the most important piece is still missing: a systematic process that will be described in what follows. A thorough use of the canonical transformations will be made in it.

From [(5), no. **3**], under a canonical transformation of the problem (1), the new Hamiltonian function will become:

$$K = H + \frac{\partial V}{\partial t}.$$

We ask whether it is possible to arrange that *the new **Hamiltonian** function K of the system will vanish* by a suitable choice of the function V. The mechanical problem will then be transformed into an equilibrium problem, in a certain sense. We would like to denote the function that makes that possible by R to distinguish it from the other generators.

Now, R should be a function of the q_k, Q_k, and t, and one will have:

$$p_k = \frac{\partial R}{\partial q_k}, \qquad P_k = -\frac{\partial R}{\partial Q_k}, \quad K = H + \frac{\partial R}{\partial t}. \tag{2}$$

The condition that R must fulfill in order for K to vanish will then read:

$$\frac{\partial}{\partial t} R(q_k, Q_k, t) + H(q_k, Q_k, t) = 0,$$

or from (2):

$$\frac{\partial R}{\partial t} + H\left(q_k, \frac{\partial R}{\partial q_k}, t\right) = 0 . \tag{3}$$

This is a *first-order partial differential equation* for R that was first found by **Hamilton**. It will arise when one replaces the p_k in the Hamiltonian function H with the derivatives of R with respect to the corresponding q_k . Since (3) must be true for all arbitrary values of the Q_k, they will play the role of integration constants.

The meaning of the partial differential equation (3) lies in the following: We assume that we have found an integral of (3) that includes f arbitrary constants $\alpha_1, \ldots, \alpha_f$:

$$R\,(q_1, \ldots, q_f, \alpha_1, \ldots, \alpha_f, t) = 0 ,$$

so a function that satisfies the differential equation for all values of the integration constants. Naturally, that is not the most general solution of the partial differential equation, which must indeed include an arbitrary function, but only a so-called *complete integral*. We can then introduce those constants α_k as new variables Q_k since R should indeed be a function of the old and new configuration parameters. The transformation formulas [(5), no. **3**] will then yield:

$$\left.\begin{aligned} p_k &= \frac{\partial R}{\partial q_k}, \\ P_k &= -\frac{\partial R}{\partial q_k} = +\beta_k, \\ K &= 0, \end{aligned}\right\} \tag{4}$$

in this case, and as a result of the third row in that, the new canonical equations will become simply:

$$\frac{dQ_k}{dt} = \frac{d\alpha_k}{dt} = 0 , \qquad \frac{dP_k}{dt} = \frac{d\beta_k}{dt} = 0 .$$

Thus, the α_k, as well as the β_k, are constant quantities for the mechanical system that can be assigned arbitrary values. They are called the *canonically-conjugate constants*. With that, the integration of the differential equations of the mechanical problem is complete since equations (4) will imply the original coordinates of the system as functions of time and the $2f$ arbitrary constants α_k and β_k.

The integration of the canonical equations is then reduced to discovery of an integral of the partial differential equation (3) that includes f constants. Initially, it would seem that we have not achieved very much since partial differential equations are harder to treat than ordinary ones, as a rule. However, in mechanics, it is shown that the partial differential equation will assume relatively-simple forms in many important cases such that its introduction would actually imply a significant advance [1].

Only one step still remains to be completed here: *If the Hamiltonian function H does not include time explicitly then the differential equation* (3) *can be simplified somewhat.* If we make the following Ansatz for R :

$$R = S(q_k, \alpha_1, \ldots, \alpha_f) - \alpha_1 t, \tag{5}$$

in which S should no longer depend upon t, and if we introduce that Ansatz into (3) then that will imply that:

$$\alpha_1 = H\left(q_k, \frac{\partial S}{\partial q_k}\right) = W, \tag{6}$$

from which time t has been eliminated. α_1 will then become the energy constant, and as such, it will be denoted by W. Now, if we have found an integral S of the partial differential equation (6) that depends upon not only α_1, but also $f - 1$ further independent constants, then the solutions of the equations of motion:

$$p_k = \frac{\partial S}{\partial q_k}, \qquad \beta_l = -\frac{\partial S}{\partial \alpha_l}, \qquad t - \beta_1 = \frac{\partial S}{\partial \alpha_1} \qquad (l = 2, \ldots, f). \tag{7}$$

Equations (3) and (6) are the simplest forms of the Hamiltonian partial differential equation, while formulas (4) and (7) include the solutions of the problem of motion in the most-transparent form. However, many variations of the process that was described will be applied in practice. Thus, in place of (3), one can also demand that the new Hamiltonian function K must be an arbitrary function of time $f(t)$, instead of vanishing. In order to do that, one must take the solution of the differential equation:

$$\frac{\partial T}{\partial t} + H\left(q_k, \frac{\partial S}{\partial q_k}\right) = f(t) \tag{8}$$

to be the generator of the canonical transformation. R will then be coupled with T by the relation:

$$R = T - \int f(t)\, dt\,. \tag{9}$$

For example, one can demand that:

$$f(t) = \text{const.} = \alpha_1\,. \tag{9.a}$$

(When one is dealing with a small perturbation of an otherwise-closed system that comes from outside of it, that is closely related to saying that the total energy would be constant in the absence of that perturbation.) Equations (8) and (9) will then imply that:

---

[1] On this topic, see the following Chap. 4 on perturbation theory.

$$\frac{\partial T}{\partial t} + H\left(q_k, \frac{\partial S}{\partial q_k}\right) = \alpha_1 , \tag{8.a}$$

$$R = T - \alpha_1 t . \tag{9.a}$$

In particular, if H does not depend upon t explicitly then one can assume that T is also independent of t and in that way, get back to (6) [(5), resp.].

Moreover, in the case of a closed system, it is indeed simplest, but not always convenient, to choose the energy constant itself to be one of the integration constants of the complete integral S. On the grounds of normalization, in the theory of constrained periodic systems (cf., Chap. 4) and its applications to quantum theory, other integration constants will be chosen (we would like to call them J_k), in terms of which, the new Hamiltonian function will be written:

$$\alpha_1 = K(J_1, \ldots, J_f) . \tag{10}$$

However, one can easily transform the variables α_k, β_k into the new constants J_k and their canonically-conjugate variables w_k with one of the generators of the form $V = \sum_k \alpha_k(J_1, \ldots, J_f)\,\beta_k$. Due to (10) and $\dot{w}_k = \partial K / \partial J_k =$ const., the latter are linear functions of time.

In all cases, the viewpoint that one must take in order to exhibit the Hamiltonian partial differential equation must remain that one must transform to new variables that represent a family of constants of the motion whose conjugate family does not occur in K then. In other words: One seeks a generator of a canonical transformation to *cyclic* variables, and that search will lead one to just the Hamiltonian partial differential equation. As an aside, it should be remarked that the form (1) of the Hamiltonian differential equation corresponds completely to the form (6), at least formally, when one treats time as also being a canonical variable, as one did in no. **4**.

13. The simplest cases of integration. – The solution of the problem of motion [(1), no. **2**] has now been reduced to the integration of the partial differential equation [(3) or (6), no. **12**]. It is a *complete* integral of it that is equipped with f integration constants α_k that one must seek. A process that will always lead to that goal cannot be given. Only two simple cases of the treatment that was given in [(6), no. **12**] will be discussed here.

The first case that admits a simple integration will occur when all of the variables are cyclic, with the exception of one of them (q_1). One will then know $f - 1$ first integrals:

$$p_k = \frac{\partial S}{\partial q_k} = \alpha_k \qquad (k = 2, \ldots, f)$$

and find that:

$$S = \sum_{k=2}^{f} \alpha_k\, q_k + S_1(q_1, \alpha_1, \alpha_2, \ldots, \alpha_f) .$$

Since H is independent of the cyclic variables $q_2, \ldots, q_f$, the differential equation [(6), no. **12**] will reduce to an ordinary one:

$$H\left(\frac{\partial S_1}{\partial q_1}, q_1, \alpha_1, \ldots, \alpha_f\right) = W = \alpha_1 ,$$

from which, S_1 can then be obtained by a quadrature.

The other case that admits a simple integration will occur when the variables p_k, q_k in the differential equation [(6), no. **12**] can be *separated*. That means that with the Ansatz:

$$S = \sum_k S_k(q_k, \alpha_1, \ldots, \alpha_f) ,$$

$$p_k = \frac{\partial S}{\partial q_k} = \frac{\partial S_k(q_k)}{\partial q_k}$$

(i.e., when S is regarded as a sum of functions that each depend upon just one of the coordinates q_k), the differential equation [(6), no. **12**] will decompose into f different differential equations for the S_k. In order to do that, it is necessary for every impulse p_k in:

$$H(p_1, \ldots, p_f, q_1, \ldots, q_f) = W$$

to be expressible as a function of the associated coordinate q_k alone, so that equation will split into f individual ones:

$$H_k(p_k, q_k) = A_k(\alpha_1, \alpha_2, \ldots, \alpha_f) .$$

The f different differential equations for the S_k will then read:

$$H_k\left(\frac{\partial S_k}{\partial q_k}, q_k\right) = A_k .$$

That will allow one to calculate the S_k by mere quadratures.

According to **Levi-Civita** ([1]), the condition for H to be separable in the coordinates being used can be written:

([1]) **T. Levi-Civita**, Math. Ann. **59** (1904), pp. 383; **F. A. Dall'Aqua**, *ibidem*, **66** (1908), pp. 398; **H. Kneser**, *ibidem* **84** (1921), pp. 277.

$$\begin{vmatrix} 0 & \dfrac{\partial H}{\partial q_j} & \dfrac{\partial H}{\partial p_j} \\ \dfrac{\partial H}{\partial q_k} & \dfrac{\partial^2 H}{\partial q_j \, \partial q_k} & \dfrac{\partial^2 H}{\partial p_j \, \partial q_k} \\ \dfrac{\partial H}{\partial p_k} & \dfrac{\partial^2 H}{\partial q_j \, \partial p_k} & \dfrac{\partial^2 H}{\partial p_j \, \partial p_k} \end{vmatrix} = 0 \quad \text{for} \quad \begin{cases} j, k = 1, 2, \dots, f \\ j \neq k. \end{cases}$$

Of course, it is mostly the separability of the function *H* that one must consider. It depends upon the coordinate system and generally requires the introduction of special separation coordinates in order to achieve the desired splitting. In many cases, the separation system is distinguished physically by the boundary of the domain of the orbits. However, that is not always the case ([1]). In fact, **Burgers** has shown ([2]) that the separation system for the motion of an electrically-charged oscillator in a magnetic field can be introduced only by a contact transformation.

Some examples of integration by separation include, among others, any central motion [as can be seen from (15), no. **2**], as well as the two-center problem, which is separable in elliptic coordinates with the two fixed centers as focal points, as **Jacobi** showed before ([3]). Moreover, **Weinacht** ([4]) succeeded in finding all systems that are separable by a point transformation for the case of a single mass-point in a conservative force field. The most-important result is that the most-general configuration coordinates that come under consideration for the separation of variables in this case are those of the three-axis ellipsoid (including its degeneracies). The associated functions for the potential energy can also be given and are obvious generalization of the aforementioned cases. Moreover, every small oscillation of an arbitrarily-composed system about a stable equilibrium configuration will admit separation by the method of eigen-oscillations. For the motion of a rigid body, the separable cases are the most-general *force-free top* (possibly with a built-in flywheel) and the *symmetric top in a gravitational field* ([5]).

14. The independence theorem of the calculus of variations. The eikonal. – In order to conclude this section on Hamilton-Jacobi mechanics, we would like to try to give a glimpse of the deep lines of reasoning that guided the creators of that theory and which have led to a fundamental advancement of mechanics in recent times by the work of **de Broglie**, **Schrödinger**, *et al.* In order to actually understand the true essence of Hamilton-Jacobi theory, it is necessary to draw upon

---

([1]) **E. Fues**, Zeit. Phys. **34** (1925), pp. 788.

([2]) **J. M. Burgers**, *Het Atoommodel van Rutherford-Bohr*, Leiden, 1918.

([3]) **W. Pauli, Jr**. gave a precise discussion and application to the H_2^+-molecule in Ann. Phys. (Leipzig) **68** (1922), pp. 177 and **K. F. Niessen**, Diss. Utrecht, 1922. A special case of the separation into elliptical coordinates is defined by the treatment of the Stark effect in parabolic coordinates by **Schwarzschild**, Berl. Ber. (1916), pp. 348 and **P. S. Epstein**, Ann. Phys. (Leipzig) **30** (1916), pp. 489.

([4]) **J. Weinacht**, Math. Ann. **91** (1924), pp. 279.

([5]) Cf., **G. Kolossoff**, Math. Ann. **60** (1905), pp. 232; **F. Reichs**, Phys. Zeit. **19** (1918), pp. 394.; **P. S. Epstein**, Verh. d. D. Phys. Ges. **17** (1916), pp. 398; Phys. Zeit. **20** (1919), pp. 289; **H. A. Kramers**, Zeit. Phys. **13** (1923), pp. 343.

some theorems of the calculus of variations again. In order to do that we start with the form (4) of no. **2** for the variational problem:

$$\int_{t_1}^{t_2} \left\{ L + \sum_k \frac{\partial L}{\partial k_k} (\dot{q}_k - k_k) \right\} dt = \text{extremum.} \tag{1}$$

The integral has the simple form:

$$J = \int_{t_1}^{t_2} \left(A + \sum_k B_k \frac{dq_k}{dt} \right) dt \tag{2}$$

here with:

$$A = L - \sum_k \frac{\partial L}{\partial k_k} k_k\,, \qquad B_k = \frac{\partial L}{\partial k_k}\,.$$

The integrand is then a linear expression in the derivatives $\dot{q}_k$ of the q_k. Along with it, the functions k_k will appear, which are varied independently of the q_k, but not their derivatives. That form recalls the complete derivative of a function Φ with respect to time:

$$\sum \frac{\partial \Phi}{\partial q_k} \dot{q}_k + \frac{\partial \Phi}{\partial t}\,.$$

That is closely related to the question of whether it is not possible to make the integral (2) independent of the path in qt-space by a special choice of the k_k as functions of q_k and t in such a way that the integral will take the same value for *all* possible functions $q_k(t)$, so it will degenerate from a function of a function, in the sense of the calculus of variations, to a pure function of the locations of the integration limits. The values of the k_k then define a covering of qt-space in such a way that every point will be associated with a well-defined value of k_k. One calls such a covering a *field*, and the question becomes that of whether there are coverings for which the integral (2) is independent of the path. A necessary and sufficient condition for that is that the B_k and A must take the form of partial derivatives of the function $\Phi(q_k, t)$:

$$A = \frac{\partial \Phi}{\partial t}\,, \qquad B_k = \frac{\partial \Phi}{\partial q_k}\,.$$

The integral:

$$J = \int_{t_1}^{t_2} \left(A + \sum_k B_k \dot{q}_k \right) dt = \int_{t_1}^{t_2} \left(\frac{\partial \Phi}{\partial t} + \sum_k \frac{\partial \Phi}{\partial q_k} \dot{q}_k \right) dt = \Phi(t_2, q_k^{(2)}) - \Phi(t_1, q_k^{(1)})$$

will then become a pure function of the limits of integration in qt-space. In order for that to be true, the A and B_k must fulfill the integrability conditions:

$$\frac{\partial A}{\partial q_k} = \frac{\partial B_k}{\partial t}, \qquad \frac{\partial B_k}{\partial q_l} = \frac{\partial B_l}{\partial q_k}.$$

The general answer to the question of how one must choose the k-field in order for those conditions to be satisfied is given by ***Hilbert****'s independence theorem:*

The integral (2) *will become independent of path when one takes any system of intermediate integrals:*

$$\frac{dq_k}{dt} = \dot{q}_k(q_1, \ldots, q_f, t)$$

of the Lagrange differential equations:

$$[L]_{q_k} = 0 \tag{3}$$

and chooses the k_k *to be equal to the corresponding* $\dot{q}_k$ *for each point* $q_1, \ldots, q_f, t$.

We shall prove that theorem here for only a system with a single degree of freedom, i.e., only one pair p, q [k, resp.]. Only a single integrability condition will exist then, namely:

$$\frac{\partial}{\partial q}\left(L - k\frac{\partial L}{\partial k}\right) = \frac{\partial}{\partial t}\left(\frac{\partial L}{\partial k}\right). \qquad (5)[sic]$$

If we perform the differentiations then we will get the condition for the independence of the integral (1) of a first-order partial differential equation for k (q, t) in the form of:

$$\frac{\partial L}{\partial q} + \frac{\partial L}{\partial k}\frac{\partial k}{\partial q} - \frac{\partial L}{\partial k}\frac{\partial k}{\partial q} - k\left(\frac{\partial^2 L}{\partial k\,\partial q} + \frac{\partial^2 L}{\partial k^2}\frac{\partial k}{\partial q}\right) = \frac{\partial^2 L}{\partial k\,\partial q} + \frac{\partial^2 L}{\partial k^2}\frac{\partial k}{\partial q}$$

or

$$\frac{\partial^2 L}{\partial k^2}\left(\frac{\partial k}{\partial t} + k\frac{\partial k}{\partial q}\right) + k\frac{\partial^2 L}{\partial k\,\partial q} + \frac{\partial^2 L}{\partial k\,\partial t} - \frac{\partial L}{\partial q} = 0\,, \tag{6}$$

which is called the *adjoint partial differential equation* to the variational problem. Now, that differential equation (that it, its statement) will be fulfilled if and only if k (q, t) is an intermediate integral of the Lagrange differential equation:

$$[L]_q \equiv \frac{\partial^2 L}{\partial \dot{q}^2}\ddot{q} + k\frac{\partial^2 L}{\partial \dot{q}\,\partial q}\dot{q} + \frac{\partial^2 L}{\partial \dot{q}\,\partial t} - \frac{\partial L}{\partial q} = 0\,. \tag{7}$$

If, if fact, $\dot{q} = k$ (q, t) is such an integral of (7), i.e., (7) is fulfilled identically, if one replaces q with the general solution:

$$q = q\,(t, \alpha) \tag{8}$$

of the differential equation $\dot{q} = k\,(q, t)$ that still includes the constant α then one will have:

$$\ddot{q} = \frac{\partial k}{\partial t} + \frac{\partial k}{\partial q}\dot{q}$$

identically in t and α. If we substitute that in (7):

$$\frac{\partial^2 L}{\partial q^2}\left(\frac{\partial k}{\partial t} + \frac{\partial k}{\partial q}\dot{q}\right) + \frac{\partial^2 L}{\partial \dot{q}\,\partial q}\dot{q} + \frac{\partial^2 L}{\partial \dot{q}\,\partial t} - \frac{\partial L}{\partial q} = 0\,,$$

and once more write k for $\dot{q}$ then we will get a relation that formally seems to be precisely the adjoint partial differential equation (6), but initially represents an ordinary differential equation in t and α that must be fulfilled identically for all values of t and α. However, if one introduces q in place of α by means of (8) then it must also be true identically in t and q, i.e., all intermediate integrals $\dot{q} = k\,(q, t)$ of the Lagrange differential equation must nonetheless satisfy the adjoint partial differential equation, as well.

If, conversely, $k\,(q, t)$ is a solution of the adjoint partial differential equation (6) and $q\,(t)$ satisfies the equation $\dot{q} = k\,(q, t)$ then we can substitute $\partial k / \partial t + k \cdot \partial k / \partial q = \ddot{q}$, and when we again write $\dot{q}$ for k, we can get back to the Lagrange differential equation, which proves our theorem completely. The theorem can be generalized for several degrees of freedom by reducing the problem to this special case ([1]).

The solutions to a variational problem, so the curves that satisfy the Lagrange differential equations, are ordinarily referred to as extremals. Thus, an independence field can always be exhibited with the help of an f-parameter family of extremals. In order to do that in the most-general possible way, so to associate every system of values $q_1, \ldots, q_f, t$ with a system of values $k_1, \ldots, k_f$ and in so doing to fulfill the condition of the independence integral, one proceeds as follows: We choose any completely-arbitrary function $F\,(q_k, t)$ that will represent an f-dimensional hypersurface in the space of q_k, t when it is set equal to zero:

$$F\,(q_1, \ldots, q_f, t) = 0\,, \tag{9}$$

and initially determine the k_k for all points of the surface from the demand that the integrand of the independence integral:

$$L + \sum_k \frac{\partial L}{\partial k_k}\,(\dot{q}_k - k_k)$$

must vanish for them. We will achieve that when we calculate the f quantities k_k from the f equations:

([1]) **D. Hilbert**, Math. Ann. **62** (1905), pp. 351.

$$\left(L - \sum_k \frac{\partial L}{\partial k_k} k_k\right) : \frac{\partial L}{\partial k_1} : \frac{\partial L}{\partial k_2} : \cdots : \frac{\partial L}{\partial k_f} = \frac{\partial F}{\partial t} : \frac{\partial F}{\partial q_1} : \frac{\partial F}{\partial q_2} : \cdots : \frac{\partial F}{\partial q_f} , \tag{10}$$

respectively, since the integrand will then be equal to:

$$\frac{\partial F}{\partial t} + \sum_k \frac{\partial F}{\partial q_k} \dot{q}_k = \frac{dF}{dt} ,$$

up to an irrelevant factor, so it will, in fact, vanish on the surface. We then let a curve $q_k = q_k(t)$ start from each point of the surface whose direction factors $\dot{q}_k$ are equal to precisely the k_k that were just determined and will satisfy the Lagrange differential equations (3) in its further evolution. That is always possible since we can indeed always find one such integral curve for an arbitrary second-order differential equation at a given point with a given direction. That means nothing besides the fact that we are taking precisely those integral curves that are *transversal* to the surface, which is a condition that is mostly identical to orthogonality in the usual sense.

Since the surface $F = 0$ is itself f-dimensional, we will then have determined an f-parameter family of curves that fill up precisely the $f + 1$-dimensional qt-space densely since in general precisely one curve will go through each point in space, except for possible singular points. We determine the values of the k_k at an arbitrary point as simply the direction of the tangent to the extremal that goes through it, so we set:

$$k_k = \dot{q}_k .$$

From the independence theorem, it is precisely that k-field that will turn the integral (2) into a pure function of position.

Now, we can see the meaning of the independence integral as follows: We imagine that all of the transversal surfaces have been drawn in the field, i.e., the surfaces F = const. that satisfy the conditions (10). When the integral J is extended between any two points of one such surface is obviously equal to zero. We now calculate it further for a path that leads from the starting point A of the actual motion to the endpoint B. Due to the independence of the path, we can choose that path as suitably as possible. We initially move forwards along the transversal surface on which the starting point lies up to the point C at which the extremal that also goes through the endpoint B meets the surface and then move further along that extremal. The first part AC makes no contribution to the integral. For the second part CB, we will have $k_k = \dot{q}_k$ everywhere, and J will then reduce to $\int_C^B L(q_k, \dot{q}_k, t)\, dt$ since the $\dot{q}_k = k_k(t)$ will indeed be determine in precisely such a way that they satisfy the Lagrange differential equations. Thus, J is nothing but the extremal value of the integral in Hamilton's principle between the two transversal surfaces that go through the starting and ending points. Since J vanishes for paths along those surfaces, they will also be surfaces of constant differences between the values of the Hamilton integral between corresponding points, i.e., points that lie on the same extremal. The quantity J, which is a function of the starting and ending points for a given extremal field, has great significance for many branches of mathematics and physics, and is usually called the *eikonal*, with the terminology that is customary in optics.

Naturally, there is a wide variety of eikonals since they depend upon an arbitrary function, namely, the starting surface $F = 0$. Among all possible starting surfaces, there is a special one that degenerates to a point, namely, the starting point of the path of integration. One will also get a field that covers all of space from it when one takes all extremals that go through it as the generator of the field. *The eikonal for a point that is reached from the starting point in the course of the motion is obviously equal to the extremal value of Hamilton's integral itself when taken over the true trajectory.*

15. Application to mechanics. The meaning of the Hamilton-Jacobi differential equation. – A partial differential equation for all possible eikonals can now be exhibited. The definition (2) in no. **14** implies immediately that the derivatives of J will be given by:

$$\left.\begin{aligned} \frac{\partial J}{\partial t} &= L - \sum_k \frac{\partial J}{\partial k_k} k_k , \\ \frac{\partial J}{\partial q_k} &= \frac{\partial J}{\partial k_k}. \end{aligned}\right\} \tag{1}$$

The right-hand sides are still functions of the k_k, so of the chosen field. However, precisely the f quantities k_k can be eliminated from those $f + 1$ relations, and what will remain then will be one condition between the derivatives of J, so a partial differential equation. That elimination can be carried out with no further analysis by using the Legendre transformation, so with a transition to canonical coordinates. Indeed, in (5) and (7), no. **2**, we had set:

$$p_k = \frac{\partial L}{\partial k_k} = \frac{\partial L}{\partial \dot{q}_k} ,$$

$$H = \sum_k \frac{\partial L}{\partial \dot{q}_k} \dot{q}_k - L ,$$

and upon eliminating p_k from (1), we will then get:

$$\frac{\partial J}{\partial t} + H\left(q_k , \frac{\partial J}{\partial q_k} , t \right) = 0 \tag{2}$$

as the partial differential equation for the eikonal. However, that is exactly the Hamilton-Jacobi differential equation (3) of no. **12**. *The interpretation of the integral of the Hamilton-Jacobi differential equation as the value of the Hamiltonian integral between the transversal surfaces of the field is revealed by that fundamental connection.*

The main theorem of no. **12** can be given a new derivation with the help of that result. We imagine that *f* parameters α_k are introduced into the defining equation [(9), no. **14**] for the starting surface such that we will have, in total, an *f*-parameter family of surfaces, one of which is our starting surface. For every other surface in that family, there is likewise an independence field that is determined by our construction such that we will also have an *f*-parameter family of such fields. That is, we take our definition of the field to be a family of intermediate integrals of the Lagrange equations that already include *f* integration constants:

$$k_k = \dot{q}_k(q_l, \alpha_l, t).$$

An eikonal will then belong to each system of values for the α_k, and the set of all those eikonals can obviously be combined into a single function $J(\alpha_k)$ that depends upon not only the starting and ending points, but also *f* parameters:

$$J = \int_A^B \left\{ L(q_l, k_l(q_l, \alpha_l, t), t) + \sum_k \frac{\partial L}{\partial k_k}(\dot{q}_k - k_k) \right\} dt.$$

However, along with *J*, its derivatives with respect to the parameters α_k must also be pure functions of position, and due to the fact that:

$$\frac{\partial L}{\partial \alpha_l} = \sum_k \frac{\partial L}{\partial k_k}\frac{\partial k_k}{\partial \alpha_l},$$

that will give simply:

$$\frac{\partial J}{\partial \alpha_l} = \int_A^B \sum_k (\dot{q}_k - k_k)\frac{\partial^2 L}{\partial k_k\,\partial \alpha_k} dt. \tag{3}$$

The integral on the right-hand side will vanish for an advance along the integral curves themselves since one will indeed always have $\dot{q}_k = k_k$ for them, i.e., the $\partial J / \partial \alpha_l$ represent functions of the q_k and *t* that are themselves constant along the integral curves. Therefore, *when they are set equal to constants* $-\beta_l$:

$$\frac{\partial J}{\partial \alpha_l} = -\beta_l, \tag{4}$$

they must be themselves integrals of the Lagrange differential equations, which was to be proved.

One will likewise get an important mechanical theorem from the converse of that theorem. *If one knows half of the integrals of a mechanical system then one can find the other half by mere quadratures*. In fact, let *f* functions be known:

$$\varphi_l(\dot{q}_k, q_k, t, \alpha_1, \ldots, \alpha_f) = 0, \qquad [\varphi_l(p_k, q_k, t, \alpha_1, \ldots, \alpha_f) = 0, \text{ resp.}]$$

then one can find the $\dot{q}_k$ as functions of q_k, t, and the first f integral constants α_f by solving them, so one can also find an f-parameter extremal field:

$$k_k = \dot{q}_k\,(q_l, t, \alpha_l) \ .$$

If one substitutes those values in the integral (1) of no. **14** then from what was said before, the integrand will be a complete differential. The associated eikonal can then be calculated by quadratures and one will then have the remaining integrals of motion in (4). If one employs the canonical form of the differential equations (so one has not found the integrals in the form of the $\dot{q}_k$ as functions of the q_k, t, and α_f, but the p_k) then one will not need to first calculate the $\dot{q}_k$, but to transform the integral J into p_k and q_k directly by means of the Legendre transformation. With (5) and (7) of no. **2**, one will find immediately that:

$$J = \int_A^B \left(\sum_k p_k\,\dot{q}_k - H \right) dt \ . \tag{5}$$

By assumption:

$$dJ = \sum_k p_k\,dq_k - H\,dt \tag{6}$$

will be a complete differential.

From the theorem that was just proved, one has that, e.g., every mechanical problem with one degree of freedom will be soluble by quadratures when it possesses the energy integral, and every problem with two degrees of freedom will be soluble by quadratures when yet another integral is known, in addition to the energy integral.

There is also a simple meaning to the integral S of the Hamilton-Jacobi partial differential equation (6) of no. **12**, when it was already integrated over time, for systems that do not include time explicitly. Namely, it is precisely the extremal value of the integral of the principle of least action, so the *action function*, and thus the integral of Jacobi's principle, which is identical to it for conservative systems. Since we have assumed the law of energy, we have:

$$2T = T - U + T + U = T - U + \alpha_1 \ ,$$

in which α_1 is the energy constant. As a result, from (5) of no. **12**, we will have:

$$2\int_A^B T\,dt \ = \ \int_A^B (T - U)\,dt + \alpha_1\,t \ = J + \alpha_1\,t = S \ , \tag{7}$$

i.e., *S has the same relationship to the principle of least action in its Jacobi form that J has to Hamilton's principle*.

The considerations of this section show that *the integration of a partial differential equation of the Hamilton-Jacobi form* (which is no essential loss of generality) *is equivalent to the integration of the corresponding canonical equations*. It is nothing but Jacobi's method of integrating first-order partial differential equations, and the extremal curves of Hamilton's principle, *so the mechanical trajectories, represent the characteristics of the partial differential equation.* In fact, when the canonical equations have been solved, so all extremals have been found, for every function $F(q_k, t) = 0$, one can find a solution of the partial differential equation that will go to F for $t = t_1$, $q_k = q_k^{(1)}$. However, as was said before, one mostly proceeds in the opposite direction by integrating the Lagrange or canonical equations with the help of the integrals of the partial differential equation (2).

That was the starting point that led **Jacobi** to his theory. The other discoverer of that connection, viz., **Hamilton**, started from the geometric meaning of the eikonal, which is, in fact, most remarkable. Namely, if we go from the representation of the eikonal in no. **14** (i.e., its description in qt-space) to a construction in f-dimensional q-space alone then we will get a system of moving eikonal surfaces $J(q_k, t) = c$, and its extremals (path curves), which are general found in a state of flux, will be its trajectories. The latter will be fixed in the aforementioned case [viz., equation (7)] of a time-independent Hamiltonian function. From the fact that $J = S - W_1 t$, the eikonal surfaces will then depart from the fixed surfaces S = const. in such a way that they will always come to cover a new S-surface again. The picture is that of the propagation of a wave, as one is probably accustomed to imagining it for optical processes.

If we take the starting surface $F = 0$ to be the generating surface of an optical process then the extremals will be the light rays in the sense of geometrical optics, and the advancing eikonal surfaces will be surfaces of equal phase, so a type of wave surface in the sense of Huygens's principle. The principle of least action will then correspond precisely to Fermat's principle of shortest light-path, when we assume that the index of refraction in q-space is proportional to the square root of the kinetic energy, which is equal to $W - U$, so it is also equal to a pure function of position. With that, the solution of the mechanical problem is reduced to the corresponding optical one. *The Hamilton-Jacobi theory then corresponds to geometrical optics*. Those considerations have recently become the foundation for the further development of quantum theorem by **Schrödinger** ([1]), which is based upon the idea that for the mechanics of atoms, the mechanics that is equivalent to ray optics will not suffice, but one must use an extension of it in the sense of true wave theory as a foundation ([2]).

([1]) **E. Schrödinger**, *Abhandlungen zur Wellenmechanik*, Leipzig, 1927.

([2]) For a more detailed analysis of this connection, which can only be touched upon briefly here, see the article "Optik und Mechanik," by **A. Landé** in Bd. XX of this *Handbuch*.